ISSN: 2301-5373
E-ISSN: 2654-5101
Volume 14 • Number 3 • February 2026

JELIKU

Jurnal Elektronik Ilmu Komputer Udayana

Informatics Study Program
Faculty of Mathematics and Natural Sciences
Udayana University

Table of Contents

Klasifikasi Mood Musik Menggunakan Support Vector Machine Berdasarkan Fitur Spektral
I Gusti Agung Ngurah Diputra Wiraguna, Luh Arida Ayu Rahning Putri, I Ketut Gede Suhartana, Gst. Ayu Vida Mastrika Giri..... 355-364

Sistem Representasi Digital Wayang Kulit Bali Menggunakan AHP Dengan Pendekatan Ontologi
I Komang Widia Pratama, Made Agung Raharja, Cokorda Pramartha, I Ketut Gede Suhartana 365-374

Klasifikasi Kategori Cerita Pendek Menggunakan XGBoost dengan Seleksi Fitur Chi-Square
M. Faisal Afandi, Ngurah Agus Sanjaya ER, Putu Gede Hendra Suputra, Luh Arida Ayu Rahning Putri 375-384

Redesign UI/UX Aplikasi Property Management Berbasis User-Centered Design dengan Evaluasi Heuristic Evaluation dan System Usability Scale
Matthew Novan Sidharta, Luh Arida Ayu Rahning Putri, I Ketut Gede Suhartana, I Gusti Ngurah Anom Cahyadi Putra..... 385-396

Optimasi Algoritma SVM dengan Ekstraksi Fitur Warna pada Klasifikasi Biji Kopi Sangrai
Gde Krishna Sankya Yogeswara, I Ketut Gede Suhartana, I Gusti Ngurah Anom Cahyadi Putra, I Wayan Supriana..... 397-406

Rancangan Sistem Monitoring Suhu, Kelembaban, Kecepatan Angin Untuk Memprediksi Keamanan Jalur Pendakian
I Wayan Wikananda Adikara, Gst. Ayu Vida Mastrika Giri, I Wayan Supriana, I Gede Surya Rahayuda..... 397-406

Sistem Informasi Manajemen Keuangan Pribadi Dengan Input Suara Dan Berbasis IoT
I Dewa Gede Partha Wijaya, I Made Widiartha, I Gede Santi Astawa, Gst. Ayu Vida Mastrika Giri..... 419-428

Perbandingan Performa CNN Inception-V3 dan ResNet50 Dalam Klasifkasi Kematangan Buah Tomat
I Putu Hanggara Diatha Putra, I Wayan Supriana, I Ketut Gede Suhartana, I Gusti Ngurah Anom Cahyadi Putra, Farhan Akmala Putra 429-438

Penentuan Entitas Tokoh Pada Satua Bali Menggunakan Algoritma Conditional Random Fields
I Made Widi Arsa Ari Saputra, I Ketut Gede Suhartana, I Gusti Ngurah Anom Cahyadi Putra, I Wayan Supriana 439-448

Implementasi Intrusion Prevention System (IPS) Menggunakan Signature – Based Detection Berbasis Suricata
I Wayan Dimas Wirahadi, I Gede Santi Astawa, Made Agung Raharja, I Made Widiartha 449-458

Integrasi Algoritma RSA dan Teknologi Kriptografi Kuantum dalam Keamanan Aplikasi Web Chatting
Albert Okario, I Putu Gede Hendra Suputra, Gst. Ayu Vida Mastrika Giri, I Made Widhi Wirawan 459-468

Analisis Sentimen Terhadap Pembatalan Piala Dunia U-20 di Indonesia Menggunakan Algorima Naïve Bayes dan Genetika
I Ketut Oning Pusparama, I Putu Gede Hendra Suputra, Gst. Ayu Vida Mastrika Giri 469-476

Pengamanan Data menggunakan Algoritma Enkripsi Advanced Encryption Standard (AES) dan Zigzag Cipher
Bagus Arimbawa I Gusti Ngurah, I Made Widiartha, I Gusti Ngurah Anom Cahyadi Putra, I Ketut Gede Suhartana 477-488

SUSUNAN DEWAN REDAKSI
JURNAL ELEKTRONIK ILMU KOMPUTER
UDAYANA (JELIKU)

Penanggung Jawab :

Dra. Ni Luh Watiniasih M.Sc., Ph.D.

Redaktur :

Ir. Ida Bagus Gede Sarasvananda, S.Kom., M.Cs.

Penyunting/Editor :

Dr. Anak Agung Istri Ngurah Eka Karyawati, S.Si.,M.Eng.

Ir. I Gusti Agung Gede Arya Kadyanan, S.Kom, M.Kom

Gst. Ayu Vida Mastrika Giri, S.Kom., M.Cs.

I Putu Praba Santika, S.Kom., M.Kom.

I Putu Satwika, S.Kom., M.Kom.

Disain Grafis :

I Nyoman Agustya Parawangsa Winantra

Alya Rayyan Tyas Azkiah

Fotografer :

Made Krishna Adiyatma

Samuel Bill Bragman Sinaga

Sekretariat :

Ni Ketut Alit Widiastuti, S.Kom.

Anak Agung Raka Darmawan, S.Kom.

I Putu Herryawan, S.Kom.

This page is intentionally left blank.

Jurnal Elektronik Ilmu Komputer Udayana
Volume 14, No 3. February 2026

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Klasifikasi *Mood* Musik Menggunakan *Support Vector Machine* Berdasarkan Fitur Spektral

I Gusti Agung Ngurah Diputra Wiraguna[a1], Luh Arida Ayu Rahning Putri[a2], I Ketut Gede Suhartana [a3], Gst. Ayu Vida Mastrika Giri [a4]

[a]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana, Jimbaran, Kuta Selatan, Badung, Bali, Indoesia

[1]gungwah1515@gmail.com
[2]rahningputri@unud.ac.id
[3]ikg.suhartana@unud.ac.id
[4]vida@unud.ac.id

Abstrak

Klasifikasi musik merupakan salah satu topik MIR (Music Information Retrieval) yang berkembang pesat. Berbagai aspek musik telah menjadi topik penelitian seputar klasifikasi musik mulai dari *genre*, instrumen, hingga *mood*. Salah satu teori taksonomi *mood* musik yang cukup terkenal adalah *circumplex model* yang memetekan *mood* musik berdasarkan dimensi *arousal* dan *valence*. *Arousal* untuk tingkat intensitas dari *mood* tersebut. Sedangkan *valence* untuk kondisi emosional positif atau negatifnya suatu *mood*. Penelitian mengklasifikasikan 4 kelas kuadran *mood* musik berdasarkan *circumplex model* menggunakan model pembelajaran mesin SVM (*Support Vector Machine*). Penelitian menggunakan jenis fitur spektral yang merupakan distribusi fitur sinyal musik pada domain frekuensi untuk mengklasifikasikan 4 kelas *mood* kuadran musik. Klasifikasi dilakukan melalui beberapa pendekatan berbeda seperti mengklasifiksikan 4 kelas secara langsung serta kombinasi model *arousal* dan *valence* terpisah. Hal ini dilakukan untuk menilai performa fitur spektral dalam klasifikasi *mood* musik berdasarkan 4 kelas kuadran dan melihat apakah ada kecenderungan fitur spektral lebih cocok dalam klasifikasi *arousal* atau *valence*. Penelitian juga mengujicobakan beberapa tahap lain seperti *parameter tuning* SVM dan mengaplikasikan seleksi fitur hingga reduksi dimensi fitur. Hasil penelitian diperoleh model SVM dengan akurasi terbaik untuk pendekatan klasifiksai 4 kelas yang memperoleh akurasi sebesar 66,67%. Hasil terbaik diperoleh setelah mengaplikasikan seleksi fitur dan reduksi dimensi fitur. Selain itu hasil menunjukkan bahwa fitur spektral cenderung mampu mengklasifikasikan *arousal* lebih baik dibandingkan *valence* dalam konteks *circumplex model* untuk *mood* musik.

Kata kunci : *Mood Musik, Circumplex Model, Fitur Spektral, Klasifikasi, Support Vector Machine*

1. Pendahuluan

Beberapa artikel meyimpulkan musik memiliki pengaruh terhadap *mood* pendengarnya [1], [2]. Pengaruh musik terhadap *mood* pendengar dapat terjadi pada siapa saja tanpa memandang usia atau jenis kelamin [1]. Beberapa aspek musik yang dinilai mampu mempengaruhi mood dapat berupa topik musik tersebut, lirik, hingga bagaimana musik tersebut dimainkan. Berdasarkan aspek - aspek tersebut musik dapat dikategorikan berdasarkan kategori tertentu seperti jenis instrumen, *genre*, hingga *mood*. *Mood* sebagai salah satu bentuk pengkategorian musik cenderung bersifat subjektif karena terkadang sebuah musik menggambarkan kondisi atau keadaan sesuai dengan keinginan musisinya. Secara psikologis, penelitian terdahulu telah mengemukakan skema pengkategorian *mood* dalam kelas tertentu. Salah satu yang cukup terkenal adalah *circumplex model* yang mengemukakan *mood* berdasarkan dimensi *arousal* dan *valence* [3], yang kemudian dijadikan landasan dalam penelitian seputar *mood* termasuk musik.

Dalam penelitian seputar musik banyak dijumpai jenis fitur musik salah satunya yaitu fitur spektral. Fitur spektral merupakan persebaran fitur sinyal suara dalam representasi domain frekuensi. Selain domain frekuensi terdapat domain lain seperti domain waktu dan *cepstral.* Namun fitur spektral dipilih karena merupakan fitur dasar sinyal suara dan representasi domain frekuensi dapat diekstraksi banyak jenis fitur spektral. Fitur spektral mencerminkan nilai karakteristik frekuensi sebuah sinyal yang

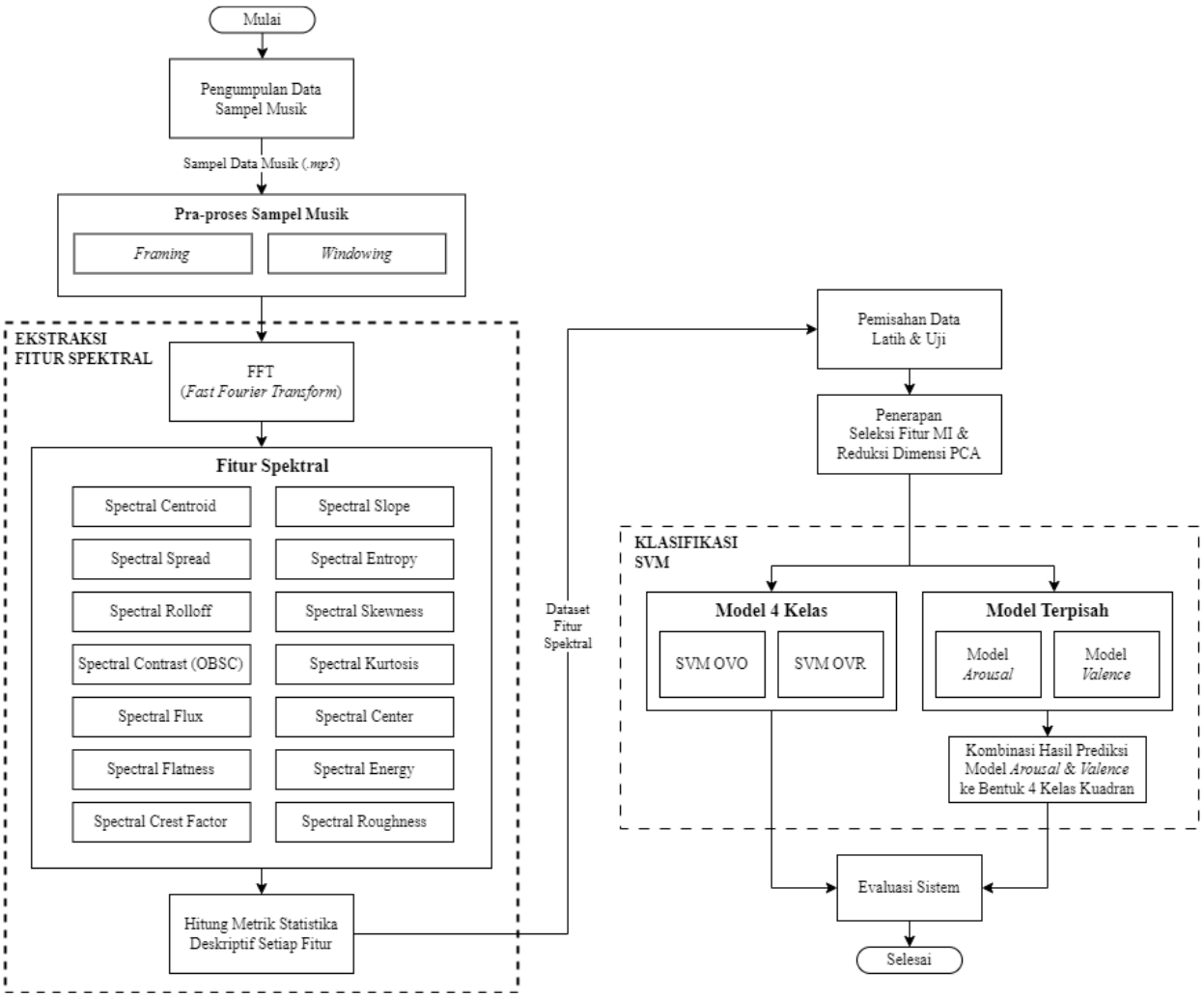
merepresentasikan harmoni, nada, hingga derau secara rasional [4]. Hal ini membuat fitur spektral lebih unggul dibandingkan fitur domain waktu dan pada beberapa kasus, fitur spektral juga mengungguli fitur domain *cepstral* [5], [6].

Menggunakan dasar teori pengkategorian *mood* circumplex model dan fitur spektral dirancang sebuah model klasifikasi *mood* musik menggunakan SVM. Beberapa penelitian sebelumnya telah mengaplikasikan SVM pada ranah musik baik dalam klasifikasi *mood*, hingga aspek lain seperti *genre* dan memperoleh performa yang baik [7], [8]. Salah satu alasan utama dipilihnya SVM adalah kemampuan klasifikasi *multiclass* meskipun pada dasarnya SVM adalah model klasifikasi biner. Hal ini sesuai digunakan pada kasus kllasifikasi *mood* 4 kuadran yang tersusun kembali oleh aspek *arousal* dan *valence*.

Penelitian ini bertujuan untuk menguj performa SVM dalam melakukan Klasifikasi musik dengan kelas *mood* berupa 4 kuadran *circumplex model* dengan beberapa pendekatan berbeda. Pengujian dilakukan melalui konfigurasi kelas dataset pada klasifikasi. Hasil penelitian diharapkan mampu menjadi acuan terkait pengkategorian *mood* dan performa fitur spektral pada ranah musik, serta menjadi motivasi untuk penelitian berikutnya baik seputar musik maupun pembelajaran mesin.

2. Metode Penelitian

Klasifikasi *mood* musik menggunakan SVM berdasarkan fitur spektral terdiri dari beberapa tahapan utama meliputi pengumpulan data sampel musik, ekstraksi fitur spektral, klasfiikasi dengan model SVM, serta pengujian dan evaluasi model. Alur penelitian dapat dilihat pada gambar 1.



Gambar 1. Alur Penelitian

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Secara umum penelitian dimulai dari mengumpulkan data sampel musik 4 kuadran *circumplex model*. Data sampel musik akan melalui tahap pra-proses melalui *framing* dan *windowing* sebelum diaplikasikan FFT (*Fast Fourier Transform*) sebagai bentuk representasi sinyal pada domain frekuensi untuk kemudian diekstraksi nilai ciri atau karakteristik sinyal berupa fitur spektral. Berbagai jenis fitur spektral diekstraksi dan dijadikan sebagai dataset fitur dalam melatih dan menguji model pembelajaran mesin SVM melalui beberapa pendekatan model Klasifikasi SVM berbeda.

2.1. Pengumpulan Data

Dalam menunjang perancangan model klasifikasi *mood* musik berdasarkan 4 kuadran *circumplex model*, diperlukan dataset berupa sampel musik yang relevan. Dataset tergolong jenis sekunder karena berasal dari sumber *open-source* yaitu *Kaggle.com* berupa kumpulan sampel musik berformat "*.mp3*" yang terbagi pada 4 kelas berbeda sesuai kuadran *circumplex model* . Selain sampel musik, dataset juga terdiri dari bebrapa file "*.csv*" yang berisikan metadata seperti anotasi kelas, judul lagu, dan musisinya [9].

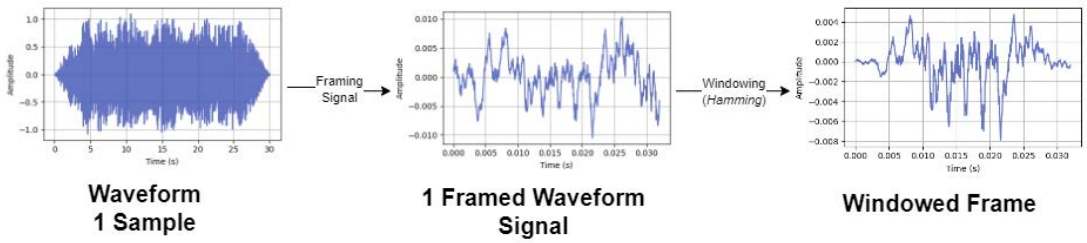
Hasil pengumpulan data memperoleh dataset sampel musik dengan jumlah seimbang untuk masing-masing 4 kelas kuadran dalam bentuk *folder*. Satu kelas berisi 225 sampel "*.mp3*" sehingga total sampel sebanyak 900 sampel musik. sampel musik berdurasi 30 detik dan bervariasi, ada yang berisi nyanyian lirik dan ada yang tidak. Setiap sampel musik akan melalui tahap pemrosesan sinyal dan ekstraksi fitur sebelum digunakan dalam melatih dan menguji model klasifikasi SVM.

2.2. Ekstraksi Fitur

Berdasarkan gambar 1, tahapan ekstraksi fitur dimulai dari pra-proses data sampel musik, penerapan FFT (*Fast Fourier Transform*), lalu hasil FFT digunakan untuk ekstraksi fitur spektral. Hasil fitur spektral yang masih berbentuk banyak *frame* kecil untuk sebuah sampel musik akan diaplikasikan metrik statistika deskriptif untuk menghitung nilai fitur spektral untuk keseluruhan 1 sampel musik**.**

a. Pra-proses Sinyal

Tahap pra-proses berfungsi mempersiapkan sampel sinyal sebelum ekstraksi fitur spektral. Pra-proses umumnya dilakukan untuk menyeragamkan data baik dari segi bentuk, tipe, hingga format disimpannya data. Tahap pra-proses yang dilakukan adalah *framing* dan *windowing*.



Gambar 2. Pra-proses Sampel Sinyal Musik

Tahap *framing* yang mengubah sinyal menjadi beberapa bagian kecil berfungsi untuk mencari karakteristik fitur sinyal musik secara detail. Mengingat sinyal terutama dalam kasus ini berupa musik merupakan gelombang non-stasioner sehingga frekuensi suara musik bisa berbeda pada durasi tertentu. Untuk tahap *windowing* menggunakan jenis *"Hamming"* yang terlampir dalam persamaan (1).

$$w(n) = 0{,}54 - 0{,}46\, cos = 2\pi\,(\frac{n}{N}) \quad (1)\ [10]$$

Fungsi *windowing* membuat transisi antar *frame* sinyal tidak terpotong secara langsung. *Windowing* mampu mengantisipasi *spectral leakage* yang terjadi jika sinyal hanya dibagi melalui *framing* saja. Setiap *frame* yang telah diaplikasikan fungsi *windowing* akan ditumpuk sebagian (*overlapping frames*) untuk mengurangi kehilangan informasi dan menjaga kontinuitas antar *frame* sesuai dengan sampel musik aslinya [10].

b. Ekstraksi Fitur Spektral

Setiap *frame* sinyal akan diaplikasikan FFT untuk merubahnya ke representasi domain frekuensi yang mengandung spektrum dan bin frekuensi penyusun sinyal. Spektrum dan bin frekuensi digunakan dalam ekstraksi berbagai jenis fitur spektral. Adapun fitur spektral yang digunakan terlampir pada tabel 1. Fitur spektral diekstraksi untuk setiap *frame* yang meggambarkan fitur representasi sinyal pada domain frekuensi [11].

Tabel 1. Fitur Spektral [11], [12], [13]

Fitur	Keterangan
Spectral Centroid	Pusat massa spektrum penyusun sinyal.
Spectral Roll-off	Frekuensi dengan spektrum dibawah persentasi tertentu.
Spectral Spread	Persebaran frekuensi di sekitar *spectral centroid.*
Spectral Contrast	Perbedaan frekuensi puncak dan lembah pada *range* frekuensi tertentu.
Spectral Flux	Perubahan spektrum antar *frame.*
Spectral Flatness	Tingkat kedataran frekuensi penyusun sinyal.
Spectral Crest Factor	Rasio spektrum puncak dengan rata – rata-nya.
Spectral Slope	Perbandingan spektrum setiap bertambahnya frekuensi.
Spectral Entropy	Ketidakteraturan distribusi spektrum sebuah sinyal.
Spectral Skewness	Tingkat simetri/asimetri distribusi spektrum.
Spectral Kurtosis	Ketajaman puncak distribusi spektrum sinyal.
Spectral Energy	Total energi spektrum dalam sebuah sinyal.
Spectral Roughness	Fluktuasi/cepatnya perubahan antar spektrum.

Untuk 1 sampel musik akan diekstraksi fitur spektral sebanyak jumlah pada tabel 1. Fitur spektral diekstraksi secara manual menggunakan deskripsi dan rumus sesuai artikel tertentu [11], [12], [13]. Setiap ekstraksi 1 jenis fitur akan menghasilkan vektor fitur karena tahap pra-proses yang membuat sebuah sampel terbagi kedalam banyak *frame*. Untuk menangani hal ini diterapkan metrik statistika deskriptif meliputi *mean, median, variance, minimum, maximum, range, standard deviation, interquartille,* dan *variance*. Metrik statistika diterapkan pada tiap jenis fitur spektral untuk setiap sampel dengan hasil berupa dataset fitur spektral. Dataset fitur kemudian akan disertakan dengan kolom nama file sampel serta kelas sampel untuk dipisahkan menjadi data latih dan uji sebagai bahan perancangan model klasifikasi SVM.

2.3. Klasifikasi SVM

Klasifikasi SVM dimulai dengan mengakses dataset fitur dan memisahkannya menjadi 80% data latih dan 20% data uji. Model SVM akan mencari *hyperplane* optimal yang mampu mengklasifikasikan data antar kelas. SVM sendiri pada dasarnya merupakan klasifikasi biner sehingga penerapan dalam kelas 4 kuadran *circumplex model* perlu dilakukan penyesuaian. Adapun persaman dasar *hyperplane* sebagai berikut.

$$\mathrm{w}\,.\,\mathrm{x} + b = 0 \qquad (2)\ [14]$$

Dari rumus dasar *hyperplane* pada persamaan (2), perhitungan *hyperplane* optimal untuk SVM dimodifikasi dalam bentuk fungsi objektif SVM yang penyelesaiannya diperoleh menggunakan *lagrange multiplier* (*α*). Bentuk persamaan fungsi objektif ini sering juga disebut fungsi *wolfe dual* [14]. Adapun rumus dari fungsi *wolfe dual* SVM sebagai berikut.

$$\begin{gathered} maximize_w\ W(a) = \sum_{i=1}^{m} a_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} a_i a_j y_i y_j x_i . x_j \\ \text{subject to, } C \geq a_i \geq 0\ ,\ for\ i = 1, \ldots, m \\ \sum_{i=1}^{m} a_i y_i = 0 \end{gathered} \qquad (3)\ [14]$$

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Menerapkan persamaan (3) akan menghasilkan nilai *α* yang dapat digunakan untuk menghitung bobot dan bias model yang kemudian digunakan dalam prediksi data. Dalam membentuk model SVM, data latih digunakan untuk menentukan nilai bobot dan bias yang dinilai menghasilkan *hyperplane* optimal sesuai persamaan (2) dan (3). Kemudian nilai bobot dan bias digunakan untuk memprediksi data uji dan dibandingkan apakah hasil prediksi sesuai dengan kelas asal.

Dalam melatih model dikenal istilah parameter tuning dengan beberapa jenis parameter seperti nilai *cost* (C), *gamma* (γ), dan *kernel trick*. Kernel khususnya membuat SVM mampu melakukan klasifikasi data dengan dimensi tinggi [14]. Beberapa variasi nilai parameter diujicobakan secara *grid-search* dalam SVM untuk mencari kombinasi parameter dengan performa terbaik. Adapun beberapa jenis kernel yang digunakan sebagai berikut.

$$\text{kernel linear,} \quad k(x_i, x_j) = (x_i \,.\, x_j) \tag{4) [14]}$$

$$\text{kernel polynomial,} \quad k(x_i, x_j) = (x_i \,.\, x_j + c)^d \tag{5) [14]}$$

$$\text{kernel RBF,} \quad k\left(x^{(i)}, x^{(j)}\right) = exp\,(\gamma||x^{(i)} - x^{(j)}||^2) \tag{6) [14]}$$

Untuk menangani klasifikasi kelas 4 kuadran, diterapkan pendekatan berupa konfigurasi jenis SVM berupa OVO, OVR, serta klasifikasi model *arousal* dan *valence* terpisah. Klasifikasi OVO dan OVR menyusun skema klasifikasi antar kelas kuadran dengan merubahnya menjadi klasifikasi biner. Skema OVO akan menyusun klasifikasi antar 1 kelas dengan 1 kelas lainnya hingga terbentuk seluruh pasangan yang mungkin dari seluruh 4 kelas. Sedangkan skema OVR langsung mengklasifikasikan 1 kelas dengan 3 kelas lainnya. Untuk pendekatan model *arousal* dan *valence* terpisah akan dibentuk 2 model klasifikasi dimana pada model *arousal* dibedakan menjadi kelas *arousal* tinggi (Q1, Q2) dan rendah (Q3, Q4). Sedangkan model *valence* dibedakan menjadi kelas *valence* positif (Q1, Q4) atau negatif (Q2, Q3). Pengujian model *arousal* dan *valence* terpisah menggunakan data uji 4 kelas kuadran melalui *mapping* hasil prediksi. Data uji dianggap diklasifikasikan benar hanya jika diprediksikan sesuai kelas asal pada kedua model *arousal* dan model *valence*.

2.4. Evaluasi Sistem

Pengujian performa sistem dilihat dari hasil prediksi data uji apakah sesuai kelas asalnya atau tidak. Hasil prediksi data uji dapat digambarkan dalam bentuk *confusion matrix* sebagai berikut.

Tabel 2. Confusion Matrix

Confusion Matrix		**Kelas Asal**	
		Positive	**Negative**
Kelas Prediksi	**Positive**	TP	FP
	Negative	FN	TN

Berdasarkan *confusion matrix* pada tabel 2 dapat disusun persamaan (7) untuk menghitung akurasi sistem. Pengukuran akurasi menghitung seberapa banyak data uji yang diklasifikasikan sesuai kelas asalnya berbanding dengan keseluruhan jumlah prediksi data uji. Salah satu alasan pemilihan metrik akurasi adalah seimbangnya jumlah sampel data musik tiap kelas.

$$akurasi = \frac{TP + TN}{TP + TN + FP + FN}\ x\ 100\% \tag{7) [8]}$$

Adapun keterangan dari *confusion matrix* pada tabel 2 serta perhitungan akurasi pada persamaan (7) dimana TP (*True Positive*) merupakan data yang diprediksikan positif dan memang kelas asalnya positif. FP (*False Positive*) merupakan data yang diprediksikan positif namun kelas sebenarnya adalah negatif. FN (*False Negative*) merupakan banyak data yang diprediksikan negatif namun kelas sebenarnya adalah positif. TN (*True Negative*) merupakan data yang diprediksi negatif dan memang kelas asalnya negatif.

Pengukuran akurasi diterapkan pada seluruh pendekatan model klasifikasi SVM. Hasil akurasi selain menunjukkan performa sistem juga menunjukkan bagaimana pengaruh fitur spektral dalam mengklasifikasikan musik berdasarkan kategori 4 kuadran menggunakan SVM. Skenario pengujian meliputi pendekatan model SVM OVO, OVR, hingga kombinasi model *arousal* dan *valence* yang dilakukan dalam bentuk *grid-search* untuk menentukan kombinasi parameter yang menghasilkan akurasi terbaik. Parameter SVM yang diujicobakan terlampir dalam tabel 3.

Tabel 3. Parameter Model SVM

Kernel	C (*Cost*)	*Gamma* (γ)
RBF	1, 10, 100	0.001, 0.01, 0.1, 0.5, 1
Linear	1, 10, 100	0.001, 0.01, 0.1, 0.5, 1
Polynomial	1, 10, 100	0.001, 0.01, 0.1, 0.5, 1

Pengujian dilanjutkan dengan mengujicobakan seleksi fitur MI dan reduksi dimensi fitur PCA. Seleksi fitur MI akan menghitung nilai MI setiap fitur dalam dataset terhadap kelas data yang ada. Nilai fitur MI memiliki rentang 0 hingga 1, Dimana nilai MI tinggi mengindikasi fitur tersebut relevan terhadap kelas data [15]. Sedangkan nilai MI kecil atau bahkan 0 berarti fitur tersebut independen dan tidak menggambarkan kelas data. Berbeda dengan seleksi fitur MI, reduksi dimensi fitur PCA akan membentuk dataset fitur baru dengan dimensi yang lebih ringkas namun mempertahankan relevansi informasi fitur data baru dengan fitur data pada dimensi asalnya [16]. Sehingga PCA cocok digunakan pada dataset berdimensi tinggi yang menyebabkan beban komputasi yang besar.

Penerapan seleksi fitur MI dan reduksi dimensi PCA bertujuan mengidentifikasi fitur spektral yang dinilai relevan dalam Klasifikasi *mood* 4 kuadran. Dengan mereduksi dimensi dataset fitur membuat pelatihan model lebih mudah dan meringankan beban komputasional. Selain itu menerapkan seleksi fitur MI dan reduksi dimensi fitur PCA diharapkan mampu meningkatkan performa model SVM yaitu peningkatan akurasi model.

3. Hasil dan Pembahasan

Hasil implementasi sistem menunjukkan beberapa perolehan akurasi model SVM untuk setiap pendekatan berbeda. Perolehan akurasi terbaik setiap pendekatan ditampilkan dengan *confusion matrix* untuk melihat karakteristik hasil klasifikasi. Pengujian dilakukan melalui *grid-search* parameter SVM yang mungkin terbentuk pada tabel 2, Pengujian model dilakukan menggunakan dataset fitur yang lengkap hingga yang telah diterapkan seleksi fitur MI dan reduksi dimensi fitur PCA.

3.1. Pendekatan Model Klasifikasi

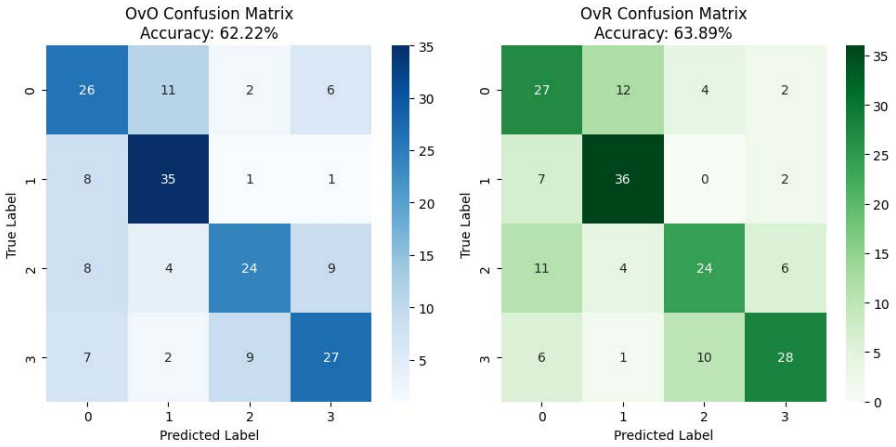
Terdiri dari pendekatan klasifikasi 4 kelas (*multi-class*) secara langsung dan klasifikasi model terpisah. Perolehan akurasi setiap pendekatan model SVM sebagai berikut.

a. Klasifikasi 4 Kelas

Klasifikasi dilakukan dengan menerapkan konfigurasi OVO dan OVR untuk menangani klasifikasi *multi-class*. Pendekatan klasifikasi OVO dan OVR diterapkan karena kasus klasifikasi yaitu 4 *mood* kuadran *circumplex model* sedangkan model SVM pada dasarnya adalah model klasifikasi biner.

Model klasifikasi OVO akan menyusun skema 4 kelas menjadi sebanyak pasangan antar 2 kelas yang mungkin terbentuk. Prediksi akan dilakukan secara *voting* untuk memilih hasil prediksi yang lebih banyak terpilih. Berbeda dengan model OVO, model klasifikasi OVR akan menyusun skema klasifikasi biner antar 2 kelas dimana kelas 1 melawan sisa kelas lainnya secara langsung. Hasil pengujian model OVO dan OVR ditunjukkan dalam bentuk *confusion matrix* berikut.

p-ISSN: 2301-5373
e-ISSN: 2654-5101

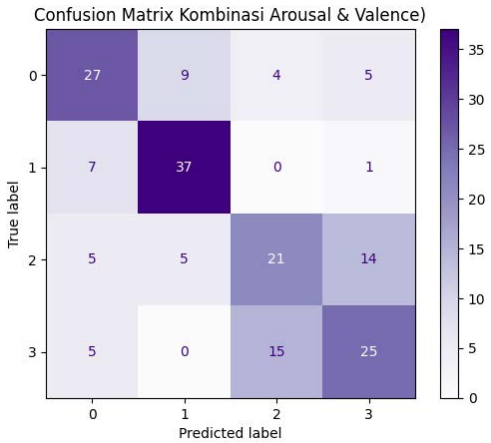


Gambar 3. *Confusion Matrix* Model OVO dan OVR

Gambar 3 memperlihatkan *confusion matrix* model SVM OVO dan OVR yang memperoleh akurasi terbaik. Model OVO memperoleh akurasi terbaik sebesar 62,22% dengan kombinasi parameter berupa nilai C senilai 10, *gamma* 0,01, dan jenis kernel RBF. Sedangkan model OVR memperoleh akurasi terbaik sebesar 63,89% dengan parameter C sebesar 1, *gamma* 0,01, dan jenis kernel RBF. Meskipun OVR memperoleh akurasi lebih baik, *confusion matrix* OVR menunjukkan banyak misklasifikasi menjadi Q1. Sedangkan model OVO persebaran hasil prediksi lebih sesuai dengan teori *circumplex model* yang menunjukkan model lebih mampu dalam klasifikasi *arousal*.

b. Klasifiksi Model *Arousal* dan *Valence* Terpisah

Klasifikasi model terpisah menyusun skema model *arousal* dan ,model *valence* dengan menyesuaikan kelas pada dataset fitur. Kedua model diuji dengan data uji *mood* 4 kelas kuadran.



Gambar 4. *Confusion Matrix* Kombinasi Model Arousal dan Valence

Gambar 4 menunjukkan *confusion matrix* kombinasi model *arousal* dan *valence* dengan perolehan akurasi terbaik sebesar 61,11%. Akurasi tersebut diperoleh melalui konfigurasi model *arousal* berparameter C senilai 10, *gamma* 0,001, dan jenis kernel RBF. Sedangkan model *valence* berkonfigurasi C senilai 100, *gamma* 0,01, dan jeis kernel RBF. Hasil kembali menunjukkan model SVM dengan fitur spektral lebih akurat dalam klasifikasi aspek *arousal.*

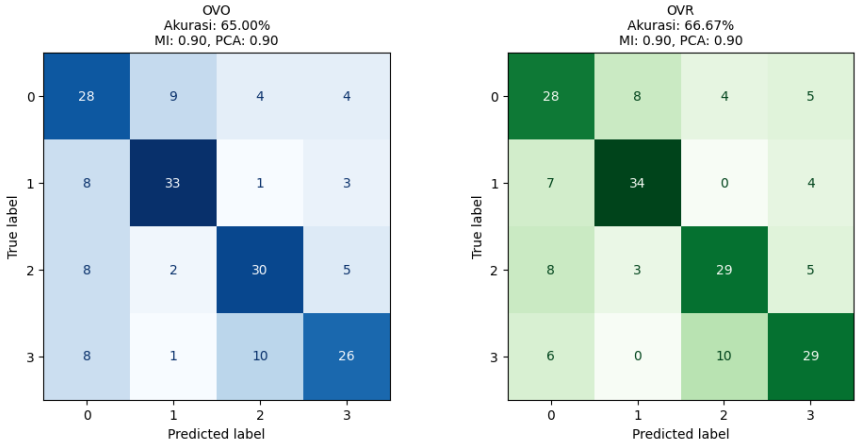
3.2. Penerapan Seleksi Fitur MI & Reduksi Dimensi Fitur PCA

Kombinasi seleksi fitur dan reduksi dimensi fitur diterapkan setelah memperoleh akurasi terbaik dari setiap pendekatan klasifikasi. Seleksi fitur yang digunakan adalah MI (*Mutual Information*) dan reduksi dimensi PCA. MI dan PCA diterapkan berpasangan untuk mengidentifikasi

kombinasi parameter MI dan PCA yang mampu meningkatkan akurasi model SVM setiap pendekatan sebelumnya. Parameter MI dan PCA yang digunakan senilai 70%, 80%, dan 90%.

a. Klasifikasi 4 Kelas

Penerapan seleksi fitur MI dan reduksi dimensi fitur PCA terbukti meningkatkan akurasi model. Karakteristik klasifikasi dan peningkatan akurasi ditunjukkan melalui *confusion matrix* model OVO dan OVR setelah seleksi fitur MI dan reduksi dimensi PCA.

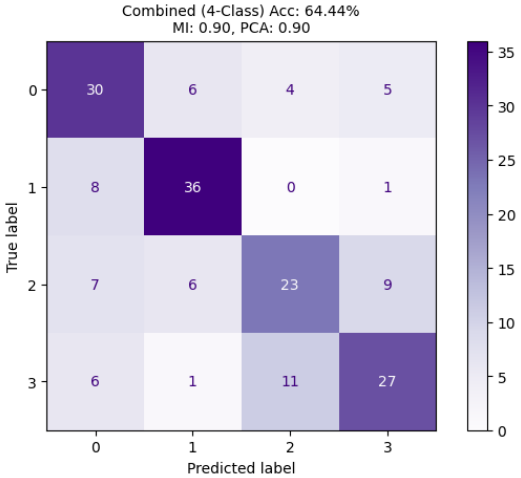


Gambar 5. *Confusion Matrix* Model OVO dan OVR (Setelah MI dan PCA)

Gambar 5 memperlihatkan peningkatan akurasi dari model OVO dan OVR setelah diterapkan seleksi fitur MI dan reduksi dimensi PCA. Akurasi terbaik masih diperoleh oleh model OVR yang meningkat menjadi 66,67%. Akurasi tersebut diperoleh oleh konfigurasi parameter C senilai 10, *gamma 0,01*, dan jenis kernel RBF. Model OVO memperoleh akurasi 65% dengan konfigurasi parameter yang sama. Terlihat bahwa penggunaan 90% fitur MI dengan jumlah *principal component* hasil PCA dengan karakteristik 90% sama – sama memperoleh akurasi terbaik di kedua model.

b. Klasifikasi Arousal dan Valence Terpisah

MI dan PCA juga diterapkan pada model *arousal* dan *valence* terpisah untuk melihat apakah terjadi peningkatan akurasi serta karakteristik hasil prediksi kedua model.



Gambar 6. *Confusion Matrix* Model Kombinasi (Setelah MI dan PCA)

Akurasi meningkat menjadi 64,44% yang diperlihatkan pada gambar 6. Akurasi terbaik diperoleh pada kombinasi model *arousal* berparameter C senilai 10, *gamma* 0,001, dan kernel RBF. Untuk model *valence* memiliki konfigurasi nilai C senilai 10, *gamma* 0,01, dan kernel RBF. Akurasi terbaik diperoleh dari penerapan seleksi fitur MI dan reduksi dimensi fitur PCA berparameter 90%. Secara keseluruhan hasil sebelum dan sesudah penerapan MI dan PCA memiliki karakteristik yang sama pada *confusion matrix* dimana model lebih mampu mengklasifikasikan aspek *arousal*.

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Hasil penelitian menunjukkan adanya pengaruh parameter tuning hingga penerapan seleksi fitur MI dan reduksi dimensi fitur PCA pada akurasi model SVM. Pendekatan model SVM berbeda menunjukkan karakteristik hasil yang cukup berbeda meskipun akurasi model tidak berbeda jauh. Pada gambar 3 model OVR memiliki akurasi yang lebih baik namun *confusion matrix* model OVO memperlihatkan hasil klasifikasi yang lebih seimbang dan sesuai dengan taksonomi *circumplex model*. Melalui pendekatan kombinasi model *arousal* dan *valence* terpisah terbukti bahwa model SVM dengan fitur spektral lebih mampu mengklasifikasikan aspek *arousal* dibandingkan *valence*.

Penerapan seleksi fitur MI dan reduksi dimensi fitur PCA pada dataset fitur mampu meningkatkan akurasi setiap pendekatan model SVM. Berdasarkan gambar 5 dan 6, karakteristik hasil Klasifikasi model SVM setelah diterapkan seleksi fitur MI dan reduksi dimensi fitur PCA masih menunujukkan kecenderungan model yang lebih akurat dalam klasifiksai *arousal*. Hasil penerapan MI dan PCA juga menunjukkan bahwa model memiliki performa baik saat parameter MI dan PCA bernilai tinggi, seperti 90%, Melihat perolehan akurasi setiap pendekatan model SVM, ditampilkan beberapa fitur dengan perolehan nilai MI tertinggi.

Tabel 4. Fitur Spektral dengan Nilai MI Tertinggi

Fitur & Metrik Statistika	Nilai MI
roughness_median	0,365027
flux_median	0,340344
contrast_band3_median	0,336179
roughness_iqr	0,335787
contrast_band3_mean	0,331809

Terlihat dari tabel 4 fitur spektral dan metrik statistika yang diekstraksi seperti *spectral roughness, spectral flux, dan spectral contrast* pada *subband* frekuensi rendah memperoleh nilai MI tertinggi di kisaran 0,3. Seleksi fitur MI memperoleh nilai MI seluruh fitur yang kemudian dapat diurutkan dari tertinggi ke terendah. Sebagai tambahan, perolehan nilai MI terendah menyentuh nilai 0 pada beberapa fitur yang mayoritas terdiri dari *spectral contrast* pada *subband* frekuensi tinggi. Dimana nilai MI 0 mengartikan fitur tidak relevan sama sekali.

Peningkatan akurasi di seluruh pendekatan model SVM sesuai dengan tujuan penerapan seleksi fitur MI yaitu memilih fitur yang relevan terhadap kelas klasifikasi berdasarkan nilai MI tiap fitur. Sedangkan metode PCA mereduksi fitur yang kompleks dengan membentuk *principal component* dengan ukuran lebih kecil namun ttetap mempertahankan informasi fitur sesuai persentase PCA yang digunakan

Berdasarkan hasil penelitian hampir keseluruhan pendekatan model klasifikasi menunjukkan karakteristik performa yang lebih akurat dalam klasifikasi *arousal* dibandingkan *valence*. Beberapa teori menjelaskan terkait hasil penelitian. Fitur spektral yang merupakan representasi domain frekuensi sebuah sinyal lebih mampu dalam klasifikasi *arousal* sesuai dengan karakteristik domain sinyal tersebut. Spektrum beserta bin frekuensi bervariasi antar *frame* sinyal lebih menunjukkan gairah atau intensitas sebuah sinyal [17]. Dalam ranah musik aspek *arousal* bisa dilihat dari cara sebuah musik dimainkan dan seberapa cepat/lambatnya. Untuk *valence* yang merupakan kondisi emosional positif atau negatif, terlihat fitur spektral kesulitan dalam melakukan klasifikasi. Hal ini menyesuaikan dengan karakteristik domain frekuensi yang terdiri dari spektrum dan bin frekuensi penyusun sinyal, secara garis besar melambangkan energi atau intensitas sebuah sinyal musik [18], [19]. Berdasarkan hasil penelitian dan penjelasan tersebut dapat dikatakan bahwa fitur spektral memang lebih cocok dalam mengenali aspek *arousal* dalam *mood* 4 kuadran *circumplex model* dibandingkan aspek *valence.*

4. Kesimpulan

Berdasarkan hasil penelitian dapat disimpulkan bahwa model SVM memperoleh akurasi terbaik sebesar 63,89% dan meningkat menjadi 66,67% setelah diterapkan seleksi fitur MI dan reduksi dimensi PCA. Hasil bentuk *confusion* matrix keseluruhan model klasifikasi menunjukkan karakteristik klasifikasi *arousal* lebih baik dibandingkan *valence*. Hasil tersebut mengimplikasikan fitur spektral lebih mampu mengklasifikasikan *arousal* atau tingkat gairah dan intensitas sebuah kategori *mood*. Berbeda dengan

valence yang merupakan kondisi emosional positif atau negatif yang cukup sulit digambarkan dalam sebuah sinyal musik, fitur spektral dapat mengidentifikasi dan mengenali aspek *arousal* lebih mudah.

Referensi

[1] A. Mamdouh and M. Ahmed, “The Impact of Music on Mood: Descriptive Observations of Listening Experiences and Their Effect on Mental Health,” 2024. [Online]. Available: https://www.researchgate.net/publication/379145461

[2] A. Tewari, “The Impact of Music on Mood Regulation,” *The International Journal of Indian Psychology*, vol. 12, no. 2, p. 5, Jun. 2024, doi: 10.25215/1202.390.

[3] J. A. Russell, “A circumplex model of affect,” *J Pers Soc Psychol*, vol. 39, no. 6, pp. 1161–1178, Dec. 1980, doi: 10.1037/h0077714.

[4] C. Constantinescu and R. Brad, “An Overview on Sound Features in Time and Frequency Domain,” *International Journal of Advanced Statistics and IT&C for Economics and Life Sciences*, vol. 13, no. 1, pp. 45–58, Dec. 2023, doi: 10.2478/ijasitels-2023-0006.

[5] F. Alías, J. C. Socoró, and X. Sevillano, “A review of physical and perceptual feature extraction techniques for speech, music and environmental sounds,” 2016, *Balkan Society of Geometers*. doi: 10.3390/app6050143.

[6] C. Prabha, S. Kaur, M. Malik, M. Uddin, and D. Nandan, “A Cutting-Edge Feature Extraction Approach for Speaker Recognition Leveraging Optimized Variance Spectral Flux and Daubechies Wavelet,” *Traitement du Signal*, vol. 40, no. 6, pp. 2845–2852, Dec. 2023, doi: 10.18280/ts.400645.

[7] C. A. Qurniaty and Kusnawi, “Ekspresi Emosi Berdasarkan Suara Menggunakan Algoritma Multi Layer Perceptron dan Support Vector Machine,” *Indonesian Journal of Computer Science Attribution*, vol. 12, no. 6, pp. 2023–4014, 2023.

[8] B. R. Ismanto, T. M. Kusuma, and D. Anggraini, “Indonesian Music Classification on Folk and Dangdut Genre Based on Rolloff Spectral Feature Using Support Vector Machine (SVM) Algorithm,” *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 15, no. 1, p. 11, Jan. 2021, doi: 10.22146/ijccs.54646.

[9] R. Panda, R. Malheiro, and R. P. Paiva, “Novel Audio Features for Music Emotion Recognition,” *IEEE Trans Affect Comput*, vol. 11, no. 4, pp. 614–626, Oct. 2020, doi: 10.1109/TAFFC.2018.2820691.

[10] O. Kamil, “Frame Blocking and Windowing Speech Signal,” 2018, [Online]. Available: https://www.researchgate.net/publication/331635757

[11] G. Sharma, K. Umapathy, and S. Krishnan, “Trends in audio signal feature extraction methods,” *Applied Acoustics*, vol. 158, Jan. 2020, doi: 10.1016/j.apacoust.2019.107020.

[12] J. Tao, D.-N. Jiang, L. Lu, H.-J. Zhang, J.-H. Tao, and L.-H. Cai, “Music type classification by spectral contrast feature MUSIC TYPE CLASSIFICATION BY SPECTRAL CONTRAST FEATURE 1,” 2002. [Online]. Available: https://www.researchgate.net/publication/313484983

[13] G. Peeters, “A Large Set Of Audio Features For Sound Description (Similarity and Classification) In The CUIDADO Project,” 2003.

[14] A. Kowalczyk, *Support Vector Machine Succinctly*. Syncfusion, 2017. [Online]. Available: www.syncfusion.com.

[15] X. Cheng, “Signals and Systems Ins,” 2024. [Online]. Available: https://soapubs.com/index.php/ICSS

[16] M. Greenacre, P. J. F. Groenen, T. Hastie, A. I. D’Enza, A. Markos, and E. Tuzhilina, “Principal component analysis,” *Nature Reviews Methods Primers*, vol. 2, no. 1, Dec. 2022, doi: 10.1038/s43586-022-00184-w.

[17] Y. H. Yang, Y. C. Lin, Y. F. Su, and H. H. Chen, “A regression approach to music emotion recognition,” *IEEE Trans Audio Speech Lang Process*, vol. 16, no. 2, pp. 448–457, Feb. 2008, doi: 10.1109/TASL.2007.911513.

[18] U. Oberst, “The fast Fourier transform,” *SIAM J Control Optim*, vol. 46, no. 2, pp. 496–540, 2007, doi: 10.1137/060658242.

[19] D. T. Kusuma, “Fast Fourier Transform (FFT) Dalam Transformasi Sinyal Frekuensi Suara Sebagai Upaya Perolehan Average Energy (AE) Musik,” *PETIR*, vol. 14, no. 1, pp. 28–35, Oct. 2020, doi: 10.33322/petir.v14i1.1022.

Jurnal Elektronik Ilmu Komputer Udayana
Volume 14, No 3. February 2026

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Sistem Representasi Digital Wayang Kulit Bali Menggunakan AHP Dengan Pendekatan Ontologi

I Komang Widia Pratama[a1], Made Agung Raharja[a2], Cokorda Pramartha[a3] , I Ketut Gede Suhartana[a4]

[a]Program Studi Informatika, Universitas Udayana
Bali, Indonesia
[1]widiapratama040@student.unud.ac.id
[2]made.agung@unud.ac.id
[3]cokorda@unud.ac.id
[4]ikg.suhartana@unud.ac.id

Abstract

This study aims to develop a digital documentation system for Balinese Wayang Kulit to support the preservation of intangible cultural heritage. The main issue addressed is the limited access to structured and comprehensible information about Wayang Kulit Bali, including its characters, stories, and cultural values. To overcome this, the study combines an ontology-based approach to construct a semantic knowledge framework and the Analytical Hierarchy Process (AHP) method to provide story recommendations based on user preferences. The developed ontology consists of six main classes Wayang, Time, Function, Role, Character Traits, and Stories, each connected through structured semantic relationships. AHP is applied to weigh user preferences based on criteria such as story complexity, moral values, duration, and popularity. System evaluation shows good consistency (CR = 0.0439) and positive usability scores in terms of learnability (4.25) and satisfaction (4.25). The proposed system is expected to enhance accessibility, understanding, and appreciation of Wayang Kulit Bali among the public through a digital platform.

Keywords: *Balinese Wayang Kulit, Analytical Hierarchy Process, Ontology, Digital Documentation, System Evaluation, Cultural Heritage*

1. Pendahuluan

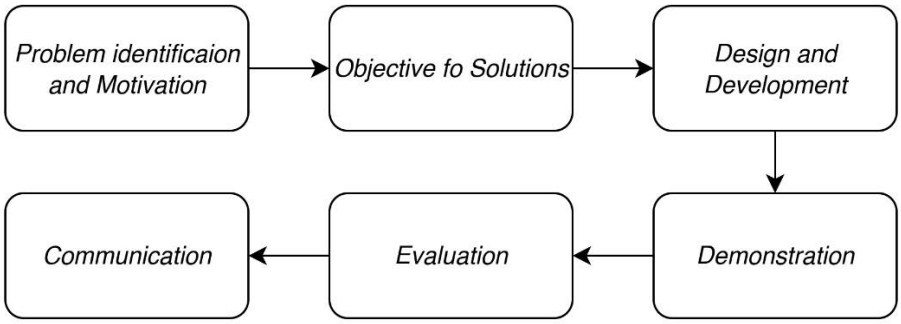
Wayang Kulit Bali merupakan salah satu warisan budaya tak benda yang mengandung nilai-nilai moral, spiritual, dan sosial yang tinggi. Namun, penyampaian pengetahuan tentang cerita, tokoh, dan filosofi dalam Wayang Kulit Bali masih bersifat lisan dan terbatas, sehingga berisiko hilang jika tidak dilakukan pelestarian berbasis teknologi [1].Teknologi *semantic web* dapat digunakan untuk menyusun pengetahuan dalam bentuk terstruktur dan bermakna. Ontologi sebagai fondasi *semantic web* mampu menggambarkan relasi antar entitas dalam domain budaya secara eksplisit dan formal [2]. Selain itu, dalam konteks interaksi pengguna, sistem dapat ditingkatkan dengan memberikan fitur rekomendasi cerita yang disesuaikan dengan preferensi pengguna. untuk menyusun pengetahuan dalam bentuk terstruktur dan bermakna. Ontologi sebagai fondasi *semantic web* mampu menggambarkan relasi antar entitas dalam domain budaya secara eksplisit dan formal. Selain itu, dalam konteks interaksi pengguna, sistem dapat ditingkatkan dengan memberikan fitur rekomendasi cerita yang disesuaikan dengan preferensi pengguna.

Untuk itu, digunakan metode *Analytical Hierarchy Process* (AHP) dalam mendukung pengambilan keputusan rekomendasi cerita secara objektif. AHP memungkinkan sistem untuk membandingkan berbagai kriteria seperti nilai moral, karakter tokoh, dan tema cerita berdasarkan bobot yang diberikan oleh pengguna [3]. Beberapa penelitian sebelumnya telah menunjukkan bahwa AHP efektif dalam konteks pengelolaan warisan budaya digital. Di Angelo et al. (2018) mengembangkan metode berbasis AHP untuk memilih pemindai 3D terbaik dalam aplikasi warisan budaya, menunjukkan bahwa AHP dapat digunakan untuk pengambilan keputusan yang sistematik dalam pelestarian budaya [4]. Selain itu, Lasorella dan Cantatore (2025) mengaplikasikan CityGML dan AHP dalam pemulihan dan pengelolaan warisan budaya terbangun, yang memberikan dasar dalam mengelola warisan budaya

secara digital dengan mempertimbangkan berbagai faktor penting [5]. Temuan-temuan ini menunjukkan bahwa AHP mampu menangani subjektivitas pengguna secara sistematis dan memberikan dasar pengambilan keputusan yang valid dalam sistem berbasis budaya. Oleh karena itu, metode ini dipilih dalam penelitian ini sebagai pendekatan yang tepat untuk mendukung fitur rekomendasi cerita Wayang Kulit Bali.

2. Metode Penelitian

Penelitian ini mengimplementasikan metodologi *Design Science Research Methodology* (DSRM), yang berfokus pada penciptaan dan penilaian artefak sebagai respons terhadap permasalahan yang telah diidentifikasi. Langkah awal dalam proses ini adalah mengidentifikasi kebutuhan, diikuti dengan perancangan dan pembuatan ontologi yang menggambarkan konsep-konsep yang berkaitan dengan wayang kulit Bali. Tahap evaluasi dilakukan untuk memastikan artefak yang dikembangkan dapat memenuhi kebutuhan dan berfungsi secara efektif. Dengan menggunakan DSRM, penelitian ini memastikan bahwa hasil yang diperoleh tidak hanya bersifat teoretis, tetapi juga aplikatif dan relevan dalam konteks praktis.



Gambar 1. Alur Metode Penelitian

2.1. Identifikasi Masalah dan Motivasi

Tantangan utama yang dihadapi adalah terbatasnya akses dan informasi mengenai Wayang Kulit Bali dalam format digital. Padahal, Wayang Kulit Bali memiliki nilai budaya yang sangat tinggi dan telah diakui oleh UNESCO sebagai Warisan Budaya Tak Benda. Informasi mengenai Wayang Kulit Bali yang terbatas dan sulit diakses dapat menghambat penyebaran pengetahuan tentang seni ini kepada masyarakat luas. Hal ini juga berpotensi mengurangi minat dan apresiasi terhadap wayang kulit Bali sebagai bagian dari budaya Bali. Oleh karena itu, diperlukan solusi untuk memperluas akses informasi tersebut melalui sistem yang memungkinkan masyarakat untuk mempelajari, memahami, dan mempromosikan Wayang Kulit Bali secara lebih efektif melalui platform digital. Sistem ini diharapkan dapat meningkatkan minat terhadap seni tradisional Bali, sekaligus mendukung pengembangan industri seni dan budaya di Bali serta berkontribusi pada perekonomian lokal.

2.2. Tujuan Solusi

Tujuan utama adalah mengembangkan sistem dokumentasi digital yang tidak hanya merepresentasikan informasi terkait Wayang Kulit Bali, tetapi juga memberikan rekomendasi cerita berdasarkan preferensi pengguna. Sistem ini akan menggunakan pendekatan ontologi untuk membangun struktur semantik yang terorganisir dan konsisten mengenai sejarah, karakter tokoh, teknik pembuatan, serta nilai budaya Wayang Kulit Bali. Selain itu, sistem ini juga akan menerapkan metode *Analytical Hierarchy Process* (AHP) untuk memberikan rekomendasi cerita yang relevan berdasarkan kriteria yang dipilih oleh pengguna. Dengan adanya sistem ini, diharapkan pengguna dapat dengan mudah menemukan informasi yang mereka butuhkan serta mendapatkan rekomendasi cerita yang sesuai dengan minat dan preferensi mereka, sehingga dapat lebih menghargai dan memahami Wayang Kulit Bali secara mendalam.

2.3. Desain dan Pengembangan Sistem

Proses desain dan pengembangan sistem ini dilakukan melalui beberapa langkah penting, yang mencakup:

p-ISSN: 2301-5373
e-ISSN: 2654-5101

a. **Analisis Kebutuhan**: Proses ini dimulai dengan identifikasi kebutuhan fungsional dan non-fungsional yang diperlukan untuk sistem dokumentasi digital Wayang Kulit Bali dan rekomendasi cerita. Kebutuhan fungsional mencakup kemampuan sistem dalam menyediakan fitur pencarian informasi terkait Wayang Kulit Bali, seperti sejarah, karakter tokoh, teknik pembuatan, dan nilai budaya. Selain itu, sistem ini juga harus mampu memberikan rekomendasi cerita berdasarkan preferensi pengguna, yang melibatkan karakter favorit, cerita yang lebih sering dicari, atau cerita dengan tema budaya yang diinginkan. Kebutuhan non-fungsional mencakup perangkat keras dan perangkat lunak yang diperlukan untuk mendukung implementasi sistem, seperti platform berbasis *website*, database untuk menyimpan informasi ontologi, serta server yang dapat mendukung pengolahan data yang besar dan memungkinkan pengambilan informasi secara efisien.
b. **Pengumpulan Data**: Pengumpulan data dilakukan melalui survei yang melibatkan masyarakat Bali untuk mengidentifikasi preferensi pengguna terkait dengan cerita Wayang Kulit Bali. Data yang dikumpulkan mencakup genre cerita, karakter tokoh, serta elemen budaya yang relevan. Selain itu, data ini juga digunakan untuk membangun model ontologi yang mencakup informasi terkait cerita Wayang Kulit Bali, seperti karakter, tokoh, tema, dan unsur budaya lainnya. Data ini akan memastikan bahwa model ontologi yang dikembangkan dapat menggambarkan dengan akurat dan komprehensif tentang Wayang Kulit Bali.
c. **Pembangunan Model Ontologi**: Tahapan ini melibatkan pembuatan model ontologi yang dapat merepresentasikan informasi Wayang Kulit Bali secara semantik dan terstruktur. Ontologi ini akan mencakup kategori dan atribut terkait cerita Wayang Kulit Bali, seperti tokoh, genre cerita, nilai budaya, teknik pembuatan, serta hubungan antar konsep dalam cerita tersebut. Model ontologi dikembangkan dengan menggunakan metodologi ontologi yang sesuai untuk memastikan bahwa data yang tersimpan dapat diorganisir dengan jelas dan mudah diakses, serta memberikan dasar yang solid untuk rekomendasi cerita berdasarkan preferensi pengguna.
d. **Implementasi Sistem**: Tahapan ini mencakup proses pengkodean dan pengembangan aplikasi berbasis *website* yang mengimplementasikan desain sistem yang telah dibuat. Sistem ini akan memungkinkan pengguna untuk mencari dan menjelajahi informasi tentang Wayang Kulit Bali yang sesuai dengan preferensi mereka. Aplikasi ini akan menggunakan metode *Analytical Hierarchy Process* (AHP) untuk memberikan rekomendasi cerita yang dipersonalisasi berdasarkan data yang ada dalam ontologi. Dengan menggunakan AHP, sistem ini akan mampu memberikan rekomendasi cerita yang relevan dan akurat, sesuai dengan kriteria yang ditentukan oleh pengguna, seperti kesukaan terhadap tokoh tertentu, jenis cerita, atau tema budaya yang diinginkan.

.

2.4. Demonstrasi dan Evaluasi

Tahap evaluasi bertujuan untuk memastikan bahwa sistem yang dikembangkan berjalan sesuai dengan perencanaan awal dan memenuhi kebutuhan pengguna. Pengujian kegunaan (usability testing) akan dilakukan untuk menilai kemudahan penggunaan dan efektivitas sistem. Evaluasi ini juga akan mempertimbangkan beberapa faktor, seperti kemudahan penggunaan, tingkat kesalahan, kecepatan akses, dan kepuasan pengguna. Kuesioner yang digunakan dalam evaluasi akan merujuk pada lima komponen utama dari *usability: learnability, memorability, efficiency, errors,* dan *satisfaction*. Hasil dari evaluasi ini akan digunakan untuk melakukan perbaikan pada sistem dan memastikan bahwa aplikasi yang dikembangkan dapat memenuhi harapan pengguna dan memberikan manfaat yang maksimal.

3. Hasil dan Pembahasan

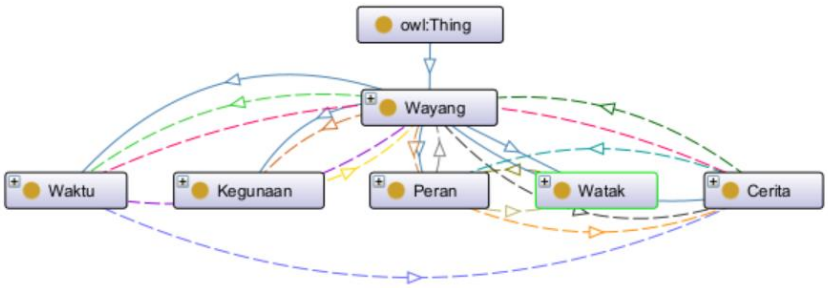
3.1. Desain dan Pembangunan Sistem

Dalam pengembangan sistem dokumentasi digital Wayang Kulit Bali dan rekomendasi cerita, tahapan desain dan pengembangan dilakukan untuk memastikan aplikasi yang efektif dan sesuai dengan tujuan penelitian. Pada tahap desain, fokus utama meliputi analisis kebutuhan fungsional dan non-fungsional, pengumpulan data relevan, pembuatan model ontologi, dan perancangan antarmuka pengguna yang mudah diakses. Sementara itu, pada tahap pengembangan, implementasi kode dilakukan untuk mewujudkan aplikasi berbasis *website* sesuai dengan desain yang telah direncanakan. Sistem ini juga

mengintegrasikan metode *Analytical Hierarchy Process* (AHP) untuk memberikan rekomendasi cerita yang dipersonalisasi berdasarkan preferensi pengguna.

a. Model Ontologi

Model ontologi yang digunakan dalam penelitian ini memiliki 6 kelas utama, yaitu Wayang, Waktu, Kegunaan, Peran, Watak, dan Cerita. Masing-masing kelas ini memiliki hubungan yang saling berkaitan melalui data properti yang mendukung pencarian semantik dalam sistem.Ontologi ini memiliki struktur grafis yang menggambarkan keterkaitan antar kelas yang terlihat pada gambar 2, di mana Wayang menjadi kelas utama yang menghubungkan berbagai kelas lainnya. Kelas Wayang berhubungan dengan Waktu, Kegunaan, Peran, Watak, dan Cerita, yang semuanya memiliki hubungan spesifik yang menunjukkan interaksi dan atribut terkait dalam konteks wayang.Grafik yang ditampilkan menunjukkan bagaimana berbagai aspek dalam dunia wayang saling terkait, membantu memperjelas hubungan antar konsep-konsep penting dalam kajian ini.



Gambar 2. Ontograf Ontologi Wayang Kulit Bali

b. Proses dengan Metode AHP

1. Struktur Hirarki



Gambar 3. *Hierarchy Process* Rekomendasi Cerita Wayang Kulit Bali

Bagan ini menunjukkan proses analisis menggunakan metode *Analytical Hierarchy Process* (AHP) untuk merekomendasikan cerita wayang kulit Bali berdasarkan empat kriteria utama: Tingkat Kompleksitas, Nilai Moral, Durasi Cerita, dan Popularitas. Cerita yang dinilai meliputi Ramayana, Mahabharata, Calonarang, Tantri, dan Cupak.

2. Matriks Perbandingan Berpasangan

Berikut adalah prosedur untuk menentukan perbandingan berpasangan menggunakan skala kualitas 1-9. Nilai eigen dan bobot prioritas kemudian dihitung pada kolom yang berjudul

p-ISSN: 2301-5373
e-ISSN: 2654-5101

"kriteria" berdasarkan nilai eigen tersebut, yang selanjutnya dikalikan dengan jumlah total nilai eigen.

Tabel 1. Matriks Perbandingan Berpasangan

	Tingkat Kompleksitas	Nilai Moral	Durasi Cerita	Popularitas
Tingkat Kompleksitas	**1**	**3**	**5**	**7**
Nilai Moral	**1/3**	**1**	**3**	**5**
Durasi Cerita	**1/5**	**1/3**	**1**	**3**
Popularitas	**1/7**	**1/5**	**1/3**	**1**

3. Menentukan Nilai Eigen

Langkah berikutnya adalah menghitung nilai eigen untuk setiap kriteria dengan mengubah matriks perbandingan berpasangan dalam tabel menjadi desimal, lalu menjumlahkan setiap baris. Setelah mendapatkan jumlah untuk setiap kolom, langkah selanjutnya adalah membagi setiap kriteria dengan jumlah pada kolom yang sesuai, kemudian menjumlahkan hasilnya untuk memperoleh nilai eigen.

Tabel 2. Nilai Eigen Kriteria

Kriteria	Eigen
Tingkat Kompleksitas	**0.5579**
Nilai Moral	**0.2633**
Durasi Cerita	**0.1219**
Popularitas	**0.0569**

4. Menguji Konsistensi

Langkah berikutnya adalah melakukan pengujian konsistensi terhadap bobot pada setiap kriteria yang ada. Proses ini bertujuan untuk memvalidasi bobot yang diberikan pada setiap kriteria. Uji konsistensi ini dilakukan dengan merujuk pada tabel *Random Consistency Index* (RI).

- Menentukan Nilai Eigen Maksimum (λ maks)

λ maks diperoleh dengan mengalikan hasil jumlah kolom yang didedikasikan dalam bentuk desimal dengan nilai eigen dari setiap kriteria.

$$\lambda \text{ maks} = \frac{(1\times0.5579) + (3\times0.2633) + (5\times0.1219) + (7\times0.0569)}{0.5579 + 0.2633 + 0.1219 + 0.0569} \quad (1)$$

$$= 4.118 \quad (2)$$

- Menghitung Indeks Konsistensi (CI)

Indeks Konsistensi (**CI**) dihitung menggunakan rumus berikut:

$$\mathrm{CI} = \frac{\lambda\text{maks} - \mathrm{n}}{\mathrm{n}-1} \quad (1)$$

Dimana n adalah jumlah kriteria (n = 4).

$$\mathrm{CI} = \frac{4.118\text{-}4}{4\text{-}1} \quad (2)$$

CI = 0.0395 (3)

- Menghitung Rasio Konsistensi (CR)
Terakhir, merupakan perhitungan Rasio Konsentrasi atau CR. Jika CR < 0,1 maka pembobotan pada setiap kriteria dapat dinyatakan konsisten. Rasio konsistensi dihitung dengan menggunakan **Random Consistency Index (RI)** untuk n = 4 (RI = 0.90):

$$\mathrm{CR} = \frac{\mathrm{CI}}{\mathrm{RI}} \quad (1)$$

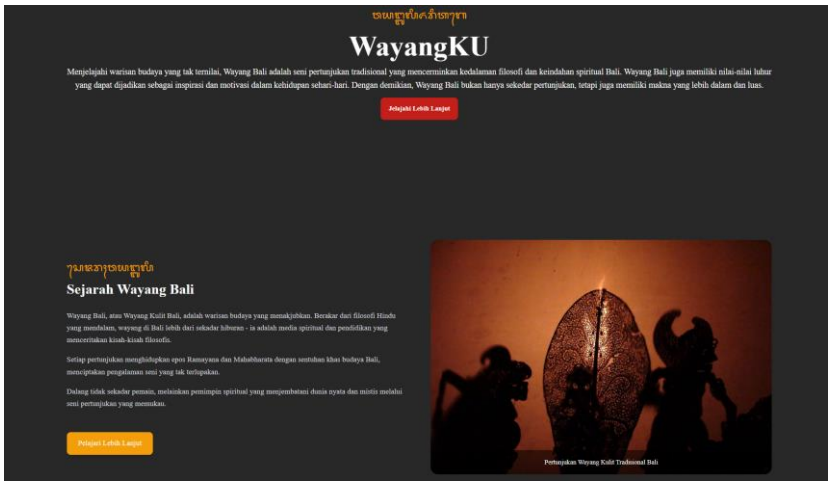
$$\mathrm{CR} = \frac{0.0395}{0.90} \quad (2)$$

CR= 0.0439 (3)

Berdasarkan nilai Rasio Konsistensi yang sebesar 0.0439, dapat disimpulkan bahwa setiap kriteria dianggap konsisten karena CR berada di bawah 0,1

c. Antarmuka Sistem

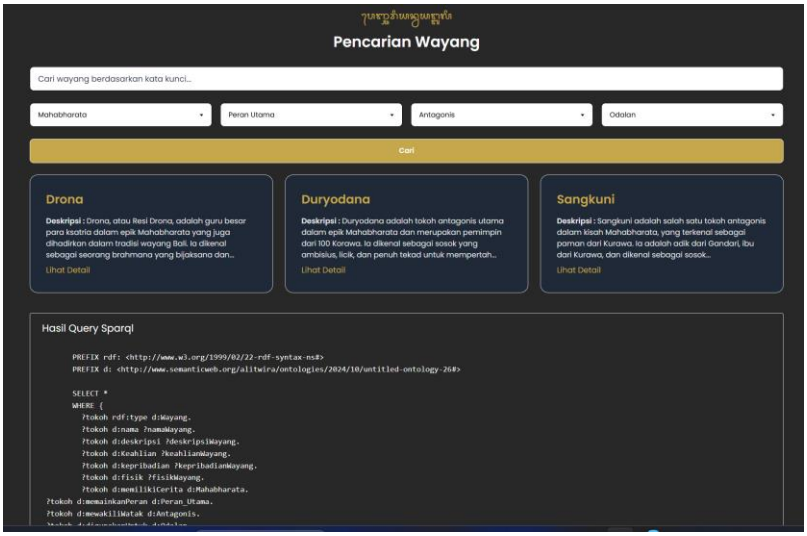
Sistem yang akan dibangun memiliki tampilan seperti pada gambar 4



Gambar 4. Tampilan *Home*

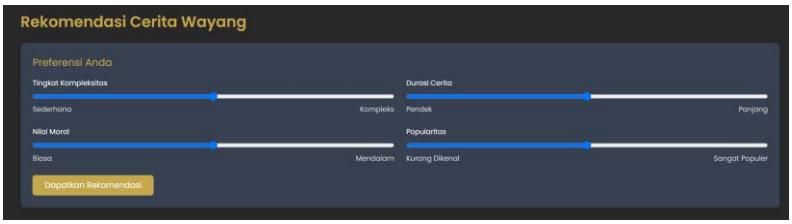
Pada Gambar 4 menampilkan halaman utama sistem Wayangku, yang memperkenalkan seni wayang kulit Bali melalui informasi sejarah, perkembangan budaya, dan peranannya dalam upacara adat. Halaman ini juga menyajikan gambar-gambar wayang kulit Bali, memberikan visualisasi yang menarik bagi pengguna.

p-ISSN: 2301-5373
e-ISSN: 2654-5101



Gambar 5. Tampilan Halaman Pencarian Wayang

Pada Gambar 5 menunjukkan halaman pencarian pada sistem, yang memungkinkan pengguna mencari tokoh wayang berdasarkan kategori seperti Cerita, Watak, dan Kegunaan. Hasil pencarian ditampilkan bersama deskripsi singkat, dengan data yang diambil melalui query SPARQL dari ontologi.



Gambar 6. Tampilan Rekomendasi Cerita Wayang

Pada Gambar 6 menampilkan fitur rekomendasi cerita wayang, di mana pengguna dapat memilih cerita berdasarkan preferensi seperti kompleksitas, durasi, nilai moral, dan popularitas.

d. Implementasi AHP Dalam Sistem *Website*

Metode *Analytical Hierarchy Process* (AHP) diintegrasikan dalam sistem web untuk menghasilkan rekomendasi cerita Wayang Kulit Bali berdasarkan preferensi pengguna. Empat kriteria utama yang digunakan yaitu tingkat kompleksitas, nilai moral, durasi, dan popularitas cerita.

Pengguna mengatur preferensinya melalui antarmuka berbasis *slider*, seperti ditunjukkan pada Gambar 6. Nilai dari slider dikonversi ke bentuk numerik dan otomatis membentuk matriks perbandingan berpasangan. Proses selanjutnya meliputi perhitungan bobot (eigen vector) dan pengujian konsistensi melalui nilai λ maks, *Consistency Index* (CI), dan *Consistency Ratio* (CR), Potongan kode berikut menggambarkan proses komputasi tersebut:

Tabel 3. *Source Code* Implementasi AHP Dalam Sistem

Source Code AHP

```
const pairwiseMatrix = [
  [1,   3,   5,   7],
  [1/3, 1,   3,   5],
  [1/5, 1/3, 1,   3],
  [1/7, 1/5, 1/3, 1]
];
```

```
const colSum = pairwiseMatrix[0].map((_, col) =>
  pairwiseMatrix.reduce((sum, row) => sum + row[col], 0)
);

const normalizedMatrix = pairwiseMatrix.map(row =>
  row.map((val, col) => val / colSum[col])
);

const eigenVector = normalizedMatrix.map(row =>
  row.reduce((sum, val) => sum + val, 0) / row.length
);

const lambdaMax = pairwiseMatrix
  .map((row, i) =>
    row.reduce((sum, val, j) => sum + val * eigenVector[j], 0)
/ eigenVector[i]
  )
  .reduce((sum, val) => sum + val, 0) / eigenVector.length;

const CI = (lambdaMax - 4) / (4 - 1);
const CR = CI / 0.90;

console.log(`CR = ${CR.toFixed(4)} → ${CR < 0.1 ? 'Konsisten' :
'Tidak Konsisten'}`);
```

Bobot akhir yang diperoleh digunakan untuk menghitung skor masing-masing cerita dalam basis data menggunakan pendekatan *weighted sum*, sehingga sistem dapat merekomendasikan cerita dengan skor tertinggi

3.2. Pengujian dan Evaluasi Sistem

a. Pengujian Validitas dan Reliabilitas

1. Uji Validitas

Uji validitas digunakan untuk menentukan apakah alat ukur (kuesioner) dapat mengekspresikan apa yang diinginkan peneliti. Berdasarkan hasil uji validitas, semua item variabel usability valid karena nilai r hitung lebih besar dari r tabel (0,349).

Tabel 4. Hasil Uji Validitas

Variabel	Kode	r Hitung	R Tabel	Keterangan
Learnability (A)	(A1)	0,502	0,349	Valid
	(A2)	0,566	0,349	Valid
	(A3)	0,580	0,349	Valid
	(A4)	0,567	0,349	Valid
	(A5)	0,784	0,349	Valid
Memorability (B)	(B1)	0,501	0,349	Valid
	(B2)	0,519	0,349	Valid
	(B3)	0,640	0,349	Valid
Efficiency (C)	(C1)	0,687	0,349	Valid
	(C2)	0,616	0,349	Valid
	(C3)	0,599	0,349	Valid
Errors (D)	(D1)	0,421	0,349	Valid
	(D2)	0,376	0,349	Valid
	(D3)	0,515	0,349	Valid
Satisfaction (E)	(E1)	0,357	0,349	Valid
	(E2)	0,477	0,349	Valid
	(E3)	0,369	0,349	Valid

p-ISSN: 2301-5373
e-ISSN: 2654-5101

	(E4)	0,427	0,349	Valid

2. Uji Reliabilitas

Uji reliabilitas mengukur konsistensi jawaban responden. Hasil uji menunjukkan nilai *Cronbach's Alpha* sebesar 0,809, yang lebih besar dari 0,70, sehingga kuesioner dinyatakan reliabel.

Tabel 5. Hasil Uji Reliabilitas

Cronbach's Alpha	r Tabel	Keterangan
0,809	0,70	Reliabel

b. Pengujian Kegunaan (*Usability Testing*)

Pengujian kegunaan dilakukan menggunakan kuesioner berbasis skala Likert yang diisi oleh 12 responden. Hasil evaluasi menunjukkan bahwa *website* memiliki performa baik dalam hal *Learnability* (rata-rata 4,25) dan Satisfaction (4,25), menandakan bahwa pengguna mudah mempelajari dan merasa puas dengan *website*. Indikator Memorability mendapatkan skor 4,14, sedangkan *Efficiency* dan *Errors* memerlukan perbaikan dengan nilai masing-masing 4,06 dan 3,92. Secara keseluruhan, *website* menunjukkan hasil positif namun masih perlu pengembangan lebih lanjut.

Tabel 6. Rata-rata Total Setiap Indikator Kegunaan

Variabel	Kode	Rata-rata	Rata-rata/Keterangan
Learnability (A)	(A1)	4,25	4.30
	(A2)	4,42	
	(A3)	4,17	
	(A4)	4,42	
	(A5)	4,25	
Memorability (B)	(B1)	4,25	4.14
	(B2)	4,00	
	(B3)	4,17	
Efficiency (C)	(C1)	4,08	4.06
	(C2)	4,17	
	(C3)	3,92	
Errors (D)	(D1)	4,17	3.92
	(D2)	3,67	
	(D3)	3,92	
Satisfaction (E)	(E1)	4,25	4.25
	(E2)	4,08	
	(E3)	4,42	
	(E4)	4,25	

4. Kesimpulan

Sistem dokumentasi digital untuk Wayang Kulit Bali berhasil menyediakan informasi terstruktur dengan pendekatan ontologi, mencakup 6 kelas utama seperti Wayang, Waktu, Kegunaan, Peran, Watak, dan Cerita. Penggunaan *Analytical Hierarchy Process* (AHP) untuk rekomendasi cerita berbasis preferensi pengguna menunjukkan hasil yang konsisten dengan Rasio Konsistensi (CR) sebesar 0.0439, menunjukkan konsistensi yang baik. Skor *Learnability* dan *Satisfaction* masing-masing 4.25 menunjukkan kemudahan penggunaan dan kepuasan pengguna. Namun, skor *Efficiency* (4.06) dan *Errors (3.92*) menunjukkan area yang perlu perbaikan. Pengembangan lebih lanjut diperlukan untuk

meningkatkan kecepatan akses dan mengurangi kesalahan guna memastikan kelanjutan pelestarian budaya Wayang Kulit Bali di era digital.

Referensi

[1] C. Pramartha, I. Koten, I. G. N. A. C. Putra, I. W. Supriana, and I. W. Arka, "Pengembangan Sistem Dokumentasi Melalui Pendekatan Ontologi untuk Praktek Budaya Bali," J. Nas. Pendidik. Tek. Inform. (JANAPATI), vol. 11, no. 3, pp. 259–268, 2022.
[2] Pramartha, I. S. Jayadi, and M. B. Atmaja, "Pengembangan Sistem Rekomendasi Melalui Pendekatan Web Semantik dan Simple Additive Weighting (SAW)," SINTECH, vol. 5, pp. 139–153, 2022
[3] M. A. Raharja and I. L. Wijayakusuma, "Implementasi Metode Analitical Hierarcy Proses (AHP) pada Sistem Penentuan Kredit Lembaga Perkreditan Desa (LPD)," J. Teknol. Inf. dan Komput. (JuTIK), vol. 6, no. 2, pp. 142–150, 2020.
[4] L. Di Angelo002C P. Di Stefano, L. Fratocchi, and A. Marzola, "An AHP-based method for choosing the best 3D scanner for cultural heritage applications," J. Cult. Herit., vol. 34, pp. 45–54, 2018.
[5] M. Lasorella and E. Cantatore, "CityGML-based model for the recovery and management of built cultural heritage: a systematic review towards digitalized processes," City Territ. Archit., vol. 12, no. 1, p. 10, 2025.
[6] M. A. Raharja, S. Purnawati, I. P. G. Adiatmika, I. N. Adiputra, and I. B. A. Swamardika, "Usability Analysis of Tembang Sekar Alit Learning (SekARAI) Applications Using the Human Computer Interaction (HCI) Model in Bali Students," in Proc. 2nd Asia Pacific Int. Conf. Ind. Eng. Oper. Manag., 202.
[7] A. R. Fadillah and I. K. A. Mogi, "Sistem Pendukung Keputusan Pemilihan Lokasi Dealer Motor Honda di Bali Berbasis Ontologi," JELIKU (J. Elektron. Ilmu Komput. Udayana), vol. 12, no. 1, pp. 23–30, Jan. 2023.
[8] M. A. Raharja and I. K. G. Suhartana, "Interaksi Manusia Komputer (IMK) Berbasis Antarmuka dan Pengalaman Pengguna", Denpasar: Pustaka Larasan, 2024.
[9] I. G. N. A. Wira Putra et al., "Perancangan Ontologi Semantik: Representasi Digital Tokoh Wayang Bali," JELIKU (J. Elektron. Ilmu Komput. Udayana), vol. 13, no. 3, pp. 715–726, Feb. 2025.

Jurnal Elektronik Ilmu Komputer Udayana
Volume 14, No 3. February 2026

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Klasifikasi Kategori Cerita Pendek Menggunakan XGBoost dengan Seleksi Fitur Chi-Square

M. Faisal Afandi[a1], Ngurah Agus Sanjaya ER[a2], Putu Gede Hendra Suputra[a3], Luh Arida Ayu Rahning Putri[a4]

[a]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus Unud, Jimbaran, Bali, 80361, Indonesia
[1]faisalafandi12345@gmail.com
[2]agus_sanjaya@unud.ac.id
[3]hendra.suputra@unud.ac.id
[4]rahningputri@unud.ac.id

Abstract

Text classification is a key challenge in natural language processing, particularly in categorizing texts by genre. This study aims to classify Indonesian short stories into three genres: romance, horror, and religion. Two ensemble machine learning algorithms, XGBoost and Random Forest, were employed. Prior to training, the data underwent preprocessing and feature extraction using TF-IDF, followed by feature selection with the Chi-Square method to enhance relevance and efficiency. Models were trained with various hyperparameter configurations and validated using 5-Fold Cross Validation. Experimental results show that Chi-Square selection significantly improved performance by reducing training time and maintaining accuracy. For instance, XGBoost with 500 features required 102 seconds compared to 427 seconds without Chi-Square, while Random Forest required only 5 seconds compared to 14.8 seconds. On the test set, XGBoost achieved the best results with 90% precision, 91% recall, and 90% F1-Score, outperforming Random Forest which obtained 86% precision, 87% recall, and 85% F1-Score. These results confirm that XGBoost generalizes better on unseen data..

Keywords: *Text Classification, Short Stories, XGBoost, Random Forest, Chi-Square, Ensemble Learning.*

1. Pendahuluan

Klasifikasi teks merupakan salah satu tantangan utama dalam *natural language processing* (NLP). Tidak seperti data numerik yang memiliki struktur yang jelas, teks memiliki karakteristik tidak terstruktur dan ambiguitas semantik, sehingga menyulitkan proses klasifikasi secara otomatis [1]. Salah satu permasalahan yang diangkat pada penelitian ini yaitu pengelompokan cerita pendek ke dalam suatu genre tertentu, seperti romantis, horor, dan agama, yang masing-masing memiliki ciri khas bahasa dan gaya naratif yang bisa saling tumpang tindih.

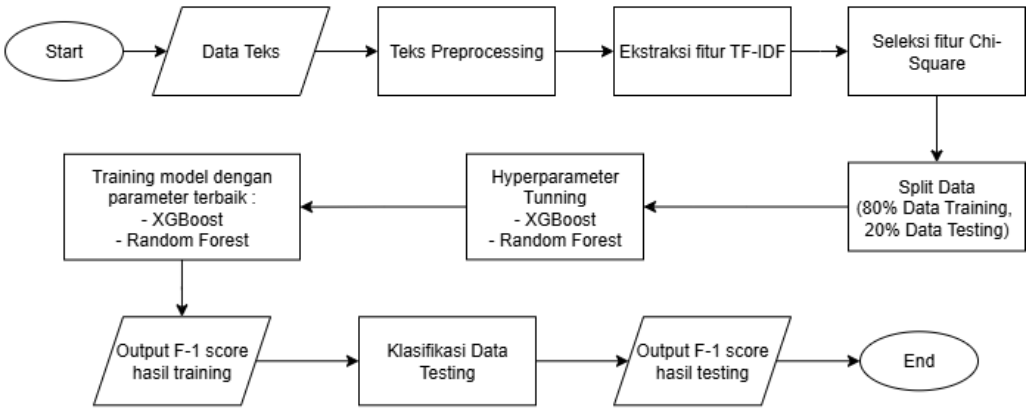
Permasalahan klasifikasi pada penelitian ini termasuk dalam kategori *multiclass classification*, di mana suatu cerita pendek akan diklasifikasikan ke dalam salah satu kategori dari banyak kategori secara bersamaan. Kompleksitas ini menuntut pemilihan algoritma yang tepat, baik dari sisi akurasi maupun efisiensi. Penelitian-penelitian sebelumnya telah mengeksplorasi berbagai pendekatan untuk klasifikasi teks, seperti penggunaan *Naïve Bayes*, *Support Vector Machine* (SVM), hingga *deep learning* berbasis *Long Short-Term Memory* (LSTM). Namun, studi yang secara spesifik membandingkan algoritma *ensemble learning* seperti XGBoost dan Random Forest dalam konteks klasifikasi genre cerita pendek masih terbatas.

XGBoost merupakan salah satu algoritma *gradient boosting* yang terkenal dengan performa tinggi dalam kompetisi *machine learning* karena efisiensi dan akurasinya yang tinggi [2]. Di sisi lain, Random Forest menggunakan pendekatan *bagging* dan cenderung lebih stabil dalam menghindari *overfitting* pada data pelatihan. Perbandingan keduanya menjadi penting mengingat keduanya merupakan pendekatan *ensemble*, tetapi dengan strategi pembentukan model yang berbeda. Dalam penelitian ini juga mengeksplorasi pengaruh seleksi fitur menggunakan metode *Chi-Square*, yang bertujuan untuk meningkatkan relevansi fitur dan memperbaiki performa model klasifikasi.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk mengevaluasi dan membandingkan performa XGBoost dan Random Forest dalam tugas klasifikasi teks cerita pendek berbahasa Indonesia berdasarkan genre. Evaluasi dilakukan berdasarkan metrik *F1-Score* dengan mempertimbangkan berbagai kombinasi *hyperparameter* serta pengaruh dari seleksi fitur *Chi-Square*.

2. Metode Penelitian

Penelitian ini dilakukan melalui serangkaian tahapan yang meliputi pengumpulan data, *preprocessing* teks, ekstraksi fitur menggunakan TF-IDF, seleksi fitur dengan Chi-Square, pelatihan model, dan evaluasi performa menggunakan metrik F1-score. Gambar 1 menunjukkan alur proses penelitian secara umum.



Gambar 1. Alur Penelitian

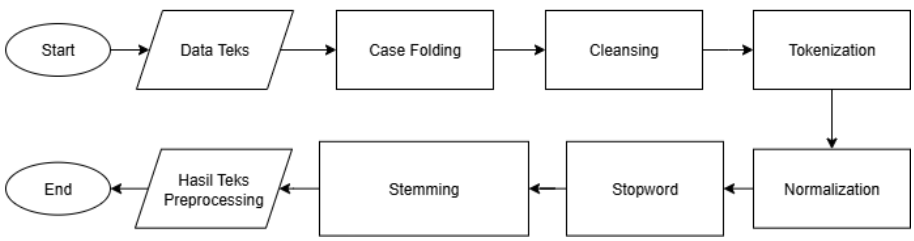
2.1. Pengumpulan Data

Data yang digunakan dalam penelitian ini bersumber dari situs *cerpenmu.com*, yang berisi berbagai cerita pendek berbahasa Indonesia. Data yang digunakan sebanyak 632 cerita pendek, dan dipilih hanya cerita pendek yang memiliki satu label genre, yaitu romantis, horor, dan agama. Data ini mewakili tiga kelas dalam skenario *multiclass classification*.

2.2. Pra-pemrosesanTeks

Sebelum dilakukan proses ekstraksi fitur, teks cerita pendek akan diproses terlebih dahulu melalui beberapa tahapan *preprocessing* seperti pada Gambar 2. Tahapan ini mencakup:

- *Case folding*: mengubah seluruh huruf menjadi huruf kecil.
- *Cleansing*: menghapus karakter khusus dan tanda baca yang tidak diperlukan.
- *Tokenisasi*: memisahkan kalimat menjadi token kata.
- *Normalisasi*: mengubah kata tidak baku menjadi kata baku, seperti "yg" menjadi "yang".
- *Stopword removal*: menghapus kata-kata umum yang tidak memiliki nilai informasi.
- *Stemming*: mengubah kata menjadi bentuk dasar.

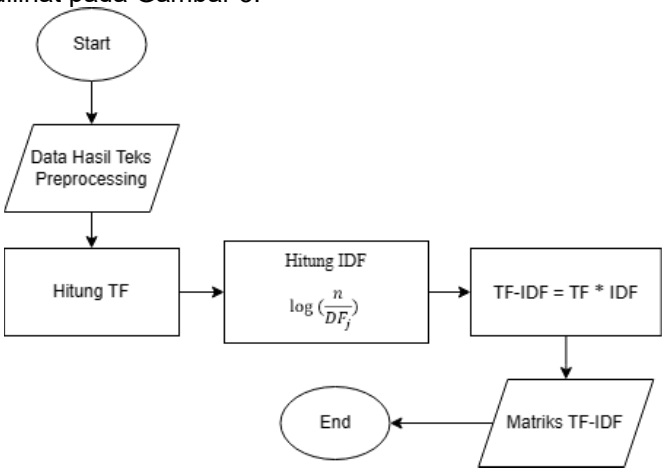


Gambar 2. Alur Pra-pemrosesanTeks

p-ISSN: 2301-5373
e-ISSN: 2654-5101

2.3. Ekstraksi Fitur TF-IDF

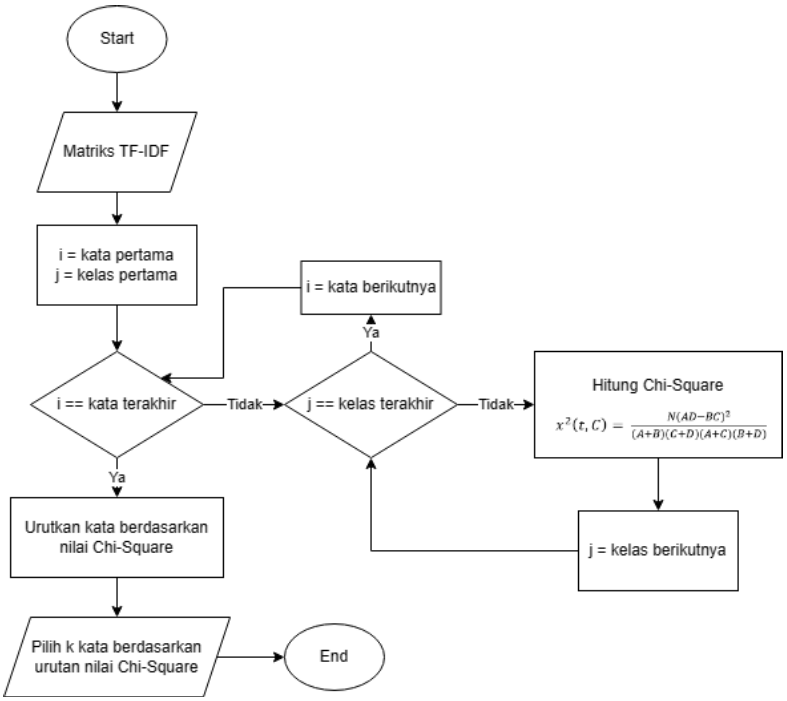
Setelah teks *preprocessing* dilakukan, tahap berikutnya adalah ekstraksi fitur menggunakan metode TF-IDF (*Term Frequency-Inverse Document Frequency*). Metode ini memberikan bobot terhadap kata-kata dalam dokumen berdasarkan frekuensi kemunculannya dan seberapa unik kata tersebut di seluruh dokumen [3]. Hasil dari proses ini adalah matriks TF-IDF yang merepresentasikan setiap dokumen dalam bentuk vektor numerik. Proses ekstraksi fitur TF-IDF dapat dilihat pada Gambar 3.



Gambar 3. Ekstraksi Fitur TF-IDF

2.4. Seleksi Fitur Chi-Square

Untuk meningkatkan performa klasifikasi dan mengurangi dimensi fitur, dilakukan seleksi fitur menggunakan metode Chi-Square. Teknik ini mengukur seberapa kuat hubungan antara setiap kata (fitur) dengan kelas target. Fitur dengan nilai Chi-Square tertinggi dipilih sebagai fitur yang paling relevan. Proses seleksi fitur Chi-Square dapat dilihat pada Gambar 4.

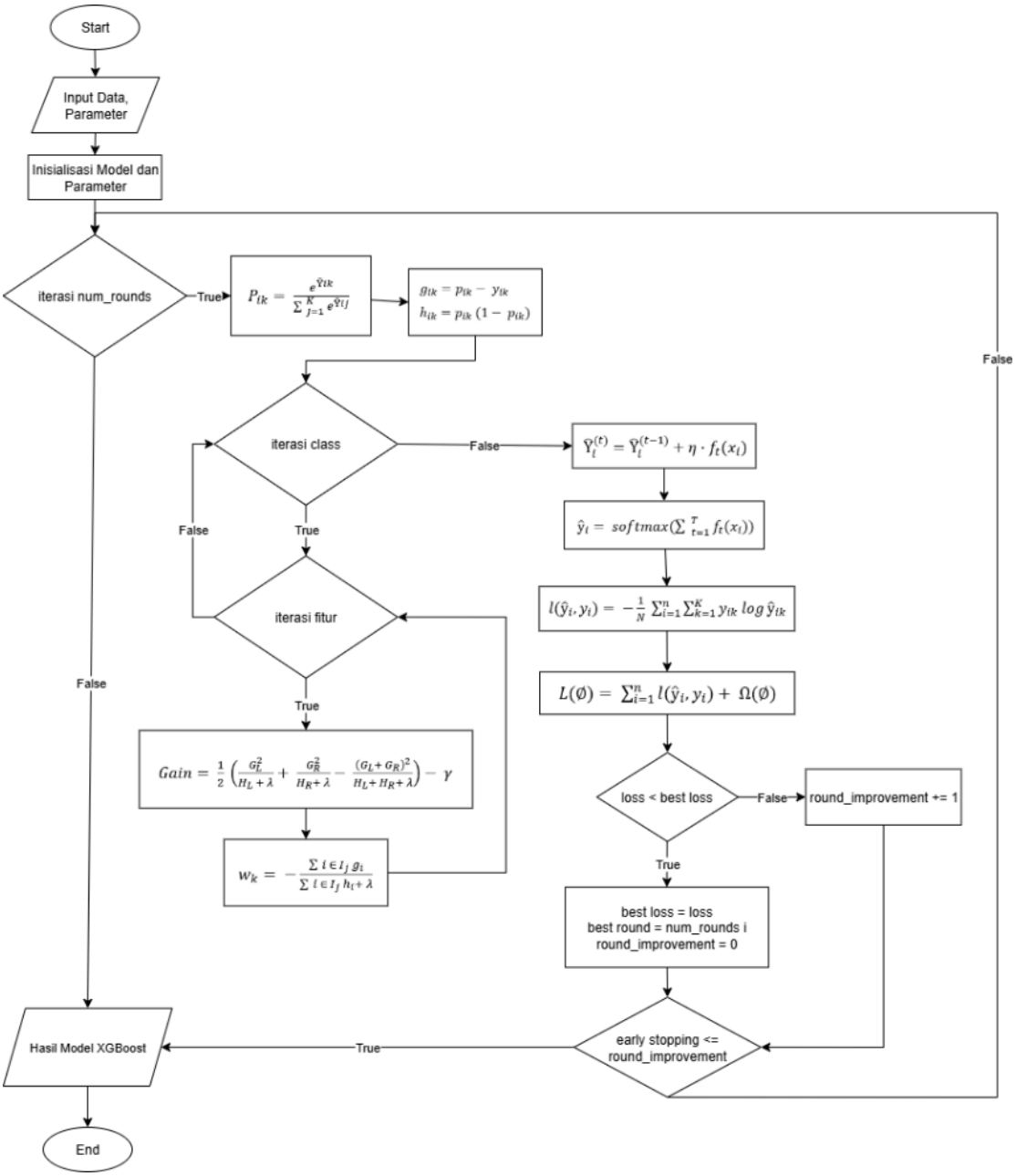


Gambar 4. Seleksi Chi-Square

2.5. Klasifikasi Teks XGBoost

XGBoost membangun model *gradient boosting decision tree* yang dioptimalkan secara bertahap untuk memperbaiki kesalahan prediksi dari model sebelumnya. Proses pelatihan

dilakukan dengan membagi data menjadi data latih dan data uji, serta menerapkan teknik *cross-validation*.



Gambar 5. Alur Pelatihan XGBoost

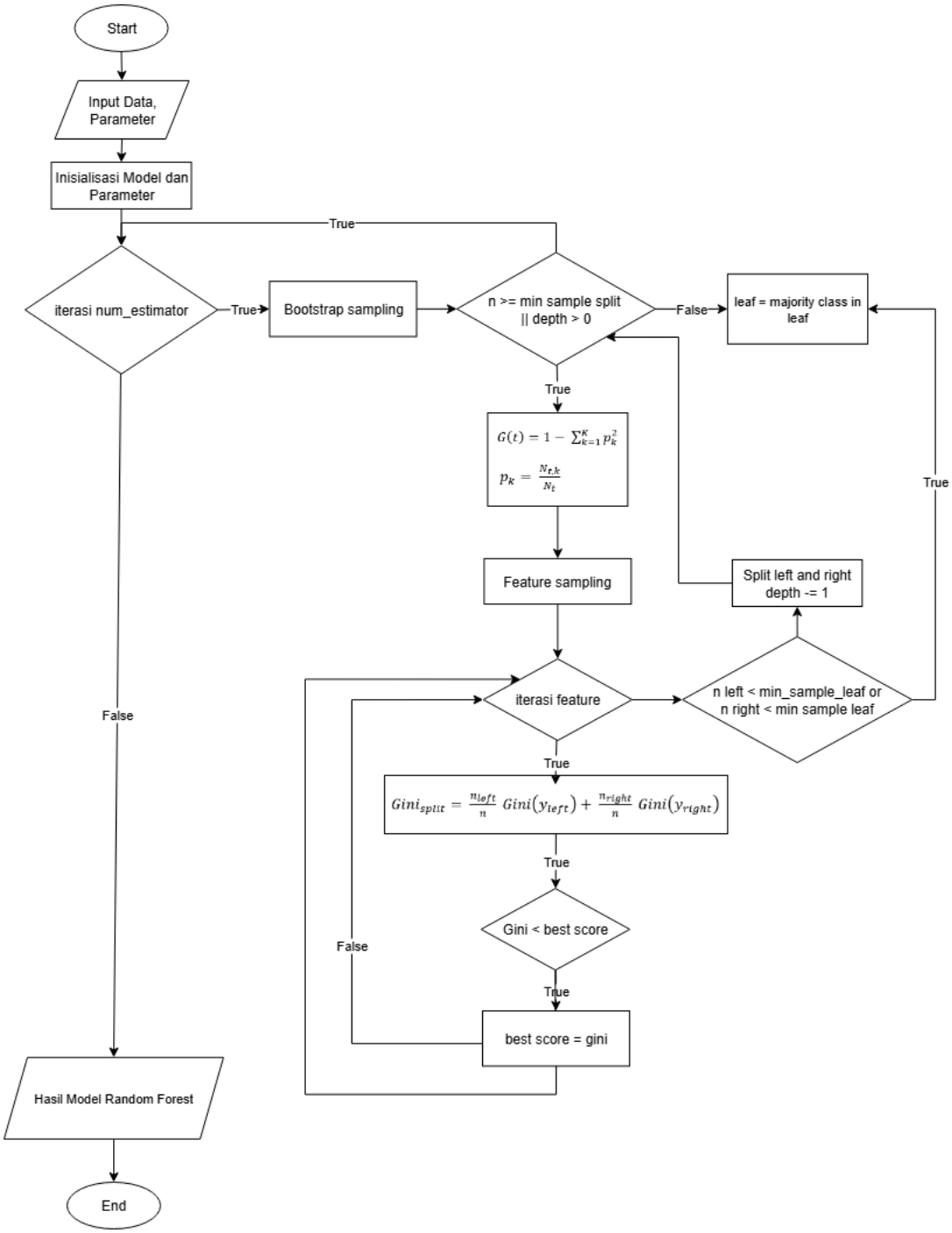
Pada gambar 5 merupakan proses pelatihan model XGBoost untuk kasus klasifikasi multiclass menggunakan pendekatan *second-order Taylor expansion*. Proses dimulai dengan memasukkan data dan parameter model, kemudian menginisialisasi nilai prediksi awal. Pada setiap iterasi *boosting round*, model menghitung probabilitas kelas menggunakan fungsi *softmax*, lalu menghitung nilai gradien dan hessian untuk setiap data dan kelas. Selanjutnya, dilakukan pencarian *split* terbaik dengan mengiterasi setiap kelas dan fitur, di mana kualitas *split* diukur menggunakan perhitungan *gain*. Bobot prediksi pada setiap *leaf* ditentukan dari rasio negatif jumlah gradien terhadap jumlah hessian ditambah parameter regularisasi. Model kemudian memperbarui prediksi dengan menambahkan kontribusi dari pohon baru yang dibangun. Setelah itu, *loss function* dihitung menggunakan *softmax cross-entropy* yang dikombinasikan dengan regularisasi, lalu dibandingkan dengan nilai *loss* terbaik sebelumnya untuk menentukan apakah terjadi perbaikan. Jika tidak ada perbaikan selama sejumlah iterasi

p-ISSN: 2301-5373
e-ISSN: 2654-5101

tertentu, proses dihentikan melalui mekanisme *early stopping*. Pelatihan berakhir ketika jumlah iterasi maksimum tercapai atau kondisi *early stopping* terpenuhi [4].

2.6. Klasifikasi Teks Random Forest

Berbeda dengan XGBoost, pada Random Forest model dilatih secara independen menggunakan teknik *bagging* untuk meningkatkan kestabilan dan akurasi prediksi. Setiap pohon dilatih pada subset data latih yang diambil secara acak dengan pengembalian (*bootstrap sample*), serta menggunakan subset fitur yang dipilih secara acak pada setiap percabangan node (*feature subsampling*). Hal ini membuat setiap pohon memiliki variasi struktur sehingga mengurangi risiko *overfitting* [5].



Gambar 6. Alur Training Random Forest

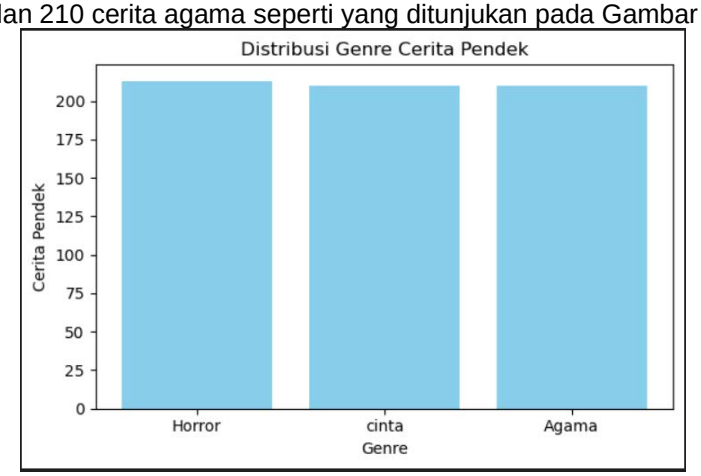
Pada gambar 6 ditunjukkan proses pelatihan model Random Forest untuk klasifikasi multiclass. Proses dimulai dengan memasukkan data dan parameter model, kemudian menginisialisasi model dengan jumlah pohon (*num_estimators*) yang akan dibangun. Pada setiap iterasi, dilakukan *bootstrap sampling* untuk membuat subset data latih acak yang digunakan membangun pohon. Proses pembentukan pohon dimulai dari root dengan memeriksa apakah jumlah sampel pada node memenuhi syarat minimum pemisahan (*min_sample_split*) dan kedalaman *tree* belum mencapai batas (*depth*). Jika syarat tidak terpenuhi, node menjadi *leaf* dengan nilai prediksi kelas mayoritas. Jika syarat terpenuhi, dihitung nilai *Gini Impurity* untuk mengukur ketidakmurnian data pada node. Selanjutnya dilakukan *feature sampling* untuk memilih subset fitur secara acak, kemudian diiterasi setiap fitur untuk mencari *split* terbaik berdasarkan nilai *Gini split*. *Split* terbaik dipilih jika menghasilkan nilai *Gini* yang lebih kecil dari skor terbaik sebelumnya. Node kemudian dipecah menjadi cabang kiri dan kanan, dan proses ini diulang secara rekursif hingga memenuhi kriteria penghentian (*stopping criteria*). Setelah semua pohon terbentuk, prediksi akhir ditentukan berdasarkan *majority voting* dari seluruh *tree* dalam ensemble.

2.7. Evaluasi Sistem

Evaluasi model dilakukan untuk mengukur performa klasifikasi berdasarkan metrik *F1-score*, yang mempertimbangkan nilai presisi dan *recall*. Evaluasi dilakukan pada data uji setelah model dilatih menggunakan *hyperparameter* terbaik

3. Result and Discussion

Untuk mengevaluasi kedua algoritma XGBoost dan Random Forest dilakukan pelatihan model menggunakan data yang sama. Dataset yang digunakan terdapat 632 cerita pendek dengan tiga kategori genre, yaitu cinta, horor, dan agama. Setiap cerita hanya mewakili satu genre, sehingga cerita ambigu atau tidak jelas kategorinya dihapus. Data disimpan dalam file Excel (.xlsx) dengan kolom Judul, Isi Cerita, dan Genre, serta diatur seimbang untuk menghindari bias, yaitu 213 cerita horor, 210 cerita cinta, dan 210 cerita agama seperti yang ditunjukan pada Gambar 7.



Gambar 7. Distribusi Data

3.1. Hyperparameter Tunning

a. Hyperparameter Tunning pada Algoritma XGBoost
Pada algoritma XGBoost dilakukan hyperparameter tunning menggunakan kombinasi yang tercantum pada Tabel 1.

Tabel 1. Hyperparameter XGBoost

Hyperparameter	Parameter
num_rounds	50, 100
learning_rate	0.1, 0.3
depth	2, 4, 6

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Nilai *subsample* = 0.8 digunakan untuk *row* dan *column sampling* secara acak pada setiap pohon. Kombinasi hyperparameter ini memberikan akurasi dan waktu pelatihan terbaik pada masing-masing konfigurasi seleksi fitur Chi-Square dengan variasi jumlah fitur: 500, 1000, 1500, dan tanpa seleksi (5000 fitur tf-idf terbaik). Evaluasi menggunakan skema 5-*Fold Cross Validation*. Detail proses hyperparameter tuning terbaik pada XGBoost dapat dilihat pada Tabel 2.

Tabel 2. Akurasi Hyperparameter Terbaik XGBoost

Jumlah Fitur	*Hyperparameter* Terbaik	*Fold* (1-5)	Rata-rata Akurasi	Rata-rata Waktu
500 Fitur Chi-Square	*num_rounds* = 100, *learning_rate* = 0,1, *depth* = 4	87,13%	88,32%	102,60 detik
		87,13%		
		88,12%		
		84,16%		
		95,05%		
1000 Fitur Chi-Square	*num_rounds* = 50, *learning_rate* = 0,1, *depth* = 6	87,13%	88,32%	104,92 detik
		87,13%		
		86,14%		
		88,12%		
		93,07%		
1500 Fitur Chi-Square	*num_rounds* = 50, *learning_rate* = 0,1, *depth* = 6	86,14%	87,33%	175,09 detik
		86,14%		
		86,14%		
		85,15%		
		93,07%		
Tanpa Chi-Square	*num_rounds* = 100, *learning_rate* = 0,3, *depth* = 4	88,12%	87,92%	427,118 detik
		88,12%		
		89,11%		
		85,15%		
		89,11%		

Pada Tabel 2 menunjukkan bahwa penggunaan seleksi fitur Chi-Square dapat sedikit meningkatkan akurasi model dan semakin sedikit fitur yang digunakan semakin cepat waktu pelatihan yang diperlukan.

b. Hyperparameter Tunning pada Algoritma Random Forest
Pada algoritma Random Forest dilakukan hyperparameter tunning menggunakan kombinasi yang tercantum pada Tabel 3.

Tabel 3. Hyperparameter Random Forest

Hyperparameter	Parameter
n_estimators	100, 200
max_depth	4, 6
min_samples_leaf	1, 3
max_features	sqrt, log2

Pada algoritma Random Forest, eksperimen dilakukan dengan kombinasi hyperparameter pada Tabel 4 dan metode seleksi fitur Chi-Square dengan variasi jumlah fitur: 500, 1000, 1500, dan tanpa seleksi (5000 fitur tf-idf terbaik). Evaluasi menggunakan skema 5-*Fold Cross Validation* untuk memperoleh akurasi tiap fold dan rata-rata keseluruhan.

Tabel 4. Akurasi Hyperparameter Terbaik Random Forest

Jumlah Fitur	*Hyperparameter* Terbaik	*Fold* (1-5)	Rata-rata Akurasi	Rata-rata Waktu
500 Fitur Chi-Square	*n_estimators* = 200, *max_depth* = 6, *min_samples_leaf* = 3, *max_features* = sqrt	85,15%	87,33%	4,97 detik
		84,16%		
		88,12%		
		90,10%		
		89,11%		
1000 Fitur Chi-Square	*n_estimators* = 100, *max_depth* = 6, *min_samples_leaf* = 3, *max_features* = sqrt	85,15%	87,33%	5,05 detik
		85,15%		
		88,12%		
		90,10%		
		88,12%		
1500 Fitur Chi-Square	*n_estimators* = 100, *max_depth* = 6, *min_samples_leaf* = 3, *max_features* = sqrt	83,17%	86,73%	6,19 detik
		86,14%		
		87,13%		
		88,12%		
		89,11%		
Tanpa Chi-Square	*n_estimators* = 100, *max_depth* = 6, *min_samples_leaf* = 3, *max_features* = sqrt	78,22%	81,58%	14,84 detik
		77,23%		
		85,15%		
		87,13%		
		80,20%		

Pada tabel 4 menunjukan bahwa model Random Forest juga mengalami peningkatan performa dan proses waktu pelatihan yang signifikan dengan penggunaan seleksi fitur Chi-Square.

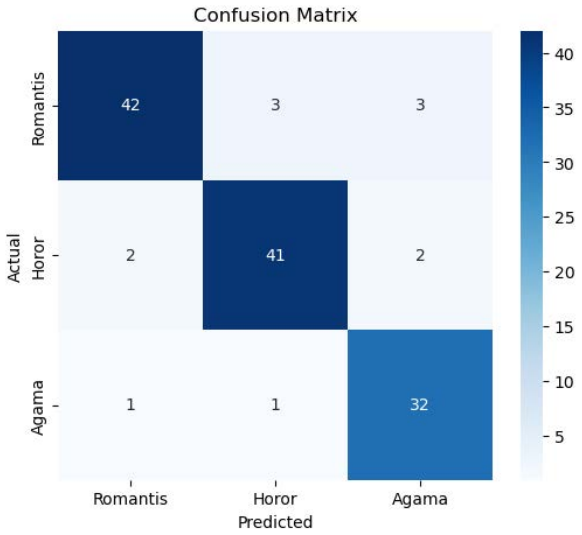
3.2. Evaluasi

Evaluasi dilakukan pada model XGBoost dan Random Forest menggunakan parameter terbaik hasil pencarian hyperparameter. XGBoost dengan num_rounds=100, learning_rate=0.1, depth=4. Sementara itu, Random Forest menggunakan parameter n_estimators=200, max_depth=6, min_samples_leaf=3, max_features=sqrt

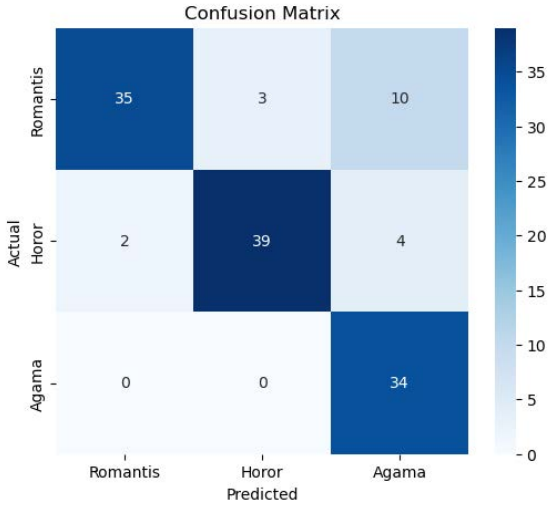
Tabel 5. Evaluasi XGboost dan Random Forest dengan Chi-Square

Algoritma	Precision	Recall	F1-Score
XGBoost	90 %	91 %	90 %
Random Forest	86 %	87 %	85 %

p-ISSN: 2301-5373
e-ISSN: 2654-5101



Gambar 8. Confusion Matrix XGBoost



Gambar 9. Confusion Matrix Random Forest

Berdasarkan Tabel 5, XGBoost mencatat F1-Score 90%, lebih unggul dibandingkan Random Forest yang mencapai 85%. Hal ini sejalan dengan karakteristik metode *boosting* yang lebih efektif dalam mengurangi bias. Hasil pada Tabel 2 dan 4 juga menunjukkan bahwa seleksi fitur Chi-Square dapat meningkatkan performa dan efisiensi model. XGBoost dengan 500 fitur mempertahankan akurasi 88% dengan waktu pelatihan 102 detik, jauh lebih cepat dibanding tanpa seleksi. Random Forest juga menunjukkan peningkatan serupa dengan akurasi 87% dan waktu pelatihan sekitar 5 detik. Secara keseluruhan, seleksi fitur membantu menurunkan kompleksitas tanpa mengorbankan akurasi, dengan XGBoost unggul dalam akurasi dan generalisasi, sementara Random Forest lebih stabil dan cepat.

4. Conclusion

Berdasarkan hasil penelitian dan implementasi yang telah dilakukan, dapat disimpulkan beberapa hal sebagai berikut:

1. Kombinasi hyperparameter berpengaruh signifikan terhadap performa algoritma. XGBoost dengan konfigurasi optimal mencapai F1-Score 90%, sementara Random Forest hanya 85%, sehingga menunjukkan keunggulan pendekatan boosting dalam menangkap pola kompleks pada data teks.
2. Seleksi fitur Chi-Square terbukti meningkatkan efisiensi tanpa mengorbankan akurasi. Pada XGBoost, 500 fitur cukup menjaga akurasi 88% dengan waktu pelatihan lebih singkat (102

detik dibanding 427 detik). Random Forest juga memperoleh waktu pelatihan lebih cepat dengan performa relatif stabil.

References

[1] I. G. A. P. Arimbawa and N. A. S. ER, "Penerapan Metode Adaboost Untuk Multi-Label Classification Pada Dokumen Teks," *Jurnal Elektronik Ilmu Komputer Udayana,* vol. 9, no. 1, pp. 127-140, 2020.

[2] I. R. Hendrawan, "PERBANDINGAN ALGORITMA NAÏVE BAYES, SVM DAN XGBOOST DALAM KLASIFIKASI TEKS SENTIMEN MASYARAKAT TERHADAP PRODUK LOKAL DI INDONESIA," *Jurnal TRANSFORMASI,* vol. 18, no. 1, pp. 1-8, 2022.

[3] A. Nurhadi, "Klasifikasi Berita Menggunakan Metode Support Vector Machine," *Jurnal Nasional Komputasi dan Teknologi Informasi,* vol. 5, no. 2, pp. 269-278, 2022.

[4] D. J. Putri, M. Dwifebri and A. , "Text Classification of Indonesian Translated Hadith Using XGBoost Model and Chi-Square Feature Selection," *Building of Informatics, Technology and Science,* vol. 4, no. 4, pp. 1732-1738, 2023.

[5] N. Husin, "Komparasi Algoritma Random Forest, Naïve Bayes, dan Bert Untuk Multi-Class Classification Pada Artikel Cable News Network (CNN)," *Jurnal Esensi Infokom Jurnal Esensi Sistem Informasi dan Sistem Komputer,* vol. 7, no. 1, pp. 75-84, 2023.

Jurnal Elektronik Ilmu Komputer Udayana
Volume 14, No 3. February 2026

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Redesign UI/UX Aplikasi *Property Management* Berbasis *User-Centered Design* Dengan Evaluasi *Heuristic Evaluation* dan *System Usability Scale*

Matthew Novan Sidharta[a1], Luh Arida Ayu Rahning Putri[a2], I Ketut Gede Suhartana[a3], I Gusti Ngurah Anom Cahyadi Putra[a4]

[a]Informatics, Udayana University
Badung, Indonesia
[1]matthewnovansidharta31@gmail.com
[2]rahningputri@unud.ac.id
[3]ikg.suhartana@unud.ac.id
[4]anom.cp@unud.ac.id

Abstrak

Perusahaan *property management* menjadi organisasi yang memiliki peran krusial untuk membantu para investor lokal maupun asing yang berinvestasi di bidang properti. Untuk memperlancar proses manajemen, maka perusahaan memanfaatkan aplikasi *property management*. Tentunya, aplikasi ini perlu dirancang agar memiliki UI/UX yang baik. Dalam penelitian ini, penulis melakukan *redesign* aplikasi *property management* berbasis UCD menggunakan metode evaluasi HE, SUS, dan pengujian 5 aspek *usability* yang masing-masing dilakukan pada tahap awal (sebelum *redesign*) dan akhir (setelah *redesign*). Terhitung ada 64 bobot masalah pada pengujian HE tahap awal. Sedangkan, skor rata-rata SUS tahap awal adalah sebesar 71. Setelah dilakukan *redesign* berdasarkan masukan dari para evaluator, maka total bobot masalah pada HE tahap akhir mengalami penurunan ke angka 30 bobot masalah. Sedangkan, skor rata-rata SUS mengalami peningkatan ke angka 79,25. Untuk 5 aspek *usability* mengalami peningkatan dengan hasil *learnability* (64% menjadi 86%)*, memorability* (64% menjadi 88%)*, efficiency* (61% menjadi 89%)*, error* (tingkat kejarangan mengalami *error* meningkat dari 54% menjadi 77%)*, satisfaction* (61% menjadi 86%).

Kata Kunci*: Property Management*, *UI/UX*, *Heuristic Evaluation, System Usability Scale, Redesign*

1. Pendahuluan

Pulau Bali dikenal oleh mancanegara oleh karena keindahan alamnya. Hal ini tidak hanya menarik para turis, namun juga para investor dari negara-negara lain yang ingin berinvestasi [1]. Salah satu jenis investasi yang dilakukan adalah investasi properti. Namun, tidak semua investor dapat memanajemen properti mereka sendiri secara optimal untuk mendapatkan *return of investment* (ROI) yang maksimal [2]. Oleh karena itu, sejumlah investor menjalin kerjasama dengan perusahaan *property management* berbasis teknologi.

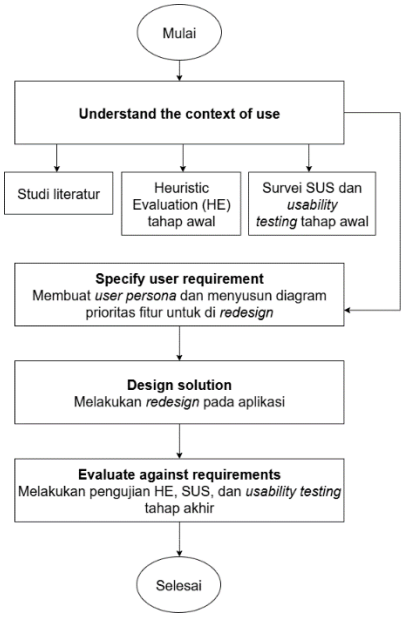
PT. Bukit Vista Nusantara adalah salah satu perusahaan yang menyediakan aplikasi *mobile* bernama BV GO sebagai *property management system*. Namun, pengembangannya sering dilakukan tanpa keterlibatan *user* secara langsung seperti melalui survei maupun wawancara. Proses ini berisiko menciptakan kesenjangan antara fitur aplikasi dengan kebutuhan *user* dan dapat menurunkan tingkat level *usability* aplikasi. Oleh karena itu, metode *user-centered design* (UCD) akan diterapkan dalam penelitian ini bersamaan dengan metode evaluasi *heuristic evaluation* (HE) dan *system usability scale* (SUS).

Metode UCD terbukti efektif dalam membangun aplikasi yang berfokus pada *user*. Dalam sebuah penelitian yang memanfaatkan UCD dan evaluasi survei SUS untuk memperbaiki *user experience* dari sistem bank, diketahui level *usability* aplikasi mengalami peningkatan yang dibuktikan dengan skor rata-rata SUS sebesar 76,5 dimana standar skor rata-rata SUS adalah 68 [3]. Di penelitian lain yang

menggunakan metode evaluasi HE, level *usability* aplikasi meningkat karena adanya penurunan nilai bobot masalah dari 48 ke 42 [4].

2. Metode Penelitian

Penelitian ini menggunakan metode UCD yang terdiri dari tahap *understand context of use, specify user requirement, design solution,* dan *evaluate against requirements* untuk mengembangkan aplikasi [3]. Dalam tahap pertama dan terakhir, penulis menggunakan metode evaluasi *Heuristic Evaluation* (HE), *System Usability Scale* (SUS), dan *usability testing* untuk mengumpulkan data awal dan data akhir, sehingga performa aplikasi sebelum dan sesudah *redesign* dapat dibandingkan. HE sendiri adalah metode evaluasi yang melibatkan ahli (evaluator) dalam prosesnya [4]. Sedangkan, SUS adalah metode pengujian *usability* yang dikembangkan oleh John Brooke pada tahun 1986 dan memiliki format 10 pertanyaan [5]. Metode HE dalam penelitian ini melibatkan 4 orang evaluator untuk menemukan jumlah temuan masalah dalam aplikasi. Masing-masing evaluator berprofesi sebagai *marketing strategist, Vice President of Engineer, guest experience manager*, dan UI/UX *designer*. Sedangkan, SUS melibatkan 5 orang *developer* aplikasi dan 5 orang *property partner* atau tim internal perusahaan. Adapun *usability testing* terhadap 5 aspek *usability* yang didasarkan pada Nielsen Model [6]. Kelima aspek tersebut adalah *learnability* (seberapa cepat *user* bisa mahir menggunakan sistem), *memorability* (kemudahan dalam mengingat posisi menu setelah jangka waktu tertentu), *efficiency* (sejauh mana sumber daya digunakan untuk mencapai tujuan dengan akurat dan lengkap), *error* (tingkat kesalahan yang dilakukan *user*), dan *satisfaction* (kebebasan dari ketidaknyamanan serta sikap positif *user* terhadap penggunaan produk). Untuk alur penelitian dapat dilihat pada gambar 1.



Gambar 1. Alur Penelitian

Tentunya, masing-masing metode memiliki rumusnya tersendiri. Untuk HE, rumus yang digunakan dapat dilihat pada persamaan (1).

$$\sum \mathrm{HX} = (1 \times \mathrm{x}) + (2 \times \mathrm{x}) + (3 \times \mathrm{x}) + (4 \times \mathrm{x}) \quad (1)\ [4]$$

Keterangan:
HX: Nilai *heuristic evaluation*
x: Jumlah masalah yang ditemukan pada tiap aspek

Untuk metode SUS, terdapat 3 langkah perhitungan dan 1 buah persamaan. Langkah perhitungan adalah sebagai berikut.

1. Setiap nilai pertanyaan pada urutan ganjil akan dikurangi 1.

2. Setiap nilai pertanyaan pada urutan genap akan digunakan untuk mengurangi 5.
3. Nilai total dikali dengan 2,5.

Untuk rumus menghitung rata-rata skor SUS dapat dilihat pada persamaan (2).

$$\bar{x} = \frac{\sum x}{n} \quad (2)\ [5]$$

Keterangan:
$\bar{x}$: Rata-rata skor SUS
∑x: Total skor SUS
n: Jumlah responden

Untuk rumus menghitung nilai rata-rata dan persentase *usability testing* dapat dilihat pada persamaan (3) dan (4).

Selain itu, metode HE menggunakan *severity ratings* dengan rentang skor 0-4 dalam penilaian. Sedangkan, metode SUS menggunakan skala likert dengan rentang skor 1-5. Detailnya dapat dilihat pada tabel 1 dan tabel 2.

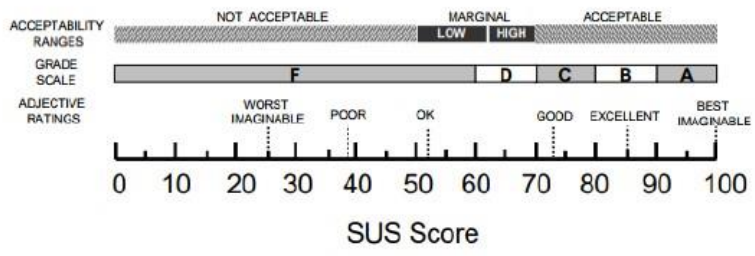
Tabel 1. Skor *Severity Ratings*

Skor	Klasifikasi
0	*Don't agree*: bukan sebuah permasalahan dan mudah untuk digunakan
1	*Cosmetic problem*: masalah ditemukan, namun tidak begitu berpengaruh pada pengalaman pengguna
2	*Minor usability problem*: perlu perbaikan, namun prioritasnya rendah karena kemungkinan hanya pengguna dari kalangan tertentu saja yang mengalami masalah tersebut
3	*Major usability problem*: pengguna kesulitan di tahap ini, sehingga penting untuk dilakukan perbaikan
4	*Usability catastrophe:* sistem perlu untuk diperbaiki karena tingkat kesalahannya sudah sangat tinggi

Tabel 2. Skor Skala Likert

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	2	3	4	5

Sebagai tambahan dalam metode SUS, digunakan skala skor SUS untuk mengetahui tingkat kelayakan aplikasi. Detailnya ada pada gambar 2.



Gambar 2. Skala Skor SUS

Terakhir adalah detail dari masing-masing pertanyaan HE, SUS, dan 5 aspek *usability testing* yang dapat dilihat pada tabel 3, gambar 3, dan gambar 4.

Tabel 3. Pertanyaan HE

Aspect Code	Question Code	Question	Score	Details of Issues Found
H1 (visibility of system status)	Q1	Does the app clearly inform you about your current page location in the app (e.g., which tab you're in)?		
	Q2	Are there clear feedback messages after actions (e.g., switching tabs, swiping newsletter cards)?		
H2 (match between system and real world)	Q3	Are the terms used in the app (e.g., "Booking", "Guest Reviews", "Contract Value") familiar and easy to understand for you?		
	Q4	Is the language consistent, clear, and understandable?		
H3 (user control and freedom)	Q5	Can you easily return to the previous screen or home when you make a mistake?		
	Q6	Are there clear confirmations for critical actions (like logout process)?		
H4 (consistency and standards)	Q7	Are the UI elements (buttons, labels, navigation) consistent across all pages?		
	Q8	Are fonts and labels consistently formatted across all pages?		
	Q9	Is the layout consistent in style, such as using the same card layout for newsletters, etc?		
H5 (error prevention)	Q10	Are error-prone actions (like		

p-ISSN: 2301-5373
e-ISSN: 2654-5101

		double booking, submitting wrong info) minimized or prevented with clear confirmations?		
	Q11	Are form fields clearly labeled to reduce user mistakes?		
H6 (recognition rather than recall)	Q12	Does the app visually highlight active menus (e.g., active tab in the bottom navigation bar)?		
	Q13	Are frequently used actions easily accessible without memorizing steps?		
H7 (flexibility and efficiency of use)	Q14	Are there features that support quick and efficient use for both new and experienced users?		
	Q15	Is the navigation in apps goes smoothly?		
H8 (aesthetic and minimalist design)	Q16	Is the UI visually clean, not cluttered with unnecessary information?		
	Q17	Are the UI elements designed for clarity and simplicity?		
H9 (help users recognize, diagnose, and recover from errors)	Q18	When there's an error (like a failed booking update), does the app show a clear, helpful message?		
	Q19	Are these error messages specific enough to guide users on how to fix them?		
H10 (help and documentation)	Q20	Is there a help section or guidance available if users need assistance with features like Booking or My Property?		

SUS		Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1.	I think that I would like to use this website frequently.	☐	☐	☐	☐	☐
2.	I found the website unnecessarily complex.	☐	☐	☐	☐	☐
3.	I thought the website was easy to use.	☐	☐	☐	☐	☐
4.	I think that I would need the support of a technical person to be able to use this website.	☐	☐	☐	☐	☐
5.	I found the various functions in this website were well integrated.	☐	☐	☐	☐	☐
6.	I thought there was too much inconsistency in this website.	☐	☐	☐	☐	☐
7.	I would imagine that most people would learn to use this website very quickly.	☐	☐	☐	☐	☐
8.	I found the website very cumbersome / awkward to use.	☐	☐	☐	☐	☐
9.	I felt very confident using the website.	☐	☐	☐	☐	☐
10.	I needed to learn a lot of things before I could get going with this system.	☐	☐	☐	☐	☐

Gambar 3. Pertanyaan SUS

Aspek	Kode Pertanyaan	Pertanyaan
Learnability	A1	(Sebelum/Setelah) di-*redesign,* apakah aplikasi BV GO dapat dipelajari dengan mudah?
	A2	Apakah Anda dapat dengan mudah melakukan navigasi antara halaman berbeda pada aplikasi?
	A3	Apakah Anda memahami isi dan konten informasi yang disajikan pada aplikasi?
	A4	Apakah tanpa instruksi tertulis atau *manual book*, Anda dapat dengan mudah mempelajari penggunaan aplikasi?
Memorability	B1	Apakah Anda dapat dengan mudah mengingat cara menggunakan fitur-fitur dari aplikasi BV GO (sebelum/setelah) *redesign*?
	B2	Apakah Anda merasa mudah kapan pun menggunakan aplikasi?
Efficiency	C1	Apakah Anda dapat dengan cepat menyelesaikan tugas-tugas tertentu di aplikasi BV GO, seperti menemukan *newsletter* dan lainnya?
	C2	Apakah Anda dengan mudah dan cepat menerima informasi secara detail dan juga spesifik terkait apa saja yang ada di dalam aplikasi?
	C3	Apakah Anda dapat secara mudah langsung menemukan informasi yang Anda ingin cari dari awal membuka aplikasi?
Error	D1	Apakah Anda tidak menemukan halaman atau fitur yang tidak sesuai dengan fungsinya pada aplikasi?

Gambar 4. Pertanyaan 5 Aspek *Usability* Bagian 1

p-ISSN: 2301-5373
e-ISSN: 2654-5101

	D2	Apakah Anda tidak mengalami kesalahan saat mencoba fitur aplikasi?
Satisfaction	E1	Apakah Anda senang dengan desain antarmuka yang ada pada aplikasi (sebelum/setelah) *redesign*?
	E2	Apakah Anda merasa nyaman dalam menggunakan aplikasi?
	E3	Apakah design aplikasi sesuai dengan ekspektasi Anda?
	E4	Apakah Anda puas dengan pengalaman penggunaan aplikasi (sebelum/setelah) dilakukan *redesign*?

Gambar 5. Pertanyaan 5 Aspek *Usability* Bagian 2

Sebagai tambahan, perbedaan antara pertanyaan 5 aspek *usability* di tahap awal dan akhir hanyalah kata “sebelum *redesign*” dan “setelah *redesign*”.

3. Hasil dan Pembahasan

3.1. Hasil Evaluasi Tahap Awal

a. Hasil HE Tahap Awal
Untuk detail dari jawaban para evaluator dapat dilihat dengan cara mengakses link spreadsheet berikut: http://bit.ly/4luyEzD. Salah satu contoh hasil perhitungan tiap aspek HE dapat dilihat pada tabel 4.

Tabel 4. Hasil Perhitungan Aspek H1 (Tahap Awal)

No.	Skor 1	Skor 2	Skor 3	Skor 4	Total
Q1	1	-	-	-	1
Q2	-	1	-	-	2
Total					**3**

Hasil total temuan masalah yang dihitung menggunakan persamaan (1) dapat dilihat pada tabel 5.

Tabel 5. Total Temuan Masalah HE Tahap Awal

Aspek	Temuan Masalah
H1	3
H2	3
H3	3
H4	8
H5	6
H6	2
H7	8
H8	5
H9	16
H10	10
Total Temuan Masalah: **64**	

b. Hasil SUS Tahap Awal
Untuk detail dari jawaban responden dapat dilihat melalui link spreadsheet berikut: http://bit.ly/4m5PODz. Hasil total skor SUS dapat dilihat pada tabel 6.

Tabel 6. Total Skor SUS Tahap Awal

No.	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Total
1.	4	4	4	4	2	4	3	4	4	4	**37**
2.	2	3	3	0	2	2	3	2	3	2	**22**
3.	3	3	3	4	3	2	3	3	3	3	**30**
4.	3	0	4	3	3	4	3	4	4	3	**31**
5.	3	2	3	3	3	3	3	3	3	3	**29**

6.	2	1	1	1	1	1	1	3	1	1	**13**
7.	4	4	4	0	4	1	3	3	4	1	**28**
8.	4	2	4	1	4	1	3	1	3	1	**24**
9.	4	4	4	0	4	4	4	4	4	0	**32**
10.	4	4	4	4	4	4	3	4	3	4	**38**
Total Nilai Akhir: 284 x 2,5 = 710											

Dengan menggunakan rumus persamaan (2), diperoleh skor rata-rata SUS sebesar 71 yang menandakan aplikasi berada di *grade scale* C akhir dan memperoleh *adjective ratings* “good”, seperti yang dapat dilihat pada gambar 2.

c. Hasil *Usability Testing* Tahap Awal
Untuk detail dari jawaban responden dapat dilihat pada link berikut: http://bit.ly/3UliYmt. Sedangkan, untuk hasil *usability testing* tahap awal dapat dilihat pada tabel 7. Sebagai catatan, TB adalah singkatan dari total bobot.

Tabel 7. Hasil *Usability Testing* Tahap Awal

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Subtotal**	**Total**
A1	3	3	4	2	3	3	2	3	4	3	**30**	**133**
A2	4	4	3	3	5	4	3	3	5	2	**36**	
A3	2	3	4	3	5	3	3	3	4	3	**33**	
A4	1	3	4	2	4	3	3	2	4	2	**34**	
B1	2	4	2	3	5	2	3	3	5	3	**32**	**64**
B2	2	4	4	2	5	3	3	2	5	2	**32**	
C1	2	2	3	1	5	3	4	3	5	4	**32**	**92**
C2	2	4	4	1	5	3	4	3	4	3	**33**	
C3	2	3	2	1	5	3	2	3	3	3	**27**	
D1	2	3	4	3	5	3	1	3	2	4	**30**	**54**
D2	2	3	2	1	4	2	3	2	3	2	**24**	
E1	3	3	4	2	5	3	2	2	4	3	**31**	**122**
E2	3	4	3	2	5	2	3	3	4	3	**32**	
E3	3	3	3	3	5	3	2	3	3	2	**30**	
E4	3	3	2	2	5	2	3	3	4	2	**29**	

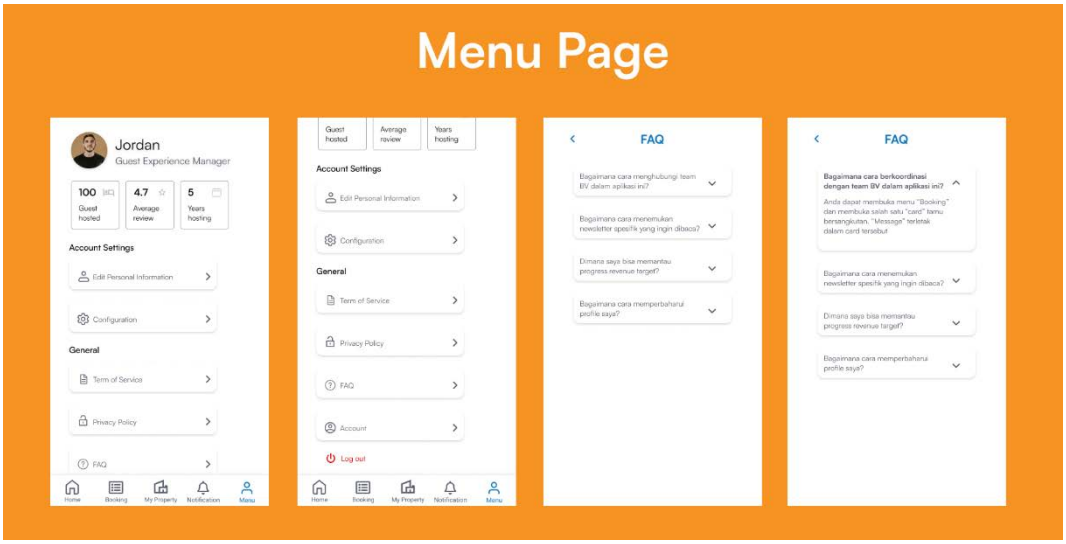
p-ISSN: 2301-5373
e-ISSN: 2654-5101

Berdasarkan Berdasarkan rumus di persamaan 3 dan 4, maka nilai rata-rata dan persentase masing-masing aspek *usability* tahap awal adalah sebagai berikut.

1. *Learnability*: 13,3 dan 64%
2. *Memorability*: 6,4 dan 64%
3. *Efficiency*: 9,2 dan 61%
4. *Error*: 5,4 dan 54%
5. *Satisfaction*: 12,2 dan 61%

3.2. Hasil *Redesign* Aplikasi

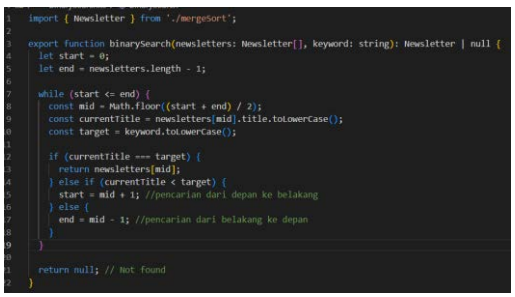
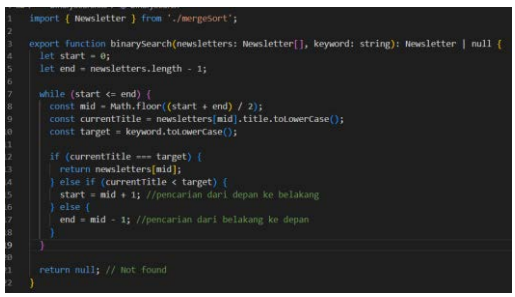
Berikut adalah salah satu contoh *redesign page* yang didasarkan pada saran para evaluator yang dapat dilihat pada gambar 6.



Gambar 6. Hasil *Redesign Menu Page*

Selain itu, adapula penambahan algoritma *merge sort* dan *binary search* yang menjadi solusi dari salah satu poin penilaian evaluator. Berikut adalah potongan kode yang dapat dilihat pada tabel 8.

Tabel 8. Potongan Kode *Merge Sort* dan *Binary Search*

Merge Sort	*Binary Search*
	

3.3. Hasil Evaluasi Tahap Akhir

a. Hasil HE Tahap Akhir
Untuk jawaban para evaluator dapat dilihat pada link berikut: http://bit.ly/4nXZlyk. Adapaun total temuan masalah pada HE tahap akhir dapat dilihat pada tabel 9.

Tabel 9. Total Temuan Masalah HE Tahap Akhir

Aspek	Temuan Masalah
H1	0
H2	2
H3	0
H4	6
H5	7
H6	1
H7	4
H8	2
H9	4
H10	4
Total Temuan Masalah: **30**	

Berdasarkan tabel 9, dapat diketahui bahwa total temuan masalah mengalami penurunan dari 64 temuan di tahap awal menjadi 30 temuan di tahap akhir.

b. Hasil SUS Tahap Akhir
Jawaban responden dapat dilihat pada link berikut: http://bit.ly/46NggNS. Hasil dari total skor SUS seperti yang tertera di tabel 10.

Tabel 10. Total Skor SUS Tahap Akhir

No.	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Total
1.	4	4	4	4	4	4	4	4	4	2	**38**
2.	4	2	3	4	4	4	3	4	4	4	**36**
3.	2	2	2	1	3	1	1	2	2	0	**16**
4.	3	3	4	3	4	3	2	3	3	3	**31**
5.	3	3	3	4	2	2	3	3	3	3	**29**
6.	4	4	4	3	4	4	4	4	4	3	**38**
7.	3	3	4	4	3	3	3	3	3	3	**32**
8.	3	3	4	4	4	3	4	2	4	3	**34**
9.	3	3	3	2	3	2	2	3	2	3	**26**
10.	4	3	4	4	4	3	4	4	4	3	**37**
Total Nilai Akhir: 317 x 2,5 = 792,5											

Dengan menggunakan rumus persamaan (2), diperoleh skor rata-rata SUS sebesar 79,25 yang menandakan aplikasi berada di *grade scale* C+ dan memperoleh *adjective ratings* “excellent”, seperti yang dapat dilihat pada gambar 2. Skor ini menunjukkan peningkatan dari skor rata-rata SUS tahap awal yang hanya sebesar 71.

c. Hasil *Usability Testing* Tahap Akhir
Untuk detail dari jawaban responden dapat dilihat pada link berikut: http://bit.ly/4mciA5T.
Untuk hasil *usability testing* tahap akhir dapat dilihat pada tabel 11. Sebagai tambahan, TB adalah singkatan dari total bobot.

Tabel 11. Hasil *Usability Testing* Tahap Akhir

	1	2	3	4	5	6	7	8	9	10	Subtotal	Total
A1	5	5	4	4	5	4	4	4	5	3	**43**	**171**
A2	5	5	4	4	5	4	4	4	5	3	**43**	
A3	5	5	4	4	5	5	4	5	4	4	**45**	
A4	4	5	4	3	5	4	4	4	4	3	**40**	
B1	5	5	4	4	5	4	3	5	5	4	**44**	**88**
B2	5	5	4	5	5	4	4	4	5	3	**44**	
C1	5	5	5	5	5	5	5	5	5	4	**49**	**134**
C2	3	5	5	4	5	5	4	5	4	3	**43**	
C3	5	5	4	4	5	5	4	4	4	2	**42**	
D1	4	5	4	3	5	3	4	4	2	4	**38**	

p-ISSN: 2301-5373
e-ISSN: 2654-5101

D2	4	4	5	4	4	4	4	4	3	3	**39**	**77**
E1	5	5	4	4	5	4	4	4	5	3	**43**	**172**
E2	5	5	5	5	5	4	4	4	5	3	**45**	
E3	5	4	5	3	5	4	3	5	4	2	**40**	
E4	5	5	5	4	5	5	3	4	5	3	**44**	

Berdasarkan rumus di persamaan 3 dan 4, maka nilai rata-rata dan persentase masing-masing aspek *usability* tahap akhir adalah sebagai berikut.

1. *Learnability*: 17,1 dan 86%.
2. *Memorability*: 8,8 dan 88%.
3. *Efficiency*: 13,4 dan 89%.
4. *Error*: 7,7 dan 77%.
5. *Satisfaction*: 17,2 dan 86%.

4. Kesimpulan

Pada akhirnya, dapat diketahui bahwa total bobot masalah pada HE tahap awal adalah 64 dan skor rata-rata SUS tahap awal adalah 71. Setelah aplikasi di *redesign* dengan pendekatan UCD dan disesuikan dengan input dari evaluator dan *user* awam, maka aplikasi diuji kembali menggunakan HE dan SUS tahap akhir (setelah *redesign*). Hasil evaluasi menunjukkan adanya penurunan total skor bobot masalah HE tahap awal sebesar 64 bobot masalah menjadi 30 bobot masalah di HE tahap akhir. Sedangkan, skor rata-rata SUS tahap awal sebesar 71 meningkat menjadi 79,25 di SUS tahap akhir. Untuk *usability testing*, kelima aspek mengalami peningkatan. Mulai dari *learnability* meningkat dari 64% menjadi 86%, *memorability* meningkat dari 64% menjadi 88%, *efficiency* meningkat dari 61% menjadi 89%, *error* (tingkat kejarangan mengalami *error*) meningkat dari 54% menjadi 77%, dan *satisfaction* meningkat dari 61% menjadi 86%. Hasil ini menunjukkan bahwa bahwa pendekatan UCD yang dilakukan dalam proses *redesign* dapat membawa *improvement* pada aplikasi.

Referensi

[1] D. Adhitiawarman, Detik Properti, 12 Agustus 2024. [Online]. Available: https://www.detik.com/properti/berita/d-7486280/bisnis-properti-di-bali-menggeliat-40-unit-villa-rp-8-m-laku-dalam-sehari. [13 Januari 2025]

[2] N.I, Arinze, A. Emininke, and A. Hillary, "*Assessment of Property Management for the Development of Real Estate Business in Enugu Urban Nigeria*" *Academic Journal of Science*, vol. 10, no. 1, p. 1-20, 2025.

[3] A. Sefian, M. Allsela, O.N. Rizky, P. Putra, and S.P. Eka, "Penerapan Metode *User Centered Design* (UCD) Dalam Perancangan *System Register* BRI *Document Management System*" *Journal of Information System Research*, vol. 8, no. 4. p. 1381-1391, 2024.

[4] F.K.S, Dewi, Y.D. Handharko, and F.V, Prasetyo, "Analisis *Usability* Menggunakan Metode *Heuristic Evaluation* dan *Web Usability Evaluation Tool* pada *Website ACC Career*" Jurnal Buana Informatika, vol. 13, no. 2, p.126-135, 2022.

[5] E. Kurniawan, Nofriadi, and A. Nata, "Penerapan *System Usability Scale* (SUS) dalam Pengukuran Kebergunaan *Website* Program Studi di STMIK Royal" *Journal of Science and Social Research*, vol. 5, no. 1. p. 43-49, 2022.

[6] H. Hanifah, Kraugusteeliana, dan I.N. Isnainiyah. "PENGUJIAN USABILITAS *USER EXPERIENCE WEBSITE UBIQUITOUS BASED LEARNING CLOUD* MENGGUNAKAN NIELSEN MODEL" Jurnal Sistem Informasi (JSI), vol. 14, no.1, p. 2585-2598, 2022.

This page is intentionally left blank.

Jurnal Elektronik Ilmu Komputer Udayana
Volume 14, No 3. February 2026

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Optimasi Algoritma SVM dengan Ekstraksi Fitur Warna pada Klasifikasi Biji Kopi Sangrai

Gde Krishna Sankya Yogeswara[1], I Ketut Gede Suhartana[2], I Gusti Ngurah Anom Cahyadi Putra[3], I Wayan Supriana[4]

[a]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam
Jalan Raya Kampus Unud, Jimbaran, Bali, 80361, Indonesia
[1] krishnasankya@gmail.com
[2] ikg.suhartana@unud.ac.id
[3] anom.cahyadi@unud.ac.id
[4]wayan.supriana@unud.ac.id

Abstract

Manual classification of roasted coffee beans is often hampered by subjectivity and inconsistency. This study aims to compare the performance of the Support Vector Machine (SVM) algorithm in classifying the ripeness of roasted coffee beans using two different feature approaches: raw pixel features and extracted hybrid features. The research focuses on finding the best hyperparameter combination for each approach and determining which method yields superior classification performance. The dataset used consists of 900 augmented coffee bean images, evenly distributed across three classes (Dark, Light, Medium). In the first approach, an SVM model was trained directly on flattened raw pixel data. In the second approach, an SVM model was trained using combined features extracted via a Convolutional Neural Network (CNN) and Color Histogram. The experimental results show a significant performance difference. The SVM model using raw pixel features achieved a maximum accuracy of 88.33% with the best parameters {kernel: 'rbf', C: 10, gamma: 0.01}. Meanwhile, the hybrid model utilizing feature extraction from the deeper_wider CNN architecture and color histograms drastically improved performance, reaching an accuracy of 98.33% with parameters {kernel: 'rbf', C: 10, gamma: 0.1}. These results demonstrate that employing high-level feature extraction through CNN is significantly superior to using raw pixels for the task of roasted coffee bean classification.

Keywords: *Support Vector Machine, Image Classification, Roasted Coffee Beans, Feature Extraction, Convolutional Neural Network, Color Histogram.*

1. Pendahuluan

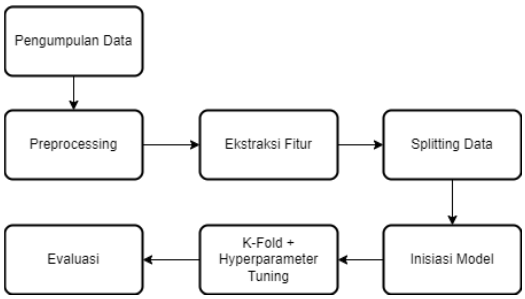
Biji kopi sangrai merupakan komoditas dengan nilai ekonomi tinggi yang kualitasnya sangat ditentukan oleh presisi proses penyangraian. Penilaian tingkat kematangan sangrai (roast level) secara konvensional mengandalkan persepsi visual manusia terhadap warna biji, sebuah metode yang rentan terhadap subjektivitas dan inkonsistensi [1]. Kesalahan penilaian dapat merugikan penyangrai dan konsumen, sehingga diperlukan solusi yang lebih objektif dan efisien.

Teknologi pengolahan citra digital menawarkan solusi menjanjikan untuk otomatisasi proses ini. Beberapa penelitian sebelumnya telah menerapkan arsitektur machine learning untuk penilaian kualitas biji kopi, salah satunya menggunakan Support Vector Machine (SVM) [6]. SVM adalah algoritma klasifikasi yang efektif dalam memisahkan data dengan mencari hyperplane optimal [2]. Namun, kinerja SVM sangat bergantung pada kualitas fitur yang diekstraksi. Di sisi lain, Convolutional Neural Network (CNN) unggul dalam mengekstraksi fitur-fitur hierarkis yang kompleks (warna, bentuk, tekstur) secara otomatis dari citra mentah [3]. Beberapa studi telah menunjukkan keberhasilan model hibrida CNN-SVM dalam berbagai domain, yang mengindikasikan potensi peningkatan kinerja dengan menggabungkan kekuatan keduanya [4].

Penelitian ini bertujuan untuk menganalisis dan membandingkan secara komprehensif kinerja dua pendekatan: (1) model SVM yang dilatih langsung menggunakan data piksel mentah, dan (2) model SVM hibrida yang menggunakan gabungan fitur hasil ekstraksi dari CNN dan Histogram Warna. Tujuannya adalah untuk menentukan metode mana yang paling unggul dan menemukan kombinasi hyperparameter terbaik untuk klasifikasi tingkat kematangan biji kopi sangrai (Dark, Light, Medium).

p-ISSN: 2301-5373
e-ISSN: 2654-5101

2. Metode Penelitian



Gambar 1. Alur Metode Penelitian

Sesuai pada Gambar 1, dalam rangka memperoleh keluaran dari penelitian ini, diterapkan sejumlah tahapan metodologis yang mencakup proses pengumpulan data, analisis, hingga tahap akhir berupa perancangan sistem sebagai bentuk demonstrasi implementasi metode yang telah dikembangkan.

2.1 Pengumpulan Data



Gambar 2. Citra Biji Kopi

Data yang digunakan adalah dataset primer yang terdiri dari 900 citra biji kopi sangrai. Dataset ini terbagi rata ke dalam tiga kelas kematangan: *Dark* (300 citra), *Light* (300 citra), dan *Medium* (300 citra). Contoh dataset seperti pada Gambar 2, pengambilan gambar dilakukan menggunakan kamera *smartphone* dengan latar belakang yang seragam untuk menjaga konsistensi.

2.2 Preprocessing

Preprocessing merupakan tahap krusial yang dapat dilakukan untuk menyelaraskan serta meningkatkan kualitas data. Beberapa tahap preprocessing dalam penelitian ini dilakukan semaksimal mungkin untuk memastikan bahwa data yang diinput terfokus secara khusus pada obyek penelitian, yaitu Biji Kopi. Berikut adalah langkah-langkah *preprocessing* yang akan diterapkan.

a. *Cropping*



Gambar 3. *Cropping* Citra Biji Kopi

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Cropping adalah langkah pertama dari proses *preprocessing* yang agar model dapat mempelajari fitur gambar tanpa adanya gangguan terkait obyek lainnya. Berdasakan contoh hasil *cropping* pada Gambar 3, hal ini akan mengurangi keberagaman objek dalam data yang diinputkan ke dalam model yang nantinya akan meningkatkan performa model dalam mengenali pola-pola penting pada citra biji kopi.

b. *Removing Background*



Gambar 4. *Removing Background* Citra Biji Kopi

Mengacu pada Gambar 4, langkah ini berfungsi sebagai lapisan pemrosesan lanjutan yang bertujuan untuk memastikan bahwa fokus utama citra input benar-benar tertuju pada objek biji kopi, tanpa adanya gangguan visual, termasuk latar belakang. Dengan demikian, model yang dikembangkan akan memiliki tingkat ketahanan (*robustness*) yang tinggi dan tetap dapat berfungsi secara optimal meskipun gambar diambil dalam kondisi atau lokasi yang berbeda-beda.

c. *Resizing*
Model memerlukan input citra dengan dimensi yang seragam agar proses pemrosesan dapat dilakukan secara konsisten dan menghasilkan output yang optimal. Oleh karena itu, pada penelitian ini dilakukan proses *resizing* citra menjadi ukuran 224x224 piksel.

2.3 Ekstraksi Fitur

Untuk model hibrida, dua jenis fitur diekstraksi dari setiap citra yang telah melalui *preprocessing*:

1. **Fitur CNN:** ebuah arsitektur CNN dirancang untuk mengekstraksi fitur visual tingkat tinggi seperti tekstur dan bentuk. Kemampuan jaringan syaraf tiruan konvolusional dalam mempelajari representasi data citra secara mendalam telah terbukti sangat efektif, seperti yang ditunjukkan pada penelitian fundamental oleh Krizhevsky, dkk. [5]. Fitur ini diambil dari output lapisan sebelum fully connected layer. Beberapa arsitektur dengan kedalaman dan lebar yang berbeda (simple, original, deeper_wider) diuji untuk menemukan ekstraktor fitur terbaik.

2. **Fitur Histogram Warna:** Histogram warna (RGB) dihitung untuk merepresentasikan distribusi warna global dari citra. Vektor histogram ini kemudian diratakan (*flattened*) untuk menjadi fitur numerik.

Kedua set fitur ini kemudian digabungkan menjadi satu vektor fitur tunggal untuk setiap citra yang akan menjadi input bagi *classifier* SVM.

2.4 Splitting Data

Keseluruhan dataset dibagi menjadi 2 bagian diantaranya data train dan data testing. Pembagian jumlah *splitting* data menggunakan proporsi yang umum yaitu 80% data train dan 20% data testing.

2.5 Iniasiasi Model

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Penelitian ini mengevaluasi dan membandingkan performa di antara tiga jenis model yang diikutsertakan. Model-model tersebut diantaranya :

a. SVM

Tabel 1. Algoritma SVM

Algoritma 1. Support Vector Machine
Input: *X*
Output: *Y (class label)*
Initialize kernel, C, gamma
Declare K
Process
If *kernel == "linear"* ***then***
$K = x.x^T$
If *kernel == "rbf"* ***then***
$K = K = (1 + x.x^T)^Q$
If *kernel == "poly"* ***then***
$K = e^{- y\|x-xT\|}$
$w \rightarrow min_w(1/2 * \|\|w\|\|^2)$ *such that* $y_n (w^T x_n + b >= 1)$
$w = \alpha \rightarrow$ *dual optimization problem*
$\alpha \rightarrow max\alpha(\alpha - 1/2 * \alpha (yy^T * K) \alpha)$ *(quadratic programming)*
$\alpha \rightarrow min\alpha(1/2 * \alpha (yy^T * K) \alpha - \alpha)$, *subject to* $y^T a = 0$, $0 <= a <= C$ *(equality in dual problem)*
partial grad $\rightarrow$ *minimize* α
End
Output
$Y = sign(sum(a * y_{isv} * K) + b)$

Algoritma *Support Vector Machine* (SVM) pada tabel 2 diawali dengan proses menerima input berupa matriks fitur X yang telah melalui tahap ekstraksi, dengan total sembilan fitur per sampel. Pada tahap awal, proses dimulai dengan inisialisasi *hyperparameter*, yaitu jenis kernel, nilai regularisasi C, dan parameter gamma yang berperan penting dalam pengendalian kompleksitas model dan fleksibilitas fungsi kernel. Selanjutnya, dilakukan pembuatan fungsi kernel untuk memproyeksikan data ke dalam ruang berdimensi lebih tinggi. Tujuan utama dari metode SVM adalah memaksimalkan margin antara kelas-kelas, yang secara matematis dicapai dengan meminimalkan norma vektor bobot w melalui formulasi *dual optimization problem*. Untuk meningkatkan efisiensi komputasi, digunakan teknik *kernel trick* dan *quadratic programming*, sehingga tidak perlu menghitung seluruh transformasi eksplisit ke ruang berdimensi tinggi. Formulasi dual ini dikonversi ke bentuk kuadratik untuk memperoleh solusi optimal terhadap parameter α. Setelah nilai-nilai α optimal diperoleh, model SVM dapat digunakan untuk melakukan prediksi dengan menghasilkan nilai label kelas yang telah di *encode* secara numerik.

2.6 K-Fold Cross Validation dan Hyperparameter Tuning

Nilai K yang akan digunakan pada proses *K-Fold Cross Validation* yaitu 5, sehingga akan terdapat lima iterasi dimana setiap iterasi akan menggunakan sebagian dataset untuk menjadi data validation dan sebagian lagi menjadi data pelatihan. Penggunaan metode validasi ini bertujuan untuk mencegah terjadinya overfitting serta memberikan performa yang valid pada sebuah model karena keseluruhan dataset akan berperan dalam proses klasifikasi dan pengujian.

Selanjutnya, *hyperparameter tuning* akan diterapkan pada setiap *fold* data untuk mencari kombinasi *hyperparameter* yang memiliki performa terbaik. Berikut adalah nilai-nilai *hyperparameter* yang akan digunakan dalam penelitian ini.

a. C = [0.01, 0.1, 1, 10, 100, 1000]
b. Kernel = [*linear*, *polynomial*, *rbf*]
c. Gamma = [0.001, 0.01, 0.1, 1, 10, 100]

2.7 Evaluasi

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Setiap hasil prediksi pada model-model yang diteliti akan melalui tahap evaluasi dengan menggunakan beberapa metrik diantaranya *precision*, *recall*, *f1-score*, dan *accuracy*. Keempat metrik tersebut dapat dikalkulasikan dengan *formula* berikut.

$$\text{Akurasi} : \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

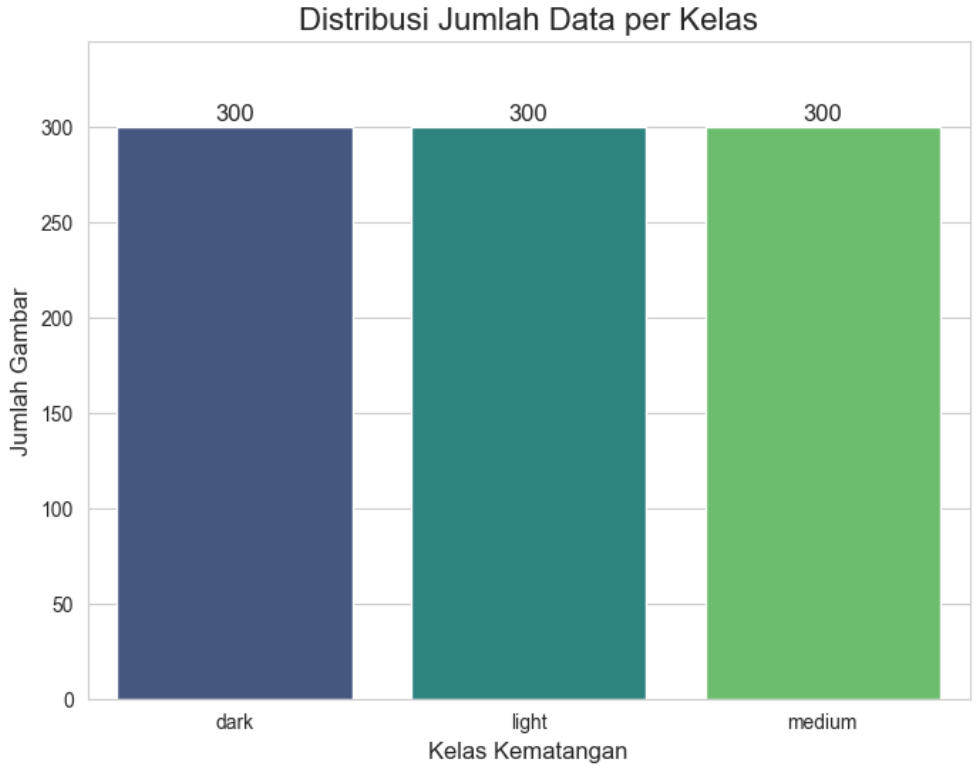
$$\text{Presisi} : \frac{TP}{TP + FP} \quad (2)$$

$$\text{Recall} : \frac{TP}{TP + FN} \quad (3)$$

$$\text{F1-score} : \frac{2 \times (recall \times presisi)}{recall + presisi} \quad (4)$$

3. Hasil dan Pembahasan

3.1. Pengumpulan Data



Gambar 5. *Barplot* Jumlah Data Citra

Setelah proses validasi manual untuk menghindari anomali dan memastikan kualitas data, dataset akhir yang digunakan terdiri dari total 900 citra. Jumlah ini terbagi rata ke dalam tiga kelas, yaitu masing-masing 300 citra untuk kopi bisa dilihat pada gambar 5.

3.2. Preprocessing

Dataset citra yang telah melalui tahap *preprocessing* akan disimpan dalam sebuah folder dengan masing-masing kelas disimpan dalam sub-folder tersendiri sesuai label kelasnya.

p-ISSN: 2301-5373
e-ISSN: 2654-5101

3.3. Ekstraksi Fitur

Peneliti melakukan ektraksi fitur khususnya berkaitan dengan fitur warna yang melibatkan rata-rata persebaran piksel (*mean*), variansi persebaran piksel (*variance*), dan warna piksel dominan (*mode/*modus). Setiap fitur ini akan diterapkan di masing-masing *channel* pada penelitian ini yaitu *Red*, *Green*, dan *Blue*. Hasil ekstraksi fitur dapat dilihat pada tabel 2.

Tabel 2. Hasil Ekstraksi Fitur

M_R	M_G	M_B	Var_R	Var_G	Var_B	Mo_R	Mo_G	Mo_B	Label
0.6504	0.48260	0.200912	0.148468	0.083375	0.021267	0.886275	0.678431	0.176471	1
0.6553	0.47399	0.17961	0.15856	0.08401	0.01702	0.91372	0.658824	0.180392	1
0.5972	0.4049	0.11463	0.13487	0.06341	0.00817	0.85098	0.53725	0.13333	1
0.6383	0.43215	0.16872	0.12292	0.05591	0.01302	0.90980	0.56470	0.18431	1
0.6242	0.42506	0.16622	0.12920	0.05972	0.01281	0.91372	0.57647	0.20000	1
...	...	...	...	...	...	...	...	...	...
0.1779	0.1589	0.12002	0.02180	0.01942	0.01588	0.20000	0.15686	0.09411	2
0.1796	0.18165	0.14832	0.01356	0.01493	0.01354	0.22352	0.19215	0.11764	2
0.1739	0.17312	0.13444	0.01117	0.01260	0.01250	0.20000	0.16078	0.09411	2
0.3203	0.30304	0.15060	0.02981	0.02835	0.01847	0.39215	0.36470	0.08235	2
0.3476	0.3246	0.14500	0.03825	0.03420	0.02168	0.41176	0.39215	0.05490	2

Rentang hasil ekstraksi fitur tidak melebihi nilai 1 karena nilai piksel pada citra telah dinormalisasi ke dalam rentang 0 hingga 1. Proses normalisasi ini dilakukan dengan membagi setiap nilai piksel dengan 255 sebelum dilakukan perhitungan untuk ekstraksi fitur.

3.4. Splitting Data

Dari total 900 citra biji kopi yang telah melalui proses pembersihan data, dilakukan pembagian data (*data splitting*) dengan proporsi 80% untuk data latih (*training data*), dan 20% untuk data uji (*testing data*).

3.5. Inisiasi Model

a. SVM

Algoritma Support Vector Machine (SVM) yang digunakan dalam penelitian ini tidak memanfaatkan library machine learning seperti scikit-learn, melainkan diimplementasikan secara mandiri (native) menggunakan library matematis seperti NumPy, statistics, dan cvxopt. Pada inisiasi awal dengan parameter {'kernel': 'linear', 'C': 1.0}, model ini dievaluasi performanya.

Berdasarkan hasil yang disajikan pada Tabel 3, model SVM dengan konfigurasi awal ini berhasil mencapai akurasi sebesar 88%, dengan rata-rata presisi 90%, recall 87%, dan F1-score 88%. Tabel ini memberikan gambaran awal mengenai kemampuan model sebelum dilakukan optimasi lebih lanjut.

Tabel 3. Matrix Evaluasi SVM

Precision	Recall	F1-Score	Accuracy
90%	87%	88%	88%

b. SVM + *Tuning*

Setelah melakukan pengujian terhadap berbagai kombinasi hyperparameter (kernel, C, dan gamma), ditemukan konfigurasi terbaik yang secara signifikan meningkatkan performa model. Hasil proses tuning yang dirangkum pada Tabel 4, menunjukkan bahwa kombinasi {kernel: 'rbf', C: 10, gamma: 0.01} menempati peringkat pertama dengan akurasi validasi

p-ISSN: 2301-5373
e-ISSN: 2654-5101

silang (CV) tertinggi, yaitu 91.11%. Tabel ini memvalidasi pentingnya proses hyperparameter tuning dalam menemukan setelan optimal untuk model.

Tabel 4. Matrix Evaluasi SVM

Peringkat	Kernel	C	Gamma	Degree	Akurasi (CV)
1	rbf	10	0.01	-	91.11%
2	rbf	1.0	0.01	-	89.44%
3	linear	10	-	-	88.33%
4	linear	1.0	-	-	88.33%
5	linear	0.1	-	-	88.33%

Model kemudian dilatih ulang menggunakan parameter terbaik tersebut. Peningkatan performa yang signifikan tercatat pada Tabel 5, di mana akurasi model pada data uji meningkat menjadi 91%. Selain itu, metrik lainnya seperti presisi, recall, dan F1-score juga mengalami kenaikan, yang menunjukkan kemampuan generalisasi model yang lebih baik setelah dioptimalkan

Tabel 5. Matrix Evaluasi SVM

Precision	**Recall**	**F1-Score**	**Accuracy**
97%	93%	95%	91%

c. SVM + *Ekstraksi fitur + K-Fold Cross Validation*

Pada tahap selanjutnya, model SVM diintegrasikan dengan pendekatan ekstraksi fitur hibrida (CNN dan Histogram Warna) dan divalidasi menggunakan K-Fold Cross Validation. Proses tuning yang komprehensif ini menghasilkan temuan penting yang disajikan pada Tabel 6. Kombinasi antara arsitektur CNN deeper_wider dengan parameter SVM {'kernel': 'linear', 'C': 0.1} terbukti memberikan rata-rata akurasi validasi silang tertinggi, mencapai 98.33%.

Tabel 6. Metrik Hasil Evaluasi SVM

Peringkat	CNN Arch	Kernel	C	Gamma	Degree	Accuracy
1	deeper_wider	linear	0.1	-	-	98.33%
2	deeper_wider	linear	1.0	-	-	98.22%
3	deeper_wider	linear	10	-	-	98.10%
4	original	linear	0.1	-	-	97.50%
5	original	linear	1.0	-	-	97.50%

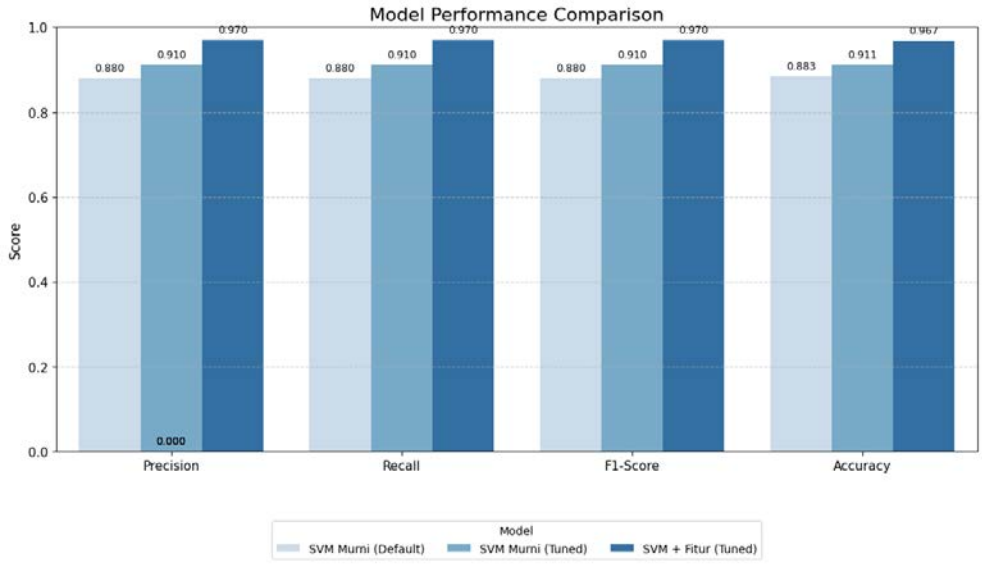
p-ISSN: 2301-5373
e-ISSN: 2654-5101

Hasil evaluasi akhir dari model hibrida terbaik ini terhadap data pengujian dapat dilihat pada **Tabel 7**. Model ini menunjukkan performa yang sangat tinggi dan konsisten di semua metrik, dengan akurasi mencapai 98% serta presisi, *recall*, dan F1-*score* masing-masing sebesar 97%.

Tabel 7. Hasil Evaluasi SVM + Ekstraksi Fitur

Precision	Recall	F1-Score	Accuracy
97%	97%	97%	98%

3.6. Evaluasi



Gambar 8. Perbandingan Performa Model

Berdasarkan hasil evaluasi pada ketiga model terhadap data pengujian, dapat diketahui bahwa ekstraksi fitur memberikan peningkatan performa yang sangat signifikan. Perbandingan performa dari ketiga pendekatan model (SVM Murni Default, SVM Murni Tuned, dan SVM + Fitur Tuned) divisualisasikan pada Gambar 8. Grafik tersebut dengan jelas menunjukkan bahwa model hibrida (SVM + Fitur) unggul secara konsisten di semua metrik evaluasi.

Untuk menegaskan hasil akhir, Tabel 9 merangkum performa tertinggi yang dicapai oleh model SVM yang dioptimalkan dengan ekstraksi fitur. Model ini mencatatkan akurasi 96.6% (nilai ini sepertinya salah ketik di teks asli, seharusnya mendekati 98% sesuai tabel sebelumnya, harap diperiksa kembali) dengan rata-rata presisi, recall, dan F1-score sebesar 97%.

Tabel 9. Performa Model SVM + Ekstraksi Fitur

Accuracy	Mean Precision	Mean Recall	Mean F1-Score
96,6%	97%	97%	97%

4. Kesimpulan

Berdasarkan hasil evaluasi model terhadap data pengujian, dapat disimpulkan bahwa pendekatan ekstraksi fitur hibrida berhasil meningkatkan performa model *Support Vector Machine*. Peningkatan ini diukur berdasarkan hasil dari proses *hyperparameter tuning* yang disertai dengan *K-Fold Cross Validation*. Proses ini memvalidasi bahwa performa terbaik model dapat dicapai dalam rentang

parameter yang telah didefinisikan, sehingga menjadi dasar yang adil untuk membandingkan pengaruh setiap pendekatan..

Peningkatan performa yang diberikan oleh pendekatan hibrida sangat signifikan dan konsisten di seluruh metrik evaluasi. Secara spesifik, model SVM yang dioptimalkan dengan ekstraksi fitur dari arsitektur CNN **deeper_wider** dan Histogram Warna dengan parameter {kernel: 'linear', C: 0.1} menunjukkan performa tertinggi dengan nilai-nilai metrik sebagai berikut:

a. Akurasi: **96,67%**
b. Mean Precision: **97%**
c. Mean Recall: **97%**
d. Mean F1-Score: **97%**

Sedangkan, model SVM Murni yang telah dioptimalkan (91,11% akurasi) juga menunjukkan peningkatan performa dari model SVM dasar (88,33% akurasi), namun hasilnya masih berada di bawah model hibrida. Hal ini menegaskan bahwa penggunaan fitur tingkat tinggi secara fundamental lebih baik untuk tugas klasifikasi biji kopi sangra.

Referensi

[1] Pires, J. B. P., et al. (2022). "Correlation of color and chemical compounds in roasted coffee." *Journal of Food Science*.

[2] Cortes, C., & Vapnik, V. (1995). "Support-vector networks." *Machine Learning, 20*(3), 273–297.

[3] LeCun, Y., et al. (1998). "Gradient-based learning applied to document recognition." *Proceedings of the IEEE, 86*(11), 2278–2324.

[4] Niu, X., & Suen, C. Y. (2012). "A novel hybrid CNN–SVM classifier for recognizing handwritten digits." *Pattern Recognition, 45*(4), 1318–1325.

[5] [5] Krizhevsky, A., et al. (2012). "ImageNet classification with deep convolutional neural networks.

[6] Ratnadewi, R., et al. (2022). "Image classification of coffee beans using SVM architecture for quality assessment." *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control, 7*(2), 181-188.

p-ISSN: 2301-5373
e-ISSN: 2654-5101

This page is intentionally left blank.

Jurnal Elektronik Ilmu Komputer Udayana
Volume 14, No 3. February 2026

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Rancangan Sistem Monitoring Suhu, Kelembaban, Kecepatan Angin Untuk Memprediksi Keamanan Jalur Pendakian

I Wayan Wikananda Adikara[a1], Gst. Ayu Vida Mastrika Giri[a2], I Wayan Supriana[b3], I Gede Surya Rahayuda[b4]

[a]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana
[1]wayannandaadikara07@email.com
[2]vida@unud.ac.id
[3]Wayan.supriana@cs.unud.ac.id
[4]igedesuryarahayuda@unud.ac.id

Abstract

Mountain climbing activities pose risks to climbers as they can cause serious injuries and often result in death, so information about conditions on climbing routes is essential for predicting safety. In this study, the Internet of Things with Decision Tree algorithms was used to build a system for monitoring conditions on mountain climbing routes. The study aims to measure the accuracy of the Decision Tree algorithm in classifying safety conditions and to develop a web-based application capable of displaying real-time data on temperature, humidity, and wind speed. Functional testing results indicate that all hardware and software components function properly. Additionally, the research results indicate that the use of the Decision Tree algorithm for safety classification on hiking trails achieves an accuracy rate of 99.27%. Thus, the classification model developed is highly effective in distinguishing between safe and hazardous conditions based on the three weather parameters. The very high accuracy demonstrates that the Decision Tree model used is highly suitable for prediction or monitoring systems, as developed in this study.

Keywords: *Decision Tree, Internet of Things, Mountain Climbing Route Safety Prediction*

1. Pendahuluan

Indonesia merupakan negara yang memiliki banyak kekayaan alam khususnya gunung. Beragamnya gunung di Indonesia mendorong tumbuh dan berkembangnya minat masyarakat dalam melakukan aktivitas mendaki gunung [1]. Akan tetapi meningkatnya jumlah pendaki di Indonesia, diikuti pula dengan meningkatnya angka kecelakaan saat mendaki gunung [2]. Aktivitas pendakian gunung memiliki risiko berdampak bahaya terhadap pendaki karena dapat menyebabkan cedera parah dan sering menimbulkan kematian, jika dibandingkan dengan olahraga lain olahraga pendakian gunung merupakan olahraga penyebab kematian terbesar [3].

Berdasarkan data Badan Nasional Pencarian dan Pertolongan (BASARNAS), kecelakaan pendakian meningkat dalam empat tahun terakhir. Pada tahun 2015 terjadi 12 kecelakaan, tahun 2016 meningkat menjadi 15 kasus, dan tahun 2017 juga terjadi 15 kasus. Kasus pendakian menjadi meningkat cukup signifikan pada tahun 2018 yaitu sebanyak 23 kejadian yang menyebabkan 6 pendaki meninggal dunia, 4 pendaki dinyatakan hilang, 7 pendaki ditemukan sakit, dan 592 pendaki ditemukan selamat. Pada tahun 2019 kejadian pendaki tersesat dan hilang juga meningkat menjadi 16 kasus jika dibandingkan tahun sebelumnya yang hanya terjadi 3 kasus [4]. Pengetahuan terhadap medan pendakian merupakan hal yang harus diketahui oleh para pendaki untuk dapat mengukur berapa panjang dan berapa lama pendakian serta dapat mengatur persiapan mendakian sesuai obyek pendakian [5].

Risiko dan bahaya pendakian gunung disebabkan oleh beberapa faktor, salah satunya adalah faktor eksternal yang meliputi cuaca atau iklim gunung yang sulit ditebak [6]. Hal ini didukung oleh peneliti lain yang menyebutkan sekitar 32% pendaki hilang dikarenakan cuaca buruk [4]. Informasi cuaca di daerah pegunungan yang menjadi tujuan pendakian sangat diperlukan sebelum melakukan pendakian agar dapat melakukan persiapan dengan lebih baik. Beberapa informasi cuaca

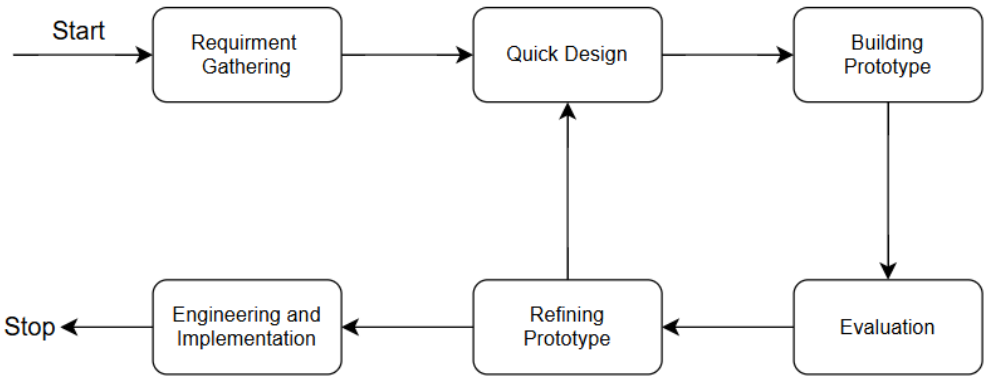
yang sering digunakan adalah suhu, kelembaban udara, dan kecepatan angin [7]. Maka dari itu penting untuk seorang pendaki mendapatkan informasi mengenai cuaca di jalur pendakian untuk mempredikasi keamanan.

Mendapatkan suatu informasi dalam waktu yang singkat semakin mudah dikarenakan adanya jangkauan internet yang luas [8]. Internet of Things merupakan sebuah konsep yang bertujuan untuk memperluas manfaat dari konektivitas internet yang tersambung secara terus-menerus [9]. Internet of Things, dapat digunakan untuk membangun sebuah sistem pemantauan cuaca secara *real-time* dan mendapatkan informasi data mengenai kondisi cuaca didaerah atau area yang dinginkan, sehingga dapat mengamati keadaan cuaca karena sering terjadi perbedaan cuaca antara suatu kawasan dengan kawasan lainnya [8].

Salah satu algoritma yang dapat digunakan untuk menganalisis data tersebut adalah algoritma *Decision Tree.* Decision Tree merupakan struktur pohon yang terdiri dari node-node yang merepresentasikan keputusan dan cabang-cabang yang merepresentasikan konsekuensi dari sebuah keputusan [10]. Peneliti lain menemukan bahwa hasil perbandingan algoritma klasifikasi *nearest neighbour, naive bayes* dan *decision tree* yang digunakan pada studi kasus pengambilan keputusan menyatakan bahwa algoritma klasifikasi *decision tree* merupakan algoritma klasifikasi yang memiliki tingkat akurasi paling tinggi dibandingkan algoritma klasifikasi naive bayes dan nearest neighbour yaitu mencapai 75.6% [11]. Sehingga dalam penelitian ini peneliti menggunakan agoritma *decision tree* untuk memprediksi suhu, kelembaban udara, dan kecepatan angin di gunung.

2. Metode Penelitian

Pada penelitian ini menggunakan data sekunder, dimana data yang digunakan berupa data suhu, kelembaban dan kecepatan angin yang didapatkan melalui data online dari website resmi milik BMKG. Penelitian ini dilakukan secara terstruktur dengan menerapkan metode prototyping [12] sebagai pendekatan dalam pengembangan sistem, guna menjaga fokus dan arah proses penelitian yang ditampilkan pada gambar 1.



Gambar 1. Metode Prototyping

Pada Gambar 1 merupakan diagram metode prototyping yang dimana terdapat enam langkah yang dilakukan. Adapun langkah-langkah pada metode prototypeing diuraikan sebagai berikut.

2.1 Analisis Kebutuhan/ Requirment Gathering

Dalam penelitian ini, analisis kebutuhan yang dilakukan adalah sesuai dengan permasalahan bagaimana sistem prediksi keamanan pada jalur pendakian berbasis IoT untuk meminimalisir terjadinya hal-hal yang tidak diinginkan pada saat melakukan pendakian gunung. Berdasarkan analisis akan kebutuhan yang diperlukan dalam membangun sistem prediksi keamanan pada jalur pendakian berbasis IoT, diperoleh gambaran mengenai komponen, perangkat, serta fungsi yang diperlukan dalam pengembangan sistem. Adapun rincian kebutuhan yang menjadi dasar dalam pembangunan sistem tersebut dijelaskan sebagai berikut:

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Tabel 1. Kebutuhan Perangkat Keras

Perangkat Keras	Keterangan
Mikrokontroler ESP 32	Mikrokontroler ESP32 berfungsi untuk memproses data pembacaan dari sensor sekaligua berfungsi untuk melakuka pengiriman data ke server
DHT 22	Sensor yang berfungsi untuk mengukur nilai suhu dan kelembaban udara
Anemometer	Sensor yang berfungsi untuk mengukur nilai kecepatan angin
Oled 0.96 inci	Layar Oled yang digunakan untuk menampilkan data pada perangkat keras
Breadboard	Breadboard berfungsi sebagai papan kerja untuk menyatukan sensor dengan mikrokontroler

Pada Tabel 1 dapat dilihat kebutuhan perangkat keras yang digunakan dalam penelitian ini. Perangkat utama yang digunakan adalah mikrokontroler ESP32 yang berfungsi untuk memproses data sensor sekaligus mengirimkan data ke server. Sensor DHT22 digunakan untuk mengukur suhu dan kelembaban udara, sedangkan sensor anemometer berfungsi untuk mengetahui kecepatan angin. Selain itu, layar OLED 0,96 inci digunakan untuk menampilkan data secara langsung, dan breadboard dimanfaatkan sebagai papan kerja untuk menyatukan sensor dengan mikrokontroler.

Tabel 2. Kebutuhan Perangkat Lunak

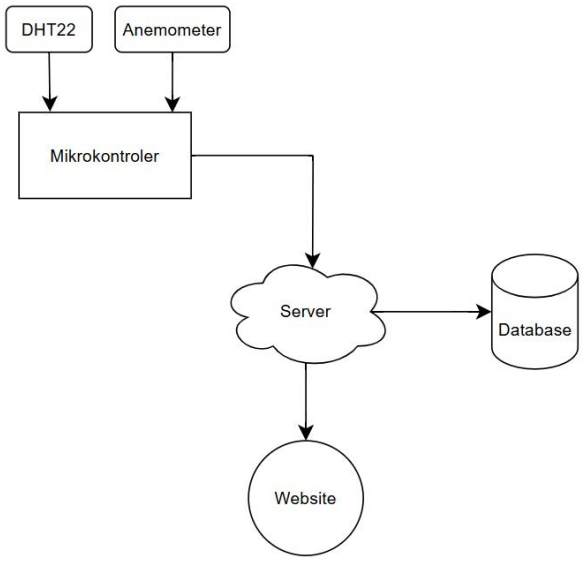
Perangkat Lunak	Keterangan
Visual Studio Code	Text Editor yang digunakan dalam pengembangangan system
Arduino IDE	Development Environment yang digunakan untuk memprogram mikrokontroler
Google Chrome	Web Browser yang digunakan untuk menjalankan aplikasi
Flask	Merupakan Server yang digunakan dalam pengembangan system

Pada Tabel 2 ditunjukkan perangkat lunak yang digunakan dalam pengembangan sistem. Visual Studio Code berfungsi sebagai text editor yang mendukung penulisan kode program. Arduino IDE digunakan sebagai environment untuk memprogram mikrokontroler ESP32. Google Chrome dimanfaatkan sebagai web browser untuk menjalankan dan menguji aplikasi yang telah dibuat. Sedangkan Flask digunakan sebagai server dalam proses pengembangan sistem, khususnya untuk menghubungkan data dari perangkat keras dengan aplikasi.

2.2 Desain Sistem/ Quick Design

Desain sistem dalam penelitian ini digambarkan ke dalam beberapa diagram yang menjelaskan alur sistem bekerja serta desain alat yang akan dibuat.

2.2.1 Alur Kerja Sistem



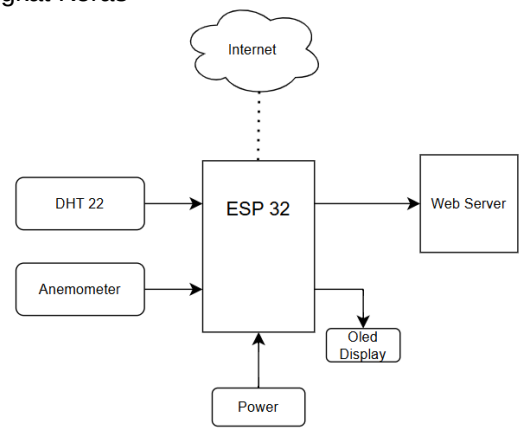
Gambar 2. Alur Kerja Sistem

Pada Gambar 2 alur kerja sistem secara umum sistem akan melakukan pembacaan data terhadap beberapa parameter, dimana parameter yang digunakan adalah suhu udara, kelembaban

udara dan kecepatan angin. Dimana hasil pembacaan akan dikirimkan oleh mikrokontroler ke web server melalui jaringan. Data yang telah dikirim ke web server nantinya akan dilakukan analisis menggunakan model yang sudah dibuat dengan algoritma *decision tree*, hasil analisis dari model dan data yang dikirimkan dari sensor akan ditampilkan di website.

Sistem yang akan dibuat dalam penelitian ini terdiri dari dua komponen, yaitu perangkat keras dan perangkat lunak yang dirancang agar saling terintegrasi. Rancangan dari perangkat keras dan perangkat lunak yang akan dibuat dijabarkan pada subbab bagian perancangan perangkat keras serta perancangan perangkat lunak.

2.2.2 Perancangan Perangkat Keras



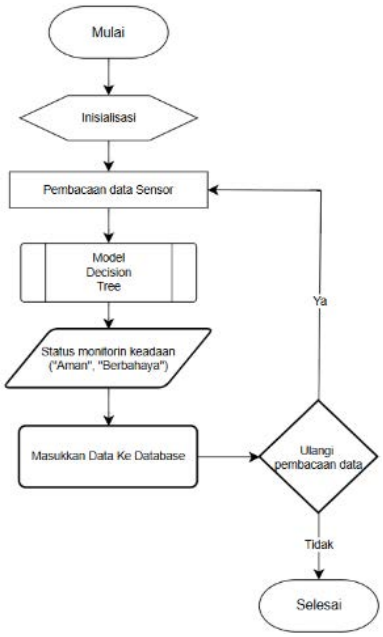
Gambar 3. Desain Perangkat Keras

Pada Gambar 3 merupakan rancangan desain perangkat keras yang akan dibuat, di mana sensor DHT22 dan sensor anemometer digunakan sebagai parameter untuk mengukur kondisi lingkungan. Sensor DHT22 digunakan untuk mengukur suhu dan kelembaban udara, kemudian sensor Anemometer digunakan untuk mengukur seberapa kencang kecepatan angin, dan layar OLED berfungsi sebagai perangkat keras yang menampilkan data dari sensor. Data dari kedua sensor tersebut nantinya akan mempengaruhi status atau keputusan dari model yang telah dibuat. Data hasil pembacaan dari sensor selanjutnya akan diolah di dalam mikrokontroler yang mana dalam hal ini digunakan ESP32 Devkit V1, data yang telah diolah ini kemudian dikirimkan ke web server sebagai data hasil monitoring yang nantinya akan disajikan dalam bentuk website. Proses pengiriman data ke server dilakukan dengan memanfaatkan modul WiFi yang telah terintegrasi dalam ESP32 Devkit V1.

2.2.3 Perancangan Perangkat Lunak

Pada tahap implementasi, sistem ini akan menggunakan sebuah website untuk memantau kondisi berdasarkan data yang diperoleh dari alat yang telah dirancang. Adapun alur kerja yang dilakukan dalam perangkat lunak tersebut dapat dilihat pada gambar 4.

p-ISSN: 2301-5373
e-ISSN: 2654-5101



Gambar 4. Flowchart Perangkat Lunak Sistem

Pada Gambar 4 merupakan desain rancangan perangkat lunak yang akan dibuat nantinya. Perangkat lunak ini dibuat dalam bentuk website. Hal yang dilakukan pertama adalah melakukan inisialisasi variabel-variabel yang akan digunakan dalam perhitungan algoritma *Decision Tree*. Kemudian melakukan pembacaan data hasil dari sensor yang dikirim oleh mikrokontroler. Data hasil pembacaan sensor kemudian dilakukan proses klasifikasi dengan algoritma *Decision Tree*. Setelah didapatkan output dari klasifikasi, maka selanjutnya data dan hasil klasifikasi tersebut disimpan ke database, selanjutnya data akan ditampilkan di dashboard sebagai informasi hasil monitoring jalur pendakian gunung serta status apakah aman atau berbahaya.

2.3 Implementasi/Building Prototype

Pembuatan alat serta perangkat lunak dilakukan pada tahap ini. Alat dibuat sesuai dengan rancangan pada perangkat keras yang terdiri dari beberapa komponen. Kemudian perangkat lunak yang dibuat sesuai dengan rancangan perangkat lunak yang telah digambarkan, dimana nantinya perangkat lunak yang dihasilkan adalah berupa sebuah website yang berisi informasi monitoring keadaan keamanan pada jalur pendakian.

2.4 Pengujian/ Evaluation

Pada tahapan ini dilakukan pengujian terhadap sistem yang telah selesai dirancang. Pengujian dilakukan untuk memastikan bahwa setiap proses pada sistem bekerja dengan baik. Kemudian dilakukan pengujian terhadap algoritma *decision tree* untuk mengetahui seberapa akurat model yang telah dibuat. Berikut adalah beberapa tahapan yang akan dilakukan dalam tahap pengujian :

- Pengujian Hardware
  Pada tahap ini dilakukan untuk memastikan sensor dan mikrokontroler berfungsi dengan baik, mampu membaca data dengan akurat, serta mengirimkan data ke server secara stabil.
- Pengujian Aplikasi
  Pada tahap ini dilakukan untuk mengevaluasi kemampuan aplikasi dalam menerima, mengolah, menyimpan, dan menampilkan data sensor secara real-time, termasuk menilai fitur grafik, klasifikasi, dan riwayat data.
- Evaluasi Model *Decision Tree*
  Tahap evaluasi terhadap metode dalam penelitian ini dilakukan menggunakan metode confusion matrix. Confusion matrix merupakan sebuah metode evaluasi yang umum digunakan untuk mengukur prestasi algoritma *Decision Tree* dalam melakukan klasifikasi. Confusion matrix memberikan gambaran seberapa baik model *Decision Tree* dalam mengklasifikasi data ke dalam kelas yang benar [13].

2.5 Desain Ulang Prototype/Refining Prototype

Tahap Refining Prototype adalah proses penyempurnaan prototype berdasarkan umpan balik yang diperoleh dari pengguna Tujuannya adalah agar prototipe semakin mendekati ekspektasi, tujuan

dan kebutuhan sistem yang sebenarnya. Proses refining dilakukan secara iteratif, artinya prototipe dapat diperbaiki dan diuji ulang berkali-kali hingga mencapai hasil akhir yang diinginkan.

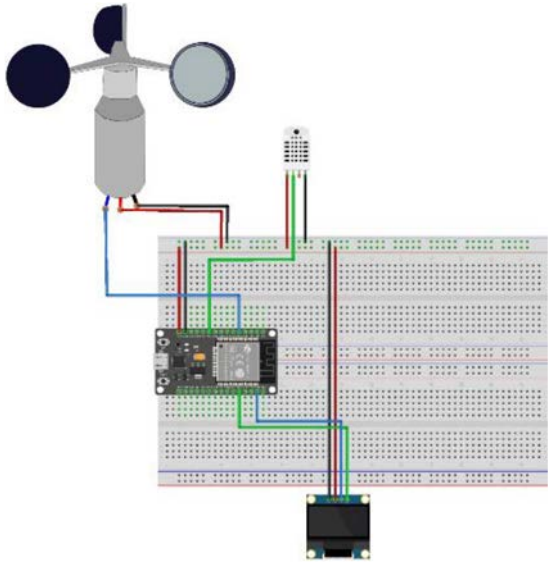
2.6 Engineering and Implementation/Rekayasa dan Implementasi

Tahap Engineering & Implementation merupakan tahap di mana prototype yang telah disempurnakan dan dikembangkan menjadi sistem akhir yang siap digunakan sebagaimana mestinya sesuai dengan tujuan awal dibangunnya sistem. Tujuan utama dari tahap ini adalah menghasilkan sistem yang tidak hanya berfungsi sesuai kebutuhan, tetapi juga aman, dan mudah dipelihara.

3. Hasil dan Pembahasan

3.1 Perangkat Keras

Perangkat keras pada penelitian ini berfungsi untuk memperoleh data kondisi lingkungan berupa suhu, kelembaban udara, dan kecepatan angin sebagai parameter prediksi keamanan. Data diperoleh melalui sensor DHT22 (suhu dan kelembaban) serta anemometer (kecepatan angin). Rangkaian skematik alat ditunjukkan pada Gambar 5.



Gambar 5. Rangkaian Skematik Perangkat Keras

Berdasarkan Gambar 5, rangkaian skematik alat yang dibuat dalam penelitian ini menunjukkan sistem berbasis IoT dengan mikrokontroler ESP32 sebagai pusat kendali utama. ESP32 bertanggung jawab dalam mengatur operasional sistem, memproses data dari sensor, dan mengirimkan informasi ke server. Pada rangkaian ini, digunakan dua buah sensor berfungsi sebagai pembaca keadaan lingkungan.

Tabel 3. Koneksi Pin pada Mikrokontroler

Komponen	Pin Komponen	Pin ESP32
DHT22	VCC	5V
	GND	GROUND
	DATA	GPIO 27
Anemometer	VCC	5V
	GND	GROUND
	DATA	GPIO 35
Layar OLED 0,96 inch	VCC	5V
	GND	GROUND
	SDA	GPIO 21
	SCL	GPIO 22

Pada Tabel 3 di atas menunjukkan koneksi komponen dengan ESP32, yaitu sensor DHT22 (VCC–5V, GND–GROUND, DATA–GPIO27) untuk membaca suhu dan kelembaban, sensor anemometer (VCC–5V, GND–GROUND, DATA–GPIO35) untuk kecepatan angin, serta layar OLED

p-ISSN: 2301-5373
e-ISSN: 2654-5101

0,96 inci dengan antarmuka I2C (VCC–5V, GND–GROUND, SDA–GPIO21, SCL–GPIO22) untuk menampilkan data.



Gambar 6. Hasil Implementasi Alat

Pada gambar 6 merupakan implementasi perangkat keras dari sistem yang telah dirancang, dalam implementasinya dapat dilihat terdapat sebuah mikrokontroler ESP32 yang terhubung dengan sensor DHT22, sensor Anemometer dan layar OLED 0,96 inci.

3.2 Tampilan Website

Tampilan dari halaman website yang telah dirancang menampilkan grafik, tabel riwayat, dan status prediksi pada dashboard. Tampilan website dapat dilihat pada Gambar 7 dan 8.



Gambar 7. Dashboard Sistem

Pada Gambar 7 menunjukkan tampilan dashboard sistem yang menampilkan informasi suhu, kelembaban, kecepatan angin, serta hasil prediksi model. Dashboard juga menyajikan nilai ideal setiap parameter, status kondisi (optimal atau danger), serta grafik tren data sensor.



Gambar 8. Tampilan Log Data

Gambar 4.6 menunjukkan tampilan halaman log data sensor yang berisi seluruh hasil monitoring dari database, termasuk catatan aktivitas sensor dan hasil klasifikasi model.

3.3 Pengujian Perangkat Keras

Pengujian perangkat keras bertujuan untuk memastikan bahwa seluruh komponen fisik yang digunakan dapat berfungsi dengan baik. Pengujian ini juga dilakukan untuk mengetahui tingkat keakuratan sensor dalam membaca data lingkungan.

Tabel 4. Pengujian Sensor

No	Fitur yang Diuji	Input Uji	Ekspektasi *Output*	Hasil Aktual	Status
1	Pengukuran suhu	Sensor didekatkan dengan sumber panas	Suhu meningkat	Suhu meningkat	✓ Sesuai
2	Pengukuran suhu	Sensor di dinginkan	Suhu menurun	Suhu menurun	✓ Sesuai
3	Pengukuran kelembaban	Disemprotkan uap	Kelembaban udara meningkat	Kelembaban meningkat	✓ Sesuai
4	Pengukuran kelembaban	Disemprotkan air berbentuk aerosol	Kelembaban udara meningkat	Kelembaban meningkat	✓ Sesuai
5	Pengukuran kecepatan angin	Sensor tidak bergerak	Data yang terbaca adalah 0	Data Sesuai	✓ Sesuai
6	Pengukuran kecepatan angin	Sensor bergerak	Data yang terbaca adalah >0	Data Sesuai	✓ Sesuai

Berdasarkan pengujian yang dilakukan pada Tabel 4 dapat disimpulkan bahwa seluruh sensor yang digunakan, yaitu DHT22 untuk pengukuran suhu dan kelembaban serta anemometer untuk pengukuran kecepatan angin, telah berfungsi dengan baik sesuai dengan rancangan. Setiap pengujian menunjukkan bahwa sensor mampu merespons perubahan kondisi lingkungan dengan benar, baik pada saat terjadi peningkatan maupun penurunan suhu, kelembaban, maupun perubahan kecepatan angin. Hal ini membuktikan bahwa perangkat keras yang digunakan layak untuk digunakan dalam membangun sistem.

Tabel 5. Pengujian Fungsionalitas Hardware

No	Fitur yang Diuji	Input Uji	Ekspektasi *Output*	Hasil Aktual	Status
1	ESP32 pembacaan data sensor	Loop pembacaan tiap 5 detik selama 5 menit	Data sensor terbaca tanpa “NaN” atau outliter ekstrime	Data Sesuai	✓ Sesuai
2	Modul koneksi ESP32	Koneksi ke SSID wi-fi	Tersambung dan memperoleh alamat IP	Terhubung ke wi-fi	✓ Sesuai
3	Menampilkan data pada layar OLED	Render data sensor pada modul OLED	Data tampil sesuai dengan serial monitor IDE	Data Sesuai	✓ Sesuai
4	Pengiriman data dari	Sensor membaca data	Data terkirim dan respon	Status : SUKSESS	✓ Sesuai

p-ISSN: 2301-5373
e-ISSN: 2654-5101

No	Fitur yang Diuji	Input Uji	Ekspektasi *Output*	Hasil Aktual	Status
	mikrokontroler ke server	(suhu 28,3°C, kelembaban 83% dan kecepatan angin 9km/j	server 200 status “SUKSESS”		
5	Koneksi internet terputus	Sensor membaca data tetapi internet terputus	Data tidak terkirim ke server	Data tidak terkirim	✓ Sesuai
6	Sumber daya hardware	Adaptor / powerbank 5V selama 6 jam	Tidak reset / burnout	Tidak terdapat kendala reset / burnout	✓ Sesuai
7	Stabilitas hardware	Running test selama 6 jam	Hardware bekerja dengan semestinya tanpa gangguan	Tidak terdapat gangguan pada hardware	✓ Sesuai

Berdasarkan pengujian yang dilakukan pada Tabel 5, dapat disimpulkan bahwa seluruh perangkat keras pada sistem, meliputi ESP32, modul koneksi WiFi, layar OLED, serta catu daya, mampu berfungsi dengan baik sesuai rancangan. ESP32 terbukti dapat membaca data sensor secara konsisten tanpa error, menampilkan hasil dengan benar pada layar OLED, dan berhasil mengirimkan data ke server. Selain itu, perangkat juga mampu menangani kondisi ketika koneksi internet terputus dengan tidak mengirimkan data, serta menunjukkan performa yang stabil tanpa terjadi reset maupun gangguan selama pengujian jangka panjang. Hal ini membuktikan bahwa perangkat keras yang digunakan layak untuk diimplementasikan pada sistem.

3.4 Pengujian Aplikasi

Pengujian pada perangkat lunak dilakukan untuk memastikan bahwa sistem yang dibangun dapat berfungsi sesuai dengan tujuan perancangan. Pengujian dilakukan untuk memastikan aplikasi dapat menerima data dari perangkat keras, memprosesnya dengan model, menyimpan ke basis data, serta menampilkan informasi secara real-time melalui dashboard, termasuk data sensor, grafik, hasil klasifikasi, dan riwayat data.

Tabel 6. Pengujian Fungsionalitas Aplikasi

No.	Fitur yang Diuji	Input Uji	Ekspektasi *Output*	Hasil Aktual	Status
1	Pembacaan data yang dikirim dari mikrokontroler	Sensor membaca data (suhu 28,3°C, kelembaban 83% dan kecepatan angin 9km/j	Respon server 200 dan mereturn status “SUKSESS”	Status : SUKSESS	✓ Sesuai
2	Load model untuk melakukan klasifikasi	Data sensor yang berhasil dikirim kemudian diklasifikasi oleh model	Berhasil terklasifikasi oleh model (Aman / Berbahaya) sesuai data	Data berhasil di klasifikasi	✓ Sesuai
3	Menyimpan data ke database	Data sensor dan hasil klasifikasi	Data tersimpan di database	Data tersimpan dengan benar	✓ Sesuai

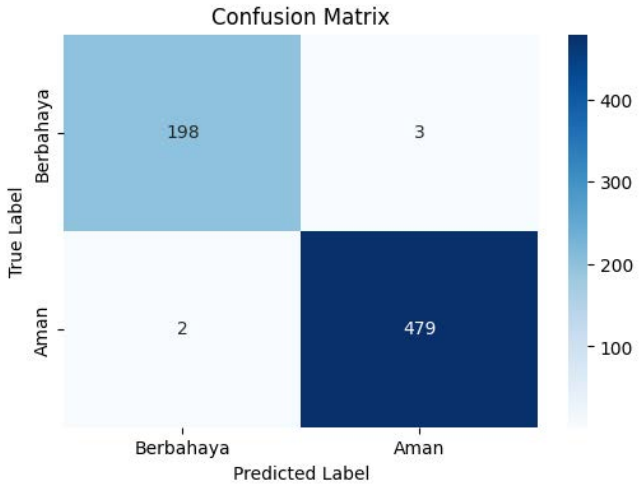
No.	Fitur yang Diuji	Input Uji	Ekspektasi *Output*	Hasil Aktual	Status
		dikirim ke database			
4	Tampilkan data sensor realtime	Mengambil data terbaru dari database	Data terbaru tampil di dashboard secara otomatis	Data tampil sesuai dengan pembacaan sensor dan hasil klasifikasi model	✓ Sesuai
5	Grafik data	Data sensor dikirim terus-menerus selama 5 menit	Grafik menampilkan perubahan nilai sesuai waktu	Grafik menampilkan kurva naik \ turun sesuai data sensor	✓ Sesuai
6	Log data	Mengambil seluruh data di database	Menampilkan seluruh data yang tersimpan di database	Data tampil sesuai database	✓ Sesuai
7	Refresh halaman	Halaman dashboard dan log data direfresh	Data tetap terbaca dan halaman berjalan normal	Data tampil kembali tanpa eror	✓ Sesuai
8	Koneksi internet terputus	Memutuskan koneksi internet mikrokontroler	Dashboard tidak mengupdate data	Data berhenti update	✓ Sesuai

Berdasarkan pengujian yang dilakukan pada Tabel 6, dapat disimpulkan bahwa seluruh fitur perangkat lunak pada sistem telah berjalan sesuai dengan rancangan. Server mampu menerima data dari mikrokontroler dengan baik, melakukan pemuatan model machine learning, serta menghasilkan klasifikasi kondisi secara tepat. Data hasil pengukuran dan klasifikasi dapat tersimpan dengan benar di dalam database, ditampilkan secara realtime di dashboard, divisualisasikan melalui grafik, dan ditampilkan kembali pada log data. Selain itu, sistem juga mampu memberikan respon yang sesuai saat koneksi internet terputus, yaitu dengan menghentikan pembaruan data di dashboard. Hal ini menunjukkan bahwa perangkat lunak yang dikembangkan telah berfungsi secara optimal.

3.5 Evaluasi Model

Evaluasi performa model terhadap seluruh data untuk mengetahui seberapa baik model mengklasifikasi seluruh sampel yang ada. Metrik yang digunakan adalah akurasi keseluruhan dan confusion matrix. Confusion matrix memberikan informasi mengenai jumlah prediksi yang benar maupun yang salah untuk setiap kategori label.

p-ISSN: 2301-5373
e-ISSN: 2654-5101



Gambar 9. Confusion Matrix

Pada Gambar 9 menampilkan hasil confusion matrix, model klasifikasi decision tree menunjukkan performa yang sangat baik dalam membedakan antara kondisi berbahaya dan aman. Dari total 682 data uji, model berhasil mengklasifikasikan 198 data berbahaya dengan benar dan 479 data aman dengan benar. Hanya terdapat 3 data berbahaya yang salah diklasifikasikan sebagai aman dan 2 data aman yang salah diklasifikasikan sebagai berbahaya.

Tabel 7. Akurasi Pengujian Model

Metric	**Nilai**
Accuracy	99.27%
Precision	0.99
Recall	0.99
F1-Score	0.99

Pada Tabel 7 ditampilkan hasil akurasi pengujian model yang digunakan dalam penelitian. Hasil pengujian menunjukkan bahwa model memiliki nilai accuracy sebesar 99,27%, dengan precision, recall, dan F1-score masing-masing bernilai 0,99.

4. Kesimpulan

Akurasi Algoritma Decision Tree dalam Klasifikasi Data: Penggunaan algoritma Decision Tree untuk mengklasifikasi keamanan pada jalur pendakian berdasarkan suhu, kelembaban udara, dan kecepatan angin menunjukkan tingkat akurasi yang sangat baik dengan akurasi 99,27%. Hasil ini menunjukkan bahwa model klasifikasi yang telah dibangun mampu sangat baik membedakan antara kondisi aman dan berbahaya berdasarkan ketiga parameter cuaca tersebut. Akurasi yang sangat tinggi menunjukkan bahwa model Decision Tree yang digunakan sangat baik dalam sistem prediksi atau pemantauan seperti yang dikembangkan dalam penelitian ini.

Perancangan sistem prediksi keamanan pada jalur pendakian: Perancangan sistem prediksi keamanan pada jalur pendakian berbasis IoT berhasil dilakukan dengan memadukan perangkat keras dan perangkat lunak secara terintegrasi. Sistem ini menggunakan mikrokontroler ESP32 yang terhubung dengan sensor DHT22 untuk membaca suhu dan kelembaban udara, serta sensor Anemometer untuk mengukur kecepatan angin. Data yang diperoleh dari sensor ditampilkan secara lokal melalui layar OLED 0,96 inci, serta dikirimkan ke server menggunakan koneksi WiFi dari ESP32. Pada sisi perangkat lunak, sistem ini dibangun dengan bahasa pemrograman python dengan framework flask sebagai webserver. Webserver menerima data dari mikrokontroler yang kemudian data tersebut langsung diklasifikasikan menggunakan model machine learning dengan algoritma Decision Tree yang telah dibangun. Hasil klasifikasi ini menentukan kondisi keamanan jalur pendakian berdasarkan parameter data, kemudian disimpan dalam database dan ditampilkan pada antarmuka website. Hasil pengujian fungsionalitas pada penelitian ini menunjukkan seluruh perangkat keras pada sistem, meliputi ESP32, modul koneksi WiFi, layar OLED, serta catu daya, mampu berfungsi dengan baik sesuai rancangan. Sejalan dengan itu, hasil pengujian fungsionalitas aplikasi juga menunjukkan bahwa seluruh perangkat keras pada sistem, meliputi ESP32, modul koneksi WiFi, layar OLED, serta catu daya, mampu berfungsi dengan baik sesuai rancangan.

Referensi

[1] A. Primanda, L. Ode, and S. Maftuh, “Perancangan Media Informasi untuk Pendaki Gunung Pemula (Studi Kasus: Jawa Barat),” *Jurnal Visualaras*, vol. 1, no. 1, pp. 1–11, 2022.

[2] U. N. Mustaniroh, Purwanto, and M. Pudail, “Analisis Daya Tarik Wisata Pendakian Gunung Andong dalam Meningkatkan Kesejahteraan Masyarakat Dusun Sawit Kabupaten Magelang: Ditinjau dari Maqashid Syariah,” *Jurnal Mirai Management*, vol. 8, no. 1, pp. 767–779, 2023.

[3] J. Parung, R. Ikawaty, and J. Untung, “Penggunaan Virtual Reality untuk Melatih Pemandu Wisata Gunung Menangani Pertolongan Pertama pada Kecelakaan di Atas Gunung,” *Jurnal Abdi Insani*, vol. 11, no. 4, pp. 1683–1690, Nov. 2024, doi: 10.29303/abdiinsani.v11i4.1972.

[4] A. A. Pamungkas, M. R. K. P. Wardana, A. D. Oktavian, D. F. Primadani, D. N. A. Tsaqif, and S. Sholiq, “Pengembangan Desain Pengalaman Pengguna Sistem Monitoring Pendaki Gunung Berbasis GPS dan LoRA dengan Metode Agile Scrum,” *Bridge : Jurnal publikasi Sistem Informasi dan Telekomunikasi*, vol. 2, no. 3, pp. 39–53, Jun. 2024, doi: 10.62951/bridge.v2i3.97.

[5] G. Aprelianto, A. Setiawan, and J. Prestilliano, “Perancangan Board Game sebagai Media Pembelajaran tentang Pendakian Gunung,” *Jurnal Desain Komunikasi Visual Asia (JESKOVSIA)*, vol. 06, no. 01, pp. 2597–4300, 2022.

[6] N. Widyayanti and A. R. Lestari, “Hubungan Kematangan Emosi Dengan Persepsi Terhadap Risiko Kecelakaan Pendakian Gunung Dan Pendaki Gunung Di Jawa Tengah,” *Jurnal Psikologi Mandiri*, vol. 5, no. 1, pp. 1–7, 2022.

[7] Mappagio, M. Ardita, and Sotyohadi, “Sistem Monitoring Multi Sensor untuk Mengetahui Perubahan Cuaca,” *Magnetika*, vol. 8, no. 1, pp. 316–324, 2024.

[8] A. S. Ratri, V. C. Poekoel, and A. M. Rumagi, “Design Of Weather Condition Monitoring System Based On Internet Of Things,” *Jurnal Teknik Informatika*, vol. 17, no. 1, pp. 1–10, 2021.

[9] Efendi, “Inernet Of Things (Iot) Sistem Pengendalian Lampu Menggunakan Raspberry Pi Berbasis Mobile,” *Jurnal Ilmu Komputer*, vol. 4, no. 2, pp. 21–27, 2018.

[10] R. N. Ramadhon, A. Ogi, A. P. Agung, R. Putra, S. S. Febrihartina, and U. Firdaus, “Implementasi Algoritma Decision Tree untuk Klasifikasi Pelanggan Aktif atau Tidak Aktif pada Data Bank,” *KarimahTauhid*, vol. 3, no. 2, pp. 1860–1874, 2024.

[11] D. Sartika and D. I. Sensuse, “Perbandingan Algoritma Klasifikasi Naive Bayes, Nearest Neighbour, dan Decision Tree pada Studi Kasus Pengambilan Keputusan Pemilihan Pola Pakaian,” *Jastisi*, vol. 1, no. 2, pp. 151–161, 2017.

[12] S. Anawar *et al.*, “Multi-Layer Privacy-Preserving (MPP) Mechanism for Protected Health Information (PHI) In A Health Campaign Management System,” *IJCSNS International Journal of Computer Science and Network Security*, vol. 21, no. 10, pp. 210–218, 2021, doi: 10.22937/IJCSNS.2021.21.10.29.

[13] O. Somantri, W. E. Nugroho, and A. R. Supriyono, “Penerapan Feature Selection Pada Algoritma Decision Tree Untuk Menentukan Pola Rekomendasi Dini Konseling,” *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 4, no. 2, pp. 263–279, Dec. 2022, doi: 10.30865/json.v4i2.5345.

Jurnal Elektronik Ilmu Komputer Udayana
Volume 14, No 3. February 2026

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Sistem Informasi Manajemen Keuangan Pribadi dengan Input Suara dan Berbasis IoT

I Dewa Gede Partha Wijaya[a1], I Made Widiartha[a2], I Gede Santi Astawa[a3], Gst Ayu Vida Mastrika Giri[a4]

[a]Informatics, Udayana University
Badung, Indonesia
[1]parthawijaya04@gmail.com
[2]madewidiartha@unud.ac.id
[3]santi.astawa@unud.ac.id
[4]vida@unud.ac.id

Abstrak

The increasing need for practical personal financial management, especially among students and young professionals, highlights the challenges of recording small daily expenses such as parking or food transactions. Manual recording often leads to inconsistencies due to forgetfulness or inconvenience. To address this issue, this study develops an Internet of Things (IoT) based personal financial management system integrating smartwatch and smartphone devices with voice input features. Voice commands are transcribed into text using Google Speech-to-Text, while transaction information such as type, description, and nominal value is extracted using the Boyer-Moore string matching algorithm. The system was implemented using Android technology with real-time synchronization through Firebase Firestore. Functional testing through blackbox methods showed that all features, including voice transcription and transaction synchronization, operated as expected. Performance testing of the Boyer-Moore algorithm was conducted on input lengths ranging from 30 to 120 characters. To minimize JVM-related bias, each experiment was repeated 100 times, and the average execution time was calculated. Regression analysis of the results produced the equation $y = 1034x + 5192$ with a coefficient of determination $R^2 = 0.992$. Furthermore, user acceptance testing using the Technology Acceptance Model (TAM) yielded positive results, with average scores above 4.00 on a 5-point scale. These results indicate that integrating IoT and voice-based input significantly improves the efficiency and user acceptance of personal financial management applications.

Kata Kunci*: Speech-to-Text, String Matching, Boyer-Moore, Internet of Things, Personal Finance*

1. Pendahuluan

Manajemen keuangan pribadi merupakan aktivitas penting yang dapat membantu individu, terutama mahasiswa, dalam mengontrol pengeluaran dan mengelola prioritas keuangan mereka. Penelitian sebelumnya menyatakan bahwa pencatatan keuangan berkontribusi terhadap pengendalian diri dalam pengeluaran sehari-hari, membantu meminimalisir pengeluaran impulsif, serta meningkatkan kesadaran akan kebutuhan finansial [5]. Namun demikian, pencatatan transaksi rutin dengan nominal kecil seperti parkir atau konsumsi harian seringkali tidak dilakukan karena alasan kerepotan maupun kelupaan.

Seiring dengan berkembangnya teknologi, input berbasis suara menjadi solusi potensial untuk meningkatkan kenyamanan dalam pencatatan transaksi. Penelitian menunjukkan bahwa kecepatan mengetik rata-rata di perangkat seluler hanya sekitar 36 kata per menit [1], sementara kecepatan bicara normal dapat mencapai 140 hingga 240 kata per menit [2]–[4]. Hal ini membuktikan bahwa penggunaan input suara secara signifikan dapat mengurangi waktu yang diperlukan untuk pencatatan transaksi harian. Selain itu, teknologi pengenalan suara berbasis Google Speech-to-Text terbukti layak diimplementasikan dalam aplikasi Android [6].

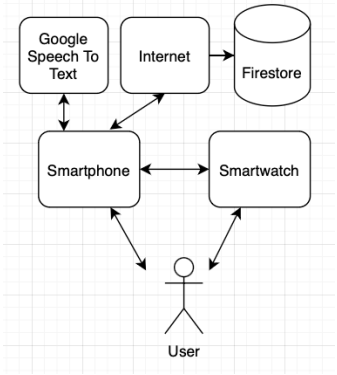
Dalam sistem ini, input suara dikonversi menjadi teks menggunakan layanan Google Speech-to-Text, kemudian diekstraksi menggunakan algoritma string matching Boyer-Moore yang dikenal cepat dalam

pencarian pola pada teks [7]. Sistem juga memanfaatkan teknologi Internet of Things (IoT) melalui integrasi smartwatch dan smartphone untuk mendukung pencatatan keuangan secara praktis dan sinkronisasi data secara real-time [8]. Dengan kombinasi teknologi ini, sistem diharapkan dapat mempermudah pengguna dalam mencatat keuangan pribadi secara lebih cepat, efisien, dan terintegrasi.

2. Metode Penelitian

Penelitian ini menerapkan metode pengembangan perangkat lunak model Waterfall, yang terdiri dari tahapan analisis kebutuhan, perancangan sistem, implementasi, dan pengujian. Sistem dirancang menggunakan platform Android dengan Kotlin serta Jetpack Compose, diintegrasikan dengan layanan cloud Firebase Firestore. Pengguna dapat melakukan pencatatan keuangan melalui smartwatch yang terhubung dengan smartphone. Proses input dilakukan menggunakan Google Speech-to-Text, sementara ekstraksi informasi dari teks dilakukan dengan algoritma Boyer-Moore.

Arsitektur sistem terdiri dari dua aplikasi utama, yaitu aplikasi smartwatch untuk menangkap suara, serta aplikasi smartphone untuk pengolahan data. Sistem mendukung sinkronisasi data secara real-time dengan Firestore. Gambar 1 berikut menjelaskan arsitektur sistem secara umum.



Gambar 1. Desain Sistem Informasi

Sistem juga diuji dengan pengujian fungsionalitas menggunakan blackbox testing. Metode ini berfokus pada pemeriksaan keluaran sistem berdasarkan masukan yang diberikan tanpa melihat proses internal. Tujuan pengujian ini adalah untuk memastikan seluruh kebutuhan fungsional dan non-fungsional sistem dapat berjalan sesuai dengan spesifikasi yang telah ditentukan.

Kebutuhan fungsional merupakan kemampuan utama yang harus dimiliki sistem untuk mendukung pencatatan keuangan berbasis suara. Daftar kebutuhan fungsional yang diuji adalah sebagai berikut:

1. Konversi Input Suara menjadi Teks – Sistem harus mampu melakukan transkripsi suara pengguna menjadi teks secara akurat menggunakan layanan Google Speech to Text.
2. Pengelolaan Data – Sistem harus menyediakan fitur pengelolaan data transaksi (penyimpanan, pembaruan, dan penghapusan) yang terintegrasi dengan basis data Firestore.
3. Konversi Teks menjadi Data Transaksi – Sistem harus mampu mengekstraksi informasi penting dari teks transkripsi, meliputi nominal serta jenis transaksi, untuk kemudian diklasifikasikan sesuai format yang ditentukan.
4. Pembuatan Laporan Keuangan – Sistem harus dapat menghasilkan laporan keuangan sederhana berdasarkan data transaksi yang telah tercatat, dengan format yang dapat diunduh pengguna.

Selain kebutuhan fungsional, sistem juga memiliki kebutuhan non-fungsional yang harus dipenuhi agar aplikasi dapat digunakan dengan baik dalam berbagai kondisi. Daftar kebutuhan non-fungsional yang diuji adalah sebagai berikut:

1. Keandalan Sistem – Sistem harus tetap mampu melakukan transkripsi dengan tingkat akurasi yang memadai dalam berbagai kondisi suara, serta memberikan opsi perbaikan manual jika terjadi kesalahan.
2. Kompatibilitas – Aplikasi harus dapat berjalan dengan lancar pada perangkat Android dengan spesifikasi rendah, sehingga dapat menjangkau lebih banyak pengguna.

p-ISSN: 2301-5373
e-ISSN: 2654-5101

3. Antarmuka Pengguna (User Interface) – Tampilan antarmuka harus mudah dipahami dan digunakan oleh rentang usia target pengguna, yaitu antara 16 hingga 32 tahun.

Lalu dilakukan juga pengujian performa algoritma Boyer-Moore dilakukan untuk mengevaluasi efisiensi pencarian pola teks pada variasi panjang input kalimat transaksi. Tujuan pengujian ini adalah untuk mengetahui hubungan antara panjang input dengan waktu eksekusi rata-rata, sehingga dapat dipastikan bahwa algoritma tetap responsif meskipun beban input bertambah.

Metode pengujian dilakukan dengan menyiapkan 10 kasus uji yang merepresentasikan variasi panjang input mulai dari 30 hingga 120 karakter. Setiap kasus uji berupa kalimat transaksi yang berisi informasi jenis transaksi, deskripsi, serta nominal dalam format rupiah. Rincian kasus uji dapat dilihat pada Tabel 1.

Tabel 1. Kasus Uji Algoritma Boyer Moore

Kasus uji	Jumlah kata
pengeluaran risol 5.000 rupiah	30
pengeluaran mie ayam bakso 10.000 rupiah	40
pengeluaran ayam geprek pedas level 5 12000 rupiah	50
pengeluaran le minerale kecil untuk minum siang 6.000 rupiah	60
pengeluaran beli cimory tiramisu di kantin kampus jimbaran 8000 rupiah	70
pengeluaran beli dua indomilk uht 950 ml di indomaret harga diskon 33.800 rupiah	80
pengeluaran beli bakso aci dengan telur di pedagang pinggir jalan depan rumah 20000 rupiah	90
pengeluaran beli mie instan, telur satu sak, sayur bayam, kol, sawi, pokcoy, dan wortel 14000 rupiah	100
pengeluaran belanja cimory pizza, pangsit, mi kuah, air mineral di kantin unud jimbaran siang ini 17000 rupiah	110
pengeluaran membeli lauk ayam, sayur, telor satu sak, rice bowl, dan buah pisang di pasar sengol batubulan 150000 rupiah	120

Data hasil pengujian selanjutnya dianalisis menggunakan regresi linear guna memperoleh hubungan matematis antara panjang input (x) dan waktu eksekusi (y).

Selain pengujian fungsional dan performa sistem, penelitian ini juga melakukan evaluasi penerimaan pengguna menggunakan pendekatan Technology Acceptance Model (TAM). Model ini dipilih karena secara luas digunakan dalam menganalisis faktor-faktor yang memengaruhi penerimaan teknologi baru oleh pengguna. Instrumen yang digunakan berupa kuesioner dengan skala Likert 1–5, di mana angka 1 merepresentasikan sangat tidak setuju dan angka 5 merepresentasikan sangat setuju.

Tabel 2. Pertanyaan TAM

Aspect Code	Question Code	Question	Score
	Q1	Saya mengetahui terdapat fitur konversi suara ke teks pada smartphone saya.	
	Q2	Saya sering menggunakan fitur konversi suara ke teks pada smartphone saya.	

Perceived Ease of Use (PEOU)	Q3	Saya merasa tidak mengalami kesulitan saat menggunakan fitur-fitur aplikasi ini.	
	Q4	Proses sinkronisasi data antara smartwatch dan smartphone berjalan dengan mudah.	
	Q5	Input transaksi melalui suara terasa lebih mudah dibandingkan metode mengetik.	
	Q6	Input transaksi melalui suara terasa lebih cepat dibandingkan mengetik.	
	Q7	Saya merasa nyaman memperbaiki transaksi yang keliru pada smartphone.	
	Q8	Sistem ini membuat proses mencatat transaksi terasa tidak merepotkan lagi.	
Perceived Usefulness (PU)	Q9	Aplikasi ini mempermudah saya dalam membuat laporan keuangan sederhana.	
	Q10	Sistem ini cukup andal untuk mencatat beberapa transaksi sekaligus melalui input suara.	
	Q11	Dengan adanya smartwatch pada sistem IoT ini, menjadikannya tetap efektif, praktis, dan cepat digunakan saat smartphone tidak dapat dijangkau.	
Attitude Toward Using (ATU)	Q12	Saya merasa puas dengan keseluruhan kinerja aplikasi ini.	
	Q13	Saya puas dengan desain antarmuka pengguna pada smartphone dan smartwatch.	
	Q14	Saya puas dengan pengalaman penggunaan aplikasi ini secara umum.	
	Q15	Saya merasa fitur-fitur yang disediakan sudah cukup untuk kebutuhan pengelolaan keuangan saya.	
Behavioral Intention to Use (BI)	Q16	Saya berencana untuk menggunakan input suara dalam mencatat keuangan pribadi.	
	Q17	Saya berencana menggunakan sistem IoT ini secara rutin dalam mencatat keuangan.	
	Q18	Saya mempertimbangkan untuk membeli smartwatch demi mendapatkan kepraktisan dalam pencatatan.	
	Q19	Saya mempertimbangkan untuk membeli smartwatch demi mendapatkan kepraktisan dalam pencatatan. Dengan asumsi harga smartwatch Wear OS adalah 2 juta rupiah	
	Q20	Saya yakin bahwa penggunaan sistem ini dapat membantu saya mencatat keuangan kapan pun transaksi terjadi sehingga meminimalisir lupa.	

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Actual Use (AU)	Q21	Saya dapat menggunakan aplikasi ini tanpa bantuan orang lain setelah sesi edukasi.	
	Q22	Saya merasa fitur-fitur yang saya coba dalam pengujian mewakili kebutuhan saya dalam mencatat keuangan pribadi.	
	Q23	Saya merasa aneh saat ingin mengucapkan input transaksi dengan suara di tempat yang cukup ramai	
	Q24	Saya merasa cukup nyaman saat mencatat dengan mendekatkan smartwatch ke mulut agar tidak banyak membuat suara	
	Q25	Saya merasa terganggu jika terdapat kesalahan pada transkripsi yang terjadi saat mencatat melalui smartwatch	
	Q26	Saya merasa mudah untuk mengingat kembali transaksi yang terjadi sebenarnya jika transkripsi salah dan harus dikoreksi	

3. Hasil dan Pembahasan

Sistem berhasil diimplementasikan dalam dua aplikasi terpisah pada perangkat smartwatch dan smartphone. Pengguna dapat dengan mudah menginput transaksi melalui suara menggunakan smartwatch, kemudian memproses dan mengelola data transaksi melalui smartphone. Pengujian blackbox menunjukkan seluruh fungsi berjalan sesuai kebutuhan, termasuk konversi suara, deteksi nominal, dan sinkronisasi data.

3.1. Hasil Pengujian Blackbox

Pengujian blackbox dilakukan untuk memastikan bahwa sistem memenuhi seluruh kebutuhan fungsional yang telah dirancang. Berdasarkan pengujian, sistem mampu menjalankan fungsi utama seperti konversi suara ke teks, pengelolaan data transaksi, ekstraksi informasi transaksi dengan algoritma string matching, serta pembuatan laporan keuangan. Hasil detail dari pengujian kebutuhan fungsional dapat dilihat pada Tabel 2.

Tabel 2. Hasil Pengujian Fungsional dengan Blackbox

Fitur yang Diuji	Skema Pengujian	Ekspektasi Hasil	Hasil Aktual
Konversi Input Suara menjadi Teks	Penulis mengumpulkan 160 data pencatatan keuangannya dengan Google Speech to Text	Teks hasil transkripsi sesuai dengan suara yang diucapkan	Mendapatkan akurasi 88% terlampir pada lampiran 2
Pengelolaan Data	Pengguna melakukan pengelolaan data dengan aplikasi	Data tersimpan di Firestore terkelola sesuai keinginan user	Sistem berhasil menyimpan dan mengelola data di Firestore sesuai keinginan user seperti yang terlampir pada lampiran 4, 5, dan 6
Konversi	Penulis menggunakan 160	Sistem mengenali	Sistem berhasil

Teks menjadi Data Transaksi	data pencatatan keuangannya untuk Menguji fungsi ini	nominal dan klasifikasi transaksi	mengonversi seluruh data uji sesuai dengan hasil yang diharapkan
Pembuatan Laporan Keuangan	Pengguna membuat laporan keuangan	Sistem menampilkan laporan dari data transaksi	Sistem berhasil membuat laporan seluruh data transaksi dengan format csv

Selain kebutuhan fungsional, pengujian juga dilakukan pada aspek non-fungsional untuk memastikan sistem memiliki keandalan, kompatibilitas, dan antarmuka yang mudah dipahami oleh pengguna. Hasilnya menunjukkan bahwa sistem mampu mempertahankan kinerja yang stabil pada berbagai kondisi, berjalan dengan baik di perangkat berspesifikasi rendah, dan mendapatkan respon positif dari pengguna dalam hal kemudahan penggunaan antarmuka. Rincian hasil pengujian kebutuhan non-fungsional dapat dilihat pada Tabel 3.

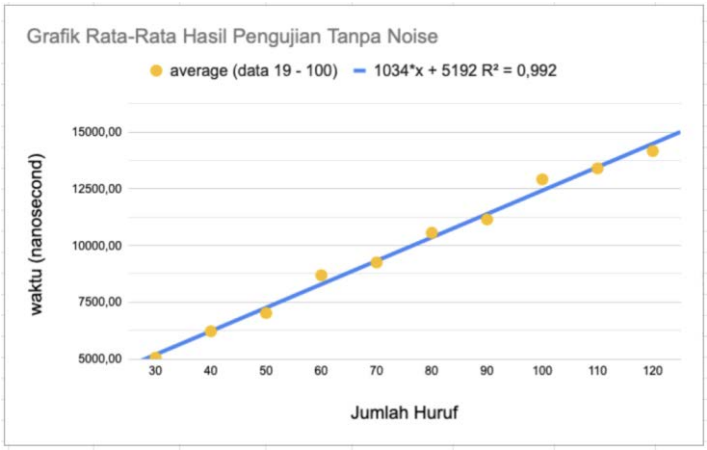
Tabel 3. Hasil Pengujian Fungsional dengan Blackbox

Aspek yang Diuji	Skema Pengujian	Ekspektasi Hasil	Hasil Aktual
Keandalan Sistem	Transkripsi dilakukan berulang kali pada berbagai kondisi suara	Sistem tetap mengenali dengan akurat, kesalahan bisa diperbaiki manual	Dapat diperbaiki manual
Kompatibilitas	Pengujian pada perangkat Android berspesifikasi rendah	Aplikasi dapat berjalan dengan lancar	Aplikasi Berjalan Lancar
Antarmuka Pengguna	Uji pada pengguna usia 16–32 tahun	UI mudah digunakan dan dipahami	Pertanyaan Q13 pengujian TAM di Tabel 4.8.2 mendapatkan hasil 4.1 dari 5

3.2. Hasil Pengujian Algoritma Boyer Moore

Pengujian performa algoritma Boyer-Moore dilakukan dengan panjang input teks bervariasi antara 30 hingga 120 karakter. Untuk mengurangi bias eksekusi akibat lingkungan JVM, setiap percobaan diulang sebanyak 100 kali, kemudian diambil nilai rata-rata dari iterasi ke-19 hingga ke-100. Hasil rata-rata eksekusi ditunjukkan pada Gambar, dengan nilai eksekusi terkecil sekitar 5.000 ns pada panjang input 30 karakter, dan meningkat hingga sekitar 14.000 ns pada panjang input 120 karakter.

p-ISSN: 2301-5373
e-ISSN: 2654-5101



Gambar 2. Grafik yang didapat dari hasil boyer moore

Analisis regresi linear terhadap data pengujian menghasilkan persamaan y = 1034x + 5192 dengan nilai koefisien determinasi R^2 = 0,992. Nilai slope sebesar 1034 menunjukkan bahwa setiap penambahan satu karakter pada panjang input akan meningkatkan waktu eksekusi rata-rata sekitar 1034 ns, sedangkan nilai intercept sebesar 5192 mengindikasikan waktu eksekusi dasar ketika panjang input berada pada titik awal pengujian, yaitu 30 karakter. Nilai R^2 yang sangat tinggi mendekati 1 menunjukkan bahwa model linear memiliki tingkat kecocokan yang sangat baik terhadap data hasil pengujian.

Sebagai contoh, prediksi waktu eksekusi untuk teks sepanjang 130 karakter dapat dihitung dengan persamaan tersebut, yaitu y = 1034(130) + 5192 = 139.614 ns. Hal ini menunjukkan bahwa performa algoritma Boyer-Moore bersifat linear dan efisien, sehingga dapat diandalkan untuk proses pencocokan string pada input teks yang relatif panjang dalam sistem pencatatan keuangan berbasis suara.

3.3. Hasil Pengujian Kuesioner TAM

Pengujian penerimaan pengguna dilakukan terhadap 44 responden dengan menggunakan kuesioner berbasis Technology Acceptance Model (TAM). Setiap responden diminta memberikan jawaban pada 26 butir pertanyaan yang telah dipetakan ke dalam lima aspek utama TAM, yaitu Perceived Ease of Use (PEOU), Perceived Usefulness (PU), Attitude Toward Using (ATU), Behavioral Intention to Use (BI), dan Actual Use (AU). Skala Likert 1–5 digunakan untuk mengukur tingkat persetujuan responden, mulai dari sangat tidak setuju hingga sangat setuju.

Tabel 3. Hasil Presentase Skor Kuesioner TAM Partisipan

Kode pertanyaan	Skor 1	Skor 2	Skor 3	Skor 4	Skor 5
Q1	1 (2,33%)	4 (9,3%)	8 (18,6%)	9 (20,93%)	21 (48,84%)
Q2	4 (9,3%)	8 (18,6%)	13 (30,23%)	8 (18,6%)	10 (23,26%)
Q3	2 (4,65%)	1 (2,33%)	10 (23,26%)	13 (30,23%)	17 (39,53%)
Q4	2 (4,65%)	5 (11,63%)	5 (11,63%)	10 (23,26%)	21 (48,84%)
Q5	2 (4,65%)	3 (6,98%)	3 (6,98%)	6 (13,95%)	29 (67,44%)
Q6	1 (2,33%)	2 (4,65%)	2 (4,65%)	11 (25,58%)	27 (62,79%)
Q7	1 (2,33%)	5 (11,63%)	11 (25,58%)	15 (34,88%)	11 (25,58%)
Q8	1 (2,33%)	3 (6,98%)	6 (13,95%)	12 (27,91%)	21 (48,84%)
Q9	2 (4,65%)	3 (6,98%)	5 (11,63%)	12 (27,91%)	21 (48,84%)
Q10	3 (6,98%)	3 (6,98%)	2 (4,65%)	16 (37,21%)	19 (44,19%)

Q11	0 (0%)	5 (11,63%)	6 (13,95%)	11 (25,58%)	21 (48,84%)
Q12	2 (4,65%)	2 (4,65%)	4 (9,3%)	19 (44,19%)	16 (37,21%)
Q13	2 (4,65%)	5 (11,63%)	6 (13,95%)	18 (41,86%)	12 (27,91%)
Q14	1 (2,33%)	3 (6,98%)	2 (4,65%)	19 (44,19%)	18 (41,86%)
Q15	1 (2,33%)	4 (9,3%)	6 (13,95%)	16 (37,21%)	16 (37,21%)
Q16	1 (2,33%)	4 (9,3%)	10 (23,26%)	13 (30,23%)	15 (34,88%)
Q17	1 (2,33%)	4 (9,3%)	12 (27,91%)	14 (32,56%)	12 (27,91%)
Q18	1 (2,33%)	8 (18,6%)	13 (30,23%)	9 (20,93%)	12 (27,91%)
Q19	2 (4,65%)	6 (13,95%)	15 (34,88%)	13 (30,23%)	7 (16,28%)
Q20	1 (2,33%)	2 (4,65%)	8 (18,6%)	15 (34,88%)	17 (39,53%)
Q21	2 (4,65%)	6 (13,95%)	9 (20,93%)	14 (32,56%)	12 (27,91%)
Q22	1 (2,33%)	3 (6,98%)	5 (11,63%)	16 (37,21%)	18 (41,86%)
Q23	1 (2,33%)	13 (30,23%)	11 (25,58%)	10 (23,26%)	8 (18,6%)
Q24	2 (4,65%)	3 (6,98%)	8 (18,6%)	19 (44,19%)	11 (25,58%)
Q25	1 (2,33%)	5 (11,63%)	15 (34,88%)	11 (25,58%)	11 (25,58%)
Q26	3 (6,98%)	1 (2,33%)	10 (23,26%)	15 (34,88%)	14 (32,56%)

Berdasarkan hasil pengolahan data (Tabel 3), terlihat bahwa sebagian besar responden memberikan skor tinggi (4 atau 5) pada mayoritas pertanyaan, terutama pada aspek kemudahan input suara dibandingkan mengetik (Q5 dan Q6) yang mencapai persentase lebih dari 60% responden memilih “sangat setuju”. Sebaliknya, pertanyaan terkait penggunaan di ruang publik (Q23) menunjukkan distribusi jawaban yang lebih beragam, dengan sebagian responden merasa canggung menggunakan input suara di tempat ramai. Hal ini menunjukkan adanya faktor situasional yang memengaruhi penerimaan teknologi.

Tabel 4. Hasil Presentase Skor Kuesioner TAM Partisipan

Aspek Uji	Skor
Perceived Ease of Use (PEOU)	4,09
Perceived Usefulness (PU)	4,08
Attitude Toward Using (ATU)	4,08
Behavioral Intention to Use (BI)	4,07
Actual Use (AU)	4,08

Secara agregat (Tabel 4), rata-rata skor pada setiap aspek TAM berada di atas 4,0 dari skala 5. Aspek dengan nilai tertinggi adalah Perceived Ease of Use (PEOU) dengan skor 4,09, diikuti oleh Perceived Usefulness (PU) sebesar 4,08, serta Attitude Toward Using (ATU), Behavioral Intention to Use (BI), dan Actual Use (AU) yang masing-masing juga berada pada kisaran 4,07–4,08. Hasil ini mengindikasikan bahwa responden menilai sistem mudah digunakan, bermanfaat, dan berpotensi besar untuk diadopsi secara rutin. Temuan ini juga mengonfirmasi bahwa integrasi IoT dengan input suara dapat memberikan pengalaman pencatatan keuangan yang lebih praktis dan efisien dibandingkan metode manual.

4. Kesimpulan

Sistem manajemen keuangan pribadi berbasis IoT dengan integrasi Google Speech-to-Text dan algoritma Boyer-Moore berhasil dikembangkan dan berfungsi dengan baik untuk mencatat transaksi secara otomatis melalui input suara. Sistem mampu mendeteksi informasi transaksi secara efisien dan menyinkronkannya ke database cloud secara real-time.

Hasil pengujian menunjukkan sistem berjalan dengan performa baik dari sisi fungsionalitas dan kecepatan algoritma string matching. Berdasarkan hasil evaluasi TAM, sistem mendapat respon positif dari pengguna, menunjukkan bahwa pendekatan ini layak untuk diterapkan dalam solusi pencatatan keuangan harian.

Referensi

[1] K. Palin et al., "How do people type on mobile devices? Observations from a study with 37,000 volunteers," in Proc. 21st Int. Conf. Human-Computer Interaction with Mobile Devices and Services, pp. 1-12, 2019.
[2] M. Brysbaert, "How many words do we read per minute? A review and meta-analysis of reading rate," J. Memory Lang., vol. 109, p. 104047, 2019.
[3] V. Kuperman et al., "A lingering question addressed: Reading rate and most efficient listening rate are highly similar," J. Exp. Psychol. Hum. Percept. Perform., vol. 47, no. 8, pp. 1103, 2021.
[4] L. Wang, "British English-speaking speed 2020," Acad. J. Humanit. Soc. Sci., vol. 4, pp. 93-100, 2021.
[5] W. W. Yanti, "Peran pencatatan keuangan harian sebagai filter diri dalam pengaturan pengeluaran pribadi," J. Dinam. Sos. Budaya, vol. 22, no. 2, pp. 257-271, 2020.
[6] S. B. Pudji, "Aplikasi konversi suara ke teks berbasis android menggunakan Google Speech API," Bianglala Informatika, vol. 2, no. 2, 2014.
[7] A. R. Chayapathi, "Survey and comparison of string matching algorithms," Turk. J. Comput. Math. Educ., vol. 12, no. 12, pp. 1471-1491, 2021.
[8] R. A. R. A. Mouha, "Internet of Things (IoT)," J. Data Anal. Inf. Process., vol. 9, no. 2, pp. 77, 2021.

This page is intentionally left blank.

Jurnal Elektronik Ilmu Komputer Udayana
Volume 14, No 3. February 2026

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Perbandingan Performa CNN Inception-V3 dan ResNet50 Dalam Klasifkasi Kematangan Buah Tomat

I Putu Hanggara Diatha Putra[a1], I Wayan Supriana[a2], I Ketut Gede Suhartana[a3], I Gusti Ngurah Anom Cahyadi Putra[a4], Farhan Akmala Putra[b5]

[a]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana Kuta Selatan, Badung, Bali, Indonesia
[b]Magister Agribisnis, Fakultas Peternakan dan Pertanian, Universitas Diponegoro, Indonesia
[1]putuhanggaradiatha@gmail.com
[2]wayan.supriana@unud.ac.id
[3]ikg.suhartana@unud.ac.id
[4]anom.cp@unud.ac.id
[5]farhanakmalaputra@gmail.com

Abstrak

Buah tomat memiliki durasi mencapai kematangan yang singkat. Sehingga jika melewati durasi kematangannya, maka buah tomat akan seketika membusuk. Para petani buah tomat melakukan proses klasifikasi kematangan buah tomat secara manual. Proses klasifikasi secara manual dapat mengakibatkan hasil klasifikasi yang tidak konsisten, yang disebabkan karena proses tersebut sangat menguras tenaga manusia jika dilakukan dalam kuantitas yang banyak. Oleh karena itu, penulis merancang sistem klasifikasi tingkat kematangan buah tomat dengan algoritma Convolutional Neural Network (CNN). Metode yang digunakan dalam penelitian ini adalah model CNN dengan arsitektur Inception-V3 dan arsitektur ResNet50. Penulis juga melakukan perbandingan performa klasifikasi dari kedua arsitektur tersebut dan membandingkan arsitektur mana yang memiliki performa klasifikasi yang lebih baik. Berdasarkan hasil pengujian yang dilakukan, nilai akurasi terbaik diperoleh ketika menggunakan model CNN dengan arsitektur ResNet50 dengan tingkat akurasi sebesar 95%. Lalu, berdasarkan pengujian menggunakan data uji menunjukkan kedua model CNN tersebut dapat melakukan klasifikasi citra dengan optimal dan tanpa kesalahan.

Kata Kunci : Sistem Klasifikasi, *Convolutional Neural Network*, Inception-V3, ResNet50, Tomat

1. Pendahuluan

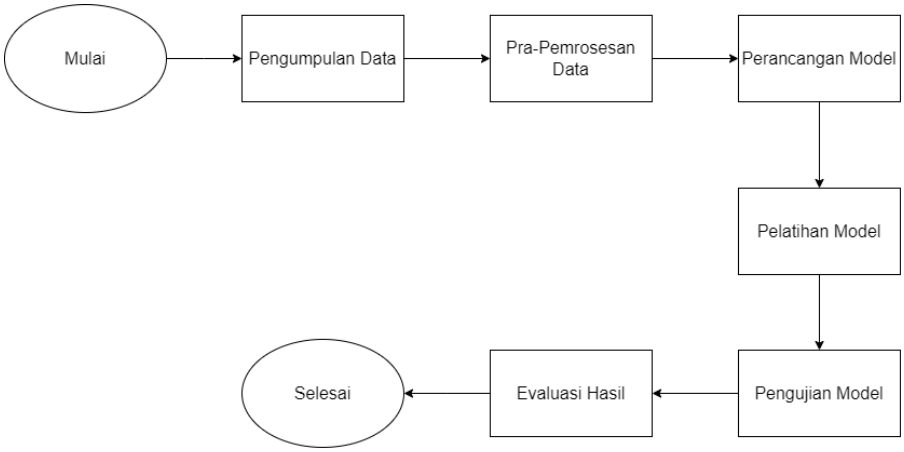
Buah tomat adalah salah satu tanaman yang potensial tumbuh di Indonesia. Hal ini dapat dibuktikan dengan produksi buah tomat yang mencapai 1.143.787,7 ton pada tahun 2023, menempati urutan keenam dari 10 besar produksi tanaman sayuran di Indonesia [1]. Buah tomat memiliki manfaat bagi tubuh karena mengandung vitamin, mineral, dan senyawa likopen sebagai antioksian [2]. Tantangan utama yang dihadapi oleh petani saat meningkatnya produksi buah tomat adalah proses klasifikasi kematangan buah tomat yang masih dilakukan secara manual. Cara tersebut tidak hanya menguras tenaga, tetapi juga rawan menghasilkan klasifikasi yang tidak konsisten, sehingga dapat menyebabkan kerugian ekonomi jika buah tomat membusuk yang disebabkan karena panen yang tidak tepat waktu [3].

Seiring berkembangnya teknologi kecerdasan buatan, implementasi *deep learning* dapat mengatasi masalah tersebut. Salah satu algoritma *deep learning* yang banyak digunakan di berbagai bidang untuk klasifikasi citra adalah *convolutional neural network* (CNN) [4]. Dalam algoritma CNN, terdapat beberapa jenis arsitektur dan arsitektur yang sudah terbukti memiliki akurasi yang tinggi adalah Inception-V3 dan ResNet50. Algoritma CNN juga banyak digunakan dalam bidang pertanian. Penggunaan CNN dalam klasifikasi kematangan buah kopi kuning menggunakan arsitektur Inception-V3 menghasilkan performa klasifikasi yang mengesankan, dengan nilai akurasi sebesar 92% [5]. Kemudian, implementasi dari CNN ResNet50 dalam mengklasifikasi kematangan buah pisang menghasilkan figur nilai akurasi sebesar 91% [6].

Berdasarkan penelitian terdahulu, penelitian ini bertujuan untuk membandingkan performa antara arsitektur Inception-V3 dan ResNet50 dalam klasifikasi tingkat kematangan buah tomat. Hasil dari penelitian ini dapat mengetahui arsitektur mana yang memiliki performa klasifikasi yang akurat dan stabil, sehingga dapat membantu petani dalam mengklasifikasi tingkat kematangan buah tomat.

2. Metode Penelitian

Penelitian ini dilakukan dengan mengikuti alur penelitian agar proses penelitian berjalan dengan terarah. Alur penelitian merupakan tahapan-tahapan proses penelitian berdasarkan pada langkah-langkah seperti pada gambar di bawah ini :



Gambar 1. Alur Penelitian

2.1. Pengumpulan Data

Dalam pengumpulan data, dataset yang digunakan untuk proses melatih model CNN bersumber dari situs Mendeley Data dan Dataset Ninja [7][8]. Total data yang digunakan adalah 530 gambar, yang terdiri dari 260 gambar pada kelas belum matang dan 260 gambar pada kelas matang. Kemudian, data tersebut dibagi dengan rasio 76% untuk data *training*, 23% untuk data *validation*, dan 1% untuk data *testing*. Berikut rincian jumlah dan contoh dari dataset yang akan digunakan. Gambar isi kiri merupakan contoh gambar dari kelas belum matang dan gambar sisi kanan adalah contoh gambar dari kelas matang.

Tabel 1. Rincian Jumlah Dataset

Belum Matang			**Matang**		
Training	*Validation*	*Testing*	*Training*	*Validation*	*Testing*
200 citra	60 citra	5 citra	200 citra	60 citra	5 citra
Total Dataset		**265 citra**	**Total Dataset**		**265 citra**
Total Keseluruhan					**530 citra**



Gambar 2. Sampel Dataset Buah Tomat [7]

p-ISSN: 2301-5373
e-ISSN: 2654-5101

2.2. Pra-Pemrosesan Data

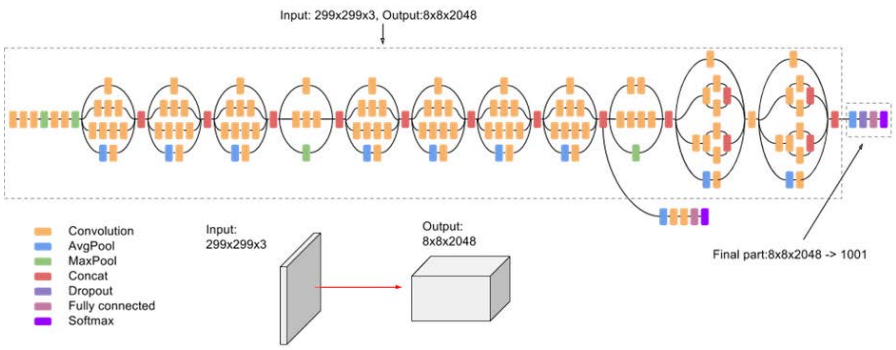
Pada tahap ini, dataset akan dilakukan pra-pemrosesan (*pre-processing*), sebuah tahap penyesuaian untuk menyeragamkan setiap data gambar buah tomat dalam satu dataset. Tahap pertama dari pra-pemrosesan adalah *rescaling*, untuk menyesuaikan nilai piksel dari [0,255] menjadi [0,1]. Kemudian, dilanjutkan ke tahap *resize* untuk menyesuaikan ukuran gambar dengan *input* standar dari masing-masing arsitektur, yaitu 299x299 piksel untuk arsitektur Inception-V3 [9]. Sedangkan ukuran standar dari arsitektur ResNet50 adalah 224x224 piksel [10]. Setelah dilakukan tahap *resize*, dilanjutkan ke tahap augmentasi data. Tahap ini dilakukan untuk menambah variasi data yang bertujuan untuk mengurangi risiko model CNN mengalami *overfitting*. Parameter yang digunakan dalam augmentasi data adalah *rotation range*, *shear range*, *zoom range*, *width shift range*, *height shift range*, *horizontal flip*, dan *fill mode*. Berikut Tabel 1 yang merupakan rincian dari parameter augmentasi data yang digunakan.

Tabel 1. Parameter Augmentasi Data

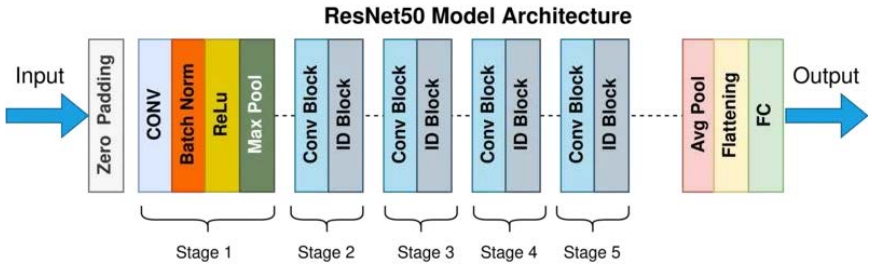
No	Parameter	Nilai
1	Rotation Range	40
2	Shear Range	0,2
3	Width Shift Range	0,2
4	Height Shift Range	0,2
5	Zoom Range	0,3
6	Horizontal Flip	True
7	Fill Mode	‘nearest’

2.3. Perancangan Model CNN Inception-V3

Setelah melalui tahap pra-pemrosesan, dilakukan tahap perancangan yang bertujuan untuk merancang model CNN yang akan digunakan sebagai model klasifikasi kematangan buah tomat. Pada tahap ini, penulis merancang model CNN dengan arsitektur Inception-V3 dan kemudian dilanjutkan dengan arsitektur ResNet50. Arsitektur Inception-V3 memiliki kernel konvolusi metode pemisahan yang berfungsi untuk memecah konvolusi besar menjadi konvolusi kecil, sehingga proses pelatihan model lebih efisien dan meningkatkan kemampuan dalam mengekstrak fitur spasial [11]. Sedangkan, arsitektur ResNet50 menggunakan konsep *residual connection*, yang memungkinkan arsitektur ResNet50 dalam menggunakan fitur dari lapisan sebelumnya agar risiko kehilangan fitur penting saat proses konvolusi dapat diminimalisir [12]. Ilustrasi dari arsitektur Inception-V3 dan ResNet50 dapat dilihat pada gambar di bawah ini.



Gambar 3. Arsitektur Inception-V3 [13]



Gambar 4. Arsitektur ResNet50 [14]

2.4. Pelatihan Model

Setelah selesai merancang kedua model CNN tersebut, dilanjutkan proses pelatihan model. Dalam proses pelatihan, penulis menggunakan beberapa skenario penggunaan *hyperparameter* dalam pelatihan model. Rincian dari nilai *hyperparameter* yang digunakan dapat dilihat pada tabel di bawah ini.

Tabel 2. Skenario Hyperparameter

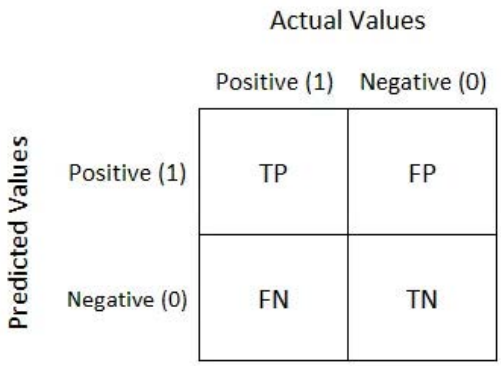
No	Parameter	Nilai
1	Epoch	10, 50, 100, 150
2	Learning rate	0,0001 dan 0,001

2.5. Pengujian Model

Setelah proses pelatihan selesai, dilanjutkan ke tahap pengujian model. Pengujian model menggunakan data *testing* yang telah disediakan di masing-masing kelas. Data *testing* terdiri dari 5 gambar dari masing-masing kelas.

2.6. Evaluasi Hasil

Tahap evaluasi hasil akan menggunakan metode confusion matrix, sebuah metode yang menghasilkan jumlah gambar beserta hasil prediksinya dari setiap kelas dengan nilai diagonal sebagai jumlah prediksi yang benar [10]. Selain itu, evaluasi hasil juga menggunakan evaluasi metrik yang terdiri dari *accuracy*, *precision*, *recall*, dan *f1-score*. Metrik *accuracy* akan mengukur persentase prediksi benar yang dihasilkan dari semua prediksi kelas, metrik *precision* akan mengukur tingkat keberhasilan suatu model dalam memprediksi kelas positif dari semua prediksi positif yang dihasilkan, metrik *recall* akan mengukur sejauh mana model mengklasifikasi kelas positif yang sebenarnya dengan akurat, lalu metrik *f1-score* akan mengukur kemampuan precision dan recall dari suatu model.



Gambar 5. Ilustrasi Confusion Matrix [15]

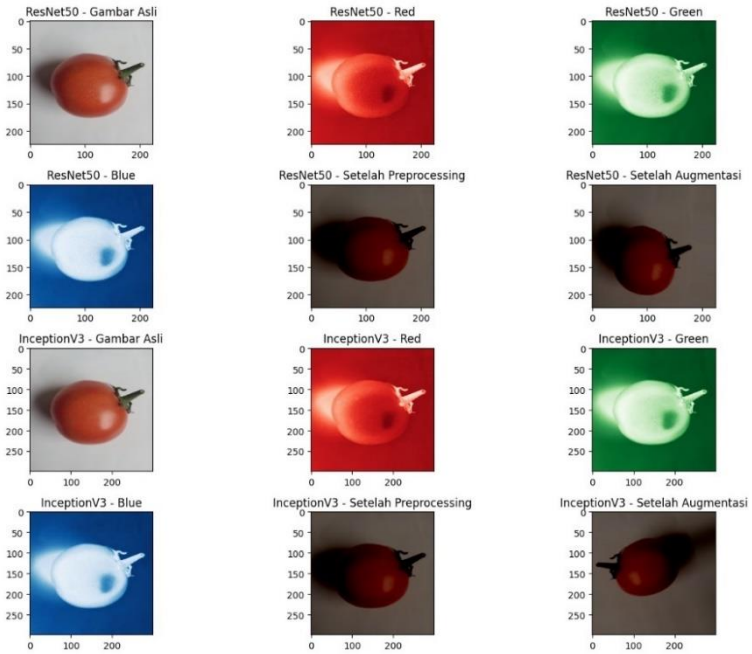
p-ISSN: 2301-5373
e-ISSN: 2654-5101

Tabel 3. Ilustrasi Hasil Metrik Evaluasi

No	Metrik	Nilai
1	*Accuracy*	0.8448
2	*Precision*	0.8479
3	*Recall*	0.8479
4	*F1-score*	0.8448

3. Hasil dan Pembahasan

Proses klasifikasi kematangan buah tomat menggunakan total data sebanyak 530 gambar, dengan persentase pembagian data 76% untuk data *training*, 23% untuk data *validation*, dan 1% untuk data *testing*. Sebelum melatih model CNN Inception-V3 dan ResNet50, dilakukan tahap pra-pemrosesan untuk menyeragamkan setiap data gambar buah tomat dalam satu dataset, agar proses pelatihan model lebih optimal. Berikut gambar dari hasil pra-pemrosesan yang terdiri dari *rescaling* dan *resize* yang dapat dilihat pada Gambar 6 di bawah ini



Gambar 6. Hasil Pra-Pemrosesan

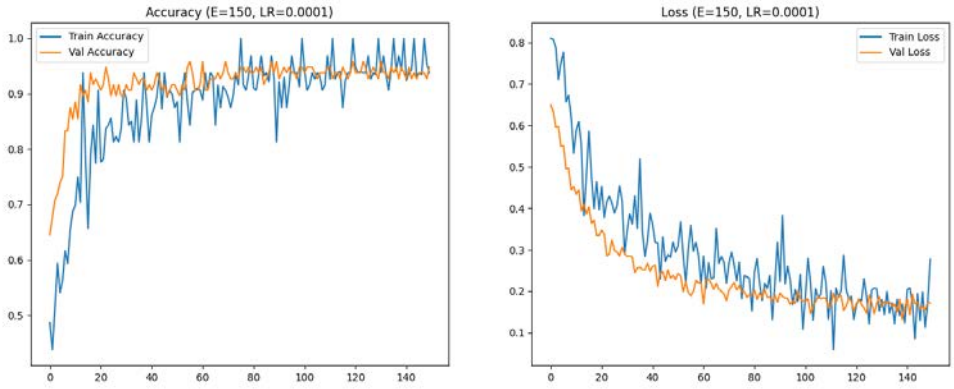
Setelah melakukan pra-pemrosesan, dilakukan tahap augmentasi data untuk memberikan variasi data agar risiko *overfitting* saat pelatihan model dapat berkurang. Parameter yang digunakan dalam proses augmentasi data terdiri dari *rotation range*, *shear range*, *zoom range*, *width shift range*, *height shift range*, *horizontal flip*, dan *fill mode*. Hasil dari proses augmentasi data dapat dilihat pada Gambar 7 di bawah ini.



Gambar 7. Hasil Augmentasi Data

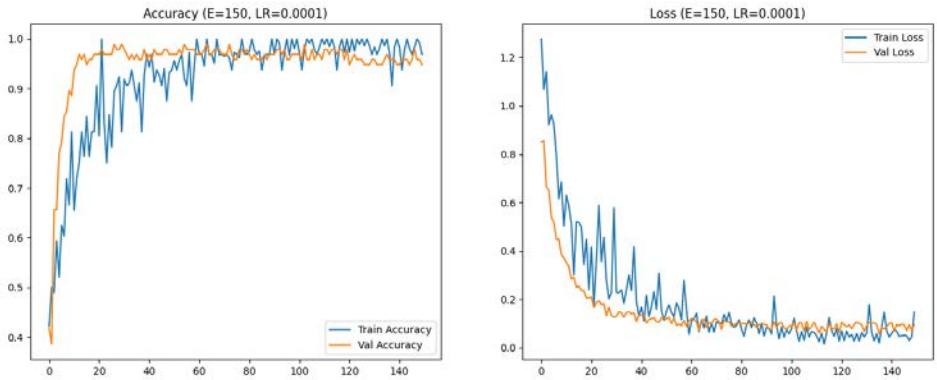
Kemudian dilanjutkan ke proses pelatihan model CNN Inception-V3 dan ResNet50. Berdasarkan proses pelatihan kedua model tersebut, diperoleh parameter yang menghasilkan tingkat akurasi paling

baik dari model CNN Inception-V3 dan ResNet50. Kedua model tersebut menghasilkan tingkat akurasi yang sama pada parameter *epoch* 150 dan nilai *learning rate* 0,0001. Model CNN Inception-V3 pada *epoch* 150 menghasilkan nilai *training accuracy* sebesar 93,75% dan nilai *validation accuracy* sebesar 94,79%. Sedangkan nilai *training loss* yang dihasilkan sebesar 27,63% dan *validation loss* sebesar 17,15%. Berikut grafik performa dari model CNN Inception-V3 yang dapat dilihat pada Gambar 8 di bawah ini.



Gambar 8. Grafik Performa Model CNN Inception-V3

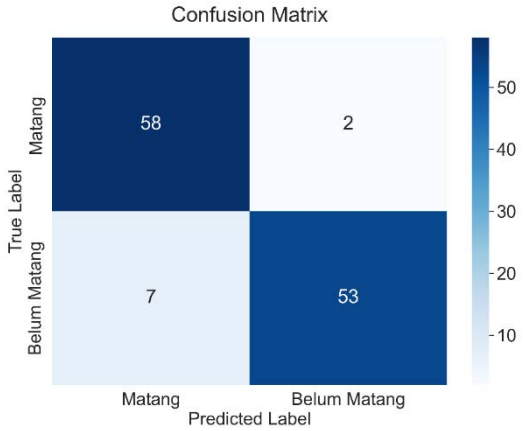
Di sisi lain, model CNN ResNet50 pada nilai *epoch* yang sama, yaitu *epoch* 150 menghasilkan nilai *training accuracy* sebesar 96,87% dan nilai *validation accuracy* sebesar 94,79%. Sedangkan nilai *training loss* yang dihasilkan sebesar 14,67% dan *validation loss* sebesar 9,46%. Berikut grafik performa dari model CNN ResNet50 yang dapat dilihat pada Gambar 9 di bawah ini.



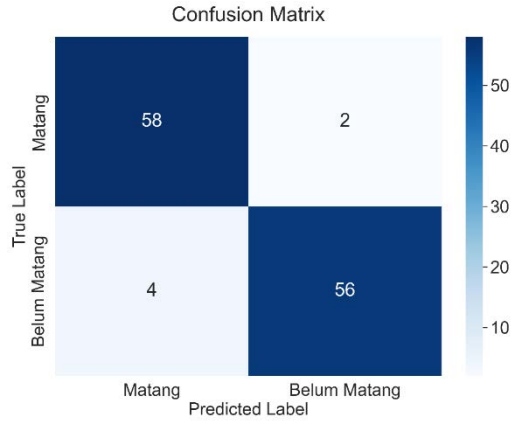
Gambar 9. Grafik Performa Model CNN ResNet50

Berdasarkan kedua grafik pada Gambar 8 dan 9, pada kurva grafik *training* terlihat bahwa arsitektur Inception-V3 mengalami fluktuasi nilai akurasi yang cukup tinggi. Dibandingkan dengan arsitektur ResNet50 yang walaupun juga mengalami fluktuasi nilai, tetapi pada pertengahan proses pelatihan model kurva bergerak lebih stabil setelah melewati *epoch* 60. Sedangkan di sisi kurva grafik *validation*, kedua arsitektur tersebut memiliki performa yang stabil. Namun, arsitektur ResNet50 memiliki performa *validation* yang lebih stabil daripada arsitektur Inception-V3. Kemudian, berdasarkan metode *confusion matrix* kedua model menunjukkan perbedaan performa yang tipis. Dari 60 gambar pada kelas belum matang dan 60 gambar pada kelas matang, terdapat 2 gambar yang salah diidentifikasi oleh model CNN Inception-V3, yaitu kelas matang yang salah diklasifikasi sebagai kelas belum matang, serta 7 gambar yang diklasifikasi sebagai kelas matang yang seharusnya adalah kelas belum matang. Sedangkan model CNN ResNet50 dengan jumlah gambar dan kelas yang sama, terdapat 2 gambar yang juga salah diklasifikasi seperti pada model CNN Inception-V3, yaitu kelas matang yang salah diklasifikasi sebagai kelas belum matang, serta 6 gambar yang diklasifikasi sebagai kelas matang yang seharusnya adalah kelas belum matang. Berikut adalah gambar dari *confusion matrix* dari kedua model tersebut yang dapat dilihat pada Gambar 10 dan 11.

p-ISSN: 2301-5373
e-ISSN: 2654-5101



Gambar 10. *Confusion Matrix* Model CNN Inception-V3



Gambar 11. *Confusion Matrix* Model CNN ResNet50

Kemudian, perbandingan antara arsitektur Inception-V3 dan ResNet50 dengan 4 metrik evaluasi (*accuracy, precision, recall,* dan *f1-score*) menunjukkan bahwa Inception-V3 menghasilkan nilai *accuracy* sebesar 94,16%, nilai *precision* sebesar 94,28%, nilai *recall* sebesar 94,17%, dan *f1-score* sebesar 94,16%. Sedangkan arsitektur ResNet50 menghasilkan nilai *accuracy* sebesar 95%, nilai *precision* sebesar 95,05%, lalu nilai *recall* serta *f1-score* yang sama-sama menghasilkan nilai 95%. Berdasarkan hasil tersebut, model CNN ResNet50 memiliki performa yang tidak berbeda jauh dengan model CNN Inception-V3. Berikut Tabel 4 dan 5 dari hasil metrik evaluasi dari kedua arsitektur tersebut.

Tabel 4. Hasil Metrik Evaluasi Model CNN Inception-V3

No	Metrik	Nilai
1	*Accuracy*	0.9417
2	*Precision*	0.9428
3	*Recall*	0.9417
4	*F1-score*	0.9416

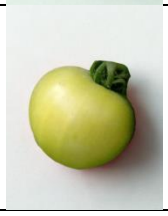
Tabel 5. Hasil Metrik Evaluasi Model CNN ResNet50

No	Metrik	Nilai
1	*Accuracy*	0.95
2	*Precision*	0.9505

3	*Recall*	0.95
4	*F1-score*	0.95

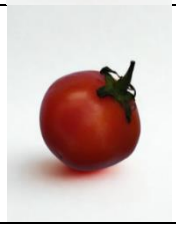
Setelah mengetahui performa dari kedua arsitektur tersebut, dilakukan pengujian menggunakan data *testing* yang telah disediakan untuk melihat implementasi kedua model tersebut dalam mengklasifikasi kematangan buah tomat. Hasil pengujian dengan data *testing* menunjukkan bahwa kedua model tersebut dapat mengklasifikasi gambar buah tomat dengan akurat. Baik arsitektur Inception-V3 maupun ResNet50 tidak mengalami kesalahan identifikasi dan sesuai dengan kelas target yang dipakai. Berikut tabel 6 dan 7 yang merupakan hasil pengujian data *testing* dari kelas belum matang dan matang.

Tabel 6. Hasil Klasifikasi Data *Testing* Kelas Belum Matang

No	Kelas Target	Gambar	Hasil Klasifikasi	
			Inception-V3	ResNet50
1	Belum Matang		Belum Matang	Belum Matang
2	Belum Matang		Belum Matang	Belum Matang
3	Belum Matang		Belum Matang	Belum Matang
4	Belum Matang		Belum Matang	Belum Matang
5	Belum Matang		Belum Matang	Belum Matang

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Tabel 7. Hasil Klasifikasi Data *Testing* Kelas Matang

No	Kelas Target	Gambar	Hasil Klasifikasi	
			Inception-V3	ResNet50
1	Matang		Matang	Matang
2	Matang		Matang	Matang
3	Matang		Matang	Matang
4	Matang		Matang	Matang
5	Matang		Matang	Matang

4. Kesimpulan

Pada penelitian ini dapat diambil kesimpulan bahwa model CNN Inception-V3 dan ResNet50 menghasilkan performa klasifikasi yang mengesankan. Berdasarkan metrik evaluasi, arsitektur Inception-V3 menghasilkan performa *accuracy* sebesar 94,17%, *precision* 94,28%, *recall* 94,17%, dan *f1-score* 94,16%. Sedangkan arsitektur ResNet50 menghasilkan performa *accuracy* sebesar 95%, *precision* 95,05%, *recall* 95%, dan *f1-score* 95%. Walaupun di atas kertas performa arsitektur ResNet50 lebih unggul daripada arsitektur Inception-V3, hasil pengujian dengan data *testing* menunjukkan kedua model tersebut dapat mengklasifikasi gambar buah tomat dengan akurat.

Referensi

[1] Badan Pusat Statistik Indonesia, “Produksi Tanaman dan Sayuran Menurut Provinsi dan Jenis Tanaman”, 8 Maret 2024, [Online], Available : https://www.bps.go.id/id/statistics-table/3/ZUhFd1JtZzJWVVpqWTJsV05XTllhVmhRSzFoNFFUMDkjMw==/produksi-tanaman-sayuran-menurut-provinsi-dan-jenis-tanaman--2023.html?year=2023 [21 Agustus 2025].

[2] A. S. Hadi, “KHASIAT BUAH TOMAT (*Solanum lycopersicum*) BERPOTENSI SEBAGAI OBAT BERBAGAI JENIS PENYAKIT”, *Empiris : Journal of Progressive Science and Mathematics*, vol. 01, no. 01, p. 7-15, 2023.
[3] H. Mubarok, S. Murni, and M. M. Santoni, “Penerapan Algoritma K-Nearest Neighbor untuk Klasifikasi Tingkat Kematangan Buah Tomat Berdasarkan Fitur Warna “, Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA), 2021.
[4] W. Vidiadivani and I. K. G. Suhartana, “Klasifikasi Jenis Sampah Menggunakan Metode *Transfer Learning* Pada *Convolutional Neural Network* (CNN)”, JELIKU (Jurnal Elektronik Ilmu Komputer Udayana), vol. 12, no. 3, 2024.
[5] U. Ungkawa and G. A. Hakim, “Klasifikasi Warna Pada Kematangan Buah Kopi Kuning Menggunakan Metode CNN Inception V3”, ELKOMIKA : Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, dan Teknik Elektronika, vol. 11, no. 3, 2023.
[6] P. D. Wardani and C. Dewi, “Klasifikasi Tipe Kematangan Buah Pisang Menggunakan Metode Ensemble Convolutional Neural Network (CNN)”, Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, vol. 1, no. 1, 2024.
[7] Mendeley Data, “Tomato Maturity Detection and Quality Grading Dataset”, 4 September 2023, [Online], Available : https://data.mendeley.com/datasets/s42kpg8h37/1 [27 Agustus 2025].
[8] Dataset Ninja, “LaboroTomato Dataset”, 14 Agustus 2020, [Online], Available : https://datasetninja.com/laboro-tomato#download [27 Agustus 2025].
[9] V. R. Pratiwi and J. Pardede, “Image Captioning Menggunakan Metode Inception-V3 dan Transformer” in Prosiding Diseminasi FTI Genap 2021/2022, Bandung.
[10] L. Trihardianingsih and H. Permatasari, “Model Transfer Learning untuk Klasifikasi Kesegaran Buah Jeruk Menggunakan Citra Digital” in Prosiding Seminar Nasional Teknologi dan Sains Tahun 2025, Kediri, 2025, vol. 4.
[11] I. Mustikasari, I. I. Tritoasmoro, and R. Purnamasari, “Identifikasi Kanker Kulit Melanoma Berbasis Inception V3 Menggunakan Pra-Pemrosesan dan Augmentasi Data pada Dataset Citra Kulit” in e-Proceeding of Engineering, Bandung, vol. 10, no. 5.
[12] I. P. Putra, Rusbandi, and D. Alamsyah, “Klasifikasi Penyakit Daun Tomat Menggunakan Metode Convolutional Neural Network”, Jurnal Algoritme, vol. 2, no. 2.
[13] Digital Ocean, “Deep Learning Architecture Explained : ResNet, InceptionV3, SqueezeNet”, 20 Maret 2025, [Online], Available : https://www.digitalocean.com/community/tutorials/popular-deep-learning-architectures-resnet-inceptionv3-squeezenet [27 Agustus 2025].
[14] Medium, “Exploring ResNet50 : An In-Depth Look at the Model Architecture and Code Implementation”, 23 Januari 2023, [Online], Available : https://medium.com/@nitishkundu1993/exploring-resnet50-an-in-depth-look-at-the-model-architecture-and-code-implementation-d8d8fa67e46f [27 Agustus 2025].
[15] Medium, “Memahami Confusion Matrix : Accuracy, Precision, Recall, dan F1-Score untuk Evaluasi Model Klasifikasi”, 11 Juni 2023, [Online], Available : https://esairina.medium.com/memahami-confusion-matrix-accuracy-precision-recall-specificity-dan-f1-score-610d4f0db7cf [27 Agustus 2025].

Jurnal Elektronik Ilmu Komputer Udayana
Volume 14, No 3. February 2026

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Penentuan Entitas Tokoh Pada Satua Bali Menggunakan Algoritma Conditional Random Fields

I Made Widi Arsa Ari Saputra[a1], I Ketut Gede Suhartana[a2], I Gusti Ngurah Anom Cahyadi Putra[b3], I Wayan Supriana[b3]

[a]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana
Jalan Raya Kampus Unud, Jimbaran,Bali, 80361, Indonesia
[1]widiarsamade@gmail.com
[2] wayan.supriana@unud.ac.id
[2]anom.cp@unud.ac.id
[2] ikg.suhartana@unud.ac.id

Abstrak

Pengenalan Entitas Bernama (NER) adalah proses untuk mengidentifikasi dan mengklasifikasi Entitas Bernama (NEs) pada suatu teks. Penelitian sebelumnya melakukan pengenalan entitas bernama pada bahasa bali menggunakan metode rule-based yang berfokus pada kelas person memberikan akurasi yang kurang memadai, karena data yang digunakan berfokus kepada fitur semantik dari kata pada kalimat(Kurniadi & ER, 2021). Tujuan dari penelitian ini adalah menerapkan fitur gramatikal dari kata berupa label part-of-speech, dan juga fitur lainnya untuk meningkatkan akurasi pada proses pengenalan entitas bernama yang berfokus pada entitas person. Metode Conditional Random Fields merupakan metode diskriminatif yang hanya berfokus kepada persebaran probabilitas bersyarat pada fitur ke label. Penelitian ini menggunakan 4 fitur, diantaranya: fitur semantik, fitur gramatikal, fitur transisi, serta fitur representasi embedding. Penelitian ini menggunakan data dari penelitian (Bimantara et al ea, 2024) yang telah dilabeli dengan label part-of-speech dan entitas bernama dengan skema labeling beginning-inside-outside (BIO). Penelitian ini menghasilkan accuracy 0.96 dengan precision sebesar 0.96, recall sebesar 0.96, dan total f1-score sebesar 0.96.

Keywords: *conditional random fields, pengenalan entitas bernama, bahasa bali, regularisasi, satua bali*

1. Pendahuluan

Bahasa Bali merupakan salah satu bahasa daerah di negara Indonesia, keberadaan bahasa Bali perlu dilestarikan, baik dalam bidang pendidikan maupun melalui teknologi digital [1]. Satua bali merupakan salah satu kebudayaan kesusastraan bali purwa pada pulau Bali. Satua bali dikenal sebagai sastra lisan yang menyebar dari mulut ke mulut tanpa diketahui pencipta aslinya [2]. Satua bali berperan dalam pembentukan karakter anak-anak di Bali, karena Satua bali sangat erat isinya dengan pendidikan moral karakter. Dalam analisis suatu cerita satua bali terdapat berbagai point analisis salah satunya adalah analisis penokohan.

Untuk melakukan analisis penokohan ini, tentunya perlu untuk mencari semua karakter yang terlibat pada suatu satua bali. Salah satu cara untuk mengekstraksi tokoh dari suatu satua bali adalah dengan melakukan ekstraksi informasi [3]. Untuk melakukan ekstraksi informasi diperlukan beberapa komponen, seperti parsing sintaksis, ekstraksi entitas, ataupun ekstraksi relasi. Ekstraksi informasi dokumen dalam bahasa Bali akan digunakan untuk mengekstraksi dokumen Satua Bali. Selain itu, NER yang dibuat dapat digunakan sebagai fitur preprocessing teks Bali yang memudahkan digitalisasi bahasa Bali yang dapat diakses dan tidak akan hilang.

Pengenalan entitas bernama (NER) adalah salah satu komponen penting dalam pengolahan bahasa alami (NLP) [4]. Istilah "Named Entity" telah banyak digunakan dalam bidang *information extraction* (IE), *question answering* (QA), dan berbagai bidang NLP lainnya [5]. NER merupakan langkah awal dalam ekstraksi informasi yang bertujuan untuk mengidentifikasi entitas yang disebutkan dalam teks dan mengkategorikannya ke dalam kategori yang telah ditentukan sebelumnya.

Conditional random fields (CRF) merupakan salah satu model probabilistik untuk melakukan segmentasi dan tagging suatu urutan data [6]. CRF hadir untuk mengatasi permasalahan ketergantungan asumsi yang besar pada model hidden markov (HMM) [7][8]. CRF dapat menentukan sendiri seberapa banyak fitur yang akan digunakan dalam suatu model CRF, berbeda dengan model HMM yang bersifat lokal dimana setiap kata (fitur) hanya bergantung pada label saat ini dan setiap label sebelumnya .CRF juga mengatasi permasalahan adanya label bias pada model maximum entropy markov (MEMM) karena distribusi kondisional label dari urutan data pada model CRF dilakukan secara keseluruhan dibandingkan model MEMM yang memformulasikan distribusi kondisional label hanya untuk setiap elemen data [9]. CRF banyak digunakan pada permasalahan natural language processing, computer vision, dan bioinformatic. salah satu permasalahan dari natural language processing adalah pengenalan entitas bernama (NER).

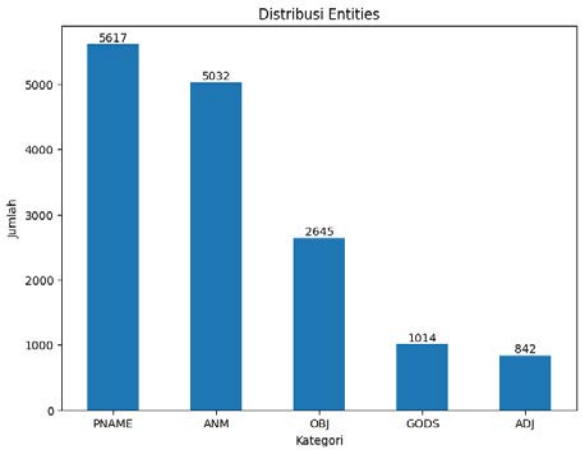
Pada penelitian yang dilakukan oleh Ekbal pada tahun 2007 [10] mengajukan pendekatan pada sistem pengenalan entitas bernama menggunakan algoritma hidden markov model untuk mengenali entitas nama pada teks berbahasa Bengali dan Hindi dan mendapatkan hasil f-score 84.50% dan 78.35%. kemudian beliau mengajukan kembali pendekatan tentang pengenalan entitas bernama menggunakan algoritma conditional random fields [11] dan mendapatkan hasil f-score 81.15%, dan 78.29% Sehingga, dari kedua penelitian tersebut beliau menyimpulkan hidden markov model lebih efektif digunakan dibanding conditional random fields. Kesimpulan ini dipertegas pada penelitian dari Chopra pada tahun 2016 [12], beliau melakukan pengenalan entitas bernama menggunakan algoritma hidden markov model pada bahasa hindi dan mendapatkan hasil f-score 97.14%. sehingga, hidden markov model menunjukan efisiensi lebih pada proses pengenalan entitas nama dibanding algoritma conditional random fields.

Namun, pada penelitian yang dilakukan oleh Jaariyah pada tahun 2017 [9], penelitian ini menggunakan algoritma conditional random fields untuk melakukan pendekatan entitas nama untuk mengenali entitas bernama pada teks berbahasa Indonesia dan mendapatkan hasil f-score 90.06%. Sedangkan pada penelitian yang dilakukan oleh Yusliani pada tahun 2020 [8], beliau mengajukan algoritma hidden markov model dalam pengenalan entitas bernama pada teks berbahasa indonesia dan mendapat f-score 86.14%. sehingga pada bahasa indonesia, algoritma conditional random fields justru lebih efektif untuk melakukan named entity recognition dibandingkan dengan algoritma hidden markov model.

2. Metode Penelitian

Penelitian ini menggunakan metode Conditional Random Fields untuk melakukan tugas pengenalan entitas bernama yang berfokus pada entitas person. Conditional Random Fields terdiri dari 3 proses: Penghitungan nilai potensial, estimasi parameter, dan inferensi.

2.1 Pengumpulan Data

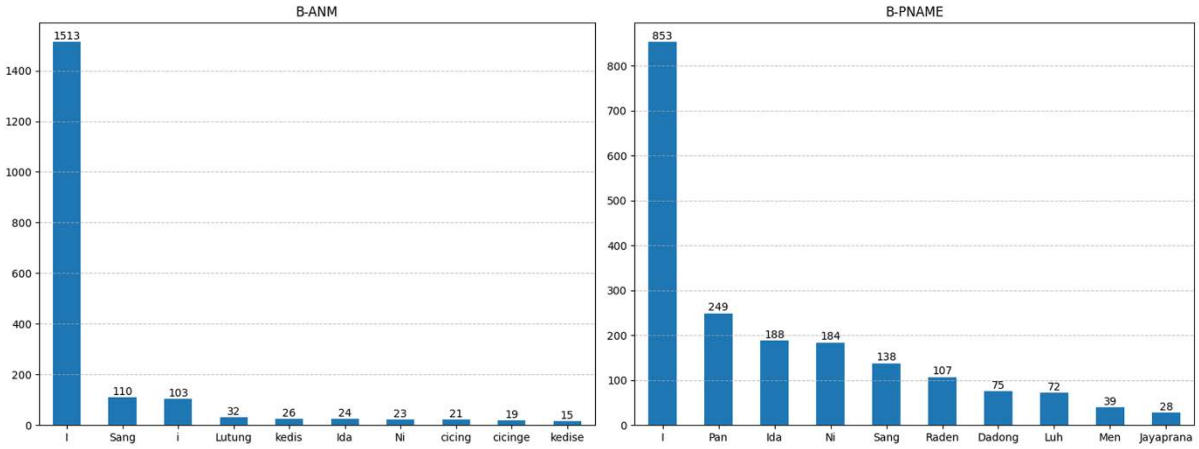


Gambar 1. Distribusi Label Dataset

Pada proses pengumpulan data, data yang akan digunakan berupa data teks berbahasa Bali seperti satua bali. Data satua bali akan diperoleh dari penelitian yang pernah dilakukan oleh (Bimantara et al ea, 2024). Terdapat 21 kelas kata yang digunakan berdasarkan penelitian sebelumnya [13] dan 11 label entitas tokoh dengan pola penamaan BIO [14]. Data ini berjumlah 89.917 fitur kata lengkap

p-ISSN: 2301-5373
e-ISSN: 2654-5101

dengan kelas kata dan label entitas bernama dengan jumlah label tokoh sebanyak 15.150 kata. Dari 15.150 kata ini, terlihat label PNAME dan ANM memberikan dominasi besar pada satua bali sebagai entitas tokoh, dapat dilihat pada **gambar 1**. Label PNAME pada dataset memiliki karakteristik penamaan daerah bali, seperti awalan I, Pan, Ida, Ni dan Sang. seperti pada **gambar 2**, atau Sang sehingga hal ini akan menjadi karakteristik dataset saat ini, dan model yang dilatih kemungkinan akan bergantung dengan keadaan ini. Label ANM juga memberikan distribusi yang besar pada dataset, mengindikasikan banyaknya tokoh binatang pada dataset saat ini, sehingga akan menjadi tantangan pada pelatihan model untuk membedakan binatang sebagai tokoh dalam cerita atau sebagai tokoh figuran.



Gambar 2. Distribusi Awalan Kata dari Tokoh pada Dataset

2.2 Perancangan Sistem

Sistem dari penelitian ini akan dimulai dengan melakukan pra-pemrosesan data untuk membersihkan data dan mengganti label sesuai keperluan penelitian. Kemudian dilakukan rekayasa fitur agar bentuk data dapat diproses kedalam model. Setelah bentuk data sudah sesuai, barulah dilakukan training menggunakan algoritma CRF dengan metode evaluasi grid-search cross validation.

1. Pra-pemrosesan data
   Data yang telah didapatkan kemudian dilakukan pra-pemrosesan berupa mengganti label B-OBJ dan I-OBJ menjadi O, dan entitas lain diruubah menjadi B-PER. Kemudian melakukan rekayasa fitur dengan merubah bentuk setiap observasi data kedalam bentuk vektor fitur, rekayasa ini merupakan hiperparameter sehingga bentuk dan jumlahnya dapat menyesuaikan dengan domain, tugas, dan keadaan dataset.

2. Rekayasa Fitur
   Fitur yang digunakan untuk melakukan training pada model CRF disini sejumlah 48 fitur. Tabel 1 memetakan vektor fitur yang digunakan pada penelitian ini sejumlah 4 tipe vektor, yaitu: vektor emisi sebagai identitas emisi dari setiap observasi kata, vektor transisi sebagai identitas transisi dari kata dan jendelanya, vektor pos sebagai identitas kelas kata dari observasi kata, dan vektor embedding sebagai identitas embedding dari kata. Ke 48 fitur ini akan diterapkan untuk setiap observasi kata sebagai representasi vektor dari data sebelum masuk ke model. Fitur transisi mendapatkan porsi jumlah yang paling banyak, yaitu dengan 20 fitur, untuk menangkap dengan jelas karakteristik urutan kata pada kalimat dalam satua bali. Hal ini penting agar model dapat mempelajari pola kemunculan dari suatu entitas tokoh diberikan suatu kalimat dalam satua bali. Karena ambiguitas tokoh dalam satua bali terutama pada tokoh binatang yang terkadang mejadi tokoh figuran, sehingga pemodelan fitur transisi dirasa penting untuk dilakukan agar model dapat mengenali tokoh utama dan tokoh figuran.

Tabel 1. Fitur Vektor

tipe	fitur	jumlah	deskripsi
emisi	bias	1	bobot dari bias yang digunakan adalah 1.0 untuk setiap fitur
	word	1	representasi karakter dari kata saat ini
	is_first	1	boolean untuk menandakan emisi awal dari suatu kalimat
	is_last	1	boolean untuk menandakan emisi akhir dari suatu kalimat
	islower	1	boolean untuk menandakan kata saat ini merupakan berhuruf kecil
	prefix	2	karakter awal dari kata saat ini dengan jendela 2
	suffix	2	karakter terakhir dari kata saat ini dengan jendela 2
	isupper	1	boolean untuk melihat kapitalisasi dari kata saat ini
	is_title	1	boolean untuk melihat kapitalisasi dari huruf pertama kata saat ini
transisi	prev_word	2	representasi kata sebelumnya dari kata saat ini dengan jendela 2
	next_word	2	representasi kata setelahnya dari kata saat ini dengan jendela 2
	prev_word_prefix	4	prefix dari kata sebelumnya yang masing masing dengan jendela 2
	prev_word_suffix	4	suffix dari kata sebelumnya yang masing masing dengan jendela 2
	next_word_prefix	4	prefix dari kata setelahnya yang masing masing dengan jendela 2
	next_word_suffix	4	suffix dari kata setelahnya yang masing masing dengan jendela 2
pos	pos	1	label kelas kata dari kata saat ini
	prev_pos	2	label kelas kata dari kata sebelumnya dengan jendela 2
	next_pos	2	label kelas kata dari kata setelahnya dengan jendela 2
embedding	emb_word	1	embedding vektor dari kata saat ini
	emb_prev_word	2	embedding vektor dari kata sebelumnya dengan jendela 2
	emb_next_word	2	embedding vektor dari kata setelahnya dengan jendela 2
	emb_prev_word_prefix	2	embedding vektor dari prefix kata sebelumya dengan jendela 2
	emb_prev_word_suffix	2	embedding vektor dari suffix kata sebelumya dengan jendela 2
	emb_next_word_prefix	2	embedding vektor dari prefix kata setelahnya dengan jendela 2
	emb_next_word_suffix	2	embedding vektor dari suffix kata setelahnya dengan jendela 2

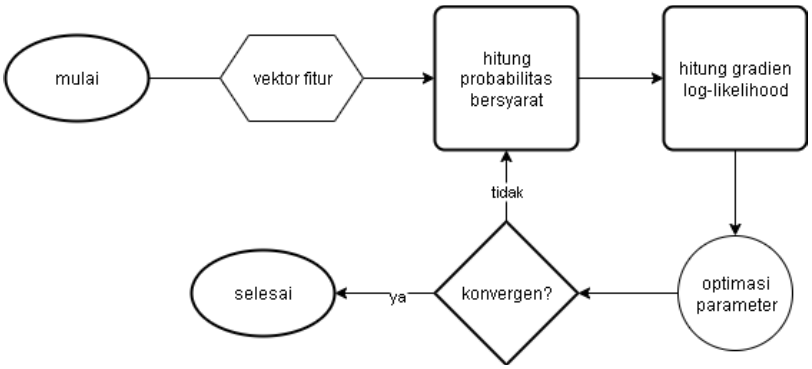
3. Pelatihan

Seperti yang terlihat pada **gambar 3**, proses pelatihan CRF diawali dengan parameter vektor fitur hasil dari rekayasa fitur. Kemudian dilakukan penghitungan probabilitas bersyarat untuk seluruh data pada training dataset.

$$p_\theta(\mathbf{y} \mid \mathbf{x}) = \frac{\prod_{i=1}^{n+1} M_i(\mathbf{y_{i-1}}, \mathbf{y_i} \mid \mathbf{x})}{(\prod_{i=1}^{n+1} M_i(\mathbf{x}))_{\text{start, stop}}} \quad (1)$$

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Numerator dari fungsi merupakan nilai potensial tiap-tiap observasi fitur (fungsi fitur setiap kata) dalam ruang lingkup kalimat saat ini. Sedangkan denominator merupakan fungsi partisi yang bertugas untuk menormalisasi probabilitas agar berjumlah 1. Z(x) atau fungsi partisi, adalah nilai dot produk untuk seluruh urutan label yang mungkin terhadap data observasi (x1, x2, x3,...,xn) yang diberikan. apabila jumlah data observasi merupakan n, sedangkan jumlah label merupakan Y, maka jumlah kombinasi urutan label yang mungkin terjadi adalah nY urutan label.



Gambar 3. Alur Pelatihan

Setelah itu, dihitung *log-likelihood* dari probabilitas bersyarat yang akan digunakan untuk menghitung nilai gradien log-likelihood yang diperlukan untuk mengetahui arah perubahan nilai dari parameter λ.

$$\mathcal{L}(\theta) = -\left(\sum_{i=1}^{n+1} \log M_i\,(y_{i-1}, y_i \mid \boldsymbol{x}) - \log Z\,(\boldsymbol{x})\right) + \alpha|\theta|_1 + \frac{\beta}{2}|\theta|_2^2 \qquad (2)$$

Setelah gradien didapatkan barulah optimasi parameter menggunakan algoritma L-BFGS.

Algoritma Limited-memory BFGS (L-BFGS) merupakan metode optimisasi numerik berbasis gradien yang digunakan untuk memaksimalkan log-likelihood dalam pelatihan model Conditional Random Fields (CRF), L-BFGS menggunakan pendekatan terbatas memori dengan menyimpan sejumlah vektor gradien dan parameter dari iterasi sebelumnya.
Alur proses:

a. Inisialisasi parameter bobot λ_0.
b. Hitung gradien awal, dari log-likelihood sebelumnya
c. Hitung perubahan nilai lambda dan gradien
d. Perbarui Hessian matriks
e. Tentukan arah pencarian d_k menggunakan estimasi inverse Hessian
f. Lakukan line search untuk menemukan panjang langkah optimal α_k yang meminimalkan fungsi objektif di sepanjang arah d_k .
g. Perbarui parameter: $\lambda_{k+1} = \lambda_k + \alpha_k\, d_k$
h. Perbarui memori terbatas: Simpan perubahan gradien Δg_k dan parameter $\Delta\lambda_k$.
i. Ulangi langkah 2–6 hingga konvergen

4. Metode Pengujian
Pada penelitian ini uji coba parameter akan dilakukan dengan metode grid-search cross validation. Tujuannya adalah mencari kombinasi parameter terbaik untuk model, serta menghindari bias pada model dengan menggunakan cross-validation. Alur dari uji coba parameter sebagai berikut:

a. Menentukan nilai hiperparameter yang akan diuji
Hiperparameter yang akan diuji disini adalah regularisasi yang digunaka pada algoritma Conditional Random Fields yaitu L1 dan L2. Nilai L1 akan dieksplorasi pada

rentang 10^{-5} hingga 10^{-1} dengan skala logaritmik basis 10 dengan kenaikan 1 pada eksponen, ini juga berlaku pada L2.

b. Membentuk Grid Kombinasi
Pada tahap ini dilakukan penyusunan kombinasi 2 parameter yang diuji yaitu L1 dan L2. Dengan 5 nilai untuk setiap parameter, berarti akan ada 25 kombinasi yang akan diujikan.

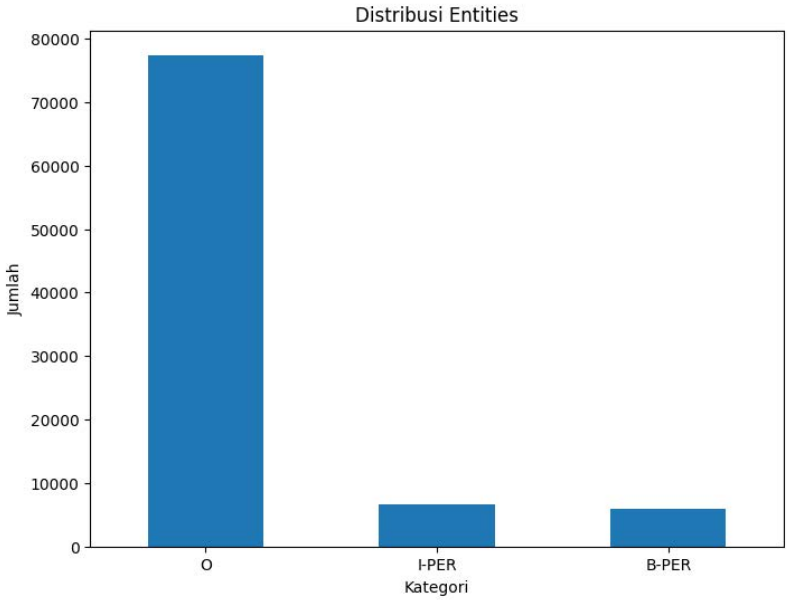
c. Menerapkan Cross-Validation
Dataset akan dibagi dengan 5 lipatan, sehingga pembagian pelatihan dan pengujian adalah 4:1. Untuk setiap kombinasi parameter akan dilatih dengan k-1 lipatan, sedangkan pengujian akan dilakukan pada lipatan yang tersisa. Proses ini akan diulangi hingga semua lipatan digunakan sebagai data uji. Kemudian setelah semua lipatan telah digunakan sebagai data uji, akan dilakukan perhitungan rata-rata akurasi dari semua lipatan yang merepresentasikan akurasi optimal pada pengujian kombinasi saat ini.

d. Memilih kombinasi terbaik
Kombinasi terbaik dipilih berdasarkan akurasi tertinggi. Kemudian model akan disimpan untuk dilakukan deployment.

3. Hasil dan Pembahasan

3.1 Data Latih

Data ini saya saring dengan mengganti label B-OBJ dan I-OBJ menjadi O, dan entitas lain saya ubah labelnya menjadi B-PER dan I-PER. Penggantian ini dilakukan untuk berfokus hanya pada entitas tokoh, sehingga label lain seperti “ANIMAL” dan “GOD” dianggap juga sebagai tokoh, label “OBJECT” tidak dibutuhkan karena ini merepresentasikan kata ganti. Data latih yang telah melalui proses pra-pemrosesan data dengan 3 label yaitu B-PNAME, I-PNAME, dan O. Rekayasa fitur dilakukan untuk setiap kata dan disimpan dalam array 2D, yang dimensi pertama merupakan kalimat dari setiap dataset. Dataset ini terdiri dari 6.6315 kalimat yang disimpan dalam array dengan kelas kata dan label entitasnya. Pada **gambar 4** ditunjukkan distribusi dari label B-PER, I-PER dan juga O.



Gambar 4. Distribusi Label Data Latih

p-ISSN: 2301-5373
e-ISSN: 2654-5101

3.2 Hasil

Implementasi CRF menggunakan library sklearn-crfsuite yang sudah secara default mengimplementasikan algoritma L-BFGS. Pelatihan model CRF menggunakan regularisasi L1 beserta L2. L1 regularization mendorong model untuk menghasilkan parameter yang lebih jarang (sparse), artinya beberapa fitur akan memiliki bobot nol sehingga diabaikan dalam model. L2 regularization mencegah overfitting dengan mendorong parameter model untuk tidak menjadi terlalu besar. Ini mengurangi sensitivitas model terhadap data pelatihan yang bervariasi. L2 regularization lebih fokus pada distribusi parameter daripada memaksakan sparsity.

Table 2. Hasil Grid-Search Cross Validation

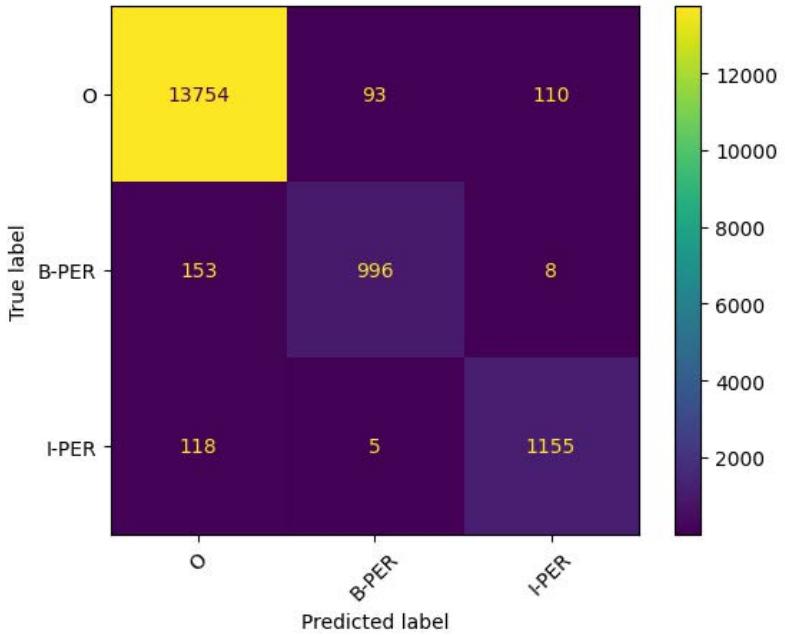
L1	L2	F1-score	accuracy	precision	recall
0.0001	0.0001	0.965360165	0.965606426	0.965258024	0.965606426
	0.001	0.966242487	0.966510067	0.96612566	0.966510067
	0.01	0.967347341	0.967619312	0.967255172	0.967619312
	0.1	0.967313309	0.96763744	0.967216799	0.96763744
0.001	0.0001	0.967051205	0.967236454	0.966973224	0.967236454
	0.001	0.966903442	0.967136791	0.966805971	0.967136791
	0.01	0.967234176	0.967497685	0.967135744	0.967497685
	0.1	0.967541741	0.967853787	0.967444056	0.967853787
0.01	0.0001	0.968308431	0.968503797	0.968219608	0.968503797
	0.001	0.968772897	0.968942969	0.968703175	0.968942969
	0.01	0.967893083	0.968106581	0.967816634	0.968106581
	0.1	0.967696585	0.968000707	0.967604502	0.968000707
0.1	0.0001	0.968594696	0.968783144	0.968510977	0.968783144
	0.001	0.968104334	0.968282621	0.96802272	0.968282621
	0.01	0.968594696	0.968783144	0.968510977	0.968783144
	0.1	0.968411335	0.968611665	0.968321444	0.968611665

Pada **tabel 2,** hasil evaluasi menunjukkan bahwa kombinasi L1 = 0.01 dan L2 = 0.001 menghasilkan kinerja terbaik, ditunjukkan oleh skor F1 tertinggi sebesar 0.968772897. Kombinasi ini juga menghasilkan nilai akurasi, presisi, dan recall yang konsisten tinggi, menandakan bahwa model tidak hanya mampu mengenali entitas secara tepat, tetapi juga konsisten dalam meminimalkan kesalahan deteksi.
Secara umum, perunurnan nilai L1 justru menurunkan kinerja model. Hal ini terlihat dari penurunan bertahap pada skor F1 ketika nilai L1 diturunkan hingga 0.0001, yang mengindikasikan bahwa regularisasi yang terlalu rendah kurang mampu memberikan penalti untuk model dalam mempelajari fitur yang tidak relevan.

Tabel 3. Matriks Evaluasi Model Terbaik L1=0.01 & L2=0.001

	precission	recall	F1-score
O	0.98	0.99	0.98
B-PER	0.91	0.86	0.88
I-PER	0.91	0.90	0.91

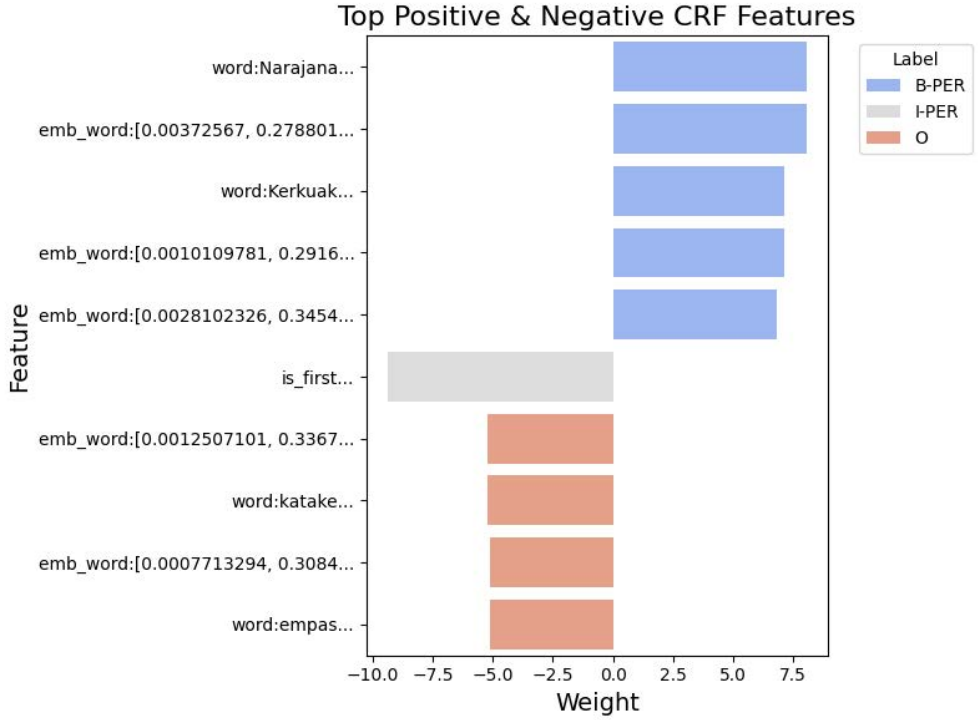
Pada **tabel 3** terlihat jelas model sangat percaya terhadap pengenalan entitas tidak bernama (O) dengan angka precission yang sangat tinggi. Sedangkan untuk label entitas bernama terbilang memiliki nilai yang cukup rendah, nilai 0.86 pada recall dari label B-PER menandakan model hanya mampu mengenali 86% dari keseluruhan entitas bernama yang ada pada data test. Begitu pula pada label I-PER yang mendapatkan persentase 90% untuk recall pada model.



Gambar 5. Confusion Matrix Model Terbaik

Pada **gambar 5** dapat dilihat, sebanyak 271 kali model gagal mengenali entitas bernama dan cenderung memberikan label O. Kecilnya angka kesalahan prediksi pada label B-PER antar I-PER menunjukkan model cukup mampu membedakan 2 entitas tokoh yang berbeda pada suatu kalimat. Walaupun seringkali model cenderung memberikan label I-PER pada observasi yang seharusnya bukan tokoh (O) sebanyak 110 kali, artinya model masih kurang mampu memprediksi akhir dari nama tokoh pada suatu kalimat.

Model CRF yang dilatih menghasilkan 24.452 bobot secara total pada model. Jumlah ini merupakan jumlah dari seluruh nilai yang mungkin pada fitur vektor dikombinasikan dengan setiap label yang ada. Pada **gambar 6** terlihat bahwa bobot tertinggi dan terendah pada model CRF yang dilatih sebagian besar berasal dari fitur word_embedding dan word itu sendiri, sedangkan fitur lain seperti transisi, kelas kata (POS), suffix, prefix, emisi, dan bias tidak muncul dalam daftar bobot ekstrem.



Gambar 6. Lima Bobot Tertinggi dan Terendah Model

3.3 Pembahasan

Dari hasil yang didapatkan menunjukkan beberapa hal penting:

1. Dominasi fitur leksikal dan embedding:
    a. Fitur word (kata itu sendiri) dan word_embedding memberikan sinyal yang paling kuat bagi model untuk memprediksi entitas, khususnya label B-PER dan O.
    b. Bobot positif tinggi pada kata-kata spesifik atau embedding tertentu menunjukkan fitur ini sangat menentukan prediksi label tertentu.
    c. Bobot negatif tinggi pada fitur embedding atau kata lain menandakan fitur itu menekan prediksi label tertentu, sehingga berperan sebagai penyeimbang.
2. Fitur transisi, kelas kata dan bias kurang berpengaruh pada bobot ekstrem:
    Walaupun CRF menggunakan transisi antar label dan bias, bobotnya tidak terlalu ekstrem, sehingga tidak muncul di 5 bobot teratas positif atau negatif. Fitur transisi sendiri bersifat penyesuaian global antar label, bukan determinan utama per token sehingga wajar tidak memberikan beban kuat. Baik transisi antar kata ataupun transisi antar kelas kata.
3. Fitur kelas kata, suffix, prefix, dan emisi:
    a. Fitur-fitur ini tampaknya memiliki bobot moderat, dekat dengan rata-rata, sehingga tidak terlalu memengaruhi prediksi secara ekstrem.
    b. Bisa berarti fitur-fitur ini berguna untuk penyesuaian halus atau menangkap pola umum, tetapi tidak sekuat kata spesifik atau embedding dalam membedakan label.

4. Kesimpulan

Berdasarkan hasil eksperimen dan evaluasi terhadap model Conditional Random Fields (CRF) dalam mendeteksi entity "person" pada teks berbahasa Bali (Satua Bali), diperoleh beberapa temuan penting. Akurasi terbaik yang berhasil dicapai oleh model adalah sebesar 96%, dengan skor F1 yang juga tinggi yaitu 96%. Hasil ini menunjukkan bahwa algoritma CRF mampu secara efektif mempelajari pola-pola linguistik yang relevan dalam teks untuk mengidentifikasi entitas nama orang.

Melalui proses tuning parameter regularisasi, kombinasi terbaik untuk parameter L1 dan L2 adalah pada nilai L1 = 0.01 dan L2 = 0.001. Kombinasi ini memberikan keseimbangan yang optimal antara sparsity dan generalisasi model, sehingga model mampu mempertahankan akurasi tinggi sekaligus menghindari overfitting. Temuan ini mengindikasikan bahwa pemilihan parameter regularisasi yang tepat memiliki pengaruh signifikan terhadap performa akhir model CRF.

Untuk penelitian selanjutnya, terdapat beberapa hal yang dapat dipertimbangkan. Pertama, penggunaan teknik preprocessing lanjutan seperti normalisasi kata dan pemanfaatan fitur linguistik tambahan (contohnya: fitur morfologi atau struktur kalimat) dapat ditelusuri lebih lanjut untuk meningkatkan akurasi model. Kedua, karena CRF bersifat supervised, pengumpulan dataset beranotasi yang lebih besar dan lebih bervariasi dapat membantu memperkuat kemampuan generalisasi model pada berbagai jenis teks berbahasa Bali.

Selain itu, akan sangat menarik untuk membandingkan performa CRF dengan pendekatan berbasis neural sequence labeling, seperti BiLSTM-CRF atau transformer-based models (misalnya BERT fine-tuned untuk NER), guna mengevaluasi keunggulan relatif dari pendekatan tradisional dan modern dalam pemrosesan bahasa daerah seperti Bali. Terakhir, implementasi model dalam aplikasi nyata seperti digitalisasi cerita rakyat Bali juga sangat direkomendasikan sebagai bentuk kontribusi terhadap pelestarian bahasa dan budaya lokal.

Daftar Pustaka

[1] I. K. Mustika, 'Pergeseran Bahasa Bali sebagai Bahasa Ibu di Era Global (Kajian Pemertahanan Bahasa)', Purwadita, vol. 2, no. 1, pp. 94–102, 2018, [Online]. Available: http://jurnal.stahnmpukuturan.ac.id/index.php/Purwadita/article/view/26.
[2] I. W. P. A. Wiguna, I. K. R. Arthana, and I. M. Putrama, 'Pengembangan Game Edukasi Satua Bali "Pan Cubling" Berbasis Android', J. Nas. Pendidik. Tek. Inform., vol. 6, no. 2, p. 192, 2017, doi: 10.23887/janapati.v6i2.11440.
[3] K. Kurniadi and N. A. S. ER, 'Person Named Entity Recognition in Balinese', JELIKU (Jurnal Elektron. Ilmu Komput. Udayana), vol. 10, no. 1, p. 99, 2021, doi: 10.24843/jlk.2021.v10.i01.p13.
[4] W. Gunawan, D. Suhartono, F. Purnomo, and A. Ongko, 'Named-Entity Recognition for Indonesian Language using Bidirectional LSTM-CNNs', Procedia Comput. Sci., vol. 135, pp. 425–432, 2018, doi: 10.1016/j.procs.2018.08.193.
[5] A. S. Wibawa and A. Purwarianti, 'Indonesian Named-entity Recognition for 15 Classes Using Ensemble Supervised Learning', Procedia Comput. Sci., vol. 81, no. May, pp. 221–228, 2016, doi: 10.1016/j.procs.2016.04.053.
[6] S. Briandoko, A. Ratna Dewi, M. Akbar Setiawan, and S. Widya Utama, 'Perbandingan Algoritma Conditional Random Field dan Hidden Markov Model pada Pos Tagging Bahasa Indonesia', J. Tek., vol. 2, no. 2, pp. 2598–294, 2018.
[7] I. N. A. R. Arintyo Archamadi, Rita Magdalena, 'Analisis dan Simulasi Identifikasi Judul Lagu dari Senandung Manusia Menggunakan Ekstraksi Ciri DCT (Discreate Cosine Transform)', vol. 3, no. 3, pp. 4575–4584, 2016.
[8] N. Yusliani, M. R. P. Sufa, A. Firdaus, Abdiansah, and S. Yoppy, 'Named-Entity Recognition Pada Teks Berbahasa Indonesia Menggunakan Metode Hidden Markov Model Dan Part-of-Speech Tagging', Linguist. Komputasional, vol. 4, no. 1, pp. 13–20, 2020.
[9] N. Jaariyah and E. Rainarli, 'Conditional Random Fields Untuk Pengenalan Entitas Bernama Pada Teks Bahasa Indonesia', Komputa J. Ilm. Komput. dan Inform., vol. 6, no. 1, pp. 29–34, 2017, doi: 10.34010/komputa.v6i1.2474.
[10] A. Ekbal and S. Bandyopadhyay, 'A Hidden Markov Model based named entity recognition system: Bengali and Hindi as case studies', Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 4815 LNCS, pp. 545–552, 2007, doi: 10.1007/978-3-540-77046-6_67.
[11] A. Ekbal and S. Bandyopadhyay, 'A Conditional Random Field Approach for Named Entity Recognition in Bengali and Hindi', Linguist. Issues Lang. Technol., vol. 2, no. 1, Nov. 2009, doi: 10.33011/lilt.v2i.1203.
[12] D. Chopra, N. Joshi, and I. Mathur, 'Named entity recognition in Hindi using hidden markov model', Proc. - 2016 2nd Int. Conf. Comput. Intell. Commun. Technol. CICT 2016, no. Table I, pp. 581–586, Aug. 2016, doi: 10.1109/CICT.2016.121.
[13] I. M. S. Bimantara, D. Purwitasari, N. A. S. ER, and P. G. S. Natha, 'Balinese story texts dataset for narrative text analyses', Data Br., vol. 56, 2024, doi: 10.1016/j.dib.2024.110781.

[4] Sang, E. F., & Buchholz, S. (2000). Introduction to the CoNLL-2000 Shared Task: Chunking. In Proceedings of the 2nd Workshop on Learning Language in Logic and the 4th Conference on Computational Natural Language Learning (CoNLL), 127-132.
[5] Explosion AI. (2022). spaCy: Industrial-strength Natural Language Processing in Python. [Online]. Available: https://spacy.io/
[6] Haddi, E., Liu, X., & Shi, Y. (2019). Deep Learning for Natural Language Processing. Springer.

Jurnal Elektronik Ilmu Komputer Udayana
Volume 14, No 3. February 2026

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Implementasi *Intrusion Prevention System* (IPS) Menggunakan *Signature – Based Detection* Berbasis Suricata

I Wayan Dimas Wirahadi[a1], I Gede Santi Astawa [a2], Made Agung Raharja [a3] I Made Widiartha [a4]

[a]Program Studi Informatika, Universitas Udayana
Bali, Indonesia
[1]iwayandimas20@gmail.com
[2] santi.astawa@unud.ac.id
[3]made.agung@unud.ac.id
[4]madewidiartha@unud.ac.id

Abstract

Network security is a crucial aspect of maintaining the integrity, confidentiality, and availability of data. One solution to strengthen network security is the implementation of an Intrusion Prevention System (IPS). This research implements a hybrid IPS that integrates Suricata as a signature-based detection engine with the Aho-Corasick algorithm for secondary log analysis and pattern matching. The system, implemented on a Windows operating system, is designed to automatically detect and block various types of attacks, such as Port Scanning, DDoS (SYN Flood), and Protocol-Specific Attacks. Testing was conducted through end-to-end attack scenarios to measure the system's effectiveness and response time. The results demonstrate that the proposed architecture is highly effective, achieving an average detection rate (recall) of 86.67% and a precision rate of 81.25%, which yields an F1-Score of 83.87%. Response time analysis revealed that Suricata detects threats within an average of 1-3 seconds, while the Aho-Corasick layer analyzes the resulting logs in a matter of milliseconds. This implementation proves that the combination of Suricata and Aho-Corasick provides a security solution that is not only effective at detecting known threats but also efficient at analyzing the results.

Keyword : *Network Security, Intrusion Prevention System, Suricata, Signature-Based Detection, Attack Detection*

1. Pendahuluan

Keamanan jaringan adalah aspek krusial untuk menjaga integritas, kerahasiaan, dan ketersediaan data. Salah satu solusi untuk memperkuat keamanan adalah penerapan Intrusion Prevention System (IPS). Seiring dengan meningkatnya digitalisasi, keamanan jaringan telah menjadi pilar utama dalam melindungi aset informasi dari berbagai ancaman siber. Kasus pelanggaran data, seperti serangan SQL Injection dan Distributed Denial-of-Service (DDoS), telah terbukti menimbulkan kerugian finansial yang signifikan dan merusak reputasi organisasi [1]. Sistem keamanan konvensional yang bersifat pasif, seperti firewall, seringkali tidak cukup untuk menangkal serangan modern yang semakin canggih dan dinamis. Oleh karena itu, diperlukan pendekatan keamanan proaktif yang tidak hanya mendeteksi, tetapi juga mampu mencegah serangan secara real-time.

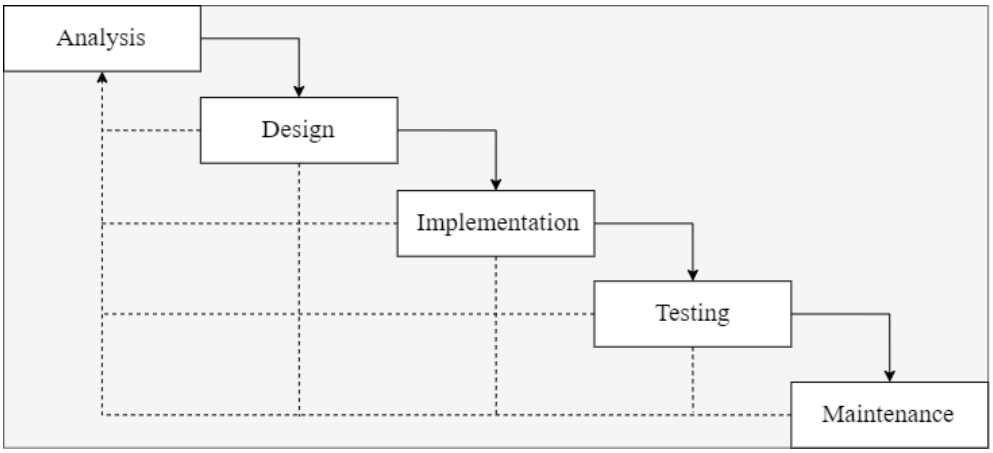
Intrusion Prevention System (IPS) muncul sebagai solusi untuk kebutuhan ini. Berbeda dengan Intrusion Detection System (IDS) yang hanya memberikan peringatan, IPS secara aktif memblokir lalu lintas berbahaya sebelum mencapai target [2]. Penelitian ini berfokus pada implementasi IPS menggunakan Suricata, sebuah engine deteksi ancaman open-source yang dikenal karena performa tingginya berkat arsitektur multi-threading [3]. Suricata mampu melakukan Deep Packet Inspection (DPI) dan menggunakan metode Deteksi Berbasis Signature untuk mengidentifikasi pola-pola serangan yang telah dikenal. Metode ini sangat efektif untuk ancaman yang polanya sudah terdokumentasi, menjadikannya garda terdepan yang andal dalam infrastruktur keamanan. Meskipun deteksi berbasis signature sangat efektif, implementasinya menghadapi tantangan skalabilitas dan performa. Seiring dengan berkembangnya ancaman baru, jumlah signature yang harus diperiksa oleh engine IPS dapat mencapai puluhan hingga ratusan ribu. Proses pencocokan pola secara linear terhadap database signature yang masif ini dapat menimbulkan latensi dan menjadi bottleneck pada

jaringan berkecepatan tinggi. Beban komputasi yang tinggi ini dapat menurunkan throughput jaringan dan bahkan membuka celah bagi penyerang untuk mengeksploitasi keterlambatan respons sistem. Oleh karena itu, diperlukan sebuah mekanisme untuk mengoptimalkan proses pencocokan signature tanpa mengorbankan kedalaman analisis.

Untuk mengatasi tantangan ini, penelitian ini mengusulkan sebuah arsitektur hibrida yang inovatif dengan mengintegrasikan Suricata dengan algoritma Aho-Corasick [4]. Algoritma ini merupakan metode pencocokan string multi-pola yang sangat efisien, mampu mencari ribuan pola secara simultan dalam satu kali proses dengan kompleksitas waktu linear. Dalam arsitektur ini, Suricata bertindak sebagai sensor utama di lapisan jaringan untuk deteksi awal dan pemblokiran, sementara Aho-Corasick berfungsi sebagai mesin analisis sekunder pada lapisan log untuk validasi, korelasi peringatan, dan threat hunting dengan kecepatan tinggi. Kontribusi utama dari penelitian ini adalah merancang dan memvalidasi sebuah model IPS yang tidak hanya efektif dalam deteksi, tetapi juga efisien dalam pemrosesan data peringatan, sehingga menciptakan siklus respons keamanan yang lebih cepat dan cerdas.

2. Metode Penelitian

Metode penelitian ini menggunakan pendekatan metodologi System Development Life Cycle (SDLC) untuk memastikan bahwa setiap tahapan implementasi dilakukan secara sistematis, terstruktur, dan dapat dievaluasi secara menyeluruh dapat dilihat pada Gambar 1. SDLC merupakan sebuah kerangka kerja terorganisir yang umum digunakan dalam pengembangan sistem informasi, yang mencakup seluruh siklus hidup sistem mulai dari tahap inisialisasi hingga pemeliharaan pasca implementasi. Pendekatan ini sangat cocok untuk penelitian yang bersifat teknis dan aplikatif, seperti pengembangan dan penerapan sistem keamanan jaringan berbasis Suricata.



Gambar 1. Alur Metode Penelitian

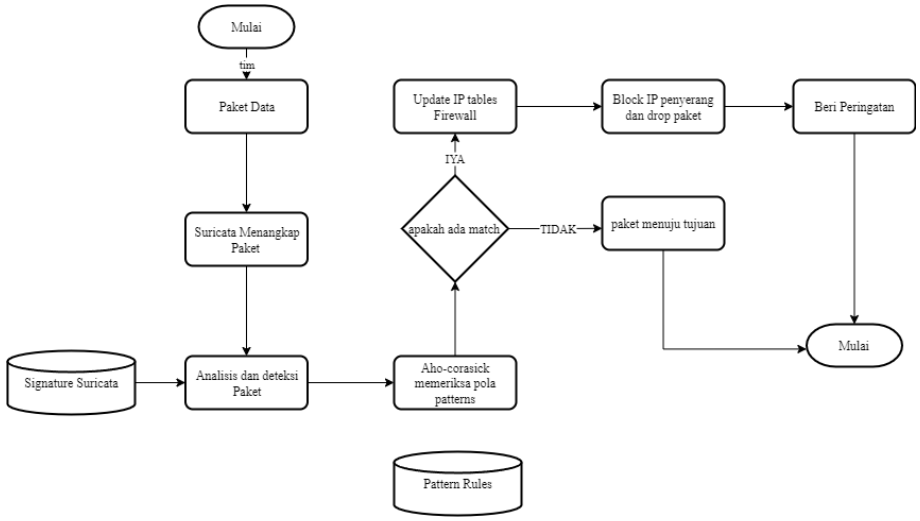
2.1. Landasan Teori

Sistem yang diusulkan dibangun di atas tiga teknologi inti. Pertama, Intrusion Prevention System (IPS), sebuah teknologi keamanan yang secara aktif memonitor dan memblokir lalu lintas berbahaya [5]. Kedua, Suricata, sebuah engine IDS/IPS open-source berkinerja tinggi yang mampu melakukan analisis paket mendalam secara paralel berkat arsitektur multi-threading [6]. Ketiga, Algoritma Aho-Corasick, sebuah metode pencocokan string multi-pola yang sangat efisien. Algoritma ini membangun finite state machine dalam bentuk struktur data Trie dengan failure links, memungkinkan pencarian ribuan pola dalam waktu linear tanpa backtracking [4]. Penelitian sebelumnya telah mengonfirmasi efektivitas Suricata dalam mendeteksi berbagai serangan [1, 7], namun integrasi dengan Aho-Corasick sebagai lapisan analisis log sekunder merupakan pendekatan baru yang ditawarkan dalam penelitian ini untuk meningkatkan efisiensi.

2.2. Arsitektur Sistem Hybrid

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Proses utama sistem divisualisasikan pada Gambar 2, yang menunjukkan alur kerja mulai dari masuknya paket data, deteksi awal oleh Suricata, analisis log oleh Aho-Corasick, hingga eksekusi blokir IP oleh firewall. Diagram ini menegaskan bahwa sistem bekerja berlapis-lapis dengan kombinasi deteksi real-time dan validasi log sekunder.



Gambar 2. Alur Kerja Sistem

Alur dimulai saat paket data masuk ke jaringan. Suricata menangkap dan menganalisis setiap paket berdasarkan ruleset signature. Jika terdeteksi potensi ancaman, log akan dihasilkan. Log ini kemudian dianalisis oleh modul Aho-Corasick untuk validasi pola. Jika pola terkonfirmasi sebagai serangan, sistem akan memblokir IP penyerang melalui firewall dan mengirimkan peringatan. Jika tidak ada kecocokan, paket diteruskan ke tujuan. Arsitektur ini terdiri dari empat lapisan utama:

1. **Lapisan Deteksi Jaringan (Suricata):** Bertugas sebagai sensor utama yang berjalan dalam mode IPS inline pada OS Windows menggunakan driver WinDivert. Lapisan ini melakukan inspeksi paket, deteksi ancaman awal berdasarkan ruleset signature, dan aksi pemblokiran instan.

2. **Lapisan Analisis Log (Python & Aho-Corasick):** Sebuah skrip Python secara real-time memonitor file log eve.json yang dihasilkan Suricata. Skrip ini menggunakan algoritma Aho-Corasick untuk melakukan pencocokan pola sekunder, validasi, dan korelasi alert dengan kecepatan tinggi.

3. **Lapisan Aksi & Manajemen (Firewall & Database):** Berdasarkan hasil analisis, skrip Python mengirimkan perintah pemblokiran IP ke Windows Firewall dan mencatat semua detail insiden ke dalam database MySQL untuk analisis historis.

4. **Lapisan Visualisasi (Flask Web App):** Sebuah dashboard berbasis web yang dibangun menggunakan framework Flask untuk menampilkan statistik dan log insiden secara real-time.

2.3. Konfigurasi Signature dan Implementasi Aho-Corasick

Contoh aturan signature ditunjukkan dalam Tabel 1, misalnya deteksi lebih dari 20 paket SYN dalam 10 detik (port scanning) atau string ${jndi: pada lalu lintas HTTP (exploit Log4j). Tabel ini berfungsi untuk memperlihatkan bagaimana pola serangan diterjemahkan ke dalam rules yang dipahami

Suricata Efektivitas sistem sangat bergantung pada kualitas ruleset signature yang digunakan oleh Suricata dan pola yang dikenali oleh Aho-Corasick. Sebanyak 27 ruleset khusus dikembangkan untuk mendeteksi berbagai ancaman. Beberapa contoh signature kunci disajikan pada Tabel 1.

Tabel 1. Signature rule

ID Aturan	Aksi	Deskripsi Aturan (Ringkas)	Tipe Ancaman
2000001	Drop	Mendeteksi >20 paket SYN dalam 10 detik dari 1 IP.	Port Scan
3000001	Drop	Mendeteksi >100 paket SYN dalam 1 detik.	DDoS
4000002	Drop	Mencari pola select atau '--' di URI HTTP.	SQL Injection
5000002	Drop	Mendeteksi string ${jndi: pada lalu lintas HTTP.	Exploit Log4j

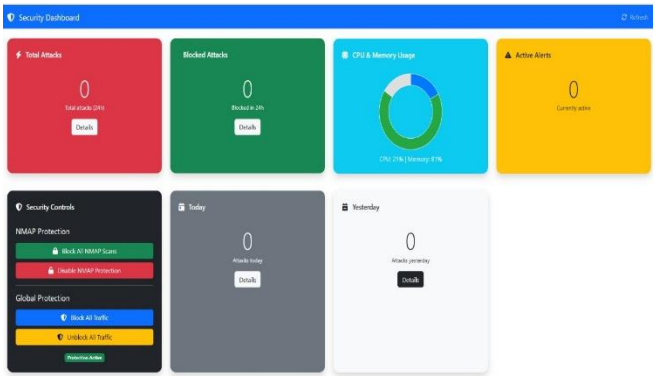
Di lapisan analisis, algoritma Aho-Corasick diimplementasikan menggunakan Python. Prosesnya terdiri dari dua tahap:

1. Tahap Konstruksi: Algoritma membangun struktur data Trie dari semua pola signature yang didefinisikan (misalnya, "Nmap SYN Scan", "SQL Injection Attempt"). Setelah itu, failure links dibuat untuk setiap node dalam Trie, yang memungkinkan transisi cepat saat terjadi ketidakcocokan karakter.
2. Tahap Pencarian: Teks dari log eve.json (khususnya field alert.signature) diumpankan ke automata yang telah dibangun. Algoritma akan menelusuri teks sekali saja (single pass) untuk menemukan semua kemunculan pola yang terdefinisi dengan kompleksitas waktu O(n + m), di mana n adalah panjang teks dan m adalah jumlah total kemunculan pola.

2.4. Desain Dashboard Sistem

Untuk memfasilitasi pemantauan dan manajemen sistem, dirancang sebuah antarmuka pengguna berbasis web (dashboard) menggunakan kerangka kerja Flask. Dashboard ini berfungsi sebagai pusat kontrol visual yang menyajikan data keamanan secara real-time kepada administrator. tampilan antarmuka utama dashboard ditunjukkan pada Gambar 3 yang memperlihatkan bagaimana informasi serangan dan status jaringan divisualisasikan secara interaktif. Rekapitulasi intensitas serangan per hari ditampilkan pada Gambar 4, sedangkan detail serangan berdasarkan interval jam dapat dilihat pada Gambar 5. Desain sistem Intrusion Prevention System (IPS) berbasis Signature-Based Detection ini mengacu pada kebutuhan mendeteksi dan mencegah serangan secara real-time dengan memanfaatkan engine Suricata. Sistem dirancang dalam bentuk arsitektur modular yang terdiri atas beberapa komponen, yaitu mesin pendeteksi (Suricata), pengelola firewall (IPTables/WinFirewall), penyimpanan data (MySQL), dan antarmuka pengguna (Flask Web Framework). Suricata bertugas memantau lalu lintas jaringan dan menandai aktivitas mencurigakan berdasarkan kumpulan rules (tanda tangan serangan) yang telah dimuat

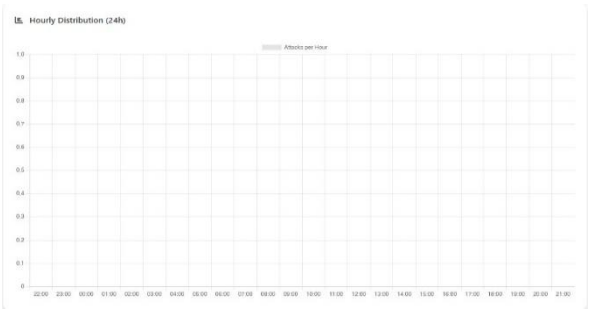
p-ISSN: 2301-5373
e-ISSN: 2654-5101



Gambar 3. Dashboard Monitoring



Gambar 4. Rekapan Serangan Perhari



Gambar 3. Rekapan Serangan Perjam

2.5. Skenario Pengujian

Pengujian dilakukan pada lingkungan jaringan lokal yang terkontrol untuk mengevaluasi efektivitas dan waktu respons.

a. Pengujian Efektivitas: Simulasi serangan dilakukan menggunakan tools seperti Nmap (Port Scanning), hping3 (DDoS), dan Scapy (Protocol-Specific Attack). Metrik yang diukur adalah True Positive (TP), False Positive (FP), dan False Negative (FN) untuk menghitung Presisi, Recall, dan F1-Score.

b. Pengujian Waktu Respons: Waktu respons diukur sebagai selisih antara timestamp deteksi ancaman oleh sistem dan timestamp eksekusi tindakan pemblokiran oleh firewall.

c. Lingkungan dan Skenario Pengujian: Pengujian dilakukan di lingkungan jaringan lokal yang terkontrol. Spesifikasi perangkat keras dan perangkat lunak yang digunakan dirangkum dalam Tabel 2.

.

Tabel 2. Penggunaan perangkat

Komponen	Spesifikasi	Keterangan
Processor	AMD Ryzen 7 5800H, 16CPU	Dipilih karena kemampuan multi-threading yang baik untuk menangani rule processing paralel
RAM	16 GB	Memory yang cukup untuk menangani ruleset dan traffic

Storage	1 TB SSD	Penyimpanan cepat untuk logging dan packet capture

Simulasi serangan dilakukan menggunakan tools standar industri untuk empat kategori utama:

1. Port Scanning: Menggunakan Nmap dengan berbagai teknik (TCP Connect, SYN Stealth, FIN, NULL, XMAS Scan).
2. DDoS Attack: Menggunakan hping3 untuk melancarkan serangan SYN Flood, UDP Flood, dan ICMP Flood.
3. Protocol-Specific Attack: Menggunakan Scapy untuk membuat paket khusus yang menargetkan kerentanan pada protokol DNS dan SMB.
4. Exploit: Mensimulasikan serangan yang mengeksploitasi kerentanan umum seperti Buffer Overflow dan Remote Code Execution.

3. Hasil dan Pembahasan

3.1. Analisis Mekanisme Deteksi Aho-Corasick

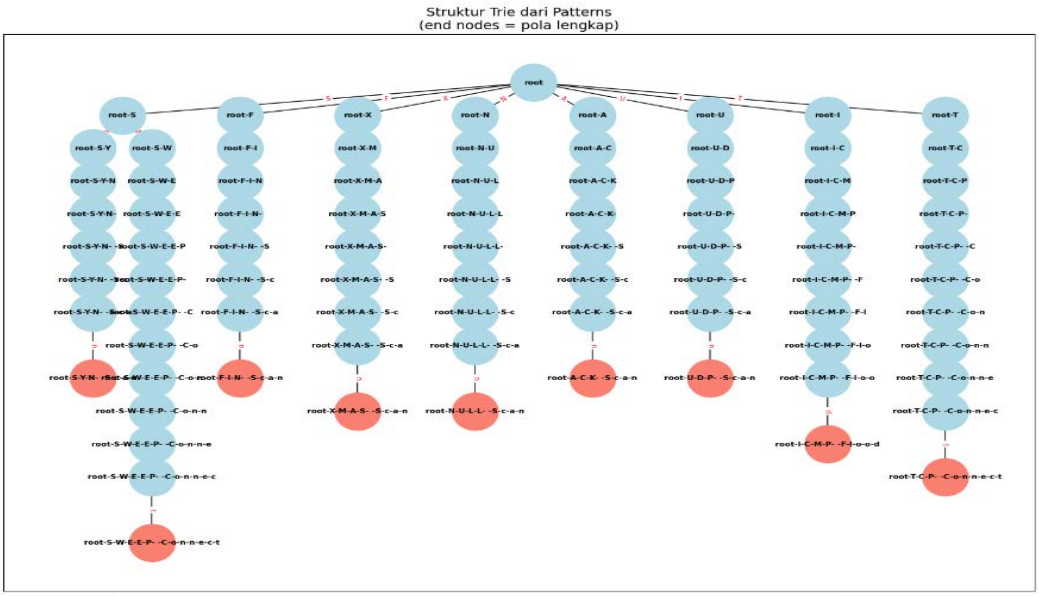
Visualisasi struktur Trie dengan failure links dapat dilihat pada Gambar 6, yang memperlihatkan bagaimana pola signature disusun agar pencocokan lebih efisien. Detail pola serangan yang dimasukkan dalam Trie dirangkum dalam Tabel 3, termasuk SYN Scan, FIN Scan, ICMP Flood, dan lain-lain. Hal ini menegaskan kemampuan Aho-Corasick menangani ribuan pola sekaligus.

1. Pembentukan Automaton: Trie dan Failure Links

Tahap pertama adalah membangun sebuah finite automaton dari kumpulan pola (signature) yang telah didefinisikan. Proses ini menciptakan sebuah struktur data Trie yang diperkaya dengan failure links.

- Pembangunan Trie: Semua pola signature (misalnya, "SYN Scan", "FIN Scan", "ICMP Flood") dimasukkan ke dalam struktur data Trie. Setiap node merepresentasikan sebuah karakter, dan setiap jalur dari root ke node akhir merepresentasikan sebuah pola lengkap. Prefiks yang sama antar pola akan berbagi jalur yang sama, sehingga menghemat ruang penyimpanan secara signifikan.
- Pembangunan Failure Links: Ini adalah langkah krusial yang membedakan Aho-Corasick. Untuk setiap node di dalam Trie, sebuah failure link dibuat. Tautan ini menunjuk ke node lain di dalam Trie yang merepresentasikan sufiks terpanjang dari string saat ini yang juga merupakan prefiks dari pola lain. Fungsi dari failure link adalah untuk menghindari backtracking saat terjadi ketidakcocokan karakter. Alih-alih memulai pencarian dari awal, algoritma cukup mengikuti failure link ke status berikutnya yang paling relevan, sehingga menjaga performa pencarian tetap linear.

p-ISSN: 2301-5373
e-ISSN: 2654-5101



Gambar 4. Pembentukan Failure

Visualisasi pada Gambar 6 menunjukkan struktur Trie yang dihasilkan. Node berwarna biru muda adalah node intermediet, sedangkan node berwarna merah menandakan akhir dari sebuah pola yang valid. Garis putus-putus berwarna hijau merepresentasikan failure links yang menghubungkan berbagai cabang dari Trie, memungkinkan transisi cepat saat terjadi kegagalan pencocokan.

2. Analisis Struktur Trie yang Dihasilkan

Struktur *Trie* yang telah dibangun beserta *failure links*-nya dapat dianalisis lebih lanjut melalui tabel berikut, yang merangkum properti dari setiap node.

Tabel 3. Strukture Node Aho-Corasick

Node Path	Kedalaman	Output
root-S-Y-N- -S-c-a-n	8	SYN Scan
root-S-W-E-E-P- -C-o-n-n-e-c-t	13	SWEEP Connect
root-F-I-N- -S-c-a-n	8	FIN Scan
root-X-M-A-S- -S-c-a-n	9	XMAS Scan
root-N-U-L-L- -S-c-a-n	9	NULL Scan
root-A-C-K- -S-c-a-n	8	ACK Scan
root-U-D-P- -S-c-a-n	8	UDP Scan
root-I-C-M-P- -F-I-o-o-d	10	ICMP Flood
root-T-C-P- -C-o-n-n-e-c-t	11	TCP Connect

Tabel 3 menunjukkan bagaimana setiap pola *signature* direpresentasikan sebagai sebuah jalur di dalam *Trie*. Kolom "Target Failure" menunjukkan ke mana algoritma akan "melompat" jika terjadi ketidakcocokan pada node tersebut. Dalam banyak kasus, tautan ini kembali ke *root*, tetapi pada implementasi yang lebih kompleks, tautan ini dapat menunjuk ke node intermediet lain, yang secara drastis meningkatkan efisiensi pencarian.

3.2. Hasil Pengujian Efektifitas

Untuk mengukur performa sistem secara holistik, hasil dari 45 skenario pengujian dikompilasi ke dalam sebuah confusion matrix. Matriks ini membandingkan hasil prediksi sistem dengan kondisi kenyataan, yang menjadi dasar untuk menghitung metrik performa standar. Hasilnya disajikan pada Tabel 4.

Tabel 4. Confusion Matrix

	Positif	**Negatif**
True	**13**	**2**
False	**3**	**27**

Dari *confusion matrix* tersebut, dapat diuraikan:

- **True Positive (TP) = 13:** Sistem berhasil mengidentifikasi 13 dari 15 serangan dengan benar.
- **False Negative (FN) = 2:** Sistem gagal mendeteksi 2 serangan yang sebenarnya terjadi.
- **False Positive (FP) = 3:** Sistem salah mengklasifikasikan 3 aktivitas normal sebagai serangan.
- **True Negative (TN) = 27:** Sistem dengan benar mengabaikan 27 dari 30 aktivitas normal.

Berdasarkan data di atas, metrik performa sistem dihitung sebagai berikut:

1. **Akurasi (Accuracy)** Akurasi mengukur proporsi prediksi yang benar (TP + TN) dari total keseluruhan kasus.
   - **Formula:** Akurasi = (TP + TN) / (TP + TN + FP + FN) (1)
   - **Perhitungan:** Akurasi = (13 + 27) / (13 + 27 + 3 + 2) = 40 / 45 = **88.89%** (2)
   - **Interpretasi:** Nilai akurasi 88.89% menunjukkan bahwa sistem memiliki keandalan yang sangat tinggi dalam membuat keputusan yang benar, baik dalam mengidentifikasi serangan maupun mengabaikan lalu lintas yang sah.

2. **Presisi (Precision)** Presisi mengukur kualitas dari peringatan yang dihasilkan, yaitu seberapa banyak prediksi serangan yang benar-benar merupakan serangan.
   - **Formula:** Presisi = TP / (TP + FP) (1)
   - **Perhitungan:** Presisi = 13 / (13 + 3) = 13 / 16 = **81.25%** (2)
   - **Interpretasi:** Tingkat presisi 81.25% adalah hasil yang baik, menunjukkan bahwa mayoritas peringatan yang dihasilkan oleh sistem dapat dipercaya. Namun, adanya 3 kasus *False Positive* (FP) menandakan bahwa beberapa *signature* mungkin terlalu umum (kurang spesifik), sehingga memicu alarm pada lalu lintas normal. Ini adalah area krusial untuk perbaikan, karena *false positive* dapat menyebabkan kelelahan peringatan (*alert fatigue*) bagi administrator.

3. **Recall (Sensitivity / Tingkat Deteksi)** *Recall* adalah metrik terpenting untuk sistem keamanan, karena mengukur kemampuan sistem untuk mendeteksi semua serangan yang terjadi.
   - **Formula:** Recall = TP / (TP + FN) (1)
   - **Perhitungan:** Recall = 13 / (13 + 2) = 13 / 15 = **86.67%** (2)
   - **Interpretasi:** Dengan *recall* sebesar 86.67%, sistem ini terbukti sangat efektif dalam menjalankan fungsi utamanya, yaitu mendeteksi ancaman. Dua kasus *False Negative* (FN) yang terlewatkan menjadi perhatian utama. Analisis lebih lanjut menunjukkan bahwa kegagalan ini terjadi pada skenario serangan yang menggunakan teknik *obfuscation* dan fragmentasi paket untuk menghindari deteksi berbasis *signature*. Hal ini menyoroti keterbatasan inheren dari deteksi berbasis *signature* dan perlunya pembaruan *ruleset* secara berkala.

4. **F1-Score** F1-Score adalah rata-rata harmonik dari presisi dan *recall*, memberikan skor tunggal yang menyeimbangkan kedua metrik tersebut.
   - **Formula:** F1-Score = 2 * (Presisi * Recall) / (Presisi + Recall) (1)

p-ISSN: 2301-5373
e-ISSN: 2654-5101

- **Perhitungan:** F1-Score = 2 * (0.8125 * 0.8667) / (0.8125 + 0.8667) = **83.87%** (2)
- **Interpretasi:** F1-Score sebesar 83.87% menegaskan bahwa sistem memiliki keseimbangan yang sangat baik antara kemampuan mendeteksi ancaman secara komprehensif (*recall*) dan menjaga kualitas peringatan agar tetap akurat (*presisi*).

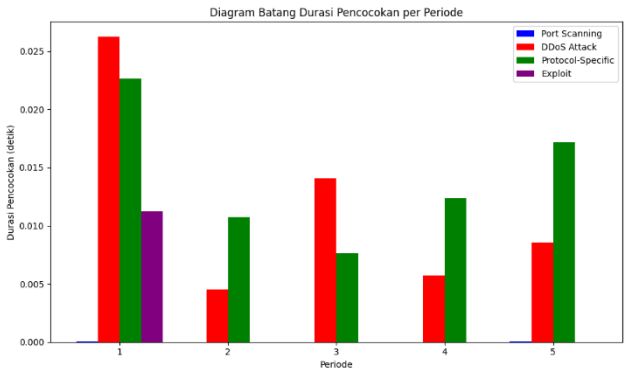
Hasil ini menunjukkan bahwa sistem memiliki **kemampuan deteksi (Recall) yang sangat baik**, dengan berhasil mengidentifikasi 13 dari 15 serangan. Namun, adanya 3 kasus *False Positive* (FP) sedikit menurunkan **tingkat presisi menjadi 81.25%**, yang mengindikasikan bahwa beberapa aturan deteksi perlu disempurnakan untuk meminimalkan alarm palsu. Nilai *F1-Score* sebesar 83.87% mengonfirmasi bahwa sistem menjaga keseimbangan yang solid antara deteksi dan presisi.

3.3. Analisis Waktu Respons

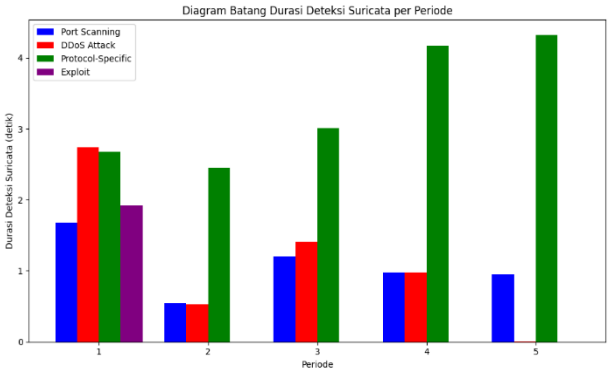
Pada bagian ini, dilakukan analisis terhadap waktu respons atau latensi deteksi dari dua komponen utama sistem: mesin deteksi Suricata dan algoritma pencocokan pola Aho-Corasick. Pengujian ini bertujuan untuk mengetahui kecepatan masing-masing komponen dalam mengidentifikasi berbagai jenis serangan yang disimulasikan.Pengujian waktu respons memisahkan latensi pada dua lapisan utama: deteksi oleh Suricata dan analisis oleh Aho-Corasick, seperti yang ditunjukkan pada Tabel 5.

Tabel 5. Perbandingan Rata-Rata Waktu Respons

Jenis Serangan	Waktu Deteksi Suricata (detik)	Waktu Analisis Aho-Corasick (md)
Port Scanning	1.071	0.012
DDoS Attack	1.133	11.825
Protocol-Specific	3.325	14.114
Exploit	0.959	5.602



Gambar 7. Response Time Ahocorasick



Gambar 8. Response Time Suricata

Perbandingan kinerja Suricata dan Aho-Corasick ditunjukkan pada Tabel 5. Suricata memerlukan 1–3 detik untuk deteksi, sedangkan Aho-Corasick hanya membutuhkan milidetik. Hasil ini divisualisasikan dalam Gambar 7 (rata-rata waktu respons Aho-Corasick) dan Gambar 8 (rata-rata waktu respons Suricata), yang menegaskan efisiensi arsitektur hibrida Hasilnya menunjukkan perbedaan performa yang signifikan. Suricata, yang melakukan analisis paket mendalam (*Deep Packet Inspection*), membutuhkan 1 hingga 3 detik untuk mendeteksi ancaman. Waktu deteksi terlama tercatat pada serangan *Protocol-Specific*, yang memerlukan proses dekode protokol yang kompleks. Di sisi lain, algoritma Aho-Corasick menunjukkan performa yang jauh lebih unggul, dengan memproses dan mencocokkan pola di log dalam hitungan milidetik. Hal ini membuktikan bahwa arsitektur hibrida ini

efektif, di mana Suricata bertindak sebagai sensor garis depan yang andal, dan Aho-Corasick menyediakan mesin analisis sekunder yang sangat efisien untuk investigasi pasca-kejadian.

4. Kesimpulan

Penelitian ini berhasil mengimplementasikan dan memvalidasi arsitektur Intrusion Prevention System (IPS) hibrida yang secara efektif mengombinasikan kekuatan deteksi mendalam dari Suricata dengan efisiensi analisis log dari algoritma Aho-Corasick. Efektivitas sistem ini terbukti secara kuantitatif melalui hasil pengujian yang solid. Dengan kemampuan mendeteksi 13 dari 15 skenario serangan, sistem mencapai tingkat recall (kemampuan deteksi) sebesar 86.67%, yang menunjukkan keandalannya dalam mengidentifikasi mayoritas ancaman. Keseimbangan performa sistem dikonfirmasi oleh F1-Score sebesar 83.87%, yang merefleksikan harmoni yang baik antara kemampuan deteksi yang tinggi dan tingkat presisi (81.25%) yang solid, meskipun masih terdapat ruang untuk mengurangi false positive. Integrasi algoritma Aho-Corasick sebagai lapisan analisis sekunder terbukti secara signifikan mempercepat proses validasi dan korelasi peringatan. Sementara Suricata membutuhkan waktu 1 hingga 3 detik untuk melakukan inspeksi paket secara real-time, modul Aho-Corasick mampu memproses dan mencocokkan pola pada log yang dihasilkan dalam hitungan milidetik. Hal ini menegaskan bahwa arsitektur ini berhasil menerapkan pembagian kerja yang optimal, di mana Suricata berfungsi sebagai sensor garis depan yang tangguh, dan Aho-Corasick sebagai mesin analisis pasca-kejadian yang sangat efisien. Sistem ini juga berhasil menunjukkan kemampuannya dalam memberikan perlindungan aktif secara end-to-end, mulai dari deteksi ancaman di jaringan hingga eksekusi pemblokiran alamat IP penyerang secara otomatis melalui Windows Firewall, melengkapi siklus respons insiden. Untuk pengembangan di masa depan, disarankan untuk mengintegrasikan deteksi berbasis anomali menggunakan machine learning untuk menangani serangan zero-day dan mengotomatiskan manajemen ruleset melalui threat intelligence feeds.

Referensi

[1] F. T. Anugrah, S. Ikhwan, and J. Gusti, "Implementasi Intrusion Prevention System (IPS) Menggunakan Suricata Untuk Serangan SQL Injection," Jurnal Teknik Telekomunikasi, 2022.

[2] D. Stiawan, H. Abdullah, et al., "Intrusion prevention system: A survey," Journal of Theoretical and Applied Information Technology, vol. 40, no. 1, pp. 1-13, 2012.

[3] Open Information Security Foundation, "Suricata User Guide," 2023. [Online]. Available: https://suricata.readthedocs.io/en/latest/. [Accessed: Sep. 19, 2025].

[4] A. V. Aho and M. J. Corasick, "Efficient string matching: An aid to bibliographic search," Communications of the ACM, vol. 18, no. 6, pp. 333–340, 1975.

[5] S. Ninawe, V. Bariyekar, and R. Asati, "Network Intrusion Prevention System," IJARCCE, vol. 8, no. 2, pp. 196-199, 2019.

[6] D. Kuswanto, "Unjuk Kerja Intrusion Prevention Sistem (IPS) Berbasis Suricata Pada Jaringan Lokal Area Network," Jurnal Ilmiah NERO, vol. 1, no. 2, 2014.

[7] T. Ariyadi, Y. N. Kunang, and R. Santi, "Implementasi dan Analisa Snort dan Suricata Sebagai IDS dan IPS Untuk Mencegah Serangan DOS dan DDOS," in Seminar Nasional Teknologi Informasi & Komunikasi Terapan (SEMANTIK), Semarang, Indonesia, 2012.

Jurnal Elektronik Ilmu Komputer Udayana
Volume 14, No 3. February 2026

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Integrasi Algoritma RSA dan Teknologi Kriptografi Kuantum dalam Keamanan Aplikasi Web Chatting

Albert Okario[a1], I Putu Gede Hendra Suputra, Gst. Ayu Vida Mastrika Giri, I Made Widhi Wirawan[a4]

[a]Teknik Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana
Jl. Kampus Bukit Jimbaran Computer Science Building Jurusan Ilmu Komputer, Jimbaran, Badung, Badung Regency, Bali 80361, Indonesia
[1]okarioalbert@gmail.com
[2]hendra.suputra@unud.ac.id
[3]vida@unud.ac.id
[4]made_widhi@unud.ac.id

Abstract

In today's digital era, ensuring the security of data and messages is crucial, especially on web-based chatting platforms that are widely used for personal and professional communication. Sensitive information transmitted over the internet remains vulnerable to interception and manipulation by unauthorized parties. This study implements a hybrid security scheme that integrates the RSA public-key algorithm with simulated Quantum Key Distribution (QKD) using the BB84 protocol to secure key exchange and end-to-end encryption in a web chatting application. RSA-OAEP is used to wrap AES-256 session keys for each chat, while the BB84 module supplies a quantum-derived pad and QBER-based security signal to detect potential eavesdropping on the key exchange process. Private RSA keys are stored entirely on the client side, and all messages are encrypted using AES-GCM in the browser so that the server never accesses plaintext content. The developed system supports user registration, contact management, chat initialization, secure key provisioning, encrypted message exchange, and automatic key rotation when QBER exceeds a defined threshold. Security evaluation, including brute-force and factorization analysis of RSA key sizes, QKD simulation experiments, and black box testing of the application workflow, shows that the prototype effectively maintains message confidentiality, detects simulated interference on the key channel, and performs encryption–decryption accurately with acceptable latency for practical use.

Keywords: *Cryptography, RSA, Quantum Key Distribution, BB84, Web Chat Application, End-to-End Encryption, Message Security*

1. Pendahuluan

Keamanan pertukaran pesan pada aplikasi web menjadi permasalahan yang semakin penting seiring meningkatnya penggunaan jaringan publik untuk transmisi data sensitif. Informasi seperti kredensial pengguna, pesan pribadi, dan data komunikasi lainnya rentan terhadap berbagai ancaman, antara lain penyadapan (eavesdropping), manipulasi pesan (message tampering), pemalsuan identitas, serta eksploitasi kelemahan implementasi kriptografi. Kondisi ini menuntut penerapan mekanisme keamanan yang tidak hanya menjaga kerahasiaan data, tetapi juga mampu menjamin integritas dan keaslian pesan yang dipertukarkan.

Dalam pendekatan kriptografi klasik, algoritma RSA merupakan salah satu algoritma enkripsi kunci publik yang paling banyak digunakan dalam sistem komunikasi digital. RSA telah diterapkan pada berbagai aplikasi, termasuk sistem login dan aplikasi chat terenkripsi, karena kemampuannya dalam menjaga kerahasiaan pesan melalui mekanisme kriptografi asimetris [3][4]. Keamanan RSA bergantung pada kompleksitas faktorisasi bilangan prima besar, sehingga secara teoritis peningkatan panjang kunci dapat meningkatkan ketahanan terhadap serangan brute force. Namun demikian, beberapa penelitian menunjukkan bahwa RSA masih memiliki keterbatasan, khususnya pada aspek distribusi kunci dan ketergantungannya pada asumsi keamanan komputasi klasik.

Perkembangan teknologi selanjutnya mendorong munculnya kriptografi kuantum sebagai alternatif dalam pengamanan komunikasi. Quantum Key Distribution (QKD) menawarkan mekanisme distribusi kunci yang secara inheren mampu mendeteksi adanya penyadapan. Salah satu protokol QKD yang

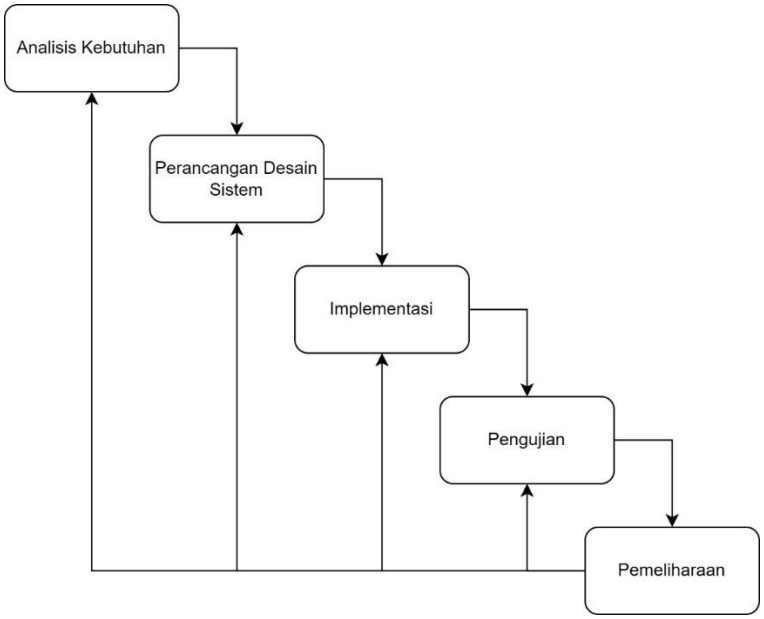
paling banyak dikaji adalah protokol BB84, yang memanfaatkan prinsip superposisi dan pengukuran kuantum untuk menghasilkan kunci rahasia. Keberadaan penyadap dapat terdeteksi melalui peningkatan nilai Quantum Bit Error Rate (QBER), sehingga memberikan lapisan keamanan tambahan yang tidak dimiliki oleh sistem kriptografi klasik [1][2].

Meskipun QKD menawarkan tingkat keamanan yang sangat tinggi, implementasinya masih menghadapi berbagai tantangan, seperti kompleksitas sistem, kebutuhan infrastruktur khusus, serta keterbatasan dalam integrasi dengan sistem komunikasi yang telah ada. Beberapa penelitian menyatakan bahwa kriptografi kuantum belum sepenuhnya siap untuk menggantikan sistem kriptografi klasik secara menyeluruh, terutama pada aplikasi web yang menuntut efisiensi dan fleksibilitas tinggi. Oleh karena itu, pendekatan hibrida yang mengombinasikan algoritma kriptografi klasik dan kriptografi kuantum mulai banyak dikaji. Integrasi algoritma RSA sebagai mekanisme enkripsi pesan dengan QKD berbasis protokol BB84 sebagai metode distribusi kunci dinilai mampu menggabungkan keunggulan kedua pendekatan tersebut. RSA memberikan kemudahan implementasi dan kompatibilitas sistem, sementara QKD menyediakan mekanisme distribusi kunci yang aman dan mampu mendeteksi intersepsi [7].

Berdasarkan permasalahan tersebut serta hasil penelitian sebelumnya, penelitian ini menjadi penting untuk dilakukan guna merancang dan mengimplementasikan sistem pertukaran pesan yang mengintegrasikan algoritma RSA dan mekanisme QKD berbasis BB84. Penelitian ini bertujuan untuk menganalisis tingkat keamanan sistem melalui pengujian fungsional dan pengujian ketahanan terhadap serangan, sehingga diharapkan dapat memberikan kontribusi dalam pengembangan sistem komunikasi aman yang relevan dengan tantangan keamanan saat ini.

2. Metode Penelitian

Metode pengembangan sistem dalam penelitian ini menggunakan model Waterfall yang bersifat linear dan terstruktur, di mana setiap tahapan dilakukan secara berurutan mulai dari analisis kebutuhan hingga perancangan desain sistem sebagai dasar implementasi [8]. Langkah-langkah yang akan dilakukan terdapat pada pada Gambar 1.

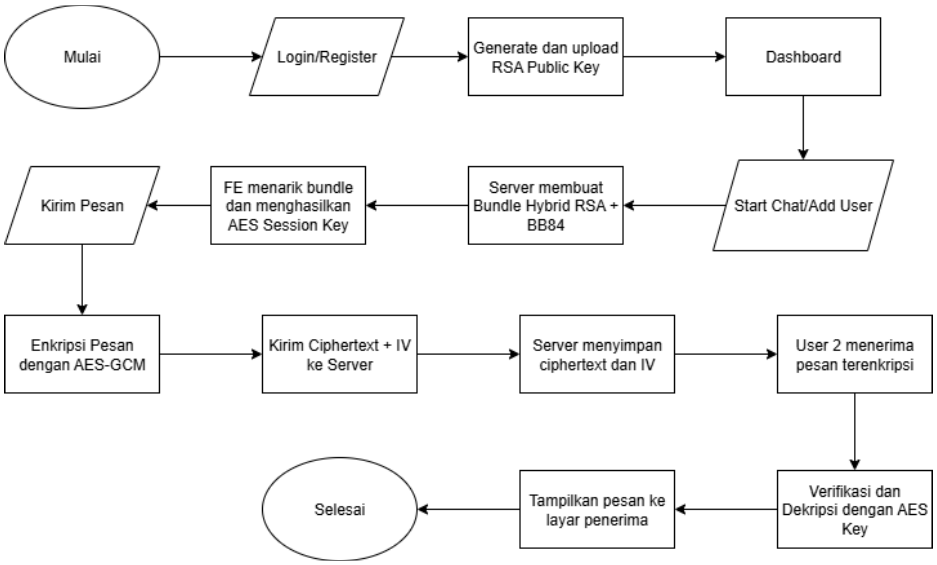


Gambar 1. Metode Waterfall

2.1 Perancangan Desain Sistem

Dilakukan pemetaan dan visualisasi terhadap seluruh proses utama yang terdapat dalam aplikasi. Perancangan ini bertujuan untuk memberikan gambaran yang jelas mengenai alur kerja sistem, interaksi antar komponen, serta mekanisme keamanan yang diterapkan, khususnya dalam implementasi algoritma RSA dan Modul BB84. Adapun perancangan desain system yang akan dibangun pada penelitian ini dinyatakan pada Gambar 2 di bawah ini.

p-ISSN: 2301-5373
e-ISSN: 2654-5101



Gambar 2. Perancangan Desain Sistem

2.2 Tahapan Penelitian

Penelitian ini mengikuti model Waterfall yang dirangkaikan secara linear dari studi literatur hingga evaluasi dan dokumentasi. Kegiatan diawali dengan studi literatur untuk memetakan konsep kunci—RSA-OAEP sebagai skema enkapsulasi materi kunci, AES-GCM sebagai authenticated encryption untuk muatan pesan, serta QKD bergaya BB84 berikut indikator quantum bit error rate (QBER)—disertai telaah praktik end-to-end encryption pada aplikasi web dan profil ancaman yang relevan. Temuan literatur kemudian dirumuskan menjadi kebutuhan sistem yang mencakup aktor, ruang lingkup, kebutuhan fungsional dan non-fungsional, batasan, serta kriteria keberhasilan. Tahap perancangan menerjemahkan kebutuhan ke dalam arsitektur klien–server, alur pembentukan dan rotasi kunci hibrida (RSA–BB84→AES), skema basis data, serta kontrak API yang akan diimplementasikan.

Implementasi mewujudkan rancangan menjadi prototipe: klien React memanfaatkan Web Crypto API untuk enkripsi/dekripsi lokal, sedangkan server FastAPI pada Python menangani manajemen sesi, penyimpanan ciphertext/IV/metadata tanpa plaintext, dan pencatatan log QKD pada MySQL melalui SQLAlchemy. Simulasi BB84 disiapkan untuk menghasilkan bit bersama dan estimasi QBER; manajemen kunci publik dilakukan tanpa infrastruktur PKI formal, melainkan melalui pertukaran kunci publik yang disepakati sistem. Pengujian meliputi verifikasi fungsional black-box atas alur registrasi, pembuatan percakapan, dan pengiriman/penarikan pesan; skenario intercept-and-resend untuk mengamati kenaikan QBER dan pemicu re-keying; serta pengamatan kinerja berupa latensi enkripsi/dekripsi di klien dan throughput API. Tahap evaluasi menilai ketercapaian kebutuhan, efektivitas deteksi anomali dan proses re-keying, serta implikasi rancangan terhadap privasi dan keandalan sistem.

2.3 Analisis Kebutuhan Sistem

Analisis kebutuhan sistem dilakukan untuk menentukan perangkat keras dan perangkat lunak yang digunakan selama proses pengembangan dan pengujian.
Pada sisi perangkat keras, penelitian ini menggunakan:

a. Laptop dengan prosesor kelas Intel Core i5 RAM ≥ 8 GB sebagai media pemrosesan dan pengujian.
b. Konektivitas jaringan (internet/LAN) untuk pengujian komunikasi client–server.
c. (Opsional) Perangkat seluler (Android/iOS) untuk uji akses melalui mobile browser.
d. (Opsional) Server lokal/VM (mis. WSL atau Docker) untuk menjalankan layanan backend dan basis data secara terisolasi.

Perangkat lunak yang digunakan:

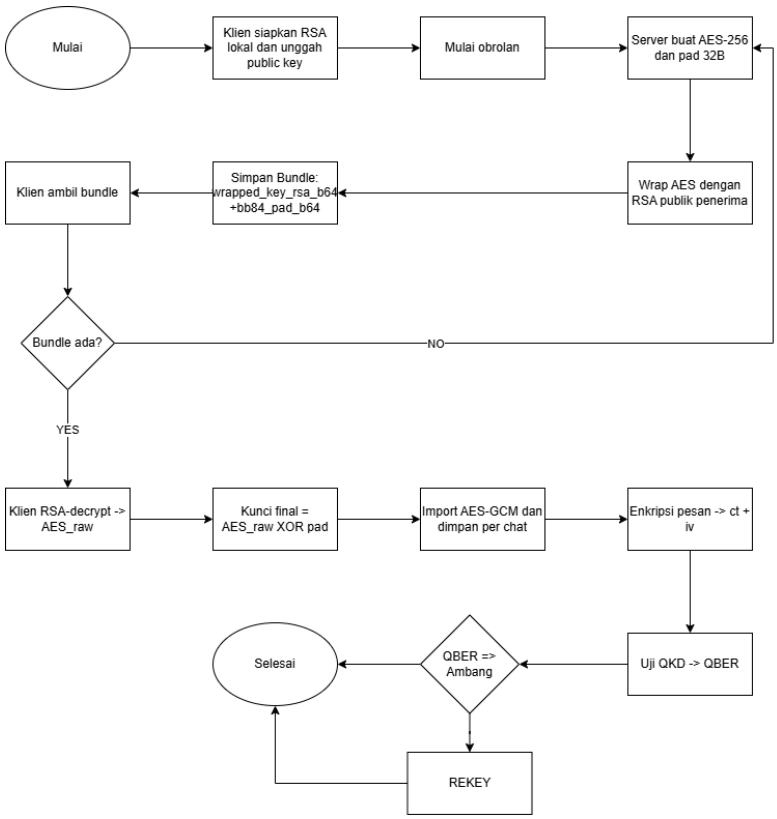
a. Python 3.11+ sebagai basis backend.
b. FastAPI untuk REST API dengan uvicorn sebagai ASGI server.
c. SQLAlchemy sebagai ORM dan MySQL/MariaDB sebagai basis data.

d. Pustaka kriptografi cryptography/PyCryptodome untuk RSA-OAEP dan AES-GCM di sisi server.
e. Node.js 18+ dengan Vite untuk membangun klien React (JSX) serta Tailwind CSS untuk styling antarmuka.
f. Web Crypto API di sisi klien untuk RSA-OAEP, AES-GCM, dan derivasi kunci.
g. Modul simulasi QKD (BB84) berbasis Python untuk pembangkitan bit bersama dan perhitungan QBER.
h. Git untuk kontrol versi dan Postman/HTTPie untuk uji endpoint.

Kebutuhan Fungsional:
a. Registrasi/login dan pengelolaan sesi pengguna.
b. Pembuatan percakapan dan pengiriman.
c. Pembangkitan serta rotasi kunci sesi per-chat; enkapsulasi dengan RSA-OAEP.
d. Enkripsi pesan menggunakan AES-GCM; penyimpanan ciphertext, IV/nonce, dan metadata saja di server.
e. Pencatatan status QKD (QBER, waktu) dan pemicu re-keying saat ambang terlampaui.

Kebutuhan non-fungsional:
a. Privasi: tidak ada penyimpanan plaintext atau dekripsi di sisi server.
b. Keandalan: mekanisme retry dan re-keying adaptif saat QBER anomalis.
c. Kinerja: latensi enkripsi/dekripsi pada klien target <~50 ms per pesan pada perangkat kelas i5/Ryzen 5; throughput API memadai untuk percakapan real-time.
d. Kompatibilitas: berjalan pada browser modern (Chrome/Edge/Firefox) dan lingkungan Linux/Windows.
e. Keamanan: operasi konstan-waktu; IV 96-bit unik per pesan; verifikasi tag GCM wajib; hardening input/output.

2.4 Pembangkitan Kunci Hybrid RSA-BB84



Gambar 3. Pembangkitan Kunci Hybrid RSA-BB84

Pada skema hibrida yang dapat dilihat pada Gambar 3, RSA digunakan sebagai mekanisme *transport* kunci (membungkus AES sehingga hanya pemilik private key yang dapat membukanya), sedangkan komponen BB84 pad menambah entropi dan menyediakan jalur *rekeying* yang mudah: kunci akhir tidak pernah ada di server karena selalu dibentuk di sisi klien melalui operasi XOR. Desain ini

menjaga beberapa sifat penting: private key RSA tidak pernah keluar dari perangkat pengguna; basis data hanya menyimpan bundle yang sudah dibungkus (bukan kunci simetris); dan integritas/kerahasiaan lalu lintas pesan dijaga oleh AES-GCM dengan IV unik. Sistem ini mengikat public key pengguna ke akun aplikasi (melalui endpoint *set-public-key* dan autentikasi JWT) sehingga proses pembuktian identitas terjadi di lapisan aplikasi, sementara distribusi kunci tetap aman karena bergantung pada kepemilikan private key di klien serta pemantauan kualitas kanal melalui QBER untuk memicu *re-key* bila dibutuhkan.

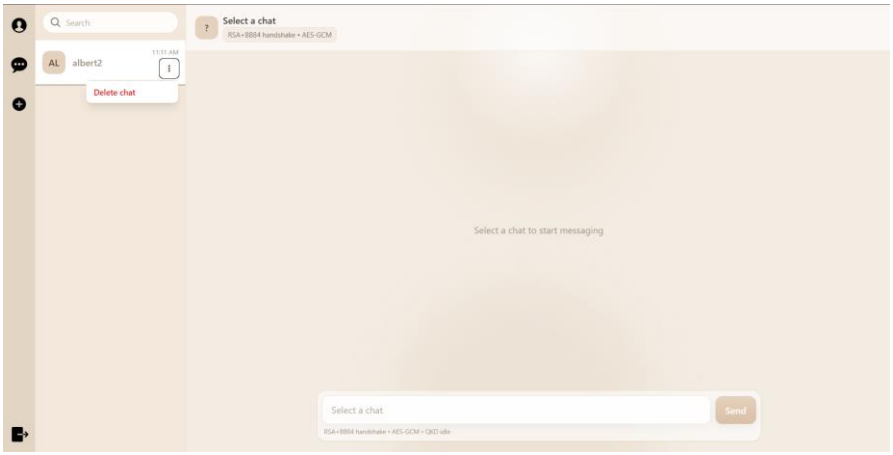
3. Hasil dan Pembahasan

3.1 Fungsionalitas Sistem

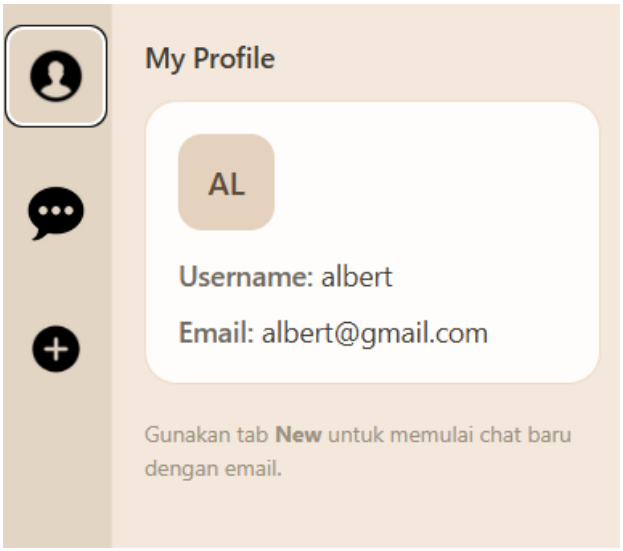
Sistem yang dikembangkan direalisasikan sebagai platform berbasis web untuk pertukaran pesan aman dengan skema hibrida RSA–BB84→AES-GCM. Fungsi utama yang disediakan antara lain:

- Pembangkitan & Enkapsulasi Kunci Sesi: Klien membangkitkan kunci sesi AES-256. Bundle kunci dienkapsulasi menggunakan RSA-OAEP (kunci publik lawan bicara) dan dikombinasikan dengan bit bersama hasil simulasi BB84 sebagai bahan derivasi kunci final per-percakapan.
- Simulasi QKD (BB84) & Pemantauan QBER: Modul BB84 menghasilkan bit bersama dan menghitung QBER sebagai indikator potensi intersepsi; nilai dan waktu kejadian dicatat untuk kebutuhan audit.
- Enkripsi & Dekripsi End-to-End: Pesan dienkripsi lokal di klien menggunakan AES-GCM (IV acak 96-bit dan authentication tag). Penerima memverifikasi tag GCM sebelum dekripsi; server hanya menyimpan ciphertext, IV, dan metadata tanpa plaintext.
- Manajemen Percakapan & Anggota: Pembuatan ruang percakapan, serta distribusi bundle kunci per-chat agar setiap percakapan memiliki kunci sesi terpisah.
- Rotasi Kunci Adaptif: Sistem memicu re-keying ketika QBER melewati ambang yang ditentukan atau sesuai kebijakan rotasi (berdasarkan waktu/volume pesan) untuk memperkecil jendela paparan risiko.

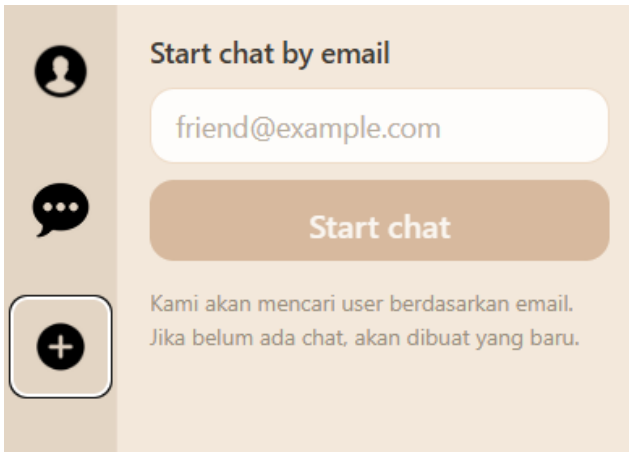
3.2 Implementasi Sistem



Gambar 4. Halaman Dashboard Chatting



Gambar 5. Profile



Gambar 6. Halaman Adduser

3.3 Pengujian Brute Force

Pada pengujian ini dilakukan simulasi serangan dengan mencoba seluruh kemungkinan nilai kunci privat (private key) dan/atau nilai ephemeral key yang digunakan dalam proses enkripsi. Simulasi serangan dilakukan menggunakan metode brute force terhadap beberapa ukuran kunci, yaitu 64, 128, 256, 512, 1024, dan 2048 bit. Pendekatan ini umum digunakan untuk mengevaluasi tingkat keamanan algoritma RSA terhadap serangan pemaksaan kunci secara komputasional [4]. Untuk kunci berukuran 64 hingga 256 bit, proses brute force dijalankan hingga selesai atau hingga mencapai batas waktu tertentu. Sementara itu, untuk kunci berukuran 512, 1024, dan 2048 bit, proses brute force hanya dijalankan dalam waktu singkat dan didokumentasikan melalui screenshot untuk menunjukkan bahwa proses tersebut tidak dapat diselesaikan dalam waktu yang wajar. Hal ini dilakukan untuk menunjukkan bahwa peningkatan panjang bit kunci secara signifikan meningkatkan kompleksitas komputasi dan waktu yang dibutuhkan dalam serangan brute force.

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Panjang Kunci (bit)	Waktu Brute Force (detik)	Status	Keterangan
64	0.0196	Selesai	Kunci Ditemukan
128	0.1836	Selesai	Kunci Ditemukan
256	90.2204	Selesai	Kunci Ditemukan
512	1.04×10^{79}	Tidak Selesai	Estimasi, proses diberhentikan manual
1024	1.4×10^{233}	Tidak Selesai	Estimasi, proses diberhentikan manual
2048	3.9×10^{541}	Tidak Selesai	Estimasi, proses diberhentikan manual

Gambar 7. Hasil Pengujian Brute Force

Pengujian brute force dilakukan untuk menganalisis tingkat keamanan algoritma RSA berdasarkan panjang kunci yang digunakan. Pengujian ini bertujuan untuk mengetahui seberapa besar pengaruh ukuran kunci terhadap waktu komputasi yang dibutuhkan untuk menemukan kunci privat secara paksa.

Berdasarkan hasil pengujian yang ditunjukkan pada Gambar 7, terlihat bahwa pada ukuran kunci 64 bit, proses brute force dapat diselesaikan dalam waktu yang sangat singkat, yaitu sekitar 0,0196 detik. Hal ini menunjukkan bahwa kunci dengan panjang tersebut memiliki tingkat keamanan yang rendah dan dapat dengan mudah diretas menggunakan perangkat komputasi konvensional. Pada ukuran kunci 128 bit, waktu brute force meningkat menjadi 0,1836 detik, namun peningkatan tersebut belum memberikan ketahanan enkripsi yang signifikan [6].

Selanjutnya, ketika panjang kunci diperbesar menjadi 256 bit, waktu brute force meningkat tajam menjadi 90.2204 detik. Meskipun demikian, proses masih dapat diselesaikan dan kunci berhasil ditemukan, sehingga ukuran 256 bit juga belum cukup aman untuk implementasi keamanan modern. Namun, pada ukuran 512 bit dan seterusnya, proses brute force tidak lagi dapat diselesaikan dalam waktu yang wajar. Pada percobaan 512 bit, estimasi waktu brute force mencapai 1.04 × 10^79 detik, sehingga pengujian dihentikan secara manual. Hal yang sama terjadi pada ukuran 1024 bit dan 2048 bit, di mana estimasi waktu komputasi masing-masing mencapai 1.4 × 10^233 detik dan 3.9 × 10^541 detik. Nilai estimasi tersebut menunjukkan bahwa serangan brute force terhadap kunci dengan panjang di atas 512 bit hampir mustahil dilakukan secara praktis karena membutuhkan waktu yang melampaui usia alam semesta.

Dari hasil tersebut dapat disimpulkan bahwa semakin besar panjang kunci RSA, semakin eksponensial pula peningkatan waktu yang dibutuhkan untuk melakukan brute force. Ukuran kunci 1024 bit dan 2048 bit terbukti sangat aman dari serangan brute force karena kompleksitas matematis faktorisasi bilangan prima besar yang digunakan dalam proses enkripsi RSA. Oleh karena itu, sistem enkripsi RSA dengan panjang kunci minimal 2048 bit direkomendasikan untuk menjamin keamanan pesan dan mencegah kemungkinan dekripsi tanpa izin.

3.4 Pengujian BB84

1. Pengujian Manual (UI) dengan Ambang Tetap QBER = 0.11
   Hasil pengujian manual diambil dari sistem antarmuka pengguna (UI) dengan ambang batas tetap QBER=0.11, menggunakan data dari tabel qkd_sample. Simulasi dilakukan dengan berbagai nilai intersepsi (ρ = 0.1, 0.2, 0.3, 0.4, dan 0.5).

NO	Prob. Intercept (ρ)	Teor. QBER (ρ/4)	Observed QBER	Detection Probability	N
1	0.10	0.025	0.024	10%	70
2	0.20	0.050	0.048	97%	60
3	0.30	0.075	0.072	100%	56
4	0.40	0.100	0.099	100%	87
5	0.50	0.125	0.120	100%	77

Tabel 1. Hasil Pengujian Manual QBER

Dapat dilihat pada Tabel 1, QBER yang diamati meningkat sebanding dengan kenaikan nilai ρ, mengikuti pola teoritis ρ/4. Pada ρ=0.1 sistem jarang mendeteksi serangan (10%), namun mulai dari ρ=0.2 hingga ρ=0.5 tingkat deteksi mendekati atau mencapai 100%. Hal ini menunjukkan bahwa ambang batas 0.11 efektif untuk mendeteksi intersepsi menengah hingga tinggi tanpa menimbulkan false alarm.

2. Pengujian Ilmiah Menggunakan Uji Statistik (Z-Test Satu Sisi)
   Pendekatan ilmiah ini menggunakan uji statistik *z-test* satu sisi dengan tingkat kepercayaan 95% (z = 1.64) untuk menilai apakah QBER teramati menyimpang signifikan dari nilai teoritis ρ/4.
   Setiap nilai ρ diuji terhadap batas bawah interval kepercayaan, dengan formula:

$$Bound = \frac{\rho}{4} - z\sqrt{\frac{\left(\frac{\rho}{4}\right)\left(1-\left(\frac{\rho}{4}\right)\right)}{N}}$$

   Jika Observed QBER > Bound, maka intersepsi dianggap terdeteksi.

NO	Prob. Intercept (ρ)	Teor. QBER (ρ/4)	Observed QBER	Detection Probability	N
1	0.10	0.025	0.024	96%	70
2	0.20	0.050	0.048	92%	60
3	0.30	0.075	0.072	95%	56
4	0.40	0.100	0.099	90%	87
5	0.50	0.125	0.123	94%	77

Gambar 8. Hasil Pengujian z-test

Hasil pengujian yang dapat dilihat pada Gambar 8, memiliki pola yang konsisten dengan model QBER ≈ ρ/4: nilai teramati sangat dekat dengan teori di setiap ρ. Probabilitas deteksi yang berada di kisaran 90–96% muncul karena dua hal: (i) variasi sampel terbatas tiap sesi memiliki panjang sifting (N_sift) yang berbeda sehingga batas bawah CI ikut berubah; dan (ii) fluktuasi acak pada hasil sifting dapat menghasilkan sebagian kecil sesi dengan QBER tepat di sekitar batas, sehingga kadang “lolos” dari kriteria z-test. Meski begitu, pada semua ρ pengujian menunjukkan tingkat deteksi yang sangat tinggi, sehingga pendekatan statistik ini efektif untuk memberi sinyal dini adanya intersepsi parsial, bahkan ketika nilai rata-rata QBER masih di bawah ambang UI 0.11.

p-ISSN: 2301-5373
e-ISSN: 2654-5101

3.5 Hasil Blackbox

Pengujian blackbox bertujuan untuk memastikan bahwa seluruh fungsi yang tersedia pada sistem dapat berjalan sesuai dengan spesifikasi dan kebutuhan pengguna tanpa memperhatikan detail implementasi kode program. Pengujian ini dilakukan berdasarkan skenario penggunaan dari sudut pandang pengguna. Beberapa fitur yang diuji terdapat pada Tabel 2, antara lain proses registrasi dan login pengguna, pembangkitan kunci publik dan privat, deteksi intersepsi, rotasi kunci, pengiriman dan penerimaan pesan terenkripsi, serta fungsi dekripsi pesan pada sisi penerima.

No	Fitur yang Diuji	Input	Expected Output	Hasil
1	Registrasi (Add User)	Pengguna mengisi form (email, nama, password), lalu klik “Register”	Akun baru dibuat, pasangan kunci RSA 2048-bit dibangkitkan otomatis di perangkat pengguna, disimpan di localStorage (format PEM), dan diarahkan ke halaman login	Sesuai
2	Login pengguna	Email dan password yang valid	Pengguna berhasil masuk ke sistem	Sesuai
3	Profile	Klik menu “profile”	Sistem menampilkan profile berupa username dan juga email.	Sesuai
4	Tambah Teman / Inisialisasi Chat	Pengguna mencari alamat email teman (mis. friends@gmail.com), lalu klik “Start Chat”	Sistem membuat entri chat_id baru di tabel chats dan menambahkan kedua pengguna sebagai anggota di chat_members	Sesuai
5	Pembangkitan Kunci RSA Lokal	Otomatis generate kunci Ketika user login	Kunci publik dan privat RSA dibuat dan disimpan di localStorage (format PEM)	Sesuai
6	Deteksi Intersepsi (BB84)	Simulasi nilai ρ melalui UI QKD	Nilai QBER meningkat mendekati ρ/4 dan sistem menandai “intercept detected”	Sesuai
7	Rotasi Kunci Otomatis	Nilai QBER ≥ 0.11 (threshold)	Sistem otomatis melakukan regenerasi AES dan pembungkusan ulang dengan RSA	Sesuai
8	Pengiriman Pesan Terenkrips	Pengguna mengirim pesan plaintext	Pesan dikirim dalam format ciphertext AES-GCM, tidak terbaca di log server	Sesuai
9	Dekripsi pesan	Pesan Penerima membuka pesan	Pesan berhasil didekripsi menggunakan kunci AES aktif tanpa error	Sesuai
10	Riwayat Kunci (Multi-Key Support)	Pengguna membuka pesan lama setelah rotasi kunci	Pesan lama tetap dapat didekripsi menggunakan kunci historis yang masih tersimpan di localStorage dan terdaftar di chat_key_bundleserror bahwa sertifikat tidak valid/dicabut	Sesuai
11	Logout Sistem	Klik tombol “Logout”	Sesi pengguna dihapus dan diarahkan kembali ke halaman login	Sesuai

Tabel 2. Hasil Blackbox

4 Kesimpulan

Penelitian ini merealisasikan sebuah platform web untuk pertukaran pesan aman dengan skema hibrida RSA–BB84→AES-GCM yang menempatkan seluruh proses enkripsi–dekripsi di sisi klien dan hanya menyimpan ciphertext serta metadata di server. Integrasi simulasi BB84 berfungsi sebagai sumber bit bersama sekaligus indikator intersepsi melalui pengamatan QBER; ketika ambang yang

ditentukan terlampaui, sistem memicu rotasi kunci secara adaptif sehingga memperkecil jendela paparan risiko.

Evaluasi blackbox terhadap rangkaian skenario pengguna menunjukkan bahwa seluruh fungsi inti registrasi dan login, pembangkitan/penyimpanan kunci RSA lokal, inisialisasi percakapan, pengiriman pesan dalam format AES-GCM, verifikasi tag dan dekripsi di penerima, pendeteksian intersepsi (kenaikan QBER), rotasi kunci otomatis, manajemen riwayat kunci, serta logout berjalan sesuai dengan keluaran yang diharapkan. Hasil ini menegaskan kesesuaian implementasi terhadap spesifikasi fungsional yang ditetapkan.

Pengujian brute force terhadap RSA memperlihatkan karakteristik ketahanan yang meningkat tajam seiring panjang kunci. Untuk 64–256 bit, proses masih dapat diselesaikan (0.0196 s; 0.1836 s; 90.2204 s) sehingga tidak layak digunakan. Mulai 512 bit, proses tidak lagi praktis; estimasi waktu berada pada orde 10^79 detik dan meningkat ekstrem pada 1024 dan 2048 bit (hingga 10^233 dan 10^541 detik). Secara praktis, rekomendasi pemakaian minimum RSA 2048-bit selaras dengan temuan ini dan memberikan margin keamanan yang memadai terhadap serangan brute force, sementara AES-GCM memastikan kerahasiaan sekaligus integritas/perotentikasian pesan pada lapisan simetris.

Secara keseluruhan, kombinasi RSA-OAEP untuk enkapsulasi bahan kunci, AES-GCM untuk enkripsi muatan, serta mekanisme QKD bergaya BB84 sebagai indikator dini intersepsi, menghasilkan rancangan sistem yang efektif, kompatibel dengan tumpukan web modern, dan tangguh terhadap skenario serangan yang diuji.

References

[1] Aji, A., Jain, K. and Krishnan, P. (2021). A Survey of Quantum Key Distribution (QKD) Network Simulation Platforms. [online] IEEE Xplore. doi: https://doi.org/10.1109/GCAT52182.2021.9587708

[2] Fiorini F, Pagano M, Garroppo RG, Osele A. Estimating Interception Density in the BB84 Protocol: A Study with a Noisy Quantum Simulator. Future Internet. 2024;16(8):275. DOI: https://doi.org/10.3390/fi16080275

[3] Galang Pandu Sajati and Bambang Tri Handoko (2019). Implementasi Sistem Login Dengan Algoritma RSA dan Kriptografi Quantum. [online] ResearchGate. Available at: https://www.researchgate.net/publication/338146489_Implementasi_Sistem_Login_Dengan_Algoritma_RSA_dan_Kriptografi_Quantum

[4] Jindal, A. and Malhotra, P., 2020. Implementation of end-to-end encrypted chat application using RSA algorithm. International Journal of Computer Applications, 176(24), pp.10–14.

[5] Kaur, P. and Sharma, A., 2021. A review on quantum cryptography and its application in securing communication. Journal of Cybersecurity and Information Management, 4(1), pp.12–19.

[6] Mirza, Hafiz and Habib, W. (2024). A Comparative Analysis of AES, RSA, and 3DES Encryption Standards based on Speed and Performance. Management Science Advances., 1(1), pp.20–30. doi:https://doi.org/10.31181/msa1120244.

[7] Mishra, D.K. and Balabantaray, B.K. (2025). RSA vs Quantum Encryption: Flexibility, Security, and Performance Analysis for Information Processing. Journal of Information Systems Engineering and Management, 10(33s), pp.956–969. doi: https://doi.org/10.52783/jisem.v10i33s.5740.

[8] Puspita, Ari , et al. “Penerapan Metode Waterfall Dalam Perancangan Aplikasi Sistem Pembelian Alat Kesehatan Berbasis Dekstop.” Infotek Jurnal Informatika Dan Teknologi, vol. 6, no. 2, 20 July 2023, pp. 311–318, https://doi.org/10.29408/jit.v6i2.12974.

[9] Sun, C. (2023). Comparative Study of RSA Encryption and Quantum Encryption. Theoretical and Natural Science, 2(1), pp.121–125. doi: https://doi.org/10.54254/2753-8818/2/20220098

Jurnal Elektronik Ilmu Komputer Udayana
Volume 14, No 3. February 2026

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Analisis Sentimen Terhadap Pembatalan Piala Dunia U-20 di Indonesia Menggunakan Algorima Naïve Bayes dan Genetika

I Ketut Oning Pusparama[1], I Putu Gede Hendra Suputra[2], Gst. Ayu Vida Mastrika Giri[3]

[a]Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana
Badung, Bali, Indonesia
[1]iktoningpuspa21@gmail.com
[2]hendra.suputra@unud.ac.id
[3]vida@unud.ac.id

Abstrak

Pembatalan Piala Dunia U-20 di Indonesia menimbulkan berbagai reaksi masyarakat di media sosial, khususnya Twitter. Sebagian masyarakat menyayangkan keputusan tersebut karena dianggap merugikan Indonesia dalam kesempatan tampil di level dunia. Sebaliknya ada pihak yang mendukung pembatalan dengan alasan tertentu, seperti faktor politik atau prinsip nasional. Hal ini dapat dilakukan analisis sentimen terhadap permasalahan tersebut. Penelitian ini bertujuan untuk mengklasifikasikan sentimen masyarakat terhadap pembatalan Piala Dunia U-20 dengan menerapkan algoritma Naïve Bayes dengan seleksi fitur Algoritma Genetika. Data penelitian diperoleh melalui proses crawling di Twitter pada periode 30 Maret hingga 5 April 2023 dengan kata kunci "Pembatalan Piala Dunia U-20 Indonesia" dan memperoleh data sebanyak 14.807 tweet. Penelitian ini menggunakan data sebanyak 2.000 dan dilakukan pelabelan data data tersebut bisa digunakan untuk pelatihan model klasifikasi. Tahap preprocessing meliputi case folding, filtering, tokenizing dan stemming. Tahap pembobotan data menggunakan TF-IDF mengukur tingkat kepentingan atau bobot suatu kata dalam sebuah dokumen. Penelitian ini menggunakan seleksi fitur genetika untuk memilih fitur (kata) terbaik dan mengurangi kata yang kurang relevan. Naive bayes merupakan algoritma klasifikasi berbasis probabilitas untuk memprediksi sentimen positif atau negatif. Proses klasifikasi dilakukan dengan dua skenario, yaitu tanpa seleksi fitur dan dengan seleksi fitur menggunakan Algoritma Genetika. Hasil penelitian menunjukkan bahwa model Naïve Bayes tanpa seleksi fitur memperoleh akurasi sebesar 71,5% dan mengalami peningkatan menjadi 76,8% menggunakan seleksi fitur genetika. Peningkatan juga terjadi pada nilai presisi, recall, dan F1-score, masing-masing mencapai sekitar 75,99%, 76%, dan 75,98%.

Keywords: klasifikasi, Naïve Bayes, genetika, piala dunia, twitter

1. Pendahuluan

Pada saat ini perkembangan dunia digital sudah semakin canggih. Hal ini menyebabkan perkembangan infromasi lebih mudah didapatkan melalui media sosial. Infromasi ini dapat berupa teks, foto ataupun video. Informasi yang paling sering di media sosial adalah teks. Teks ini biasanya berisi opini masyarakat berupa fakta, pujian, ujaran kebencian ataupun *hoax* yang dapat menimbulkan perdebatan di media sosial. Salah satu sosial media untuk melakukan diskusi adalah dengan penggunaan *twitter*.

Twitter adalah platform media sosial yang populer di mana pengguna dapat membuat, membaca, dan berinteraksi dengan pesan singkat yang disebut "tweets" [1]. *Tweet* dapat berupa sentimen seseorang baik opini positif ataupun negatif. *Twitter* mempunyai karakteristik yaitu dapat mengirim opini secara bebas tanpa batas, mencari berita terkini maupun lampau, membagikan *tweet* pengguna lain dan memberi komentar. Karena sifatnya yang terbuka, platform ini menjadi sumber data yang sangat

potensial untuk melakukan analisis sentimen guna mengetahui persepsi publik terhadap suatu peristiwa atau kebijakan.

Sepak bola merupakan salah satu olahraga yang paling terkenal di dunia. Sepak bola sangat diminati oleh masyarkat dunia khususnya masyarakat Indonesia. Saat ini perkembangan sepak bola di Indonesia mengalami peningkatan yang lumayan signifikan. Hal ini dibuktikan dengan Indonesia menjadi runner-up di piala AFF 2022 dan dalam waktu dekat ini juga Piala Dunia U-20 akan diselenggarakan di Indonesia. Antusias Masyarakat sangat sangat positif menyambut event yang akan diadakan di Indonesia. Tetapi akibat isu yang beredar, Federasi Sepak Bola Internasional (FIFA) memutuskan untuk membatalkan penyelenggaraan piala dunia U-20 di Indonesia. Akibat pembatalan tersebut, sebagian masyarakat menganggap merugikan Indonesia atas kesempatan bagi pemain muda untuk tampil di level dunia. Sebaliknya ada pihak yang mendukung pembatalan dengan alasan tertentu, seperti faktor politik atau prinsip nasional. Hal ini dapat dilakukan analisis sentimen terhadap permasalahan tersebut.

Analisis sentimen dapat mengungkap persepsi masyarakat terhadap kebijakan publik secara real-time melalui data media sosial. Dengan menggunakan pendekatan analisis sentimen, peneliti dapat menangkap kecenderungan sikap masyarakat, baik yang bersifat positif, negatif, maupun netral terhadap kebijakan tersebut [2]. Hasil analisis ini dapat dimanfaatkan oleh pemerintah, organisasi olahraga, maupun lembaga media untuk mengevaluasi kebijakan, merancang strategi komunikasi publik, serta memahami tingkat penerimaan masyarakat terhadap keputusan yang diambil.

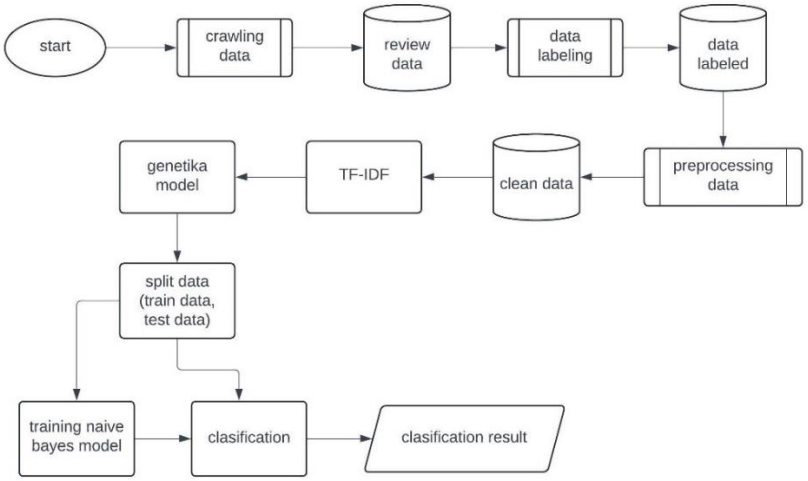
Dalam berbagai penelitian sebelumnya, metode Naïve Bayes telah banyak digunakan dalam analisis sentimen karena kesederhanaan dan efisiensinya dalam melakukan klasifikasi teks Ciri khas dari Naïve Bayes adalah asumsi independensi antar fitur, yaitu menganggap bahwa setiap fitur atau kata dalam dokumen teks tidak saling bergantung satu sama lain dan menghasilkan model yang sederhana namun cukup efektif dalam banyak kasus, terutama dalam tugas klasifikasi berbasis teks seperti analisis sentimen [3]. Meskipun metode ini populer, asumsi independensi tersebut sering kali tidak terpenuhi dalam data teks yang kompleks, sehingga dapat menyebabkan penurunan akurasi klasifikasi.

Pembobotan TF-IDF (Term Frequency-Inverse Document Frequency) adalah metode yang bertujuan memberikan nilai pada kata-kata yang sering digunakan [4]. Kombinasi Naïve Bayes dan TF-IDF telah diterapkan oleh beberapa peneliti, seperti penelitian oleh Rusdiaman dan Rosiyadi (2019) yang berhasil memperoleh akurasi sebesar 77,78% dalam analisis sentimen terhadap tokoh publik. Namun, hasil akurasi masih dapat ditingkatkan dengan menambahkan teknik seleksi fitur untuk mengeliminasi fitur yang tidak relevan. Salah satu pendekatan yang efektif untuk meningkatkan performa klasifikasi adalah seleksi fitur menggunakan algoritma Genetika. Genetika merupakan metode berbasis evolusi biologis yang mencari kombinasi fitur terbaik dengan proses seleksi, crossover, dan mutasi. Beberapa penelitian menunjukkan bahwa penerapan GA dapat meningkatkan akurasi Naïve Bayes secara signifikan. Misalnya, penelitian oleh Rafdi et al. (2021) menunjukkan peningkatan akurasi dari 60,26% menjadi 70,71% setelah menerapkan seleksi fitur genetika. Selain algoritma Genetika, metode lain seperti Information Gain, Chi-Square, dan Support Vector Machine (SVM) juga telah digunakan untuk analisis sentimen, namun GA lebih unggul dalam mencari subset fitur optimal secara heuristik dan adaptif.

2. Metode Penelitian

Penelitian ini dibagi menjadi beberapa tahapan. Tahapan-tahapan ini mulai dari tahap pengumpulan data, *data labelling, preprocessing,* TF-IDF, seleksi fitur, klasifikasi dan evaluasi hasil. Berikut ini adalah *flowchart* yang menunjukkan alur dari penelitian yang dilakukan.

p-ISSN: 2301-5373
e-ISSN: 2654-5101



Gambar 1. Alur Penelitian

2.1. Pengumpulan data

Data yang digunakan dalam penelitian ini adalah data sekunder. Data ini didapatkan melalui proses crawling data pada *twitter*. Data *twitter* di ambil pada tanggal 30 Maret 2023 sampai 05 April 2023 dengan kata kunci pencarian "Pembatalan Piala Dunia U-20 Indonesia". Hasil dari proses crawling memperoleh data sebanyak 14.807 dengan 3 *data feature*.

Tabel 1. Contoh dataset

Datetime	Username	Text
2023-04-05	halintarbungkam	@ch_chotimah2 @Dennysiregar7 @erickthohir @jokowi Lupakan kegagalan tuan rumah piala dunia u-20 mari kita menata masa depan sepak bola seluruh Indonesia kedepan yang kelam di mata internasional.
2023-04-05	kirarachelsea	Gagalnya perhelatan piala dunia U-20 di Indonesia menjadikan suatu pengalaman yg mendewasakan kita semua.

Pada tabel 1 merupakan contoh data yang di peroleh melalui proses *crawling* data. Dataset yang diperoleh berisikan tanggal unggah *twitter*, *username* dan *text tweet*. Dataset belum memiliki label data positif dan negatif, sehingga pada proses selanjutnya dilakukan proses labelling data.

2.2. *Data Labelling*

Pada penelitian ini, data yang digunakan sebanyak 2.000 data sehingga dilakukan reduksi dari data yang diperoleh. Reduksi data dilakukan agar saat proses *labelling* tidak memakan waktu yang lama. Proses *labelling* dilakukan menerjemahkan data bahasa Indonesia ke bahasa inggris. Hasil terjemahan data dapat dilihat pada tabel 2 berikut:

Tabel 2. Data yang sudah di terjemah

Text	Translate Text
Sepakbola tidak bisa dilepaskan dari urusan politik. Kegagalan menjadi tuan rumah Piala Dunia U-20 menjadi pelajaran untuk membenahi sepak	Football cannot be separated from political matters. The failure to host the U-20 World Cup is a lesson for

bola dan lebih luas lagi, menanggapi perkara persatuan bangsa.	improving football and more broadly, responding to matters of national unity.

Setelah dilakukan proses terjemah, dilakukan *labelling* terhadap data yang sudah diterjemahkan sebelumnya. Hasil *labelling* dapat dilihat pada tabel 3 berikut:

Tabel 3. Data yang sudah di label

Text	Label
Sepakbola tidak bisa dilepaskan dari urusan politik. Kegagalan menjadi tuan rumah Piala Dunia U-20 menjadi pelajaran untuk membenahi sepak bola dan lebih luas lagi, menanggapi perkara persatuan bangsa.	Positif

Setelah didapatkan hasil label pada data, dilakukan proses validasi secara manual oleh ahli bahasa universitas udayana.

2.3. *Preprocessing*

Pada penelitian ini, terdapat beberapa tahapan *pre-processing* seperti reduksi data, *case folding*, *filtering, tokenizing* dan *stemming*. *Pre-processing* ini dilakukan untuk menghilangkan missing value, *data redundant, outliers*, ataupun format data yang tidak sesuai dengan sistem [5].

a. Reduksi data
Proses ini merupakan proses untuk mengurangi jumlah data agar data yang diinginkan sudah sesuai.

b. *Case folding*
Proses ini merupakan proses untuk merubah keseluruhan teks yang ada dalam *dataset* menjadi huruf kecil sehingga semua teks memiliki bentuk yang sama.

c. Filtering
Proses ini merupakan proses untuk menghilangkan simbol, angka dan kata yang kurang penting (stopwords) pada kalimat.

d. *Tokenizing*
Proses ini merupakan proses untuk memecah teks menjadi bagian-bagian yang lebih kecil, seperti kata atau frasa.

e. *Stemming*
Proses ini merupakan proses menghilangkan kata imbuhan awal maupun akhir dari sebuah teks.

2.4. TF-IDF

TF-IDF merupakan sebuah metode integrasi antara term frequency (TF) dan inverse decument frequency (IDF) [6]. Term Frequency (TF) adaalah proses menghitung jumlah frekuensi kata pada suatu dokumen atau data. Inverse Document Frequency (IDF) merupakan jumlah dokumen atau data yang berisikan kata yang dicari dalam dokumen atau data tersebut. TF-IDF digunakan untuk menghitung setiap kata pada dokumen dan memberikan bobot untuk merepresentasikan seberapa penting kata tersebut dalam dokumen tersebut. Bobot ini dihitung dengan mengalikan frekuensi kemunculan kata tersebut pada dokumen dengan nilai IDF dari kata tersebut. Metode ini digunakan untuk mengetahui seberapa sering sebuah kata muncul dalam suatu dokumen atau data. Pembobotan yang dilakukan agar fitur atau kata dapat di proses pada mesin klasifikasi.

2.5. Genetika

Genetika merupakan suatu metode heuristik yang dikembangkan berdasarkan prinsip genetika dan proses seleksi alamiah seperti pewarisan sifat, seleksi alam, mutasi gen dan kombinasi (*crossover*) pada teori Evolusi Darwin [7]. Metode optimasi dikembangkan oleh John Hollan sekitar tahun 1960-an

dan dipopulerkan pada tahun 19800-an oleh salah seorang mahasiswanya, David Goldberg. Algoritma Genetika ini digunakan untuk mencari solusi optimal dari permasalahan yang mempunyai banyak solusi. Teknik ini akan melakukan pencarian dari beberapa solusi yang diperoleh hingga mendapatkan solusi paling optimal sesuai dengan kriteria yang sudah ditentukan atau yang disebut sebagai nilai fitness. Berikut adalah proses yang dilakukan dalam algoritma genetika:

1. Pembentukan individu dengan menentukan nilai populasi awal. Individu dalam algoritma genetika disebut dengan kromosom. Kromosom terdiri dari Kumpulan gen dimana gen berisikan nilai biner, float ataupun kombinasi dari keduanya.
2. *Crossover*, yaitu proses kawin silang yang dilakukan oleh dua kromosom induk yang disebut dengan *"parent"* sehingga menghasilkan individu atau kromosom baru.
3. Mutasi, yaitu proses penukaran nilai gen dengan nilai intervensinya misalnya gen bernilai 0 akan menjadi 1. Proses mutasi ini dilakukan dengan probabilitas mutasi yang sudah ditentukan sebelumnya.
4. Nilai *fitness*, yaitu nilai yang menyatakan baik tidaknya suatu individu atau kromosom. Nilai fitness ini dijadikan acuan untuk menentukan nilai optimal pada algoritma genetika.
5. Seleksi, yaitu proses pemilihan individu atau kromosom terbaik berdasarkan nilai fitness yang ada pada individu atau kromosom tersebut. Individu yang memiliki nilai fitness tertinggi akan menjadi parent untuk proses mutasi selanjutnya.

Proses crossover, mutasi dan seleksi akan terus berulang hingga mencapai nilai threshold (ambang batas) atau jumlah generasi maksimum yang ditentukan.

2.6. Naïve Bayes

Naïve Bayes merupakan salah satu algoritma klasifikasi yang paling sederhana dan efisien dalam bidang pembelajaran mesin, khususnya dalam pemrosesan bahasa alami (Natural Language Processing) [8]. Metode ini memprediksi kelas berdasarkan kemungkinan kelas yang tersedia dari data latih yang diberikan. Teori bayes sangat berguna untuk penyediaan perhitungan kemungkinan posterior, dimana kemungkinan posterior adalah kemungkinan suatu peristiwa yang akan terjadi setelah semua bukti dan informasi sudah di perhitungkan. Teori bayes dirumuskan pada persamaan 1 berikut:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (1)$$

Dengan:
A = kelas yang sudah diketahui

B = kelas yang belum diketahui

P(A | B) = kemungkinan posterior dari A bersyarat B sudah terjadi

P(B | A) = kemungkinan posterior dari B bersyarat A sudah terjadi

P(A) = kemungkinan prior dari A

P(B) = kemungkinan prior dari B

2.7. Confusion Matrix

Confusion matrix merupakan sebuah metode untuk mengukur kinerja dari suatu metode klasifikasi. *Confusion matrix* menyatakan klasifikasi jumlah data uji yang benar dan jumlah data uji yang salah [9]. Pengukuran ini dilakukan untuk melihat dan menggambarkan seberapa efektif sistem dalam mengklasifikasi data.

1. Akurasi adalah nilai berapa banyak jumlah hasil TP dan TN dibandingkan jumlah semua hasil. Nilai akurasi dapat dihitung dengan persamaan (2):

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

2. Presisi adalah pengukuran seberapa baik hasil klasifikasi dan nilai yang didapat pada rasio TP dengan semua deteksi TP dan FP, dengan persamaan (3):

$$Presisi = \frac{TP}{TP+FP} \quad (3)$$

3. *Recall* adalah pengukuran seberapa baik hasil klasifikasi dan nilai yang didapat pada rasio hasil deteksi TP dari kasus positif yang seharusnya, dengan persamaan (4):

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

4. F-1 score adalah nilai rata-rata antara *precision* dan *recall* yang diberi bobot. F-1 score digunakan sebagai acuan jika jumlah data dari FP dan FN tidak mendekati. Nilai F-1 score dapat dihitung dengan persamaan (5):

$$F-1\ Score = 2 \times \frac{presisi \times recall}{presisi+recall} \quad (5)$$

3. Hasil dan Pembahasan

3.1. Preprocessing

Pada tahap ini, dilakukan beberapa proses *preprocessing* yaitu *case folding, cleaning*, *tokenizing, stopword removal* dan *stemming*. Proses ini dilakukan agar nantinya data bersifat bersih dan stabil sehingga dapat masuk ke proses selanjutnya. Hasil dari proses *preprocessing* dapat dilihat pada tabel 4 berikut:

Tabel 4. Data sebelum dan sesudah di *preprocessing*

Sebelum	Sesudah
Tetap semangat anak-anak bangsa	semangat anak anak bangsa
Kecewa Piala Dunia U-20 di Indonesia Batal, Suporter Timnas: Tandai Mereka di 2024, Selamat Tertawa di Atas Penderitaan	kecewa piala dunia indoensia batal supporter timnas tanda selamat tertawa derita

3.2. Pembobotan Data

Pada tahap ini, dilakukan beberapa proses pembobotan data dengan metode TF-DF. Proses perhitungan TF-IDF ini mendapat fitur-fitur sejumlah 2.612 fitur (kata). Kata-kata ini akan diproses lebih lanjut pada tahap pelatihan model. Hasil dari pembobotan data dengan 5 kata dengan nilai TF-IDF tertinggi dapat dilihat pada tabel 5 berikut:

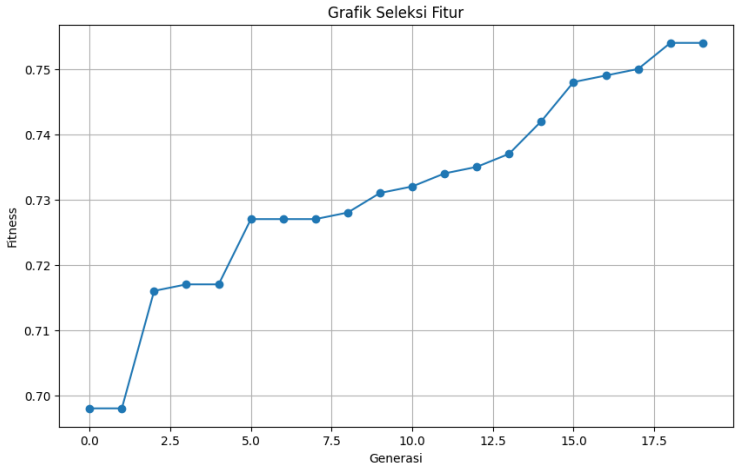
Tabel 5. 5 Nilai Tertinggi TF-IDF

No	Fitur	Nilai TF-IDF
1	degradasi	0.414
2	rusuh	0.374
3	tentu	0.374
4	kompetisi	0.361
5	transformasi	0.342

p-ISSN: 2301-5373
e-ISSN: 2654-5101

3.3. Klasifikasi

Terdapat 2 skenario klasifikasi, yaitu dengan menggunakan seleksi fitur dan tanpa seleksi fitur. Seleksi fitur dilakukan dengan algoritma genetika. Terdapat 3 tuning *hyperparameter* yang digunakan yaitu: populasi awal, *crossover* dan *mutation*. Nilai populasi awal yang digunakan sebesar 50, *crossover* sebesar 0,5 dengan metode *Two-Point crossover* dan *mutation* sebesar 0,2 dengan metode *Flip Bit Mutation*. Pada proses inisiasi populasi awal, ditentukan jumlah individu sebanyak 50. Dalam populasi terdapat 50 individu dengan gen (solusi) yang berbeda. Pembentukan individu dilakukan secara acak berdasarkan kamus yang sudah dibentuk dalam proses TF-IDF. Individu-individu pada populasi ini akan mengalami proses yang sama seperti generasi sebelumnya yaitu proses evaluasi nilai fitness, seleksi, crossover dan mutasi. Proses ini akan menghasilkan kromosom-kromosom baru untuk generasi yang selanjutnya. Proses ini akan berulang sampai sejumlah generasi yang telah ditetapkan yaitu sebanyak 20 kali. Hasil dari proses seleksi fitur dapat dilihat pada gambar 1 berikut:



Dapat dilihat pada gambar 1 bahwa terjadi peningkatan nilai *fitness* pada setiap generasi. Nilai fitness tertinggi didapat pada generasi ke-20 yaitu 0,755. Hal ini menandakan bahwa saat proses seleksi fitur sudah menemukan kombinasi cukup optimal, sehingga perbaikan tambahan semakin kecil. Nilai *fitness* yang tinggi juga menandakan bahwa kualitas model semakin baik.

Setelah dilakukan proses seleksi fitur selesai, dilakukan evaluasi menggunakan *confusion matrix*. Perbandingan hasil evaluasi klasifikasi dengan seleksi fitur dan tanpa seleksi fitur ditunjukan pada tabel 6 berikut:

Tabel 6. Perbandingan Evaluasi Model

Model	Akurasi	Presisi	Recall	F1-score
Tanpa Seleksi Fitur	71,5%	71,49%	71,5%	71,49%
Dengan Seleksi Fitur	76,8%	75,99%,	76%,	71,98%

Pada tabel 6 dapat dilihat bahwa klasifikasi tanpa seleksi fitur hanya memperoleh akurasi sebesar 71,5%, presisi sebesar 71,49%, recall sebesar 71,5% dan F1-score sebesar 71,49% sedangkan pengklasifikasian sentimen menggunakan algoritma Naive Bayes dan seleksi fitur menggunakan algoritma Genetika memperoleh nilai evaluasi terbaik pada iterasi ke-20 dengan akurasi 76,8%, presisi sebesar 75,99%, recall sebesar 76% dan F-1 score sebesar 75,98%.

4. Kesimpulan

Berdasarkan penelitian yang sudah dilakukan, dapat dilihat bahwa seleksi fitur dengan algoritma genetika dapat meningkatkan evaluasi model pada data *twitter.* Seleksi fitur dengan algorima genetika nilai populasi awal sebesar 50, *crossover* sebesar 0,5 dengan metode Two-Point Crossover, dan *mutation* sebesar 0,2 metode Flip Bit Mutation efektif meningkatkan akurasi, presisi, *recall* dan F1-score. Klasifikasi menggunakan *naïve bayes* tanpa seleksi fitur memperoleh akurasi sebesar 71,5%, presisi sebesar 71,49%, recall sebesar 71,5% dan F-1 score sebesar 71,49% sedangkan pengklasifikasian sentimen menggunakan algoritma Naive Bayes dan seleksi fitur menggunakan algoritma Genetika memperoleh nilai evaluasi terbaik pada generasi ke-20 dengan akurasi 76,8%, presisi sebesar 75,99%, recall sebesar 76% dan F-1 score sebesar 75,98%. Keaikan akurasi menggunakan seleksi fitur meningkatkan peforma klasifikasi pada data twitter sebesar 6-7%.

Referensi

[1] Yusupa, A., and Tarigan, V. 2024. Perbandingan Algoritma Machine Learning dalam Analisis Sentimen Mobil Listrik di Indonesia pada Media Sosial Twitter/X JIP (Jurnal Informatika Polinema), vol. 10, no. 4, pp. 479–490.

[2] Yanuar Laik, A. A., Nabilla, A., Diah, A., Sumanto, Indra, A., and Arya, Y. 2025. Komparasi Naive Bayes dan SVM untuk Analisis Sentimen Pada E-Commerce Seller Center, Jurnal Sains dan Teknologi (JSIT), vol. 5, no. 3, pp. 199–208.

[3] Setiawan, D., Umar, N., and M. Adnan Nur. 2024. Optimasi Ekstraksi Fitur untuk Meningkatkan Akurasi Naïve Bayes dalam Analisis Sentimen Objek Wisata Bulukumba, Sistemasi: Jurnal Sistem Informasi, vol. 13, no. 5, pp. 2209–2221.Apriani, R., and Gustian, D. 2019. Analisis Sentimen dengan Naïve Bayes Terhadap Komentar Aplikasi Tokopedia, vol. 6, no. 1, pp. 54-62.

[4] Jiang, Z., Gao, B., He, Y., Han, Y., Doyle, P., and Zhu, Q. 2021. Text Classification Using Novel Term Weighting Scheme-Based Improved TF-IDF for Internet Media Reports, Mathematical Problems in Engineering, vol. 2021.

[5] Wulan, S. T., Bettiza, M., and Hayaty, N. 2017. Optimasi Seleksi Fitur Klasifikasi Naïve Bayes Menggunaan Algoritma Genetika Untuk Prediksi Risiko Kredit Konsumen.

[6] Ginting, A., Purba, D., Sinaga, L. M., and Sagala, M. 2025. Analisis Sentimen Masyarakat Twitter terhadap Emas Digital Menggunakan Algoritma Naïve Bayes, KAKIFIKOM (Kumpulan Artikel Karya Ilmiah Fakultas Ilmu Komputer), pp. 8–14, 2025.

[7] Normati, D., and Prayogi S. A. 2021. Implementasi Naïve Bayes Clasifier dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter, vol. 5, no. 2, pp. 697-711.

[8] Rusdiaman, D., and Rosiyadi, D. 2019. Analisa Sentimen Terhadap Tokoh Publik Menggunakan Metode Naïve Bayes Classifier Dan Support Vector Machine. CESS (Journal of Computer Engineering System and Science), vol. 4, no. 2, pp. 230-235.

[9] Rafdi, A., Mawengkang, H., and Efendi, S. 2021. Sentiment Analysis Using Naive Bayes Algorithm with Feature Selection Particle Swarm Optimization (PSO) and Genetic Algorithm, International Journal of Advances in Data and Information Systems, vol. 2, no. 2, pp. 96–104.

Jurnal Elektronik Ilmu Komputer Udayana
Volume 14, No 3. February 2026

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Pengamanan Data menggunakan Algoritma Enkripsi Advanced Encryption Standard (AES) dan Zigzag Cipher

I Gusti Ngurah Bagus Arimbawa [a1], I Made Widiartha [a2] , I Ketut Gede Suhartana [a3] , I Gusti Ngurah Anom Cahyadi Putra. [a4]

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,

Universitas Udayana

Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia

[1]dosman10m5@gmail.com

[2]madewidiartha@unud.ac.id

[3]ikg.suhartana@unud.ac.id

[4]anom.cp@unud.ac.id

Abstrak

Masalah keamanan dan kerahasiaan data merupakan aspek penting dalam komunikasi modern. Kriptografi adalah metode utama untuk mengamankan informasi dari suatu data dengan mengubahnya menjadi data yang tidak dapat dibaca menggunakan suatu algoritma enkripsi. Penelitian ini bertujuan untuk meningkatkan keamanan enkripsi data, khususnya dalam format video, dengan memodifikasi algoritma Advanced Encryption Standard (AES-256) menggunakan penambahan langkah enkripsi Zigzag Cipher. Penelitian dilakukan dengan menguji performa algoritma zigzag cipher, AES-256, AES-256 + zigzag pada input data, AES-256 + zigzag cipher pada output data, dan AES-256 + zigzag cipher pada input password, digunakan untuk memproses 30 file video yang dipotong ke durasi tertentu dari 3 video sumber dan 1 file video dengan tampilan monolitik (1 warna), dengan menggunakan 4 password per algoritma. Pengujian diterapkan menggunakan bahasa pemrograman Python 3.8.10. Hasil penelitian menunjukkan bahwa penambahan Zigzag Cipher pada AES-256 dapat meningkatkan kompleksitas dan keamanan enkripsi. Algoritma zigzag sendiri diketahui lebih cepat sekitar 20 kali dibandingkan AES atau variasinya. Meskipun penambahan zigzag cipher menyebabkan penambahan waktu pemrosesan, peningkatan ini relatif kecil. Algoritma AES Modifikasi 1 (Zigzag pada plaintext AES) secara konsisten mencapai nilai entropy yang lebih tinggi, sementara semua algoritma yang diuji mampu mencapai koefisiensi korelasi yang rendah, mendekati 0.

Kata Kunci: Kriptografi, Enkripsi, Video, Algoritma Zigzag, Algoritma AES, AES-256, Python

1. Pendahuluan

Perkembangan teknologi informasi dan komunikasi telah meningkatkan ancaman terhadap keamanan data, seperti kasus kebocoran data pribadi [1], [2]. Kriptografi menjadi solusi utama untuk melindungi data. Salah satu algoritma enkripsi yang unggul adalah Advanced Encryption Standard (AES) karena fleksibilitas panjang kunci dan kecepatan pemrosesan yang cukup tinggi [1], [3], [4], [5]. Namun, AES juga memiliki kelemahan, seperti ketergantungan pada keamanan penyimpanan kunci dan penggunaan sumber daya komputasi yang lebih besar untuk data yang besar [3], [6]. Untuk mengatasi kelemahan keamanan dan penggunaan suber daya komputasi ini, penelitian sebelumnya telah mengusulkan solusi berupa kombinasi algoritma, seperti

penelitian dari Fardianto dkk. yang mengusulkan kombinasi zigzag ciphere dengan vigenere cipher [7], penelitian Handoko dkk. yang mengkombinasi vigenere cipher dengan AES dan hashing [8], dan peneltian Triansyah yang mengkombinasi vigenere cipher dengan AES [9]. Pendekatan ini menginspirasi penelitian ini, yang mengusulkan penggabungan AES-256 dengan Zigzag Cipher, sebuah algoritma transposisi sederhana yang dapat meningkatkan kompleksitas ciphertext.

Penelitian ini mengembangkan hasil dan metode penelitian sebelumnya yang dilakukan oleh Al Tamimi, Mandal, dan Handoko untuk mengevaluasi performa algoritma-algoritma yang diuji, yaitu dari segi kecepatan pemrosesan data / data throughput [3], [4], dan dari segi keamanan enkripsi berdasarkan analisa statistik entropi dan koefisiensi korelasi [8]. Tujuan penelitian ini adalah untuk mengembangkan algoritma enkripsi dan dekripsi dengan memodifikasi AES-256 menggunakan Zigzag Cipher, dan menguji efektivitas modifikasi algoritma dalam mempengaruhi efisiensi proses serta entropi dan koefisien korelasi hasil enkripsi dan dekripsi [3], [4],[8], [10].

2. Metodologi Penelitian

Penelitian ini memfokuskan pada modifikasi algoritma AES-256 dengan penambahan langkah enkripsi Zigzag Cipher. Perbandingan algoritma yang dimodifikasi hanya dilakukan dengan AES-256, bukan algoritma lain. Pengujian dilakukan dengan simulasi menggunakan data video format .mp4, dan implementasi program dibuat dengan bahasa pemrograman Python. Pengujian melibatkan total 5 algoritma, 43 file plaintext, dan 4 password.

Tabel 1 Daftar File Plainteks

File Sumber	File Di uji
Colortest1.mp4	Colortest1.mp4
	colortest1_black.mp4
	colortest1_white.mp4
	colortest1_red.mp4
	colortest1_green.mp4
	colortest1_blue.mp4
	colortest1_grey.mp4
	colortest1_greyscale_brightness.mp4
	colortest1_greyline_horizontal.mp4
	colortest1_greyline_vertical.mp4
	colortest1_grey_checkerboard.mp4
	colortest1_rgb_brightness.mp4
	colortest1_gradient.mp4
Testvid1.mp4	testvid1_0s_to_2s_2s.mp4
	testvid1_0s_to_4s_4s.mp4
	testvid1_0s_to_6s_6s.mp4
	testvid1_0s_to_8s_8s.mp4
	testvid1_0s_to_10s_10s.mp4
	testvid1_0s_to_15s_15s.mp4
	testvid1_0s_to_20s_20s.mp4
	testvid1_0s_to_25s_25s.mp4
	testvid1_0s_to_30s_30s.mp4
	testvid1_0s_to_60s_60s.mp4
Testvid2.mp4	testvid2_0s_to_2s_2s.mp4
	testvid2_0s_to_4s_4s.mp4
	testvid2_0s_to_6s_6s.mp4
	testvid2_0s_to_8s_8s.mp4
	testvid2_0s_to_10s_10s.mp4
	testvid2_0s_to_15s_15s.mp4
	testvid2_0s_to_20s_20s.mp4

p-ISSN: 2301-5373
e-ISSN: 2654-5101

	testvid2_0s_to_25s_25s.mp4
	testvid2_0s_to_30s_30s.mp4
	testvid2_0s_to_60s_60s.mp4
Testvid3.mp4	testvid3_0s_to_2s_2s.mp4
	testvid3_0s_to_4s_4s.mp4
	testvid3_0s_to_6s_6s.mp4
	testvid3_0s_to_8s_8s.mp4
	testvid3_0s_to_10s_10s.mp4
	testvid3_0s_to_15s_15s.mp4
	testvid3_0s_to_20s_20s.mp4
	testvid3_0s_to_25s_25s.mp4
	testvid3_0s_to_30s_30s.mp4
	testvid3_0s_to_60s_60s.mp4

Tabel 2. Daftar Password

Password
password
passwoad
12345678
IlkomUdayana2025

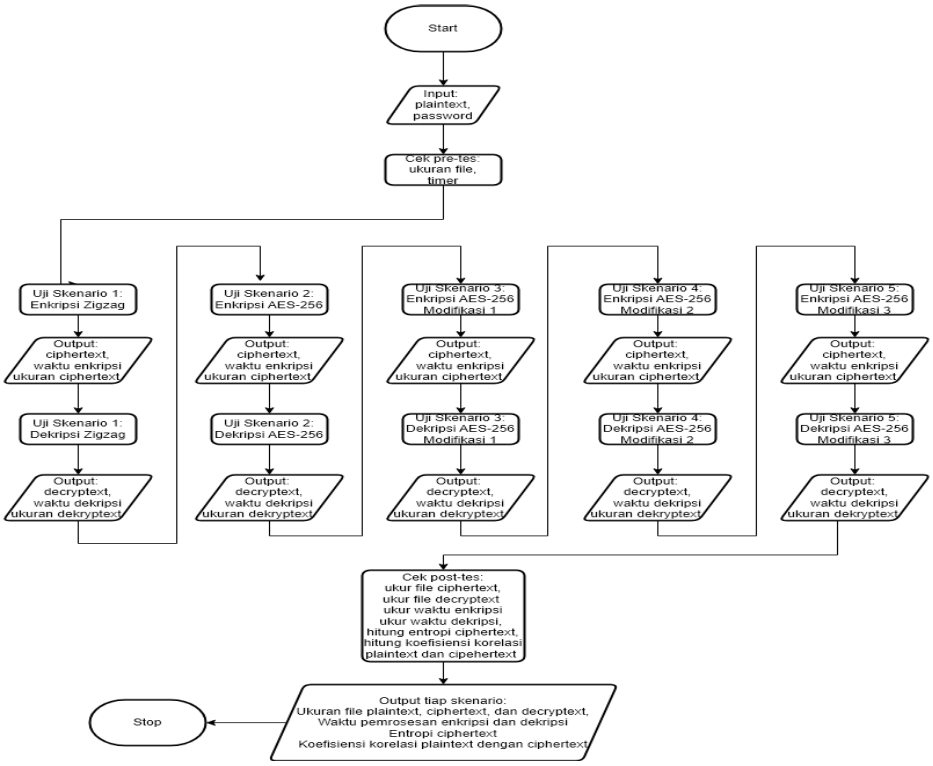
Tabel 3. Daftar Algoritma

Nama Algoritma	Penjelasan
Zigzag	Algoritma zigzag cipher dasar
AES	Algoritma AES-256 dasar
AES Mod 1	Zigzag pada input data AES-256
AES Mod 2	Zigzag pada output data AES-256
AES Mod 3	Zigzag pada input password AES-256

Tabel 4. Data Pengujian

File Plainteks	Password	Jumlah Algoritma	Total Cipherteks
31	4	5	620

Pengujian dilakukan dengan mengikuti langkah berikut:



Gambar 1. Flowchart proses pengujian

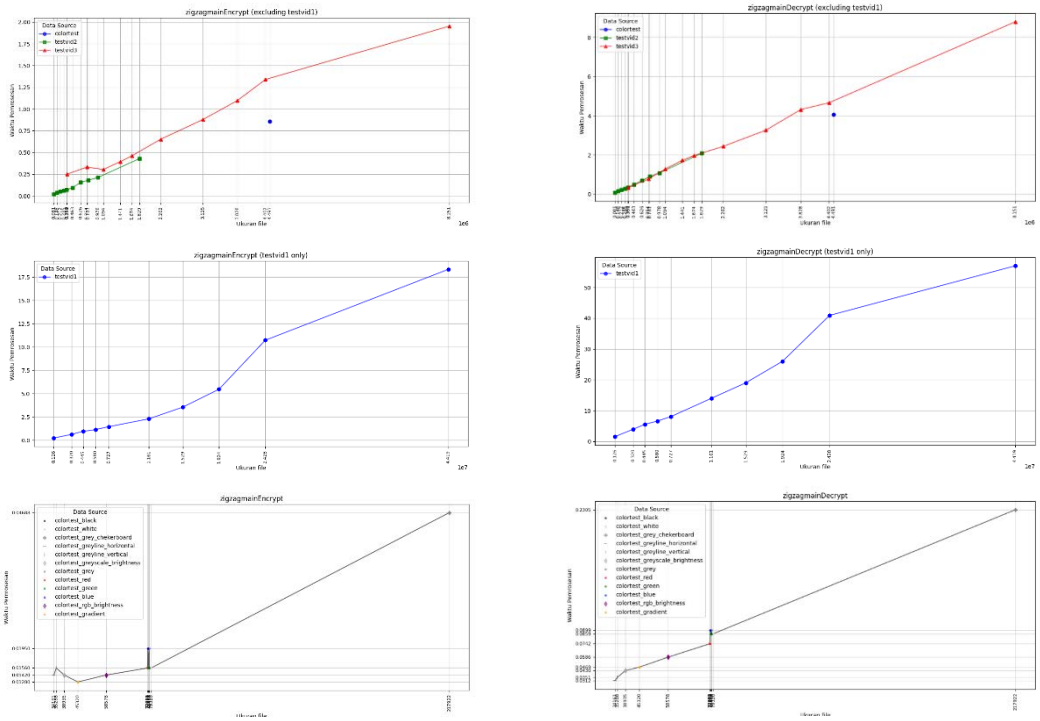
Berdasarkan flowchart di gambar 1 di atas, proses pengujian akan dilakukan dengan cara meng-iterasi setiap skenario pengujian terhadap data input. Setiap pasangan input plaintext dan password akan digunakan untuk menguji 5 skenario enkripsi dan dekripsi. Hasil setiap skenario yang disimpan meliputi ukuran file plaintext (sebelum diproses), ciphertext, dan decryptext (plaintext hasil dekripsi), waktu pemrosesan enkripsi dan dekripsi, pemakaian memory dari pemrosesan enkripsi dan dekripsi, nilai entropi ciphertext dan plaintext, nilai koefisiensi korelasi ciphertext dan plaintext.

3. Hasil dan Pembahasan

Berikut adalah graf hasil pengujian. Data dalam graf adalah nilai dari tiap titik data video yang diproses, dikelompokkan berdasarkan algoritmanya, dan di rata-ratakan berdasarkan passwordnya. Karena rentang datanya sangat berbeda dari rentag data sumber-sumber lainnya, data dari sumber testvid1 dan colortest dipisahkan.

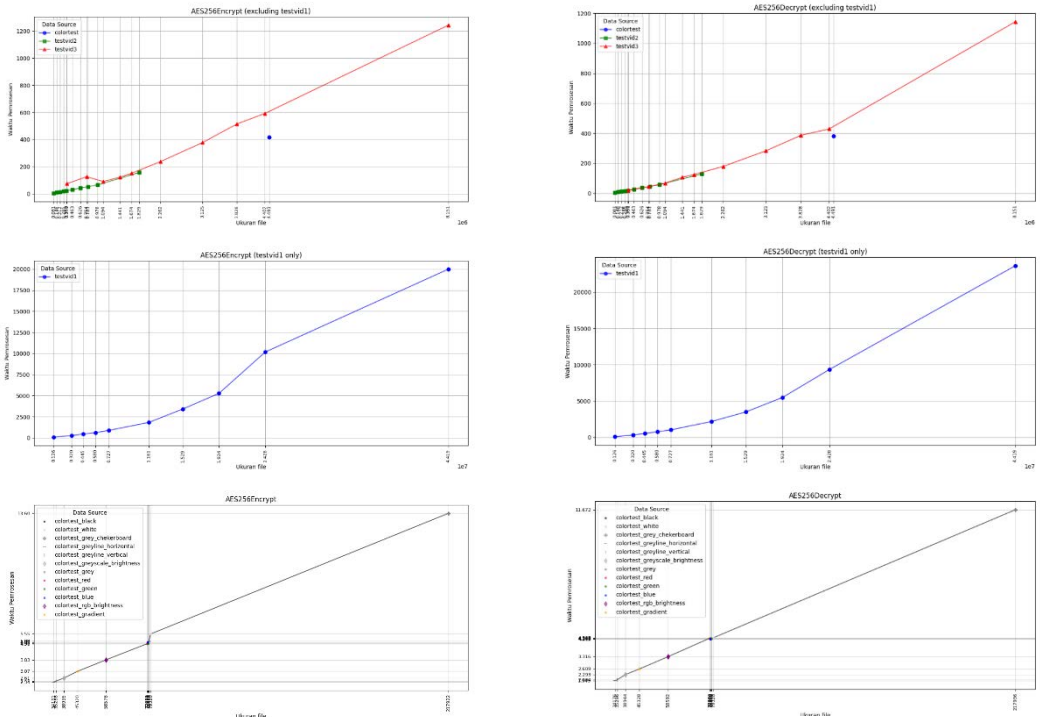
3.1 Pemrosesan Enkripsi dan Dekripsi

p-ISSN: 2301-5373
e-ISSN: 2654-5101



Gambar 2 Graf Waktu Pemrosesan zigzag

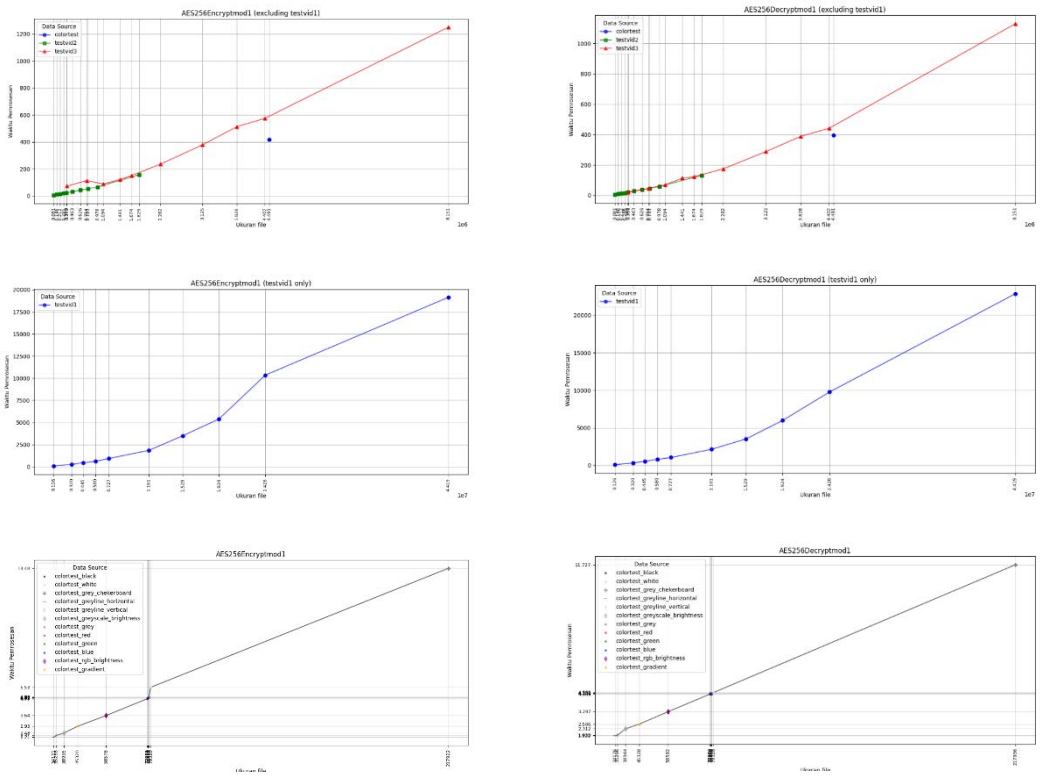
Dapat dilihat pada gambar 2 bahwa peningkatan waktu pemrosesan cenderung konsisten untuk ukuran file berbeda. Waktu pemrosesan meningkat secara linier seiring dengan meningkatnya ukuran file. Untuk video – video dari testvid1, peningkatan waktu pemrosesan sedikit melambat diantara klip 10 detik dan klip 30 detik, atau antara file berukuran 10 megabyte dan 25 megabyte, sebelum kembali ke pola linier sebelumnya.



Gambar 3 Graf Waktu Pemrosesan AES

Dapat dilihat dari gambar 3 diatas, bahwa Algoritma AES bahwa perubahan waktu pemrosesan cenderung konsisten untuk ukuran file berbeda. Waktu pemrosesan meningkat secara linier seiring dengan meningkatnya ukuran file. Untuk klip dari testvid1, peningkatan waktu pemrosesan sedikit melambat diantara klip 10 detik dan klip 30 detik, atau antara file berukuran 10 megabyte dan 25 megabyte, sebelum kembali ke pola penigkatan sebelumnya yang lebih stabil.

Kecepatan pemrosesan mulai rendah, lalu meningkat dan memuncak pada ukuran file 0,9 – 1,2 megabyte, kemudian mulai menurun. Untuk video klip yang mulai dengan ukuran file di atas 1 megabyte, kecepatan pemrosesan nampak menurun seiring dengan penigkatan ukuran file.

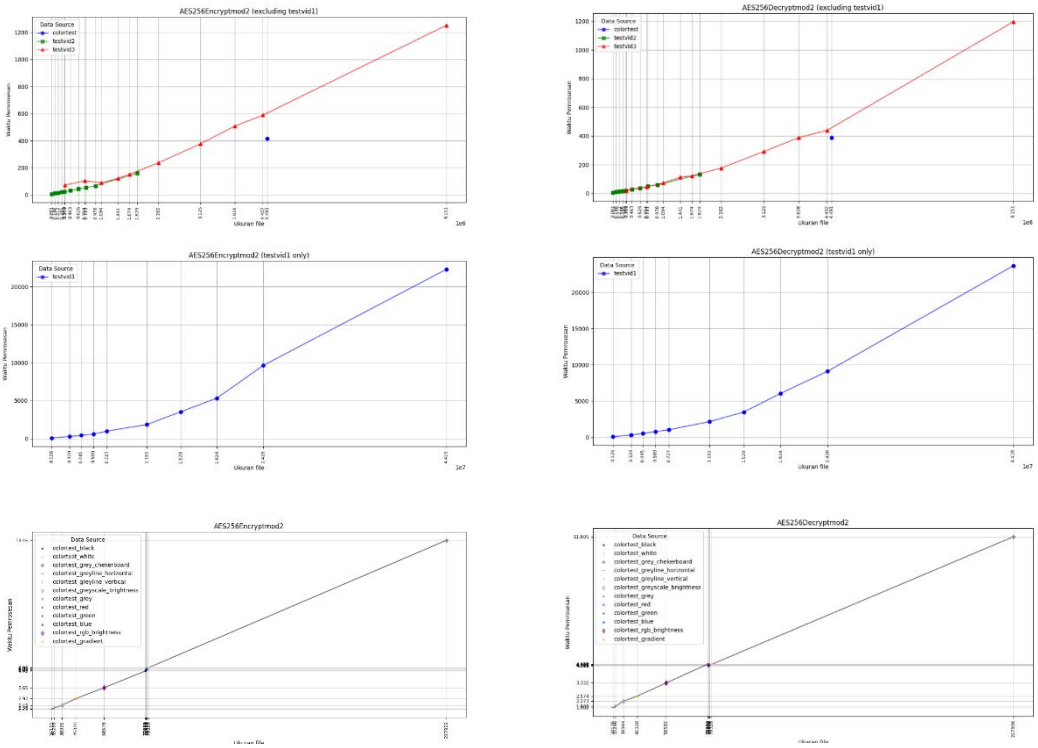


Gambar 4 Graf Waktu Pemrosesan AES Mod 1

Dapat dilihat dari Gambar 4 diatas, bahwa waktu pemrosesan cenderung konsisten untuk ukuran file berbeda. Waktu pemrosesan meningkat secara linier seiring dengan meningkatnya ukuran file. Untuk klip dari testvid1, peningkatan waktu pemrosesan sedikit melambat diantara klip 10 detik dan klip 30 detik, atau antara file berukuran 10 megabyte dan 25 megabyte, sebelum kembali ke pola penigkatan sebelumnya yang lebih stabil.

Kecepatan pemrosesan mulai rendah, lalu meningkat dan memuncak pada ukuran file 0,9 – 1,2 megabyte, kemudian mulai menurun. Untuk video klip yang mulai dengan ukuran file di atas 1 megabyte, kecepatan pemrosesan nampak menurun seiring dengan penigkatan ukuran file.

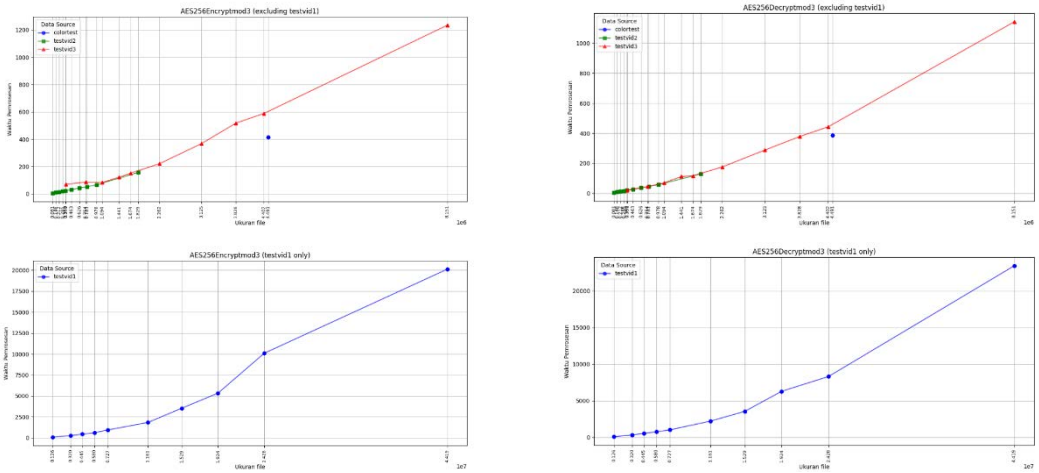
p-ISSN: 2301-5373
e-ISSN: 2654-5101

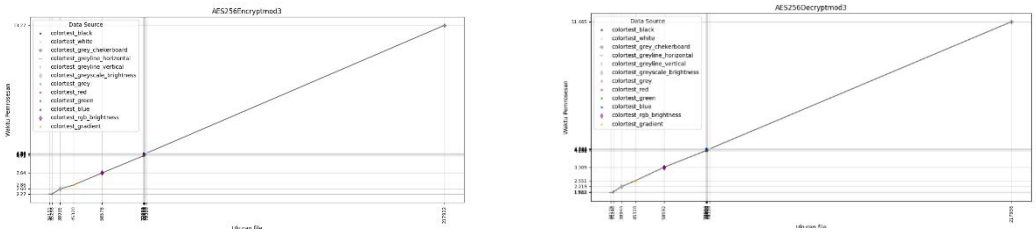


Gambar 5 Graf Waktu Pemrosesan AES Mod 2

Dapat dilihat pada Gambar 5 diatas bahwa waktu pemrosesan cenderung konsisten untuk ukuran file berbeda. Waktu pemrosesan meningkat secara linier seiring dengan meningkatnya ukuran file. Untuk klip dari testvid1, peningkatan waktu pemrosesan sedikit melambat diantara klip 10 detik dan klip 30 detik, atau antara file berukuran 10 megabyte dan 25 megabyte, sebelum kembali ke pola penigkatan sebelumnya yang lebih stabil.

Kecepatan pemrosesan mulai rendah, lalu meningkat dan memuncak pada ukuran file 0,9 – 1,2 megabyte, kemudian mulai menurun. Untuk video klip yang mulai dengan ukuran file di atas 1 megabyte, kecepatan pemrosesan nampak menurun seiring dengan penigkatan ukuran file.





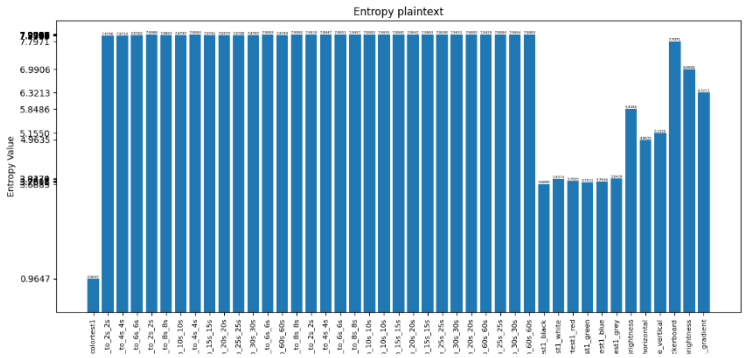
Gambar 6 Graf Waktu Pemrosesan AES Mod 3

Dapat dilihat pada gambar 6 diatas bahwa waktu pemrosesan cenderung konsisten untuk ukuran file berbeda. Waktu pemrosesan meningkat secara linier seiring dengan meningkatnya ukuran file. Untuk klip dari testvid1, peningkatan waktu pemrosesan sedikit melambat diantara klip 10 detik dan klip 30 detik, atau antara file berukuran 10 megabyte dan 25 megabyte, sebelum kembali ke pola penigkatan sebelumnya yang lebih stabil.

Kecepatan pemrosesan mulai rendah, lalu meningkat dan memuncak pada ukuran file 0,9 – 1,2 megabyte, kemudian mulai menurun. Untuk video klip yang mulai dengan ukuran file di atas 1 megabyte, kecepatan pemrosesan nampak menurun seiring dengan penigkatan ukuran file.

3.2 Entropi.

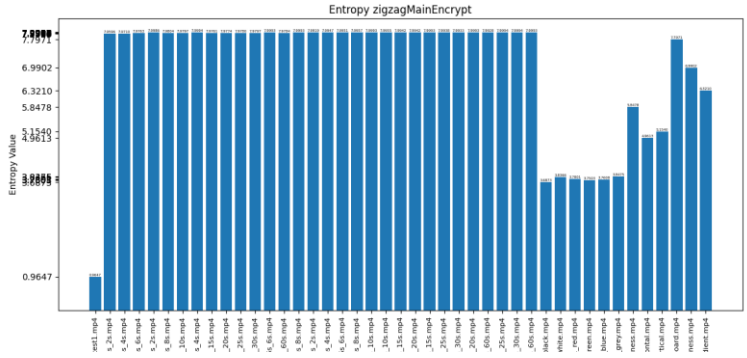
Berikut adalah graf data entropi dari semua plainteks dan cipherteks.



Gambar 7 Entropi Plainteks

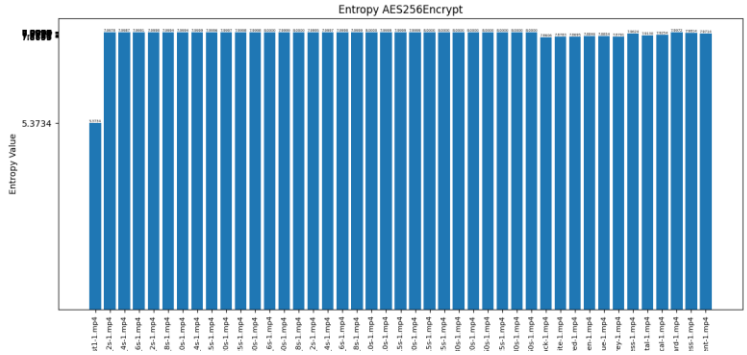
Sebagaimana dapat pada gambar 7, nilai entropi plainteks bervariasi sesuai dengan kompleksitas vide plainteks yang digunakan. Colortest dan variasinya memiliki nilai entropi yang cenderung lebih rendah (>7), sedangakan testvid dan variasinya memiliki entropi yang dekat dengan maksimal (7,99)

p-ISSN: 2301-5373
e-ISSN: 2654-5101



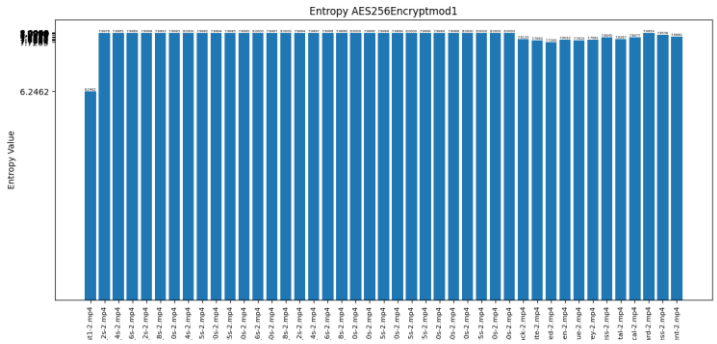
Gambar 8 Entropi Zigzag

Sebagaimana dapat dilihat pada gambar 8, nilai entropi cipherteks hasil enkripsi zigzag adalah sama dengan entropi plainteks.



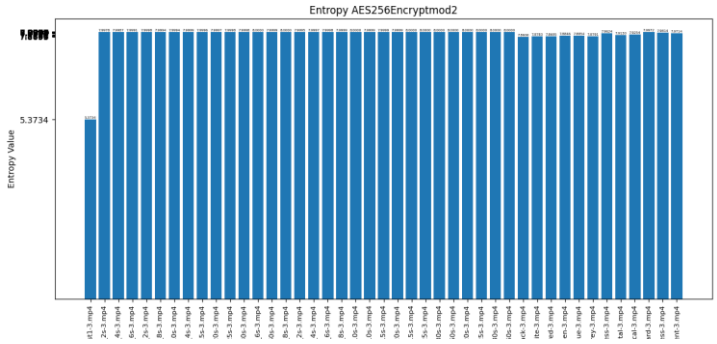
Gambar 9 Entropi AES

Dapat dilihat pada gambar 9, bahwa nilai entropi semua cipherteks meningkat mendekati nilai maksimal, kecuali cipherteks colortest1. Untuk colortest1, nilai entropinya menigkat secara signifikan dari 0,96 menjadi 5,37.



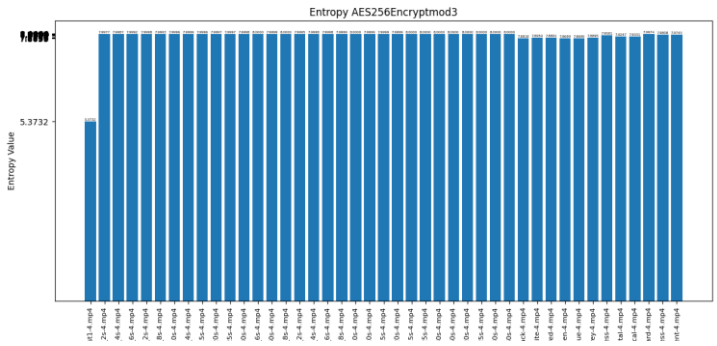
Gambar 10 Entropi AES Mod 1

Dapat dilihat bahwa pada gambar 10, nilai entropi semua cipherteks meningkat mendekati nilai maksimal, kecuali cipherteks colortest1. Untuk colortest1, nilai entropinya menigkat secara signifikan dari 0,96 menjadi 6,24.



Gambar 11 Entropi AES Mod 2

Dapat dilihat pada gambar 11, bahwa nilai entropi semua cipherteks meningkat mendekati nilai maksimal, kecuali cipherteks colortest1. Untuk colortest1, nilai entropinya menigkat secara signifikan dari 0,96 menjadi 5,37.



Gambar 12 Entropi AES Mod 3

Dapat dilihat pada gambar 12, bahwa nilai entropi semua cipherteks meningkat mendekati nilai maksimal, kecuali cipherteks colortest1. Untuk colortest1, nilai entropinya menigkat secara signifikan dari 0,96 menjadi 5,37.

3.3. Koefisien Korelasi.

Berikut adalah tabel rekapitulasi rata rata nilai koefisiensi dari semua sumber dan algoritma.

Tabel 5 Rerata Koefisien Korelasi per Sumber dan Algoritma

Sumber file	Zigzag	AES	AESMod1	AESMod2	AESMod3
Colortest	0,02039739	-0,00640786	0,00757968	0,00016432	0,00444395
Testvid1	0,00020646	-0,00015818	-0,00002606	-0,00003165	0,00004864
Testvid2	0,00117669	0,00114971	0,00026359	-0,00068181	-0,00004546
Testvid3	0,00159160	0,00025032	-0,00032084	-0,00009882	0,00028400
Rata rata	0,00130971	-0,00315938	0,00343016	-0,00007300	0,00207219

Dari tabel diatas, dapat dilihat bahwa AESMod2 mengahsilkan cipherteks dengan nilai korelasi paling dekat dengan nol, artinya paling jauh korelasinya dengan plainteks.

p-ISSN: 2301-5373
e-ISSN: 2654-5101

3.4. Pembahasan

Hasil penelitian yang sudah dilakukan menunjukkan bahwa:

a. Waktu pemrosesan file meningkat secara linier seiring dengan meningkatnya ukuran file.

b. Zigzag Cipher memiliki waktu pemrosesan yang jauh lebih cepat (sekitar 20 kali) dibandingkan AES dan variasinya. Penambahan proses enkripsi Zigzag pada AES, dari AES Mod 1, hingga Mod 3, hanya menambahkan waktu pemrosesan yang relatif kecil (± 0,05% – 2%).

c. Algoritma Zigzag tidak mengubah distribusi frekuensi nilai byte (nilai Shannon Entropy sama dengan plaintext), namun AES dapat menghasilkan nilai entropy lebih tinggi dibandingkan plainteks.

d. Semua algoritma yang diuji berhasil mengubah korelasi antara plaintext dan ciphertext mendekati nilai 0, hingga tiga digit pecahan desimal (0,00x).

e. Algoritma AES Mod 1 adalah algoritma yang menghasilkan peningkatan entropi terbesar (rata rata +1,18%, peningkatan terbesar +67,01%), dan dapat menghasilkan koefisien korelasi dan waktu pemrosesan yang setara dengan algoritma AES dan Modifikasi AES lainnya.

4. Kesimpulan

Penelitian ini menunjukkan bahwa salah satu metode modifikasi algoritma AES yang diusulkan, yaitu AES Mod 1, berhasil meningkatkan nilai Shannon Entropy dari cipherteks dibandingkan algoritma AES dasar, dengan rata rata peningkatan sebesar 1,18% lebih besar dari entropi AES dasar. Semua algoritma AES dan modifikasinya memiliki waktu pemrosesan dan koefisien korelasi yang mirip (variasi waktu pemrosesan ± 0,05% – 2%, dan koefisien korelasi dibawah tiga digit pecahan desimal (0,00x))

Untuk penelitian selanjutnya, disarankan untuk melakukan penelitian lebih dalam mengenai integrasi algoritma AES dan Zigazg, serta meneliti kombinasi algoritma-algoritma lainnya. Penelitian berikutnya juga dapat menambah jumlah dan jenis data yang diuji, tidak hanya video, dan menggunakan metode atau algoritma-algoritma khusus untuk memproses jenis data tertentu. Penelitian berikutnya juga dapat meneliti dampak modifikasi algoritma dengan cara menggabungkan dua atau lebih algortima yang sudah ada dengan menggunakan metode evaluasi yang lebih lengkap, tidak hanya terbatas pada entropi dan koefisien korelasi.

References

[1] Almadira, A., Pratama, Y., & Purwani, F. (2024). Melindungi data di dunia digital: Peran strategis enkripsi dalam keamanan data. Journal of Scientech Research and Development, 6(2), 540–549.
[2] Cyberlands. (2022). Top security breaches in Indonesia. Diakses dari https://www.cyberlands.io
[3] Al Tamimi, A. K. (2006). Performance analysis of data encryption algorithms (Master's thesis). Washington University in St. Louis.
[4] Mandal, P. C. (2012). Evaluation of performance of the symmetric key algorithms: DES, 3DES, AES and Blowfish. Journal of Global Research in Computer Science (JGRCS), 3(8), 67–70.
[5] Sharma, N., Prabhjot, & Kaur, H. (2017). A review of information security using cryptography technique. International Journal of Advanced Research in Computer Science, 8(4), 323–326.
[6] Wiharto, Y., & Irawan, A. (2018). Enkripsi data menggunakan Advanced Encryption Standard 256. Jurnal Kilat, 7(2), 91–98.

[7] Fardianto, F. A. E., Yanto, F., Iskandar, I., & Pizaini. (2023). Kombinasi algoritma kriptografi Vigenere Cipher dengan metode Zig-Zag dalam pengamanan pesan teks. Jurnal Computer Science and Information Technology (CoSciTech), 4(1), 182–192.
[8] Handoko, L. B., & Umam, C. (2022). Kombinasi Vigenere-AES256 dan fungsi hash dalam kriptografi aplikasi chatting. Prosiding Seminar Nasional Sains dan Teknologi, 12(1), 390–397.
[9] Triansyah, H., Pratama, A., Syahputra, F., & Gunawan, I. (2019). Kombinasi kriptografi algoritma Vigenere Cipher dan algoritma AES untuk pengamanan pesan teks. TECHSI, 11(3), 409–418.
[10] Wahyudi, E. N., Ardhianto, E., Handoko, W. T., Murti, H., Supriyanto, E., Lestariningsih, E., & Redjeki, R. S. (2024). Peningkatan keamanan data melalui teknik super enkripsi menggunakan algoritma Vigenere dan Caesar. Jurnal Informatika Polinema, 10(3), 315.



ISSN



E-ISSN