

TESIS
ANALISIS KINERJA MODEL SUPPORT VECTOR MACHINE
DAN K-NEAREST NEIGHBORS DENGAN EKSTRAKSI
FITUR WINNOWING PADA SISTEM AUTOMATED ESSAY
SCORING



disusun oleh

MUHAMMAD FAHMI LUTHFI

22.55.2296

Konsentrasi : Business Intelligence

FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2026

TESIS
ANALISIS KINERJA MODEL SUPPORT VECTOR MACHINE
DAN K-NEAREST NEIGHBORS DENGAN EKSTRAKSI
FITUR WINNOWING PADA SISTEM AUTOMATED ESSAY
SCORING

PERFORMANCE ANALYSIS OF SUPPORT VECTOR
MACHINE AND K-NEAREST NEIGHBORS MODELS WITH
WINNOWING FEATURE EXTRACTION IN AUTOMATED
ESSAY SCORING

Diajukan untuk memenuhi salah satu syarat mencapai derajat Pascasarjana
Program Studi S2 Teknik Informatika



disusun oleh

MUHAMMAD FAHMI LUTHFI

22.55.2296

Konsentrasi : Business Intelligence

FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2026

HALAMAN PERSETUJUAN

ANALISIS KINERJA MODEL SUPPORT VECTOR MACHINE DAN K-NEAREST NEIGHBORS DENGAN EKSTRAKSI FITUR WINNOWING PADA SISTEM AUTOMATED ESSAY SCORING

PERFORMANCE ANALYSIS OF SUPPORT VECTOR MACHINE AND K-NEAREST NEIGHBORS MODELS WITH WINNOWING FEATURE EXTRACTION IN AUTOMATED ESSAY SCORING

yang disusun dan diajukan oleh

Muhammad Fahmi Luthfi
22.55.2296

telah disetujui oleh Dosen Pembimbing Tesis
pada tanggal 03 Februari 2026

Dosen Pembimbing,



Prof. Dr. Kusriani, M.Kom.
NIK. 190302106

HALAMAN PENGESAHAN

ANALISIS KINERJA MODEL SUPPORT VECTOR MACHINE DAN K-NEAREST NEIGHBORS DENGAN EKSTRAKSI FITUR WINNOWERING PADA SISTEM AUTOMATED ESSAY SCORING

PERFORMANCE ANALYSIS OF SUPPORT VECTOR MACHINE AND K-NEAREST NEIGHBORS MODELS WITH WINNOWERING FEATURE EXTRACTION IN AUTOMATED ESSAY SCORING

yang disusun dan diajukan oleh

Muhammad Fahmi Luthfi

22.55.2296

Telah dipertahankan di depan Dewan Penguji
pada tanggal 03 Februari 2026

Susunan Dewan Penguji

Nama Penguji

Dr. Ferry Wahyu Wibowo, S.SI, M.Cs.

NIK. 190302235

I Made Artha Agastya, S.T., M.Eng., Ph.D.

NIK. 190302352

Prof. Dr. Kusriani, S.Kom., M.Kom.

NIK. 190302106

Tanda Tangan



Tesis ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Magister Komputer
Tanggal 03 Februari 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.

NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN TESIS

Yang bertandatangan di bawah ini,

Nama mahasiswa : **Muhammad Fahmi Luthfi**
NIM : **22.55.2296**

Menyatakan bahwa Tesis dengan judul berikut:

Analisis Kinerja Model Support Vector Machine Dan K-Nearest Neighbors Dengan Ekstraksi Fitur Winnowing Pada Sistem Automated Essay Scoring

Dosen Pembimbing : **Prof. Dr. Kusriani, M.Kom.**

1. Karya tulis ini adalah benar-benar **ASLI** dan **BELUM PERNAH** diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian **SAYA** sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab **SAYA**, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini **SAYA** buat dengan sesungguhnya, apabila di kemudian hari terdapat **penyimpangan** dan **ketidakbenaran** dalam pernyataan ini, maka **SAYA** bersedia menerima **SANKSI AKADEMIK** dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 03 Februari 2026

Yang Menyatakan,



Muhammad Fahmi Luthfi

HALAMAN PERSEMBAHAN

Segala puji bagi Allah SWT atas limpahan rahmat, taufik, dan hidayah-Nya sehingga penulis dapat menyelesaikan tesis ini.

Dengan penuh rasa syukur dan hormat, karya sederhana ini penulis persembahkan kepada:

- Ayahanda KH. Zainuddin Rusydan, S.Ag. dan Ibunda Hj. Sri Kholidah, S.Pd.I tercinta, atas segala doa, kasih sayang, dan dukungan yang tiada henti.
- Adik-adikku M. Ghozali Adnan dan M. Nawawi Rusydan, yang selalu memberi semangat dan keceriaan.
- Istri tercinta Naili Fauziyah dan anakku M. Abdullah Ibnu Mubarak, sumber kebahagiaan, kekuatan dan inspirasi dalam setiap langkah.
- Keluarga besar Alm. H. Dja'far, atas teladan dan doa yang senantiasa menyertai.
- SMK NU Al Hidayah Kudus, sebagai tempat pengabdian dan pengembangan ilmu yang memberi makna tersendiri dalam perjalanan akademik penulis.
- Prof. Dr. Kusriani, M.Kom., selaku dosen pembimbing yang dengan sabar membimbing, mengarahkan, dan memberikan ilmu yang begitu berharga.
- Rekan sejawat, khususnya Gustian Herlambang, yang bersama-sama berjuang menyelesaikan tugas akhir ini.

Semoga karya ini menjadi amal yang bermanfaat dan memberikan kontribusi bagi pengembangan ilmu pengetahuan serta kemajuan pendidikan.

KATA PENGANTAR

Puji syukur kehadirat Allah SWT yang telah melimpahkan rahmat, taufik, dan hidayah-Nya sehingga penulis dapat menyelesaikan tesis yang berjudul “Analisis Kinerja Model Support Vector Machine Dan K-Nearest Neighbors Dengan Ekstraksi Fitur Winnowing Pada Sistem Automated Essay Scoring” dengan baik.

Penyusunan tesis ini dimaksudkan sebagai salah satu syarat untuk memperoleh gelar Magister pada Program Pascasarjana PJJ Informatika. Dalam proses penyusunan tesis ini, penulis mendapatkan bimbingan, dukungan, dan bantuan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis menyampaikan ucapan terima kasih dan penghargaan yang setulus-tulusnya kepada:

1. Prof. Dr. Kusriani, M.Kom., selaku dosen pembimbing yang telah memberikan bimbingan, arahan, dan motivasi dengan penuh kesabaran serta keikhlasan selama proses penyusunan tesis ini.
2. Tim Dosen Penguji, yang telah memberikan kritik, saran, dan masukan yang sangat berharga dalam penyempurnaan karya ini.
3. Ayahanda KH. Zainuddin Rusydan, S.Ag. dan Ibunda Hj. Sri Kholidah, S.Pd.I., atas doa, dukungan, dan kasih sayang yang tiada terhingga.
4. Istri tercinta, Naili Fauziyah, serta anakku, M. Abdullah Ibnul Mubarak, yang menjadi sumber semangat dan inspirasi dalam setiap langkah perjuangan.
5. Adik-adikku, M. Ghozali Adnan dan M. Nawawi Rusydan, atas dorongan dan doa yang senantiasa menyertai.
6. Rekan sejawat Gustian Herlambang, yang bersama-sama berjuang menyelesaikan tesis dengan semangat dan saling mendukung.
7. Seluruh keluarga besar SMK NU Al Hidayah Kudus, atas kesempatan, dukungan, dan fasilitas yang telah diberikan dalam pelaksanaan penelitian ini.

Penulis menyadari bahwa tesis ini masih jauh dari sempurna. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan untuk perbaikan di masa mendatang. Semoga tesis ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan dan dunia pendidikan, khususnya dalam bidang sistem penilaian otomatis berbasis komputer.

Yogyakarta, 03 Februari 2026

Penulis

DAFTAR ISI

HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERNYATAAN KEASLIAN TESIS	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR TABEL.....	ix
DAFTAR GAMBAR.....	x
DAFTAR LAMPIRAN.....	xi
DAFTAR LAMBANG DAN SINGKATAN.....	xii
DAFTAR ISTILAH.....	xiii
INTISARI.....	xiv
ABSTRACT.....	xv
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Batasan Masalah	4
1.4 Tujuan Penelitian.....	5
1.5 Manfaat Penelitian	5
BAB 2 TINJAUAN PUSTAKA	6
2.1 Tinjauan Pustaka.....	6
2.2 Keaslian Penelitian.....	10
2.3 Landasan Teori.....	16
BAB 3 METODE PENELITIAN.....	28
3.1 Jenis, Sifat, dan Pendekatan Penelitian.....	28
3.2 Metode Pengumpulan Data.....	28

3.3	Metode Analisis Data	29
3.4	Alur Penelitian	30
BAB 4	HASIL PENELITIAN DAN PEMBAHASAN	35
4.1	Deskripsi Data dan Preprocessing	35
4.2	Ekstraksi Fitur Menggunakan Winnowing	44
4.3	Penerapan Algoritma Machine Learning	50
4.4	Hasil Evaluasi Performa Model	57
4.5	Analisis Hasil	59
BAB 5	PENUTUP	66
5.1	Kesimpulan	66
5.2	Saran	68
DAFTAR PUSTAKA		69
LAMPIRAN		72



DAFTAR TABEL

Table 2.1 Matriks literatur review dan posisi penelitian.....	10
Table 4.1 Data Set	36
Table 4.2 Hasil sebelum dan sesudah preprosesing	39
Table 4.3 Alur Input dan Output Setiap Tahapan Proses	41
Table 4.4 Contoh Perubahan Data dari Input ke Output	42
Table 4.5 Hasil Ekstraksi n-gram	46
Table 4.6 Hasil Fingerprint	50
Table 4.7 Perbandingan Model	59
Table 4.8 Perbandingan Performa Model pada Berbagai Skenario	61



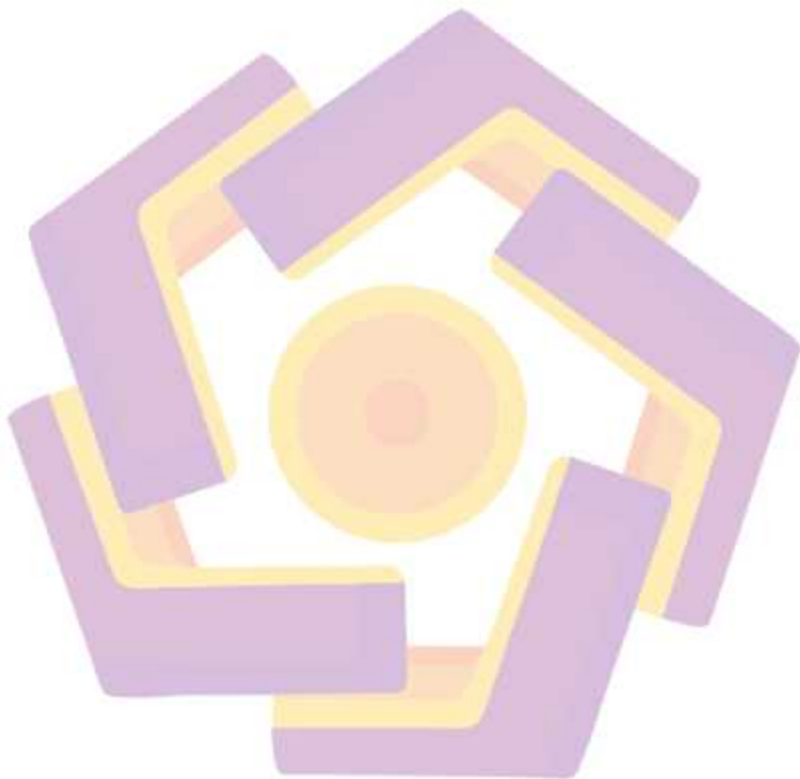
DAFTAR GAMBAR

Gambar 2.1 Konsep Algoritma WInnowing	20
Gambar 3.1 Alur Penelitian	31
Gambar 4.1 Teknik N-Gram	44
Gambar 4.2 Preprocessing Teks.....	47
Gambar 4.3 Seleksi Fingerprint	49
Gambar 4.4 Preprocessing, N-gram, Hashing, dan WInnowing	52
Gambar 4.5 KNN Training dan Evaluasi.....	53
Gambar 4.6 Preprocessing, N-gram, Hashing, dan WInnowing	55
Gambar 4.7 SVM Training dan Evaluasi.....	56
Gambar 4.8 Confusion Matrix	59



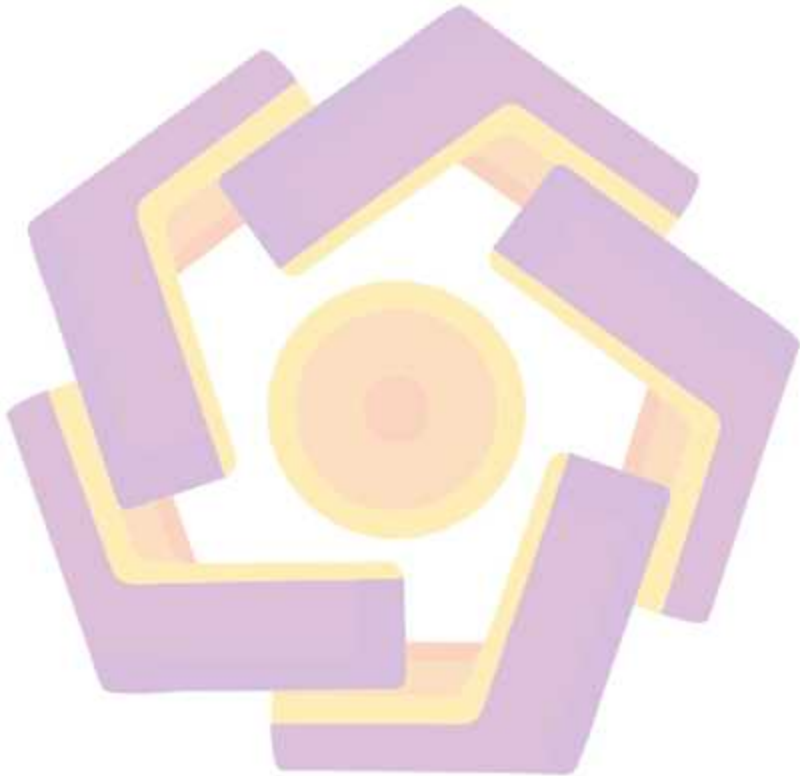
DAFTAR LAMPIRAN

Lampiran 1 Dataset.....	72
Lampiran 2 Source Code.....	72



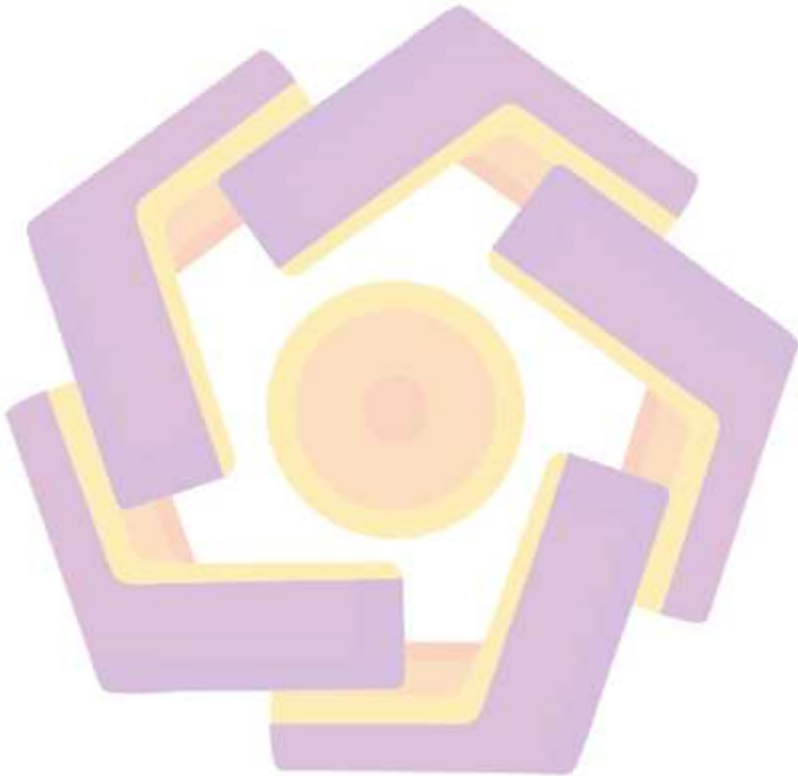
DAFTAR LAMBANG DAN SINGKATAN

SVM	Support Vector Machines
KNN	K-Nearest Neighbors
AES	Automated Essay Scoring
ASAP	Automated Student Assessment Prize
CBT	Computer Based Testing



DAFTAR ISTILAH

Vektor besaran yang mempunyai arah
Eigen Value akar akar persamaan
Essay (dalam penelitian ini): Jawaban tertulis siswa terhadap suatu prompt atau pertanyaan ujian dalam konteks sistem pendidikan Amerika Serikat, sebagaimana digunakan dalam dataset Automated Student Assessment Prize (ASAP).



INTISARI

Penelitian ini berjudul “Analisis Kinerja Model Support Vector Machine Dan K-Nearest Neighbors Dengan Ekstraksi Fitur Winnowing Pada Sistem Penilaian Esai Otomatis” yang bertujuan untuk mengukur dan membandingkan performa kombinasi metode ekstraksi fitur Winnowing dengan algoritma klasifikasi Support Vector Machine (SVM) dan K-Nearest Neighbors (KNN) dalam penilaian esai otomatis berbasis teks. Penelitian ini menggunakan pendekatan kuantitatif dengan jenis penelitian eksperimental. Dataset yang digunakan adalah Automated Student Assessment Prize (ASAP) Automated Essay Scoring (AES) dari Kaggle yang terdiri atas 12.976 jawaban tertulis siswa berbahasa Inggris terhadap prompt ujian, yang dalam konteks sistem pendidikan Amerika Serikat disebut sebagai essay. Tahapan penelitian meliputi preprocessing teks (pembersihan, tokenisasi, dan stemming), ekstraksi fitur menggunakan Winnowing dengan teknik k-gram dan fingerprinting, serta klasifikasi menggunakan algoritma SVM dan KNN. Analisis kinerja dilakukan berdasarkan metrik akurasi, presisi, recall, F1-Score, dan waktu komputasi.

Hasil penelitian menunjukkan bahwa kombinasi SVM dengan fitur hasil ekstraksi Winnowing menghasilkan performa yang lebih baik dibandingkan KNN dengan fitur hasil ekstraksi Winnowing, dengan akurasi sebesar 87,6%, F1-Score 0,868, dan precision 0,891. Meskipun SVM memerlukan waktu komputasi lebih lama (5,18 detik) dibandingkan KNN (0,07 detik), model SVM terbukti lebih akurat dan stabil dalam mengklasifikasikan skor esai.

Kesimpulannya, algoritma SVM lebih direkomendasikan untuk sistem penilaian esai otomatis yang mengutamakan akurasi, sementara KNN cocok untuk aplikasi yang membutuhkan efisiensi waktu. Penelitian selanjutnya disarankan untuk mengeksplorasi metode ekstraksi fitur lain dan optimasi parameter model guna meningkatkan performa sistem secara keseluruhan.

Kata kunci: Winnowing; K-Nearest Neighbors; Support Vector Machine; Feature Extraction; Automated Essay Grading.

ABSTRACT

This research aims to evaluate the application of Winnowing feature extraction technique combined with K-Nearest Neighbors (KNN) and Support Vector Machine (SVM) machine learning models in a text-based automated essay grading system. This topic is important because automated scoring can speed up the essay evaluation process and reduce the potential for human bias. In this study, essays—defined as student-written responses to assessment prompts in the U.S. education context—were classified using N-grams and Winnowing fingerprinting techniques to improve accuracy and feature representation of the text. The evaluation results showed that the SVM dengan fitur hasil ekstraksi Winnowing model produced the most superior performance with an accuracy of 0.876 and F1-score of 0.868430. Despite requiring higher computation time (5.181894 seconds), SVM consistently classified the essays with better precision and stability. In contrast, the KNN dengan fitur hasil ekstraksi Winnowing model showed lower performance with an accuracy of 0.707 and F1-score of 0.686699, but had the advantage of much faster computation time (0.077723 seconds). The use of Winnowing proved effective in improving the performance of both models, with a more significant impact on SVM. In conclusion, SVM is more recommended for automated essay scoring systems that prioritize accuracy, while KNN is more suitable for applications that prioritize time efficiency.

Keyword: Winnowing; K-Nearest Neighbors; Support Vector Machine; Feature Extraction; Automated Essay Grading.

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi dan komunikasi di era digital saat ini telah membawa perubahan besar dalam berbagai bidang kehidupan, termasuk pendidikan. Salah satu inovasi yang banyak diadopsi adalah penerapan Computer Based Testing (CBT), yaitu sistem evaluasi berbasis komputer yang menggantikan metode ujian konvensional berbasis kertas. CBT memiliki sejumlah keunggulan, antara lain efisiensi waktu, penghematan biaya operasional, serta kemampuan untuk memberikan umpan balik secara cepat dan akurat kepada peserta didik. Penerapan CBT juga memungkinkan pelaksanaan evaluasi dalam skala besar dengan tingkat akurasi yang lebih tinggi.

Meskipun demikian, salah satu tantangan utama dalam implementasi CBT terletak pada penilaian otomatis untuk soal berbentuk esai. Berbeda dengan soal pilihan ganda yang dapat dinilai secara langsung berdasarkan kunci jawaban. Dalam penelitian ini, istilah esai merujuk pada jawaban tertulis peserta didik terhadap suatu prompt atau pertanyaan ujian, sebagaimana digunakan dalam dataset Automated Student Assessment Prize (ASAP) yang berasal dari konteks pendidikan di Amerika Serikat, dan bukan esai dalam pengertian karangan bebas sebagaimana umum dipahami dalam sistem pendidikan Indonesia. Oleh karena itu, dibutuhkan pendekatan komputasional yang mampu menilai esai secara objektif, konsisten, dan efisien (Pradeep & Kowsalya, 2022).

Salah satu pendekatan yang digunakan dalam sistem penilaian esai otomatis adalah algoritma WInnowing, yaitu metode perhitungan kemiripan teks (text similarity algorithm) yang berfungsi untuk mengekstraksi fitur berbasis fingerprint teks. Algoritma WInnowing bekerja dengan membentuk fingerprint dari teks menggunakan teknik n-gram dan rolling hash, kemudian memilih nilai hash minimum dalam setiap window untuk merepresentasikan ciri khas dokumen. Fingerprint yang dihasilkan menjadi dasar dalam membandingkan tingkat kemiripan antar teks secara cepat dan efisien (M. Fitri et al., n.d.).

Di sisi lain, perkembangan Artificial Intelligence dan Machine Learning telah mendorong munculnya konsep Automated Essay Scoring (AES), yaitu sistem yang secara otomatis menilai jawaban tertulis peserta terhadap prompt atau pertanyaan tertentu, dengan memanfaatkan model-model pembelajaran komputer. AES dirancang untuk memberikan penilaian yang konsisten dan bebas bias manusia, sehingga proses evaluasi menjadi lebih cepat dan objektif. Penerapan AES dalam lingkungan CBT diharapkan dapat meningkatkan efektivitas evaluasi pembelajaran sekaligus mengurangi beban kerja pengajar (Ifenthaler, 2022).

Komponen utama dalam sistem AES adalah ekstraksi fitur teks, yaitu tahap transformasi teks esai menjadi representasi numerik agar dapat diproses oleh model pembelajaran mesin (Ifenthaler, 2022). Keberhasilan sistem penilaian otomatis sangat bergantung pada kualitas representasi fitur teks serta kemampuan model klasifikasi dalam memanfaatkan fitur tersebut secara optimal. Meskipun berbagai metode seperti bag-of-words, TF-IDF, dan word embedding telah banyak digunakan, pendekatan tersebut sering kali belum cukup efektif dalam

merepresentasikan kesamaan semantik antar teks dengan variasi panjang dan gaya bahasa yang tinggi.

Untuk mengatasi keterbatasan tersebut, algoritma Winnowing dapat diadaptasi sebagai metode ekstraksi fitur teks yang efisien dalam sistem AES. Melalui proses k-gram dan fingerprinting, Winnowing mampu menangkap pola karakteristik dari teks esai secara kompak dan bermakna. Namun demikian, efektivitas algoritma Winnowing sangat dipengaruhi oleh model pembelajaran mesin yang digunakan dalam tahap klasifikasi. Oleh karena itu, diperlukan analisis kinerja model pembelajaran mesin yang dikombinasikan dengan ekstraksi fitur Winnowing, khususnya Support Vector Machine (SVM) dan K-Nearest Neighbors (KNN), untuk mengetahui karakteristik performa masing-masing model dalam sistem penilaian esai otomatis. Kedua model ini memiliki keunggulan berbeda, di mana SVM dikenal lebih unggul dalam presisi dan kestabilan klasifikasi, sementara KNN memiliki waktu komputasi yang lebih cepat.

Berdasarkan uraian tersebut, penelitian ini berfokus pada analisis kinerja model Support Vector Machine (SVM) dan K-Nearest Neighbors (KNN) dengan menggunakan ekstraksi fitur Winnowing dalam sistem Automated Essay Scoring berbasis Computer Based Testing (CBT). Tujuan penelitian ini adalah untuk menganalisis dan membandingkan kinerja model SVM dan KNN yang menggunakan ekstraksi fitur Winnowing dalam proses klasifikasi jawaban esai secara otomatis. Hasil yang diharapkan dari penelitian ini adalah terciptanya sistem penilaian esai berbasis teks yang mampu memberikan hasil penilaian secara objektif, akurat, dan efisien, sehingga dapat menjadi kontribusi nyata dalam

pengembangan sistem evaluasi digital berbasis kecerdasan buatan di lingkungan pendidikan.

1.2 Rumusan Masalah

Berdasarkan latar belakang tersebut, maka rumusan masalah yang bisa dibuat penulis adalah sebagai berikut:

1. Bagaimana kinerja model Support Vector Machine (SVM) dan K-Nearest Neighbors (KNN) yang menggunakan ekstraksi fitur Winnowing dalam sistem Automated Essay Scoring?
2. Model manakah yang menunjukkan kinerja lebih optimal berdasarkan metrik akurasi, precision, recall, F1-score, dan waktu komputasi?
3. Bagaimana perbedaan karakteristik performa antara model SVM dan KNN dalam menilai jawaban esai berbasis teks menggunakan fitur Winnowing?

1.3 Batasan Masalah

Adapun yang menjadi batasan masalah dari penelitian ini adalah:

1. Penelitian ini difokuskan pada penerapan algoritma Winnowing sebagai metode ekstraksi fitur teks yang dikombinasikan dengan algoritma Support Vector Machine (SVM) dan K-Nearest Neighbors (KNN) untuk proses klasifikasi skor esai otomatis.
2. Dataset yang digunakan merupakan dataset Automated Student Assessment Prize (ASAP) dari Kaggle, yang terdiri atas esai berbahasa Inggris.
3. Penilaian hanya difokuskan pada aspek akurasi klasifikasi skor esai otomatis, precision, dan F1-Score, tanpa menganalisis aspek semantik mendalam atau kualitas isi esai secara manual.
4. Penelitian ini tidak membahas optimasi parameter model seperti grid search, hyperparameter tuning lanjutan, atau teknik ensemble learning.
5. Bahasa yang digunakan dalam dataset adalah bahasa Inggris, sehingga hasil penelitian ini belum mencakup penilaian esai berbahasa Indonesia.
6. Evaluasi sistem hanya dilakukan berdasarkan kinerja algoritma pada data uji tanpa implementasi langsung dalam sistem CBT atau LMS.

1.4 Tujuan Penelitian

Adapun yang menjadi tujuan dari penelitian ini adalah:

1. Untuk menerapkan metode ekstraksi fitur Winnowing yang dikombinasikan dengan algoritma Support Vector Machine (SVM) dan K-Nearest Neighbors (KNN) sebagai solusi dalam meningkatkan akurasi sistem penilaian esai otomatis berbasis teks.
2. Untuk menganalisis dan membandingkan performa kedua model (SVM dengan fitur hasil ekstraksi Winnowing dan KNN dengan fitur hasil ekstraksi Winnowing) berdasarkan nilai akurasi, precision, dan F1-Score, guna menentukan model yang paling optimal dalam klasifikasi skor esai otomatis.

1.5 Manfaat Penelitian

Adapun yang menjadi manfaat dari penelitian ini adalah:

1. Penelitian ini memberikan kontribusi dalam pengembangan ilmu pengetahuan, khususnya pada bidang Natural Language Processing (NLP) dan machine learning, dengan menunjukkan efektivitas algoritma Winnowing dalam proses ekstraksi fitur teks serta perbandingan performa antara model SVM dan KNN dalam penilaian esai otomatis. Hasil penelitian ini dapat menjadi referensi akademik untuk penelitian lanjutan terkait optimasi algoritma penilaian teks berbasis kecerdasan buatan.
2. Sistem yang dihasilkan dari penelitian ini memiliki kemampuan untuk menilai esai secara otomatis dengan akurasi tinggi dan waktu komputasi yang efisien. Dengan kombinasi metode Winnowing dan algoritma pembelajaran mesin, sistem ini mampu mengurangi subjektivitas manusia dalam penilaian serta mempercepat proses evaluasi hasil ujian berbasis teks.
3. Jika diterapkan dalam lingkungan sekolah atau lembaga pendidikan berbasis Computer-Based Testing (CBT), sistem ini dapat membantu guru atau penguji dalam menilai esai siswa secara lebih cepat, objektif, dan konsisten. Hal ini akan meningkatkan efisiensi proses penilaian, terutama dalam ujian berskala besar, serta mendukung penerapan sistem evaluasi digital yang transparan dan akurat.

BAB 2

TINJAUAN PUSTAKA

2.1 Tinjauan Pustaka

Dalam Penelitian pertama, Implementasi Deteksi Plagiarisme Menggunakan Metode N-Gram dan Jaccard Similarity Terhadap Algoritma Winnowing, membahas tentang implementasi gabungan metode N-Gram dan Jaccard Similarity pada algoritma Winnowing kedalam sistem untuk mendeteksi plagiarisme (Yudhana et al., n.d.). Dari penelitian tersebut dapat disimpulkan bahwa N-Gram sangat mempengaruhi hasil dari similarity dan dalam penggunaannya pemilihan nilai N-Gram yang tepat sangat diperlukan. Tingkat kesamaan pada dua dokumen akan menghasilkan nilai similarity yang berbeda jika N-Gram dan nilai window yang dimasukkan berbeda-beda nilainya. Jika nilai yang dimasukkan kecil maka hasil kesamaan bernilai tinggi. Semakin besar nilai N-Gram dan window yang dimasukkan maka nilai similarity atau kesamaan antar dokumen akan semakin rendah. Mendeteksi plagiarisme menggunakan metode N-Gram dan Jaccard Similarity terhadap algoritma Winnowing cukup baik digunakan untuk membandingkan kesamaan antara dua dokumen dan cukup baik digunakan untuk meminimalisir tindakan plagiarisme.

Penelitian kedua, rancang bangun sistem penilaian essay otomatis menggunakan metode Cosine Similarity (R. Fitri et al., 2015). Tujuan dari pembangunan sistem ini adalah menganalisa mekanisme penilaian essay yang selama ini telah dilaksanakan pengajar, serta merancang dan mengaplikasikan

penilaian ujian essay otomatis menggunakan metode Cosine Similarity. Proses penilaian dari sistem ini dibagi menjadi lima tahap yaitu, tahap pertama adalah input data yang dilakukan oleh pengajar dan siswa berupa pertanyaan soal, kunci jawaban, dan jawaban siswa. Tahap kedua adalah pemrosesan teks dengan melakukan proses tokenisasi, stopwords dan stemming. Tahapan ketiga adalah menghitung term frequency dan pembobotan kata, tahapan keempat adalah perhitungan kemiripan menggunakan Cosine Similarity. Tahap terakhir merupakan konversi nilai kemiripan menjadi nilai ujian essay.

Penelitian ketiga, menganalisis sentimen ulasan aplikasi SIREKAP 2024 (Hidayat et al., 2025a) menggunakan algoritma Naive Bayes (NB) dan Support Vector Machine (SVM) dengan pendekatan kuantitatif. Data dikumpulkan melalui web scraping, kemudian diproses dengan tahap preprocessing yang meliputi pembersihan teks, tokenisasi, dan stopwords removal, serta ekstraksi fitur menggunakan TF-IDF. Hasil penelitian menunjukkan bahwa SVM mencapai akurasi 95%, lebih tinggi dibandingkan NB yang memperoleh akurasi 94%, dengan SVM juga memiliki recall dan F1-score yang lebih baik, menjadikannya lebih efektif dalam mengidentifikasi sentimen yang lebih kompleks. Kelebihan dari penelitian ini terletak pada kecepatan implementasi Naive Bayes dan kemampuan SVM dalam menangani data berdimensi tinggi, namun terdapat kelemahan terkait ketidakseimbangan kelas yang mempengaruhi kinerja pada kelas minoritas. Selain itu, penelitian ini tidak mengeksplorasi algoritma Winnowing, yang berpotensi meningkatkan analisis kualitas teks dalam Automated Essay Scoring (AES). Teknik preprocessing seperti TF-IDF yang digunakan dalam penelitian ini dapat diadaptasi

untuk persiapan data esai dalam AES. Keunggulan SVM dalam klasifikasi teks yang kompleks mendukung penerapannya dalam AES untuk menilai kualitas esai dengan lebih efektif. Gap yang ditemukan adalah peluang untuk mengintegrasikan Winnowing dalam analisis fingerprinting esai, yang dapat meningkatkan akurasi dalam penilaian esai otomatis.

Penelitian keempat, Penelitian berjudul "Penilaian Esai Pendek Otomatis Berdasarkan Similaritas Semantik dengan SBERT" oleh Nurul Chamidah (Chamidah et al., 2022), Universitas Pembangunan Nasional Veteran Jakarta, bertujuan untuk mengevaluasi kinerja model penilaian esai otomatis dengan membandingkan esai jawaban mahasiswa dengan kunci jawaban evaluator berdasarkan kesamaan semantik. Penelitian ini menggunakan dataset bahasa Indonesia dari ujian mata kuliah Pengantar Basis Data dan memanfaatkan model Siamese-BERT (SBERT) untuk menghasilkan representasi numerik dari teks, yang kemudian dianalisis menggunakan Cosine Similarity untuk mengukur kesamaan semantik. Evaluasi model dilakukan dengan menggunakan metrik Mean Absolute Error (MAE) dan Pearson Correlation, dengan hasil MAE rata-rata 0.26 dan korelasi rata-rata 0.78. Model menunjukkan kinerja yang baik, meskipun ada tantangan dalam mengurangi error, yang tercermin dalam nilai MAE yang masih cukup besar. Penelitian ini menyimpulkan bahwa penggunaan SBERT efektif untuk penilaian esai otomatis dan menyarankan pelatihan ulang model dengan dataset yang lebih besar untuk meningkatkan akurasi.

Penelitian ketiga, Penelitian berjudul "Membandingkan Kinerja Metode Ekstraksi Fitur dan Model Pembelajaran Mesin dalam Penilaian Esai" oleh Lihua

Yao (Yao & Jiao, 2023), bertujuan untuk mengevaluasi kinerja berbagai metode ekstraksi fitur dan model pembelajaran mesin dalam sistem penilaian esai otomatis (AES). Penelitian ini menggunakan data dari Kaggle dan dataset ASAP, serta menerapkan NLP dan BERT untuk pemrosesan korpus dan ekstraksi fitur. Metodologi penelitian melibatkan empat langkah utama: pra-pemrosesan data, ekstraksi fitur, pelatihan model, dan evaluasi kinerja model menggunakan metrik seperti presisi, recall, akurasi, F-score, dan Quadratic Weighted Kappa (QWK). Hasil penelitian menunjukkan bahwa model jaringan saraf tiruan (ANN) yang dilatih pada semua data prompt memberikan kinerja terbaik dengan QWK keseluruhan 0.9724. Penambahan fitur seperti skor keterbacaan dari Spacy Textstat secara signifikan meningkatkan hasil prediksi, sementara model NLP menunjukkan kinerja lebih baik dibandingkan BERT. Namun, penelitian ini menghadapi keterbatasan terkait waktu komputasi dan kualitas data, serta tantangan dalam mengoptimalkan parameter fine tuning. Penulis menyarankan eksplorasi lebih lanjut terhadap jumlah fitur, teknik penambahan data, dan penerapan model transformer seperti GPT-2, GPT-3, T5, dan Roberta untuk meningkatkan akurasi penilaian esai otomatis. Penelitian ini memberikan wawasan penting mengenai penggabungan data pelatihan yang lebih komprehensif dan pemilihan model yang tepat untuk sistem penilaian esai otomatis.

2.2 Keaslian Penelitian

Table 2.1 Matriks literatur review dan posisi penelitian
Optimalisasi Algoritma Winnowing Untuk Penilaian Esai Otomatis Pada Sistem Computer Based Testing (CBT)

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
1	Analisis Sentimen Ulasan Pengguna pada Aplikasi SIREKAP 2024 Menggunakan Machine Learning	Taufik Hidayat, Nurchim, Numalitasari. Smart Comp: Jurnalnya Orang Pintar Komputer. Tahun 2024	Untuk menganalisis sentimen ulasan pengguna aplikasi SIREKAP menggunakan algoritma Naive Bayes (NB) dan Support Vector Machine (SVM) dengan pendekatan kuantitatif berbasis teks.	SVM menghasilkan akurasi 95%, lebih tinggi dibanding NB (94%). SVM juga memiliki recall dan F1-score lebih baik, sehingga lebih efektif untuk sentimen kompleks.	Kelemahan utama adalah ketidakseimbangan kelas yang mempengaruhi performa pada kelas minoritas, serta belum menerapkan teknik Winnowing dalam ekstraksi fitur teks.	Menunjukkan keunggulan SVM dalam klasifikasi teks, namun belum menerapkan Winnowing untuk ekstraksi fitur. Gap: peluang untuk mengintegrasikan SVM dengan fitur hasil ekstraksi Winnowing guna meningkatkan akurasi dan efisiensi dalam sistem Automated Essay Scoring (AES).
2	Penilaian Esai Pendek Otomatis Berdasarkan Similaritas Semantik dengan SBERT	Nurul Chamidah1, Mayanda Mega Santoni2, Helena Nurramdhani Irmada3, Ria Astriratma4, Yulnelly5.	Mengevaluasi kinerja model penilaian esai otomatis berdasarkan kesamaan semantik antara jawaban mahasiswa dan kunci jawaban menggunakan Siamese-BERT	Model SBERT menghasilkan MAE rata-rata 0.26 dan Pearson Correlation 0.78, menunjukkan performa baik dalam menilai kesamaan semantik antar teks.	Tantangan terdapat pada nilai MAE yang masih cukup besar, serta kebutuhan dataset lebih besar untuk meningkatkan akurasi model.	Fokus pada analisis semantik berbasis deep learning, namun belum mengombinasikan Winnowing dan algoritma Machine Learning yang lebih ringan untuk penilaian berbasis teks umum. Gap: perlunya eksplorasi integrasi

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
		Jurnal Techno.com Tahun 2022.	(SBERT) dan Cosine Similarity.			Winnowing untuk efisiensi dan generalisasi tinggi.
3	Comparing Performance of Feature Extraction Methods and Machine Learning Models in Automatic Essay Scoring	Lihua Yao, Hong Jiao. Journal of Educational Measurement and Evaluation. Tahun 2023.	Mengevaluasi kinerja berbagai metode ekstraksi fitur dan model Machine Learning dalam sistem penilaian esai otomatis (AES) menggunakan dataset Kaggle dan ASAP.	Model Artificial Neural Network (ANN) yang dilatih dengan semua prompt menunjukkan QWK 0.9724, dengan fitur keterbacaan dari Spacy Textstat meningkatkan hasil prediksi. Model NLP juga lebih baik dibanding BERT.	Kelemahan meliputi waktu komputasi tinggi, kualitas data terbatas, dan optimasi fine tuning yang kompleks.	Memberikan wawasan penting terkait penggabungan data pelatihan besar dan pemilihan model yang tepat. Gap: belum ada integrasi Winnowing sebagai metode ekstraksi fitur ringan yang dapat meningkatkan efisiensi dan interpretabilitas model AES.
4	Implementasi Natural Language Processing (NLP) dan Algoritma Cosine Similarity dalam Penilaian Ujian Esai Otomatis	Daniel Oktodeli Sihombing, Jurnal Sistem Komputer dan Informatika (JSON), Tahun 2022	Penelitian ini bertujuan mengimplementasikan Natural Language Processing (NLP) dan algoritma Cosine Similarity untuk melakukan penilaian ujian esai secara otomatis.	Penelitian ini menunjukkan penggunaan NLP dan metode cosine similarity dapat membantu dalam penilaian soal esai secara otomatis. Hasil perhitungan cosine similarity menunjukkan kemiripan jawaban mahasiswa dengan kunci jawaban dengan persentase tertinggi		Penelitian ini menggunakan Cosine Similarity untuk menghitung kemiripan jawaban, sedangkan penulis menggunakan algoritma Winnowing dengan tambahan Jaccard Coefficient untuk memperbaiki pengukuran kemiripan antara jawaban dan kunci jawaban.

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
				90,58% dan terendah 76,70%		
5	Aplikasi Penilaian Jawaban Esai Otomatis Menggunakan Metode Synonym Recognition Dan Cosine Similarity Berbasis Web	Imam Mufiid, Sri Lestanti, Ni'ma Kholila, Jurnal MNEMONIC, Tahun 2021	Penelitian ini bertujuan untuk menciptakan sebuah aplikasi yang dapat secara otomatis menilai jawaban esai siswa, sehingga dapat mengurangi beban kerja guru dalam proses koreksi.	Penelitian menunjukkan bahwa penilaian otomatis terhadap jawaban esai melibatkan penggunaan metode cosine similarity dan pengenalan sinonim. Selanjutnya, metode ini diuji menggunakan uji akurasi Root Mean Squared Error (RMSE) dan menghasilkan tingkat error yang cukup tinggi. Salah satu penyebabnya adalah kesalahan pengetikan oleh siswa yang mengurangi nilai similaritas.	Saran untuk penelitian selanjutnya meliputi penambahan analisis similaritas antara teks berdasarkan jenis kata, pengembangan aplikasi yang lebih dinamis untuk berbagai mata pelajaran atau bahasa, serta peningkatan pengujian akurasi terhadap metode yang digunakan.	Penelitian ini menggunakan Cosine Similarity dan Synonym Recognition untuk penilaian, sementara penulis menggunakan algoritma Winnowing untuk mengidentifikasi kemiripan esai.
6	Penerapan Sistem Penilaian Otomatis Pada Jawaban Tes Uraian Menggunakan Algoritma Rabin	Dimas Tifli Irhami Hamir, Ekohariadi, Jurnal IT - EDU, Tahun 2022	Penelitian ini memilih algoritma Rabin Karp karena dianggap efektif untuk mencari pola kata atau frasa dalam teks yang	Penelitian ini menunjukkan bahwa Sistem Penilaian Otomatis menggunakan Algoritma Rabin Karp telah dikembangkan	Untuk penelitian selanjutnya, disarankan agar memilih bentuk soal dan rubrik penilaian yang lebih sesuai, serta menyempurnakan	Penelitian ini menggunakan Algoritma Rabin Karp untuk mendeteksi pola dalam teks, sementara penulis menggunakan algoritma Winnowing, yang lebih fokus

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
	Karp Untuk Mata Pelajaran Informatika Di Ma Kanjeng Sepuh		panjang, sedangkan pendekatan Synonym Recognition digunakan untuk memperhitungkan sinonim kata dalam proses penilaian, sehingga sistem dapat lebih fleksibel dalam menilai variasi jawaban siswa.	dengan metode R&D model Plomp dan divalidasi dengan hasil sangat valid. Pengujian pada 55 siswa menunjukkan rata-rata nilai manual 74,15 dan nilai otomatis 76,28. Meskipun uji kesepakatan menggunakan Koefisien Kappa menunjukkan hasil rendah, uji reliabilitas menggunakan ICC menunjukkan hasil tinggi, mengindikasikan kesepakatan rendah tetapi reliabilitas tinggi antara penilaian manual dan otomatis.	konversi nilai otomatis. Selain itu, penerapan sistem penilaian otomatis menggunakan Algoritma Rabin Karp dan pendekatan Synonym Recognition sebaiknya diperluas ke mata pelajaran lain selain Informatika.	pada deteksi kemiripan dengan teks
7	Sistem Penilaian Esai Otomatis Menggunakan Algoritma Stemming Nazief dan Adriani	Ahmad, R. Sasuc, R. RO, Jurnal Teknologi Transportasi dan Logistik	Melakukan penilaian otomatis pada e-learning menggunakan metode stemming Nazief & Adriani.	Penggunaan metode stemming Nazief & Adriani untuk penilaian otomatis menghasilkan tingkat akurasi yang tinggi.	Diharapkan sistem penilaian dapat memperhatikan sinonim serta mendukung penggabungan bahasa	Penelitian ini menggunakan metode stemming untuk pengolahan teks, sementara penulis mengembangkan algoritma Winnowing dan

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
		Volume 1:40 2, December 2020, Hal 101-108, 2020			Inggris untuk kata benda.	jaccard coefficient untuk penilaian esai otomatis.
8	An Analysis of Automated Essay Grading Systems	Srivastava, Kshitiz, Dhanda, Nansrata Shrivastava, Anurag. International Journal of Recent Technology and Engineering, 2020	Melakukan kajian dan perbandingan terhadap berbagai sistem penilaian esai otomatis, seperti Project Essay Grade (PEG), Intelligent Essay Assessor (IEA), Educational Testing Service (ETS) Electronic Essay Rater (E-Rater), C-Rater, BETSY, Intelligent Essay Marking System, SEAR, Paperless School Free Text Marking Engine, dan Automark, serta lainnya.	Sebagian besar sistem yang dikembangkan berfokus pada aspek gaya, konten, atau keduanya. Hanya SAGrade dan SAGE yang menitikberatkan pada fitur semantik tingkat rendah, namun keduanya masih dalam tahap pengembangan.	Sistem penilaian saat ini belum mampu mendeteksi kebenaran isi esai. Oleh karena itu, diperlukan pendekatan berbeda untuk memeriksa konsistensi esai.	Penelitian ini membandingkan berbagai sistem penilaian esai otomatis, sementara penulis lebih fokus pada penerapan algoritma Winnowing dan Jaccard Coefficient.
9	Implementasi Algoritma Winnowing Pada	Wahyudi, Indra Bahri, Syamsul	Menerapkan algoritma winnowing untuk melakukan	Algoritma Winnowing dapat berfungsi dengan baik untuk penilaian		Penelitian ini menggunakan Algoritma Winnowing untuk penilaian otomatis, namun

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
	Sistem Penilaian Otomatis Jawaban Esai Pada Ujian Online Berbasis Web	Handayani, Popon. Jurnal Teknik Komputer, AMIK BSI, 2020	penilaian otomatis pada ujian online, khususnya untuk soal esai.	otomatis ujian online, dengan selisih rata-rata sebesar 5,68%.		penulis menambahkan Jaccard Coefficient untuk meningkatkan analisis kemiripan antara jawaban dan kunci jawaban, memperbaiki hasil perhitungan yang lebih presisi..

2.3 Landasan Teori

2.3.1 Sistem Computer Based Testing (CBT)

Sistem Computer Based Testing (CBT) merupakan metode evaluasi berbasis komputer yang menggantikan sistem ujian konvensional berbasis kertas. CBT memungkinkan pelaksanaan ujian secara lebih efisien, objektif, serta memberikan hasil penilaian yang cepat dan akurat. Menurut (Pradeep & Kowsalya, 2022), CBT tidak hanya meningkatkan efisiensi waktu dan biaya, tetapi juga mempermudah proses penyimpanan dan analisis data hasil tes. Namun, salah satu tantangan utama dalam penerapan CBT adalah penilaian esai secara otomatis. Berbeda dengan soal objektif seperti pilihan ganda, esai menuntut kemampuan analisis, berpikir kritis, serta kemampuan berbahasa yang kompleks. Oleh karena itu, diperlukan sistem Automated Essay Scoring (AES) yang mampu menilai esai secara objektif, akurat, dan cepat.

2.3.2 Automated Essay Scoring

Automated Essay Scoring (AES) atau penilaian esai otomatis merupakan sistem berbasis komputer yang dirancang untuk menilai jawaban berbentuk teks terbuka secara otomatis dengan menggunakan metode statistik, machine learning, serta teknik Natural Language Processing (NLP). Sistem ini meniru cara penilaian manusia terhadap tulisan dengan menganalisis berbagai atribut linguistik seperti isi, gaya bahasa, struktur, tata bahasa, serta mekanika penulisan (Ifenthaler, 2022)

AES memungkinkan proses evaluasi tulisan dilakukan secara cepat, objektif, dan konsisten, sehingga dapat mengurangi bias penilaian manusia serta memberikan umpan balik secara instan kepada peserta didik (Ifenthaler, 2022).

a. Komponen dan Proses Kerja AES

Secara umum, sistem AES bekerja melalui beberapa tahap utama (Ifenthaler, 2022), (Burrows et al., 2015) yaitu:

1. **Persiapan Data (Preparation)** – Teks esai dibersihkan dari karakter yang tidak relevan agar dapat dianalisis secara terstruktur.
2. **Tokenisasi (Tokenizing)** – Teks dipecah menjadi unit-unit kecil seperti kata, kalimat, dan tanda baca.
3. **Pelabelan (Tagging)** – Setiap token diberi label morfologis atau sintaktis (misalnya kata benda, kata kerja, dsb.).
4. **Stemming/Lemmatization** – Mengubah setiap kata ke bentuk dasarnya.
5. **Analisis (Analytics)** – Teks dianalisis menggunakan pendekatan NLP berdasarkan atribut linguistik dan semantik.
6. **Prediksi Skor (Prediction)** – Algoritma menghasilkan skor berdasarkan kesamaan antara teks uji dan teks referensi.
7. **Verifikasi (Validation)** – Hasil skor dibandingkan dengan data pembandingan atau hasil penilaian manusia untuk menguji keakuratan

b. Atribut Penilaian dalam AES

Zupanc dan Bosnic (Zupanc & Bosnic, 2017) mengelompokkan atribut penilaian dalam AES menjadi tiga kategori utama:

- a. Atribut Linguistik – mencakup panjang kalimat, jumlah kata, struktur sintaksis, variasi kosakata, dan kesalahan tata bahasa.
- b. Atribut Gaya (Style Attributes) – mencakup gaya penulisan, konsistensi nada, dan keterpaduan antarparagraf.
- c. Atribut Isi dan Semantik (Content Attributes) – meliputi kesesuaian isi dengan topik, kedalaman analisis, serta kemiripan semantik dengan teks acuan.

Pendekatan berbasis semantik seperti Latent Semantic Analysis (LSA) dan Pattern Matching Techniques (PMT) digunakan untuk mengukur tingkat kesamaan makna antar teks (Wild, 2016)

c. Metodologi dan Algoritma dalam AES

Sistem AES modern banyak menggunakan kombinasi machine learning dan Natural Language Processing. Beberapa metode yang sering digunakan antara lain (Ifenthaler, 2022) (Wild, 2016):

- Latent Semantic Analysis (LSA): menganalisis kesamaan makna antar kata dan kalimat dalam ruang semantik.
- Pattern Matching: mencocokkan pola teks peserta dengan referensi yang telah diberi skor oleh ahli.
- Neural Network dan Regression Models: digunakan untuk memprediksi skor akhir berdasarkan atribut linguistik yang terdeteksi.

- **Similarity-Based Matching:** seperti Jaccard Similarity atau Cosine Similarity untuk mengukur kedekatan antara teks jawaban dan teks acuan.

Dalam konteks penelitian, algoritma WInnowing digunakan sebagai metode document fingerprinting untuk menghitung tingkat kesamaan antar teks secara efisien berdasarkan k-gram dan rolling hash. Hal ini menjadikan WInnowing relevan sebagai komponen utama dalam penerapan AES untuk penilaian esai otomatis pada sistem CBT.

d. Implikasi AES dalam Pembelajaran Digital

Integrasi AES dalam sistem pembelajaran digital seperti Computer-Based Testing (CBT) dan Learning Management System (LMS) memberikan manfaat besar, antara lain:

- Menyediakan umpan balik formatif secara real-time kepada siswa.
- Mendukung penilaian berbasis data (data-driven assessment).
- Meningkatkan motivasi belajar melalui evaluasi yang cepat dan objektif.

Dengan dukungan analisis semantik dan algoritma seperti WInnowing, sistem AES dapat diterapkan untuk menilai kemampuan menulis dalam Bahasa Indonesia secara lebih akurat dan adaptif terhadap konteks pendidikan nasional.

2.3.3 Algoritma WInnowing

WInnowing adalah algoritma yang digunakan untuk melakukan proses document fingerprinting. Algoritma winnowing melakukan

penghitungan nilai- nilai hash dari setiap n-gram, untuk mencari nilai hash selanjutnya digunakan fungsi rolling hash. Kemudian dibentuk window dari nilai-nilai hash tersebut. Dalam setiap window dipilih nilai hash minimum. Jika ada lebih dari satu hash dengan nilai minimum, dipilih nilai hash yang paling kanan. Kemudian semua nilai hash terpilih disimpan untuk dijadikan fingerprint dari suatu dokumen (Yudhana et al., n.d. 2018).

Fingerprint ini yang akan dijadikan dasar pembandingan kesamaan antara teks yang telah dimasukkan. Syarat dari algoritma deteksi penjiplakan whitespace insensitivity, yaitu pencocokan teks file seharusnya tidak terpengaruh oleh spasi, jenis huruf kapital, tanda baca dan sebagainya, noise suppression yaitu menghindari pencocokan teks file dengan panjang kata yang terlalu kecil atau kurang relevan dan bukan merupakan kata yang umum digunakan, dan position independence yaitu pencocokan teks file seharusnya tidak bergantung pada posisi kata-kata sehingga kata dengan urutan posisi berbeda masih dapat dikenali jika terjadi kesamaan (Yudhana et al., n.d.).

Konsep Algoritma Winnowing ini ditunjukkan pada Gambar 2.1



Gambar 2.1 Konsep Algoritma Winnowing

Winnowing telah memenuhi syarat-syarat tersebut dengan cara membuang seluruh karakter-karakter yang tidak relevan misal: tanda baca, spasi dan juga karakter lain, sehingga nantinya hanya karakter-karakter yang berupa huruf atau angka yang akan diproses lebih lanjut.

a. *Whitespace Insensitive*

Whitespace insensitive adalah salah satu tahapan dari algoritma deteksi kemiripan dimana pada tahap ini akan membuang tanda baca, spasi dan karakter-karakter yang tidak relevan (@, #, \$, %, ^, & dan lainnya)

Contoh : aku cinta universitas amikom !!!

Akan diubah menjadi : akucintauniversitasamikom

b. Pembentukan rangkaian k-gram

Pada tahap ini akan dibentuk rangkaian karakter sepanjang k dari hasil proses *whitespace insensitive*. Jumlah data pengelompokan k-gram ini bisa dimulai dari 2, 3, 5, 7 dan seterusnya.

Contoh berdasarkan teks sebelumnya akan di ubah ke deretan k semisal dengan ukuran $k = 5$ maka akan menjadi "akuci kucin ucint cinta intaun tauni auniv unive niver ivers versi ersit rsita sitas itasa tasam asami samik amiko mikom".

c. Perhitungan Nilai Hash pada setiap n-gram

Pada tahapan ini akan dilakukan perhitungan nilai hash dari teks yang terbentuk dari proses sebelumnya dengan melakukan rolling hash. Rolling hash merupakan suatu cara untuk mentransformasikan sebuah string menjadi suatu nilai yang unik dengan panjang tertentu dimana

berfungsi sebagai penanda tersebut. Fungsi untuk mendapatkan nilai disebut fungsi hash sedangkan nilai yang dihasilkan disebut nilai hash. Berikut adalah fungsi hash. Seperti yang ditunjukkan pada persamaan (2.1).

$$HC_k = C_1 * b^{(k-1)} + C_2 * b^{(k-2)} + \dots + C_k * b^{(k-k)} \quad (2.1)$$

Keterangan :

C = nilai ascii karakter

B = basis (bilangan prima)

K = banyak karakter

- d. Pembentukan window dari nilai hash.
- e. Pemilihan nilai terkecil dari setiap window untuk dijadikan fingerprint.

2.3.4 Support Vector Machine

Support Vector Machine (SVM) merupakan salah satu algoritma dalam pembelajaran mesin (machine learning) yang berfungsi untuk melakukan klasifikasi maupun regresi. Prinsip utama SVM adalah mencari hyperplane optimal yang mampu memisahkan data ke dalam dua kelas dengan jarak maksimum atau maximum margin. Garis batas atau hyperplane ini ditentukan berdasarkan data yang paling dekat dengan bidang pemisah, yang dikenal sebagai support vectors (Alwanda, 2022).

Dalam proses klasifikasinya, SVM bekerja dengan memetakan data ke ruang berdimensi lebih tinggi menggunakan fungsi kernel. Pendekatan ini memungkinkan data yang tidak dapat dipisahkan secara linear dalam ruang asli menjadi dapat dipisahkan dengan baik pada ruang fitur baru.

Beberapa jenis kernel yang sering digunakan meliputi linear kernel, polynomial kernel, radial basis function (RBF), dan sigmoid kernel.

Secara matematis, SVM mencari fungsi keputusan. Seperti yang ditunjukkan pada persamaan (2.2):

$$f(x) = w^T x + b \quad (2.2)$$

di mana w adalah vektor bobot dan b adalah bias. Tujuannya adalah meminimalkan fungsi objektif. Seperti yang ditunjukkan pada persamaan (2.3):

$$\min \frac{1}{2} \|w\|^2 \quad (2.3)$$

dengan syarat bahwa setiap data memenuhi ketentuan. Seperti yang ditunjukkan pada persamaan (2.4):

$$y_i(w^T x_i + b) \geq 1 \quad (2.4)$$

(Schölkopf & Smola, 2001), dalam Alwanda, 2022).

Untuk kasus di mana data tidak dapat dipisahkan sempurna, SVM menggunakan parameter regularization CCC untuk menyeimbangkan antara margin yang besar dan tingkat kesalahan klasifikasi. Nilai CCC yang besar akan meminimalkan kesalahan klasifikasi, sementara nilai yang kecil akan memperlebar margin dengan risiko peningkatan kesalahan (Alwanda, 2022).

SVM juga memiliki kemampuan untuk menangani data yang bersifat non-linear dengan memanfaatkan konsep kernel trick. Teknik ini

memungkinkan SVM melakukan pemetaan data ke ruang fitur berdimensi tinggi tanpa harus menghitung koordinat aktual dari ruang tersebut. Pendekatan ini membuat SVM efektif dalam mengenali pola-pola kompleks pada data berdimensi tinggi (Cristianini et al., 2001)

Kelebihan utama dari SVM terletak pada kemampuannya menghasilkan performa klasifikasi yang baik meskipun pada dataset dengan ukuran kecil, serta kemampuannya menghindari overfitting bila parameter dan fungsi kernel diatur dengan tepat. Namun demikian, SVM juga memiliki kelemahan seperti waktu komputasi yang tinggi untuk dataset berukuran besar, dan kesulitan dalam menentukan parameter kernel yang optimal (Alwanda, 2022).

Dalam konteks penelitian ini, SVM digunakan sebagai model klasifikasi untuk menilai tingkat kesamaan teks hasil ekstraksi fitur dari algoritma Winnowing. Kombinasi Winnowing dan SVM diharapkan dapat meningkatkan akurasi serta konsistensi sistem penilaian esai otomatis berbasis Computer-Based Testing (CBT), dengan meminimalkan subjektivitas dan meningkatkan efisiensi proses evaluasi.

2.3.5 K-Nearest Neighbors

Algoritma K-Nearest Neighbor (KNN) merupakan salah satu metode klasifikasi non-parametrik yang bekerja dengan prinsip kesamaan jarak antar data. Algoritma ini digunakan untuk mengelompokkan suatu data baru berdasarkan kedekatannya dengan sejumlah data pelatihan (training data) yang sudah memiliki label kelas. Dengan kata lain, KNN

mengasumsikan bahwa data yang memiliki kemiripan karakteristik akan berada pada kelas yang sama (Amri, 2024), (Cover & Hart, 1967).

Kinerja KNN sangat bergantung pada kualitas dan relevansi fitur yang digunakan dalam proses klasifikasi. Apabila fitur tidak terdistribusi secara optimal, maka akurasi hasil klasifikasi dapat menurun. Oleh karena itu, penelitian-penelitian terkini yang menerapkan algoritma KNN sering kali menekankan pentingnya pemilihan fitur dan pembobotan nilai atribut agar algoritma dapat bekerja lebih efisien dalam mengenali pola data (Amri, 2024).

Secara umum, KNN termasuk dalam kategori instance-based learning, di mana seluruh data pelatihan disimpan sebagai acuan untuk mengklasifikasikan data baru. Proses klasifikasi dilakukan dengan menghitung jarak antara data uji dan seluruh data pelatihan, kemudian memilih sejumlah k tetangga terdekat sebagai dasar penentuan kelas. Representasi data pelatihan biasanya berada dalam ruang multidimensi yang merefleksikan setiap fitur dari data tersebut (Han et al., 2012), (Amri, 2024).

Rumus umum untuk menghitung jarak Euclidean dalam algoritma KNN adalah. Seperti yang ditunjukkan pada persamaan (2.5):

$$d = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \quad (2.5)$$

Keterangan:

- d : jarak antar dua titik data

- x : vektor fitur data pelatihan
- y : vektor fitur data uji

Setelah jarak antara semua pasangan data dihitung, sistem akan mengurutkan hasil perhitungan dari jarak terkecil hingga terbesar. Selanjutnya, KNN akan memilih k jarak terdekat, lalu melakukan proses voting untuk menentukan kelas data uji berdasarkan kelas mayoritas dari tetangga terdekat tersebut (Amri, 2024).

Kelebihan algoritma KNN terletak pada kesederhanaan konsep dan efektivitasnya dalam mengklasifikasikan data dalam jumlah besar dengan tingkat akurasi yang cukup tinggi. Namun demikian, KNN juga memiliki beberapa keterbatasan. Salah satu kelemahan utamanya adalah kesulitan dalam menentukan nilai k yang optimal. Nilai k yang terlalu kecil dapat menyebabkan model terlalu sensitif terhadap noise, sedangkan nilai k yang terlalu besar dapat menurunkan ketepatan klasifikasi karena mempertimbangkan terlalu banyak data dari kelas lain (Larose & Larose, 2014) (Amri, 2024).

Selain itu, algoritma ini memerlukan waktu komputasi yang tinggi karena harus menghitung jarak setiap data uji terhadap seluruh data pelatihan. Kondisi ini menjadi semakin kompleks apabila ukuran dataset besar dan jumlah fitur tinggi. Oleh sebab itu, dalam implementasi praktis, perlu dilakukan kompromi antara kecepatan komputasi dan tingkat akurasi agar penggunaan algoritma KNN tetap efisien (Amri, 2024).

Dalam konteks penelitian ini, KNN berperan sebagai salah satu metode pembelajaran mesin yang digunakan untuk mengklasifikasikan hasil ekstraksi fitur teks dari algoritma Winnowing. Kombinasi Winnowing dan KNN diharapkan dapat mendukung sistem penilaian esai otomatis berbasis Computer-Based Testing (CBT) dengan kinerja yang efisien dan hasil klasifikasi yang memadai.



BAB 3 METODE PENELITIAN

3.1 Jenis, Sifat, dan Pendekatan Penelitian

Jenis penelitian ini adalah penelitian eksperimental dengan pendekatan kuantitatif. Penelitian ini menguji performa algoritma Winnowing dalam penilaian esai otomatis berbasis teks menggunakan data esai berlabel skor numerik dari dataset Kaggle, yang selanjutnya dikonversi ke dalam kategori kelas (low, medium, dan high) untuk kepentingan analisis kinerja model klasifikasi. Proses penelitian meliputi preprocessing teks (pembersihan, tokenisasi, dan stemming), ekstraksi fitur menggunakan Winnowing dengan teknik k-gram dan fingerprinting, serta klasifikasi menggunakan algoritma SVM dan KNN. Evaluasi kinerja dilakukan berdasarkan akurasi, F1-score, dan waktu komputasi untuk menentukan algoritma dengan performa terbaik.

3.2 Metode Pengumpulan Data

Dataset yang digunakan dalam penelitian ini adalah Automated Student Assessment Prize (ASAP) Automated Essay Scoring Dataset yang diperoleh dari Kaggle. (<https://www.kaggle.com/competitions/learning-agency-lab-automated-essay-scoring-2>) Dataset ini terdiri dari 12,976 jawaban tertulis siswa (student-written responses) terhadap berbagai prompt ujian berbahasa Inggris dalam konteks pendidikan formal di Amerika Serikat.

Setiap jawaban telah dinilai oleh penilai manusia (human rater) dan dilengkapi dengan skor numerik pada rentang 1–6, yang diberikan berdasarkan

rubrik penilaian standar sesuai dengan karakteristik masing-masing prompt. Skor numerik tersebut digunakan sebagai ground truth dalam penelitian ini. Selanjutnya, skor numerik dikategorikan ke dalam tiga kelas, yaitu low, medium, dan high, untuk memformulasikan permasalahan penilaian esai sebagai tugas klasifikasi, sehingga memungkinkan analisis dan perbandingan kinerja model Support Vector Machine (SVM) dan K-Nearest Neighbors (KNN) dalam menentukan model dengan performa terbaik.

3.3 Metode Analisis Data

Analisis data dalam penelitian ini dilakukan untuk membandingkan kinerja dua model klasifikasi, yaitu Support Vector Machine (SVM) dan K-Nearest Neighbors (KNN), dalam sistem Automated Essay Scoring dengan memanfaatkan fitur teks hasil ekstraksi algoritma Winnowing. Data yang dianalisis berupa jawaban tertulis siswa yang telah dinilai oleh penilai manusia dan diberikan skor numerik pada dataset Automated Student Assessment Prize (ASAP). Data tersebut terlebih dahulu dipersiapkan melalui proses preprocessing teks dan pelabelan kelas sebagaimana dijelaskan pada subbab sebelumnya, sehingga siap digunakan dalam proses analisis model klasifikasi.

Pada tahap analisis, setiap jawaban teks direpresentasikan dalam bentuk fitur fingerprint hasil ekstraksi Winnowing. Representasi fitur ini digunakan secara konsisten sebagai masukan pada kedua model klasifikasi, SVM dan KNN, untuk memastikan bahwa perbedaan kinerja yang dihasilkan semata-mata disebabkan oleh karakteristik algoritma klasifikasi yang digunakan. Analisis data difokuskan

pada kemampuan masing-masing model dalam mengklasifikasikan jawaban esai ke dalam kategori performa berdasarkan label kelas yang telah ditentukan.

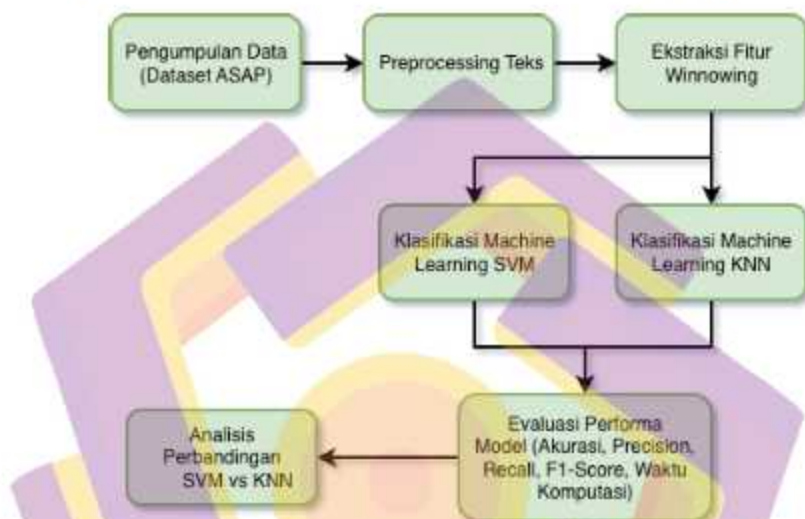
Evaluasi terhadap hasil analisis dilakukan dengan mengukur kinerja model menggunakan metrik akurasi, precision, recall, F1-Score, serta waktu komputasi. Teknik 5-fold cross-validation diterapkan untuk menguji kestabilan dan kemampuan generalisasi model terhadap variasi data latih dan data uji. Selain itu, analisis statistik menggunakan uji paired t-test digunakan untuk menilai signifikansi perbedaan kinerja antara model SVM dan KNN, sehingga kesimpulan yang diambil tidak hanya didasarkan pada nilai metrik semata, tetapi juga pada validitas statistik dari perbedaan tersebut.

Hasil analisis data ini selanjutnya menjadi dasar dalam pembahasan kinerja masing-masing model pada Bab IV serta penarikan kesimpulan mengenai model klasifikasi yang paling optimal untuk digunakan dalam sistem penilaian esai otomatis berbasis ekstraksi fitur Winnowing.

3.4 Alur Penelitian

Tahapan penelitian dimulai dari pengumpulan data jawaban esai berbahasa Inggris yang dilengkapi skor numerik asli dari dataset Automated Student Assessment Prize (ASAP) yang diperoleh melalui Kaggle, dilanjutkan dengan preprocessing teks. Selanjutnya dilakukan proses diskretisasi skor numerik ke dalam label kelas (low, medium, dan high) sebagai persiapan data untuk tugas klasifikasi. Ekstraksi fitur dilakukan menggunakan algoritma Winnowing dengan teknik n-gram dan fingerprinting. Fitur yang dihasilkan kemudian digunakan pada dua model klasifikasi secara paralel, yaitu Support Vector Machine (SVM) dan K-

Nearest Neighbors (KNN). Tahap akhir penelitian adalah evaluasi performa model menggunakan metrik akurasi, precision, recall, F1-score, serta waktu komputasi, yang selanjutnya dianalisis untuk menentukan pendekatan yang paling optimal.



Gambar 3.1 Alur Penelitian

3.4.1 Pengumpulan Data

Sumber data sekunder dalam penelitian ini adalah Dataset Automated Student Assessment Prize (ASAP) Automated Essay Scoring (AES) dari Kaggle yang terdiri dari 12.976 jawaban tertulis siswa (student-written responses) terhadap prompt ujian berbahasa Inggris, yang dikumpulkan dari konteks pendidikan formal di Amerika Serikat dan dinilai menggunakan rubrik penilaian standar.

<https://www.kaggle.com/competitions/learning-agency-lab-automated-essay-scoring-2>

3.4.2 Preprocessing Teks

Preprocessing dalam penelitian ini mengadaptasi berbagai teknik yang telah terbukti efektif dalam penelitian sebelumnya. Proses tokenisasi dan penghapusan stopwords dilakukan menggunakan library NLTK (Bird et al., 2009) dengan daftar stopwords bahasa Indonesia yang dikembangkan oleh (Tala, 2009). Untuk stemming, diterapkan dua pendekatan berbeda: algoritma Porter Stemmer (Porter, 1980) untuk teks bahasa Inggris dan library Sastrawi (Team, 2016) untuk teks bahasa Indonesia guna mereduksi kata ke bentuk dasarnya. Selain itu, penelitian ini memanfaatkan SymSpell (Garbe, 2017) untuk melakukan koreksi ejaan otomatis guna menormalisasi variasi penulisan kata yang mengandung kesalahan ketik.

Tahap preprocessing ini bertujuan untuk menyeragamkan dan menyiapkan teks jawaban agar siap diproses pada tahap ekstraksi fitur dan klasifikasi, serta tidak menjadi fokus utama evaluasi dalam penelitian ini.

3.4.3 Ekstraksi Fitur Winnowing

Proses ekstraksi fitur dilakukan dengan menggunakan algoritma Winnowing. Algoritma ini memanfaatkan teknik n-gram untuk membagi teks esai menjadi potongan-potongan kecil, yang kemudian diubah menjadi nilai hash menggunakan metode polynomial rolling hash. Nilai hash ini kemudian dipilih berdasarkan window untuk menghasilkan fingerprint yang mewakili fitur esai dalam bentuk vektor numerik (Prabowo & Setiawan,

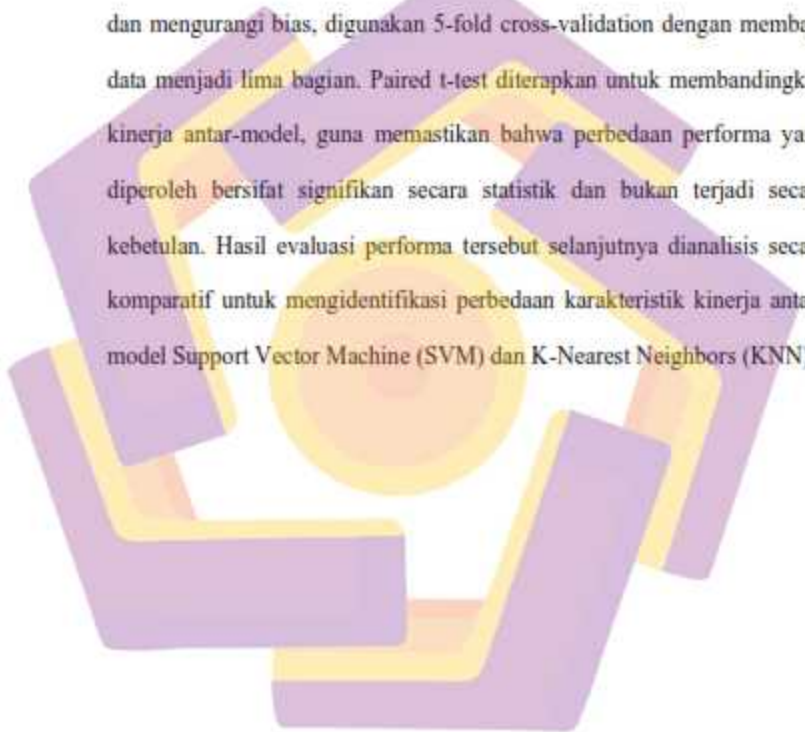
2022), (Sari & Rahman, 2021). Dalam penelitian ini, algoritma WInnowing digunakan sebagai metode ekstraksi fitur yang bersifat tetap (fixed feature extraction), sehingga fokus penelitian tidak terletak pada optimasi WInnowing, melainkan pada analisis kinerja model klasifikasi yang memanfaatkan fitur hasil ekstraksi tersebut.

3.4.4 Klasifikasi Machine Learning

Tahap klasifikasi dilakukan dengan membangun dua model pembelajaran mesin secara terpisah, yaitu Support Vector Machine (SVM) dan K-Nearest Neighbors (KNN), yang keduanya menggunakan fitur hasil ekstraksi WInnowing yang sama. Pendekatan ini dilakukan untuk memastikan bahwa perbedaan kinerja yang dihasilkan sepenuhnya berasal dari karakteristik masing-masing model klasifikasi, bukan dari perbedaan representasi fitur. Support Vector Machine (SVM) menggunakan kernel linear atau Radial Basis Function (RBF) dengan optimasi parameter C dan gamma melalui teknik GridSearchCV untuk meningkatkan performa model (Goyal, 2021). Selanjutnya, K-Nearest Neighbors (KNN) diimplementasikan dengan pengoptimalan parameter K menggunakan validasi silang untuk menentukan nilai optimal berdasarkan akurasi prediksi (Goyal, 2022). Algoritma ini dipilih karena keunggulannya dalam menangani berbagai karakteristik data.

3.4.5 Evaluasi Performa

Evaluasi performa model dilakukan untuk menjawab rumusan masalah penelitian terkait perbandingan kinerja model SVM dan KNN. Pengukuran kinerja dilakukan menggunakan metrik akurasi, precision, recall, F1-Score, serta waktu komputasi. Waktu komputasi dihitung untuk menilai efisiensi proses training dan testing. Untuk memastikan generalisasi dan mengurangi bias, digunakan 5-fold cross-validation dengan membagi data menjadi lima bagian. Paired t-test diterapkan untuk membandingkan kinerja antar-model, guna memastikan bahwa perbedaan performa yang diperoleh bersifat signifikan secara statistik dan bukan terjadi secara kebetulan. Hasil evaluasi performa tersebut selanjutnya dianalisis secara komparatif untuk mengidentifikasi perbedaan karakteristik kinerja antara model Support Vector Machine (SVM) dan K-Nearest Neighbors (KNN).



BAB 4 HASIL PENELITIAN DAN PEMBAHASAN

4.1 Deskripsi Data dan Preprocessing

4.1.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini berasal dari Automated Student Assessment Prize (ASAP) Automated Essay Scoring (AES) yang tersedia di Kaggle. <https://www.kaggle.com/competitions/learning-agency-lab-automated-essay-scoring-2> Dataset ini terdiri dari 12.976 jawaban tertulis siswa (student-written responses) berbahasa Inggris yang telah diberi skor numerik asli (human rater score) berdasarkan rubrik penilaian yang ditetapkan oleh kompetisi. Setiap esai memiliki informasi berupa ID esai (Essay_Id), teks esai lengkap (Full_Text), dan skor (Score).

Skor numerik pada dataset ASAP merupakan hasil penilaian manusia (human rater) dan menjadi dasar dalam proses diskretisasi ke dalam kategori kelas pada penelitian ini.

Dalam penelitian ini, untuk kebutuhan eksperimen, digunakan 5.000 esai yang diambil secara acak (random sampling) dengan random seed 42 untuk memastikan reproduksibilitas hasil. Dari 5.000 esai tersebut, dilakukan pembagian menjadi:

- Data pelatihan (training) sebanyak 4.000 esai, digunakan untuk membangun model penilaian otomatis.
- Data pengujian (testing) sebanyak 1.000 esai, digunakan untuk mengevaluasi performa model.

Proses sampling ini memastikan bahwa data yang digunakan representatif dari keseluruhan dataset dan distribusi skor tetap terjaga. Untuk memudahkan pembaca memahami struktur data, berikut ditunjukkan pada Tabel 4.1:

Table 4.1 Data Set

No	essay_id	full text	score
0	000d118	Many people have car where they live. The thing they don't know is tha...	3
1	000fc60	I am a scientist at NASA that is discussing the "face" on mars. I will...	3
2	001ab80	People always wish they had the same technology that they have seen in...	4
3	001bdc0	We all heard about Venus, the planet without almost oxygen with earthq...	4
4	002ba53	Dear, State SenatorThis is a letter to argue in favor of keeping the...	3
5	0030e86	If I were to choose between keeping the electoral college or abolishin...	4
6	0033037	The posibility of a face reconizing computer would be very helpful for ...	2
7	0033bf4	What is the Seagoing Cowboys program?It was to help many countries th...	3
8	0036253	The challenge of exploring VenusThis storie is about the challeng of...	2
9	0040e27	There are many reasons why you should join seagoing cowboys program. Y...	3

Selain itu, distribusi skor pada data training dan testing juga telah diperiksa untuk memastikan representasi kategori skor, sebagai berikut:

- Data training (4.000 esai): setiap kategori skor Low, Medium, dan High terdistribusi secara proporsional.
- Data testing (1.000 esai): mengikuti distribusi skor pada data training untuk evaluasi yang adil terhadap performa model.

Dengan pengaturan ini, dataset siap untuk tahap preprocessing dan ekstraksi fitur.

Dalam penelitian ini, skor numerik asli pada dataset ASAP tidak digunakan secara langsung, melainkan dilakukan proses diskretisasi ke dalam tiga kategori kelas, yaitu Low, Medium, dan High. Pelabelan ini bertujuan untuk mengubah permasalahan penilaian esai menjadi tugas klasifikasi guna kepentingan analisis kinerja model SVM dan KNN. Kategori tersebut tidak dimaksudkan untuk menggantikan nilai numerik yang digunakan oleh guru atau dosen dalam penilaian akademik.

4.1.2 Tahapan Preprocessing

Setelah data dikumpulkan, tahap berikutnya adalah preprocessing untuk mempersiapkan teks esai agar dapat diproses oleh algoritma Winnowing. Preprocessing dilakukan agar teks lebih bersih, konsisten, dan representatif untuk ekstraksi fitur. Tahapan preprocessing yang dilakukan penelitian adalah sebagai berikut:

1. Pembersihan Teks (Text Cleaning)

Setiap teks esai diubah menjadi huruf kecil (lowercasing) untuk memastikan keseragaman. Selanjutnya, karakter non-alfanumerik, seperti tanda baca, simbol, dan angka yang tidak relevan dihapus menggunakan ekspresi reguler. Spasi ganda juga diganti menjadi spasi tunggal, sehingga teks menjadi lebih rapi.

2. Tokenisasi

Teks yang sudah dibersihkan kemudian dipecah menjadi token atau kata-kata individual menggunakan tokenizer dari NLTK.

Tokenisasi mempermudah proses ekstraksi fitur berbasis kata maupun n-gram.

3. Penghapusan Stopwords

Kata-kata umum yang tidak memiliki makna penting dalam penilaian esai, seperti "the", "is", "and", dihapus. Stopwords yang digunakan diambil dari file custom stopwords yang diunduh melalui Kaggle, sehingga lebih sesuai dengan konteks esai berbahasa Inggris.

4. Stemming / Lemmatization (Opsional)

Penelitian ini, stemming tidak dilakukan. Namun, kata-kata tetap dibersihkan dan tokenisasi memastikan teks siap untuk langkah berikutnya, yaitu ekstraksi fitur dengan Winnowing.

5. Kategori Skor (Score Categorization)

Setiap skor numerik pada esai dikategorikan menjadi tiga kelas:

- Low : skor < 2
- Medium : $2 \leq \text{skor} < 4$
- High : skor ≥ 4
- Rentang kategori skor tersebut ditentukan berdasarkan distribusi skor numerik pada subset data yang digunakan, dengan tujuan menjaga keseimbangan kelas dan memungkinkan proses klasifikasi berjalan secara efektif.

Tahapan preprocessing ini menghasilkan teks esai yang lebih ringkas dan bebas kata-kata yang tidak relevan, sehingga model machine learning dapat lebih mudah mengenali pola dan kemiripan antar esai.

Berikut adalah hasil sebelum dan sesudah preprocessing ditunjukkan pada

Tabel 4.2:

Table 4.2 Hasil sebelum dan sesudah preprosesing

Raw Text	Preprocessed Text
What is the electoral college? The electoral college process plays a key role in the U.S. Presidential election.	electoral college electoral college process plays key role us presidential election
Driverless cars are not a good idea for the future. They may pose a safety risk and cause job losses.	driverless cars good idea future create safety risks cause job losses
Dear State Senator, I'm NOT in favor of the current healthcare plan.	dear state senator im favor keeping electoral health care plan
I'm going to tell you some positive and some negative ways that smart cars will affect the world.	im going tellin positive neative ways smart cars affect world
Is the face on Mars an alien artifact or just a natural landform?	face mars alien artifact natural landform far unknown
Driverless cars could be a good thing. I totally agree with the idea.	driverless cars could good thing totally agree idea
There are many advantages of limiting car usage. It decreases pollution and can improve overall health.	many advantages limiting car usage decreasing pollution improve health
The face on Mars is a landform in the region of Mars called Cydonia.	face mars landform region mars called cydonia
The challenge of exploring Venus. Exploring Venus would help us understand its atmosphere and geology.	challenge exploring venus exploring venus would help understand atmosphere geology
Though it may seem amazing to think about having driverless cars on the road, the technology is still not quite ready.	though may seem amazing think driveless cars technology not ready

Dengan preprocessing ini, teks esai menjadi lebih bersih, padat, dan siap untuk diekstraksi fitur menggunakan metode Winnowing. Tahap selanjutnya adalah melakukan ekstraksi fingerprint menggunakan n-gram

dan algoritma WInnowing untuk membangun representasi numerik yang akan digunakan dalam model klasifikasi.

4.1.3 Ringkasan Alur Input dan Output Setiap Tahapan Proses

Untuk memastikan bahwa setiap tahapan penelitian dapat dipahami dan ditelusuri secara runtut, penelitian ini mendefinisikan input dan output pada setiap proses secara eksplisit. Pendekatan ini bertujuan agar alur transformasi data, mulai dari teks esai mentah hingga keluaran model klasifikasi, dapat dijelaskan dengan jelas dan tidak menimbulkan ambiguitas pada tahap evaluasi hasil.

Proses penelitian dimulai dari data mentah berupa teks esai (`full_text`) dan skor numerik (`score`) yang diperoleh dari dataset Automated Student Assessment Prize (ASAP). Teks esai tersebut selanjutnya melalui tahap preprocessing untuk menghasilkan representasi teks yang lebih bersih dan konsisten. Hasil preprocessing ini kemudian digunakan sebagai masukan pada tahap ekstraksi fitur menggunakan algoritma WInnowing.

Pada tahap ekstraksi fitur, teks hasil preprocessing dipecah menjadi `n`-gram, kemudian diubah menjadi nilai hash dan diseleksi menggunakan mekanisme windowing pada algoritma WInnowing. Proses ini menghasilkan fingerprint teks yang disimpan dalam bentuk fitur `wInnowing_features`, yang merepresentasikan karakteristik utama dari setiap esai secara ringkas. Fitur inilah yang menjadi output utama dari tahap WInnowing dan digunakan sebagai masukan pada proses selanjutnya.

Meskipun fingerprint hasil WInnowing berupa nilai hash numerik, fitur tersebut masih bersifat diskrit dan belum memiliki bobot kepentingan

relatif. Oleh karena itu, penelitian ini menerapkan TF-IDF pada `winnowing_features` untuk mengonversi fingerprint menjadi vektor numerik berbobot. Tahap ini tidak dimaksudkan untuk mengekstraksi fitur baru dari teks mentah, melainkan untuk membentuk representasi vektor yang sesuai agar dapat diproses oleh algoritma klasifikasi.

Vektor hasil TF-IDF kemudian digunakan sebagai input model klasifikasi, yaitu Support Vector Machine (SVM) dan K-Nearest Neighbors (KNN). Keluaran dari tahap klasifikasi berupa label prediksi kategori skor (Low, Medium, dan High) yang selanjutnya dievaluasi menggunakan metrik akurasi, precision, recall, F1-Score, serta confusion matrix.

Ringkasan hubungan antara input dan output pada setiap tahapan proses penelitian disajikan pada Tabel 4.3.

Table 4.3 Alur Input dan Output Setiap Tahapan Proses

Tahap Proses	Input	Proses	Output
Input Data	full_text, score	Pembacaan dataset menggunakan <code>pd.read_csv()</code>	Data esai mentah
Preprocessing	full_text	Lowercase, pembersihan karakter, tokenisasi	processed_text
Kategorisasi Skor	score	Diskretisasi skor	Label kelas (Low, Medium, High)
N-gram	processed_text	Pembentukan 5-gram	Daftar n-gram
Hashing	n-gram	SHA-1 hashing	Nilai hash
Winnowing	Hash n-gram	Windowing dan seleksi minimum hash	Fingerprint

Fitur WInnowing	Fingerprint	Penggabungan fingerprint	winnowing_features
TF-IDF	winnowing_features	Pembobotan dan vektorisasi	Vektor numerik
Klasifikasi	Vektor TF-IDF	SVM dan KNN	Label prediksi
Evaluasi	Label aktual & prediksi	Perhitungan metrik evaluasi	Nilai performa model

4.1.4 Contoh Perubahan Data dari Input ke Output

Untuk memberikan gambaran yang lebih konkret mengenai alur pemrosesan data dalam penelitian ini, disajikan contoh perubahan data dari teks esai mentah hingga menjadi fitur yang digunakan oleh model klasifikasi.

Contoh ini bertujuan untuk memperjelas bagaimana setiap tahapan proses mengubah bentuk data secara bertahap, sehingga tidak terjadi loncatan proses dari input langsung ke hasil klasifikasi. Contoh yang ditampilkan merupakan ilustrasi dari satu data esai yang diproses menggunakan tahapan dan parameter yang sama dengan eksperimen utama, yaitu preprocessing teks, pembentukan n-gram, hashing, seleksi fingerprint menggunakan WInnowing, serta vektorisasi menggunakan TF-IDF

Table 4.4 Contoh Perubahan Data dari Input ke Output

Tahap Proses	Contoh Data
Teks Esai Asli (Raw Text)	"Driverless cars are not a good idea for the future."
Teks Setelah Preprocessing	driverless cars not good idea future
n-gram (n = 5)	driverless cars not good ideacars not good idea future
Hash n-gram	2345678998765432
Fingerprint (WInnowing)	23456789

Fitur Winnowing (winnowing_features)	"23456789"
Vektor TF-IDF	[0.00, 0.73, 0.00, ...]
Input Model Klasifikasi	Vektor numerik hasil TF-IDF
Output Model	Label kelas: Medium

Berdasarkan Tabel 4.4, dapat dilihat bahwa teks esai mentah terlebih dahulu dibersihkan melalui tahap preprocessing untuk menghasilkan representasi teks yang lebih ringkas dan konsisten. Selanjutnya, teks tersebut dipecah menjadi n-gram untuk menangkap pola lokal antar kata. Setiap n-gram kemudian diubah menjadi nilai hash dan diseleksi menggunakan algoritma Winnowing, sehingga hanya fingerprint yang paling representatif yang dipertahankan.

Fingerprint hasil Winnowing inilah yang disimpan sebagai fitur winnowing_features, yang merepresentasikan karakteristik utama dari esai dalam bentuk yang ringkas. Untuk memungkinkan proses klasifikasi menggunakan SVM dan KNN, fingerprint tersebut selanjutnya dikonversi menjadi vektor numerik berbobot menggunakan TF-IDF. Vektor inilah yang digunakan sebagai input model untuk menghasilkan prediksi kategori skor esai.

Dengan penyajian contoh ini, alur perubahan data dari input hingga output dapat ditelusuri secara jelas, sekaligus menunjukkan peran masing-masing tahapan dalam membentuk representasi data yang digunakan pada proses klasifikasi.

4.2 Ekstraksi Fitur Menggunakan WInnowing

WInnowing merupakan salah satu teknik penting dalam pemrosesan teks yang digunakan untuk mengekstrak fitur-fitur esensial dari dokumen. Secara umum, WInnowing dirancang untuk meningkatkan efisiensi proses pencocokan teks dengan memfokuskan perhatian pada elemen-elemen teks yang paling relevan, sehingga dapat mendukung deteksi kemiripan dokumen atau plagiarisme. Dalam penelitian ini, WInnowing diterapkan untuk mengekstraksi fitur dari esai yang kemudian digunakan sebagai input bagi model klasifikasi otomatis. Pendekatan ini memungkinkan sistem untuk bekerja dengan representasi teks yang lebih ringkas dan informatif, sekaligus mengurangi kompleksitas komputasi dibandingkan jika harus menganalisis seluruh teks secara langsung.

Teknik WInnowing yang diterapkan dalam penelitian ini menggunakan dua komponen utama: n-gram dan fingerprinting. Kedua teknik ini saling melengkapi dalam membangun representasi teks yang robust dan mudah dibandingkan.

4.2.1 Teknik N-gram

N-gram merupakan salah satu teknik fundamental dalam ekstraksi fitur teks yang digunakan untuk menangkap pola urutan kata dalam sebuah dokumen. Secara konseptual, n-gram memecah teks menjadi subsekuen kata sebanyak n token secara berurutan. Teknik ini mempertahankan konteks lokal antar kata, yang sangat penting dalam analisis semantik teks.

Dalam penelitian ini, n-gram diterapkan untuk menyiapkan teks agar dapat diproses lebih lanjut dengan algoritma WInnowing. Setiap esai yang telah melalui tahap preprocessing (pembersihan teks, tokenisasi,

penghapusan stopwords, dan normalisasi huruf) dipecah menjadi n-gram dengan panjang $n=5$ kata, ditunjukkan pada Gambar 4.1 berikut:

Penjelasan Gambar 4.1:

1. `text.split()` membagi teks menjadi token kata.
2. Perulangan `for i in range(len(words)-n+1)` digunakan untuk membentuk subsekuen kata sepanjang n token.
3. `'join(words[i:i+n])` menggabungkan token menjadi string n-gram.

Penerapan n-gram memungkinkan sistem untuk menangkap urutan kata penting yang muncul dalam esai, sehingga pola-pola kata yang signifikan tetap terjaga. Misalnya, frasa "driverless cars are not a good idea for the future" akan menghasilkan beberapa 5-gram seperti berikut:

- "driverless cars are not a good"
- "cars are not a good idea for"
- "are not a good idea for the"
- "not a good idea for the future"

Menurut (Ramli et al., 2021), penggunaan n-gram dalam algoritma Winnowing untuk deteksi plagiarisme dapat meningkatkan efektivitas sistem dalam mengenali dokumen yang memiliki kesamaan. Sementara itu, (Narasimhulu et al., 2024), menjelaskan bahwa Winnowing memungkinkan seleksi fitur yang efisien dalam klasifikasi biner, yang berguna dalam

berbagai aplikasi deteksi plagiarisme dan pemrosesan teks. Pendekatan ini memungkinkan sistem untuk fokus pada elemen-elemen yang paling berarti dalam teks, sehingga meningkatkan akurasi pencocokan kata dan efisiensi proses pencarian. Contoh hasil ekstraksi n-gram ditunjukkan pada Tabel 4.5 berikut:

Table 4.5 Hasil Ekstraksi n-gram

Raw Text	Preprocessed Text	n-gram (n=5)
What is the electoral college? The electoral college process plays a key role in the U.S. Presidential election.	electoral college electoral college process plays key role us presidential election	'electoral college electoral college process plays', 'college electoral college process plays key', 'electoral college process plays key role', 'college process plays key role us', 'process plays key role us presidential', 'plays key role us presidential election'
Driverless cars are not a good idea for the future. They may pose a safety risk and cause job losses.	driverless cars good idea future create safety risks cause job losses	'driverless cars good idea future create', 'cars good idea future create safety', 'good idea future create safety risks', 'idea future create safety risks cause', 'future create safety risks cause job', 'create safety risks cause job losses'
Dear State Senator, I'm NOT in favor of the current healthcare plan.	dear state senator im favor keeping electoral health care plan	'dear state senator im favor keeping', 'state senator im favor keeping electoral', 'senator im favor keeping electoral health', 'im favor keeping electoral health care', 'favor keeping electoral health care plan'

4.2.2 Fingerprinting

Fingerprinting adalah teknik untuk menghasilkan representasi numerik ringkas dari teks dengan tujuan mempercepat pencocokan teks atau deteksi duplikasi. Dalam algoritma WInnowing, fingerprinting digunakan

untuk mengubah setiap segmen teks (n-gram) menjadi nilai hash unik, sehingga dibandingkan teks dapat dilakukan pada level hash, bukan kata atau kalimat penuh. Pendekatan ini meningkatkan efisiensi komputasi sekaligus tetap mempertahankan informasi penting dalam teks. tahapan fingerprinting adalah sebagai berikut:

1. Preprocessing Teks

Setiap teks pada dataset melalui preprocessing:

- Lowercase: mengubah semua huruf menjadi huruf kecil.
- Pembersihan karakter: menghapus tanda baca dan karakter non-alfanumerik.
- Tokenisasi: memisahkan teks menjadi kata-kata.

```
def preprocess_text(text):  
    text = text.lower()  
    text = re.sub(r'[^\w-9]+', '', text)  
    text = re.sub(r'(\w+)', '\n', text.strip())  
    tokens = word_tokenize(text)  
    return ' '.join(tokens)
```

Gambar 4.2 Preprocessing Teks

2. Ekstraksi n-gram

Setelah teks dibersihkan, teks dipecah menjadi n-gram (misalnya $n=5$), yaitu subsekuen kata berurutan sepanjang n kata. Contoh teks preprocessed, "electoral college electoral college process plays key role us presidential election". Menghasilkan 5-gram: "electoral college electoral college process plays" "college electoral college process plays key" "electoral college process plays key role".

3. Hashing n-gram

Setiap n-gram diubah menjadi nilai hash integer menggunakan SHA-1 untuk menghasilkan fingerprint unik. Contoh: "electoral college electoral college process plays" → 43567281.

"college electoral college process plays key" → 89234512

4. Seleksi Fingerprint (Winnowing)

Setelah setiap n-gram diubah menjadi nilai hash unik, tahap selanjutnya adalah seleksi fingerprint menggunakan algoritma Winnowing. Tujuannya adalah untuk memilih subset hash yang paling representatif dari seluruh teks sehingga dapat digunakan sebagai fitur untuk klasifikasi, sambil tetap mengurangi redundansi.

Prinsip Kerja:

- **Windowing.** Algoritma Winnowing menggunakan konsep window dengan ukuran tertentu (`window_size`). Window adalah kelompok berturut-turut dari beberapa hash n-gram. Misalnya, jika `window_size = 4`, maka setiap window terdiri dari 4 hash n-gram berturut-turut.
- **Pemilihan Minimum Hash.** Dalam setiap window, winnowing memilih hash minimum sebagai fingerprint. Pemilihan minimum hash ini memastikan bahwa fingerprint yang dipilih stabil, bahkan jika terjadi perubahan minor pada teks (misalnya, typo atau kata tambahan). Dengan demikian, Winnowing menekankan robustness terhadap variasi teks sekaligus mempertahankan informasi yang paling signifikan.

- Pergerakan Window. Window bergerak secara bertahap ke seluruh urutan hash. Jika hash minimum pada window saat ini berbeda dari hash minimum sebelumnya, maka hash baru ini ditambahkan ke fingerprint. Hal ini mencegah duplikasi fingerprint yang berdekatan dan menghasilkan representasi teks yang lebih ringkas.
- Keuntungan Seleksi Fingerprint. Efisiensi komputasi, Sistem hanya perlu membandingkan subset hash, bukan seluruh teks. Reduksi redundansi, Hash yang berdekatan dengan nilai yang sama tidak disimpan berulang kali. Ketahanan terhadap perubahan minor. Fingerprint tetap konsisten meskipun teks mengalami sedikit modifikasi.

```
def windowlog(hashes, window_size=4):
    if not hashes:
        return []

    fingerprints = []
    min_hash = float("inf")
    min_idx = -1

    for i in range(len(hashes)):
        window = hashes[i:i+window_size]
        if len(window) < window_size:
            break
        local_min = min(window)
        local_idx = i + window.index(local_min)
        if (local_min, local_idx) != (min_hash, min_idx):
            fingerprints.append((local_min, local_idx))
            min_hash, min_idx = local_min, local_idx

    return fingerprints
```

Gambar 4.3 Seleksi Fingerprint

5. Hasil Fingerprint

Fingerprint akhir adalah urutan hash n-gram terpilih. Data ini digunakan sebagai fitur untuk klasifikasi skor esai menggunakan SVM dan KNN.

Table 4.6 Hasil Fingerprint

Raw Text	n-gram (n=5)	Fingerprint (Hash)
What is the electoral college?	"electoral college electoral college process plays"	43567281
	"college electoral college process plays key"	89234512
	"electoral college process plays key role"	67238945
Driverless cars are not a good idea for the future	"driverless cars not good idea for"	23456789
	"cars not good idea for future"	98765432
	"not good idea for future"	12345678

Pada Tabel 4.6:

- **Raw Text:** Menunjukkan teks asli dari dataset setelah dipilih sebagai contoh.
- **n-gram (n=5):** Segmen kata berurutan sepanjang 5 kata yang dihasilkan dari teks preprocessed. N-gram ini menjadi unit dasar fingerprint.
- **Fingerprint (Hash):** Nilai hash integer dari setiap n-gram, yang dipilih melalui algoritma WInnowing sebagai representasi teks yang ringkas dan efisien.

4.3 Penerapan Algoritma Machine Learning

4.3.1 Algoritma K-Nearest Neighbors (KNN)

Algoritma K-Nearest Neighbors (KNN) diterapkan dalam penelitian ini untuk melakukan klasifikasi esai berdasarkan kategori skor yang telah ditentukan (Low, Medium, High). Proses penerapan KNN sepenuhnya bergantung pada hasil ekstraksi fitur menggunakan algoritma WInnowing, yang sebelumnya menghasilkan representasi fingerprint sebagai fitur

karakteristik setiap esai. Dengan demikian, model KNN tidak bekerja pada teks mentah, tetapi pada fitur terstruktur berupa nilai hash fingerprint hasil ekstraksi `Winnowing`.

Tahap awal proses dilakukan dengan memuat dataset menggunakan perintah `pd.read_csv`, kemudian mengambil 5.000 sampel acak. Dataset ini dibagi menjadi training set (4.000 esai) dan test set (1.000 esai). Dataset ini diproses melalui tahap preprocessing yang meliputi normalisasi huruf, penghapusan karakter non-alfanumerik, serta tokenisasi menggunakan fungsi `preprocess_text`. Setelah proses pembersihan teks, setiap esai diubah menjadi sekumpulan `n-gram` dengan fungsi `generate_ngrams`, dan setiap `n-gram` di-hash menggunakan fungsi `hash_ngram`. Seluruh nilai hash kemudian diproses melalui algoritma `Winnowing` untuk menghasilkan fingerprint yang disimpan dalam kolom `winnowing_features`. Seluruh proses ini ditunjukkan pada Gambar 4.4 berikut:

```

def preprocess_text(text):
    text = text.lower()
    text = re.sub(r'[^a-z0-9]+', '', text)
    text = re.sub(r'(\w+)', '%s ' % text.strip(),
                text)
    tokens = word_tokenize(text)
    return ' '.join(tokens)

def generate_ngrams(text, n=3):
    words = text.split()
    return [' '.join(words[i:i+n]) for i in range(len(words)-n+1)]

def hash_ngram(ngram):
    return int(hashlib.md5(ngram.encode('utf-8')).hexdigest(), 16) % (2**64)

def winnowing(hashes, window_size=4):
    if not hashes:
        return []

    fingerprints = []
    min_hash_index = -1
    last_min_hash = -1

    window = [(hashes[i], i) for i in range(min(window_size, len(hashes)))]
    min_val, min_idx = min(window)

    fingerprints.append((min_val, min_idx))
    min_hash_index = min_idx
    last_min_hash = min_val

    for i in range(window_size, len(hashes)):
        if window[-1][1] == i - window_size:
            window.pop()

        window.append((hashes[i], i))

        if min_hash_index == i - window_size or hashes[i] < last_min_hash:
            min_val, min_idx = min(window)
            if min_val < last_min_hash or min_idx < min_hash_index:
                fingerprints.append((min_val, min_idx))
            min_hash_index = min_idx
            last_min_hash = min_val

    return fingerprints

def extract_winnowing_features(text, n=3, window_size=4):
    ngrams = generate_ngrams(text, n)
    hashes = [hash_ngram(ngram) for ngram in ngrams]
    fingerprints = winnowing(hashes, window_size)
    return ' '.join('%d' % h for h, idx in fingerprints)

```

Gambar 4.4 Preprocessing, N-gram, Hashing, dan Winnowing

Setelah fitur fingerprint diperoleh, tahap selanjutnya adalah mengonversi fitur tersebut menjadi vektor numerik melalui TfidfVectorizer. Hasil vektorisasi inilah yang digunakan sebagai input model KNN. Pembagian data menjadi 80% pelatihan dan 20% pengujian dilakukan menggunakan train_test_split, dan model KNN dilatih menggunakan parameter n_neighbors=3, ditunjukkan pada Gambar 4.5 berikut:

```

X_train, X_test, y_train, y_test = train_test_split(
    train_df['winnowing_features'],
    train_df['score_category'],
    test_size=0.2,
    random_state=42
)

vectorizer = TfidfVectorizer()
X_train_vec = vectorizer.fit_transform(X_train)
X_test_vec = vectorizer.transform(X_test)

models = {
    'SVM': SVC(kernel='linear'),
    'KNN': KNeighborsClassifier(n_neighbors=3)
}

for name, model in models.items():
    model.fit(X_train_vec, y_train)
    y_pred = model.predict(X_test_vec)
    print(classification_report(y_test, y_pred))

```

Gambar 4.5 KNN Training dan Evaluasi

Pada tahap klasifikasi, KNN bekerja dengan mengukur kedekatan satu esai uji terhadap seluruh esai pelatihan berdasarkan jarak vektor TF-IDF fingerprint. Setiap esai kemudian diberi label sesuai dengan mayoritas kelas dari tiga tetangga terdekatnya ($k = 3$). Proses ini memungkinkan model untuk mengelompokkan esai dengan pola tulisan yang mirip berdasarkan kesamaan nilai hash fingerprint, sehingga lebih relevan dibandingkan pencocokan teks secara langsung. Hasil evaluasi kemudian ditampilkan dalam bentuk akurasi, F1-Score, precision, recall, serta confusion matrix.

Model KNN berhasil melakukan klasifikasi dengan cukup stabil ketika fitur yang digunakan adalah fingerprint Winnowing, karena fingerprint mampu merangkum struktur konten secara ringkas namun tetap mempertahankan karakteristik penting dalam setiap esai. Dengan demikian,

penggabungan WInnowing dan KNN memberikan proses klasifikasi yang efisien.

4.3.2 Support Vector Machine (SVM)

Algoritma Support Vector Machine (SVM) diterapkan dalam penelitian ini untuk melakukan klasifikasi esai berdasarkan kategori skor yang telah ditentukan (Low, Medium, dan High). Proses penerapan SVM sepenuhnya bergantung pada hasil ekstraksi fitur menggunakan algoritma WInnowing, yang menghasilkan representasi fingerprint sebagai fitur karakteristik setiap esai. Dengan demikian, model SVM tidak bekerja pada teks mentah, tetapi pada fitur terstruktur berupa nilai hash fingerprint hasil ekstraksi WInnowing. Teknik SVM dipilih dengan kernel linear, karena kemampuannya yang efektif dalam menangani data dengan dimensi tinggi dan pola kompleks pada teks (Contreras et al., 2021).

Tahap awal proses dilakukan dengan memuat dataset menggunakan perintah `pd.read_csv`, kemudian mengambil 5.000 sampel acak. Dataset dibagi menjadi training set (4.000 esai) dan test set (1.000 esai) menggunakan fungsi `train_test_split`. Seluruh data kemudian diproses melalui tahap `preprocessing`, meliputi normalisasi huruf, penghapusan karakter non-alfanumerik, tokenisasi, dan penghapusan stopwords menggunakan fungsi `preprocess_text`. Setelah teks dibersihkan, setiap esai diubah menjadi n-gram menggunakan fungsi `generate_ngrams`, setiap n-gram di-hash menggunakan fungsi `hash_ngram`, dan seluruh nilai hash diproses melalui algoritma WInnowing untuk menghasilkan fingerprint.

Fingerprint ini disimpan dalam kolom `winning_features`. Seluruh proses ini ditunjukkan pada Gambar 4.6 berikut:

```
def preprocess_text(text):
    text = text.lower()
    text = re.sub(r'[0-9]+', '', text)
    text = re.sub(r'[^a-z ]+', '', text).strip()
    tokens = word_tokenize(text)
    return ' '.join(tokens)

def generate_ngrams(text, n=1):
    words = text.split()
    return [' '.join(words[i:i+n]) for i in range(len(words)-n+1)]

def hash_ngram(ngram):
    return int(hashlib.sha1(ngram.encode('utf-8')).hexdigest(), 16) % (2**64)

def winnowing_hashes(window_size=1):
    if not hashes:
        return []

    fingerprints = []
    min_hash_index = -1
    last_min_hash = -1

    window = [hashes[i], 0] for i in range(min(window_size, len(hashes)))
    min_val, min_idx = min(window)

    fingerprints.append((min_val, min_idx))
    min_hash_index = min_idx
    last_min_hash = min_val

    for i in range(window_size, len(hashes)):
        if window and window[0][1] == i - window_size:
            window.pop(0)

        window.append((hashes[i], i))

        if min_hash_index == i - window_size or hashes[i] < last_min_hash:
            min_val, min_idx = min(window)
            if min_val < last_min_hash or min_idx != min_hash_index:
                fingerprints.append((min_val, min_idx))
            min_hash_index = min_idx
            last_min_hash = min_val

    return fingerprints

def extract_winning_features(text, n=1, window_size=1):
    ngram = generate_ngrams(text, n)
    hashes = [hash_ngram(ngram) for ngram in ngrams]
    fingerprints = winnowing_hashes(window_size)
    return ' '.join(str(h) for h, idx in fingerprints)
```

Gambar 4.6 Preprocessing, N-gram, Hashing, dan Winnowing

Setelah fingerprint diperoleh, model SVM dilatih menggunakan dataset yang telah diproses, dan evaluasi dilakukan menggunakan data uji internal maupun eksternal. Proses pelatihan dan evaluasi ini ditunjukkan pada Gambar 4.7:

```

def evaluate_on_external_data(test_path):
    print("Evaluating model on external data")
    vectorizer = joblib.load('vect_vectorizer.pkl')
    svm_model = joblib.load('svm_classifier.pkl')
    svm_model = joblib.load('svm_classifier.pkl')

    print("Loading data into external...")
    test_df = pd.read_csv(test_path)
    test_df['processed_text'] = test_df['text'].apply(lambda x: preprocess_text(x))
    test_df['winnowing_features'] = test_df['processed_text'].apply(lambda x: winnowing.extract(winnowing_features))
    test_df['score_category'] = test_df['score'].astype('category')

    X_test = vectorizer.transform(test_df['winnowing_features'])
    y_test = test_df['score_category']

    models = [
        {'name': svm_model,
         'desc': 'svm_model'}
    ]

    results = []

    for name, model in models.items():
        print(f"Model: {name}")
        start_time = time.time()
        y_pred = model.predict(X_test)
        end_time = time.time()

        acc = accuracy_score(y_test, y_pred)
        f1 = f1_score(y_test, y_pred, average='weighted')
        precision = precision_score(y_test, y_pred, average='weighted')
        recall = recall_score(y_test, y_pred, average='weighted')
        computation_time = end_time - start_time

        print(f"Accuracy: {acc:.4f}")
        print(f"F1 Score: {f1:.4f}")
        print(f"Precision: {precision:.4f}")
        print(f"Recall: {recall:.4f}")
        print(f"Computation Time: {computation_time:.4f} sec")
        print(f"Classification Report:")
        print(classification_report(y_test, y_pred))

```

Gambar 4.7 SVM Training dan Evaluasi

Pada tahap klasifikasi, SVM menghitung hyperplane optimal yang memisahkan esai ke dalam tiga kategori skor berdasarkan vektor fingerprint. Setiap esai uji diklasifikasikan sesuai posisi relatif terhadap hyperplane tersebut. Hasil evaluasi ditampilkan dalam bentuk akurasi, F1-Score, precision, recall, serta confusion matrix, menunjukkan bahwa model SVM mampu melakukan klasifikasi dengan akurasi tinggi dan stabil.

Kombinasi Winnowing dan SVM menghasilkan proses klasifikasi yang efisien dan relevan, karena fingerprint merangkum konten esai secara ringkas namun tetap mempertahankan karakteristik utama setiap esai (Contreras et al., 2021).

4.4 Hasil Evaluasi Performa Model

4.4.1 Evaluasi Model KNN berbasis fitur hasil ekstraksi Winnowing

Model K-Nearest Neighbors (KNN) yang menggunakan metode ekstraksi fitur Winnowing menunjukkan performa yang relatif moderat dibandingkan model lainnya. Dalam hal akurasi, model ini mencapai nilai sebesar 0.707, yang berarti sekitar 70,7% prediksi kategori skor berhasil diklasifikasikan dengan benar. Nilai F1-Score yang diperoleh sebesar 0.686699 menunjukkan bahwa keseimbangan antara precision dan recall masih tergolong cukup, namun belum optimal. Nilai precision sebesar 0.686783 menandakan bahwa dari seluruh prediksi yang dikategorikan ke dalam suatu kelas, sekitar 68,7% adalah benar, sedangkan nilai recall sebesar 0.707 mengindikasikan bahwa sekitar 70,7% dari seluruh data aktual berhasil dikenali oleh model.

4.4.2 Evaluasi Model SVM berbasis fitur hasil ekstraksi Winnowing

Model Winnowing + Support Vector Machine (SVM) menunjukkan performa yang unggul dalam evaluasi klasifikasi kategori skor. Dengan akurasi sebesar 0.876, model ini mampu mengklasifikasikan sekitar 87,6% data uji secara tepat. F1-Score yang mencapai 0.868430 mencerminkan keseimbangan yang sangat baik antara nilai precision dan recall. Nilai precision yang tinggi, yaitu 0.891291, menunjukkan bahwa sebagian besar prediksi yang dilakukan oleh model adalah benar, sementara nilai recall sebesar 0.876 menunjukkan bahwa sebagian besar data aktual berhasil dikenali dan diklasifikasikan sesuai kelasnya.

4.4.3 Perbandingan Kinerja Model KNN dan SVM berbasis fitur hasil ekstraksi Winnowing

Berdasarkan hasil yang ditunjukkan dalam Tabel 4.5, terlihat adanya perbedaan kinerja yang cukup signifikan antara model KNN dengan fitur hasil ekstraksi Winnowing dan SVM dengan fitur hasil ekstraksi Winnowing. Model SVM dengan fitur hasil ekstraksi Winnowing menunjukkan performa yang lebih unggul dalam hampir seluruh metrik evaluasi. Akurasi model SVM mencapai 0,876, jauh di atas model KNN yang hanya memperoleh akurasi sebesar 0,707. Begitu pula pada metrik F1-Score, precision, dan recall, SVM konsisten menunjukkan keunggulan yang mencerminkan kemampuannya dalam menangkap pola klasifikasi yang kompleks secara lebih efektif. Hasil ini menunjukkan bahwa SVM lebih andal untuk digunakan dalam konteks klasifikasi kategori skor esai berbasis fitur winnowing.

Dari segi efisiensi komputasi, model KNN dengan fitur hasil ekstraksi Winnowing memiliki waktu komputasi yang jauh lebih singkat, yakni hanya 0,077723 detik, dibandingkan dengan SVM dengan fitur hasil ekstraksi Winnowing yang memerlukan waktu sekitar 5,181894 detik. Namun, trade-off antara kecepatan dan ketepatan menjadi pertimbangan penting, terutama dalam sistem penilaian otomatis yang mengutamakan akurasi. Dengan demikian, meskipun KNN lebih ringan secara waktu proses, model SVM lebih direkomendasikan karena memberikan hasil klasifikasi yang lebih presisi dan stabil sebagaimana juga diungkapkan oleh (Gomiasti et al., 2024). yang menunjukkan kelebihan SVM dalam

mengatasi data dengan banyak fitur, terutama dalam pengolahan data teks dengan representasi fitur yang kompleks seperti WInnowing.

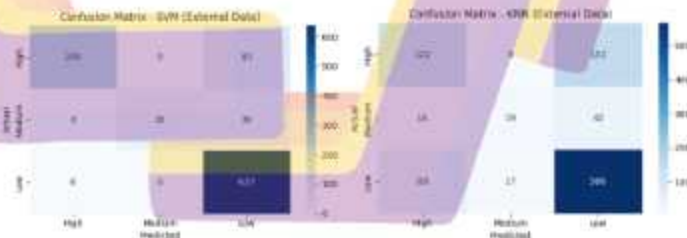
Table 4.7 Perbandingan Model

Model	Akurasi	F1-Score	Precision	Recall	Waktu Komputasi (s)
SVM dengan fitur hasil ekstraksi WInnowing	0.876	0.868430	0.891291	0.876	5.181894
KNN dengan fitur hasil ekstraksi WInnowing	0.707	0.686699	0.686783	0.707	0.077723

4.5 Analisis Hasil

4.5.1 Perbandingan Antara Model dan Metode

Confusion matrix pada bagian ini menggambarkan hasil klasifikasi jawaban esai ke dalam kategori performa Low, Medium, dan High sebagai representasi kelas eksperimen, bukan sebagai nilai akademik akhir sebagaimana digunakan dalam penilaian oleh pendidik.



Gambar 4.8 Confusion Matrix

Gambar 4.1 confusion matrix dari dua model klasifikasi, yaitu SVM dengan fitur hasil ekstraksi WInnowing dan KNN dengan fitur hasil ekstraksi WInnowing, yang masing-masing diterapkan pada data uji eksternal. Berdasarkan visualisasi ini, terlihat bahwa model SVM dengan

fitur hasil ekstraksi Winnowing menunjukkan distribusi klasifikasi yang lebih konsisten dan akurat dibandingkan model KNN dengan fitur hasil ekstraksi Winnowing. Pada confusion matrix SVM, prediksi untuk kategori "Low" sangat dominan dan akurat dengan 637 prediksi yang benar. Selain itu, kategori "High" juga memiliki tingkat keberhasilan tinggi, yaitu 200 prediksi benar, dengan sedikit kesalahan klasifikasi ke kategori "Low" sebanyak 82 kasus. Sementara itu, untuk kelas "Medium", SVM berhasil memprediksi 39 data secara tepat.

Sebaliknya, confusion matrix dari model KNN menunjukkan lebih banyak kesalahan klasifikasi. Meskipun prediksi untuk kelas "Low" masih cukup tinggi (566 benar), terdapat penyimpangan yang cukup besar, seperti kelas "High" yang sebagian besar justru diklasifikasikan ke dalam kelas "Low" (132 salah klasifikasi), serta kelas "Medium" yang tidak hanya memiliki prediksi benar lebih rendah (19), tetapi juga salah dialokasikan ke kelas lain secara signifikan. Hal ini mengindikasikan bahwa KNN cenderung memiliki keterbatasan dalam membedakan fitur antar kelas ketika menggunakan representasi teks berbasis Winnowing.

Berdasarkan hasil confusion matrix pada Gambar 4.8, perbedaan kinerja antara SVM dan KNN tidak hanya terlihat dari nilai akurasi, tetapi juga dari pola kesalahan klasifikasi. Model KNN menunjukkan kecenderungan salah mengklasifikasikan kelas 'High' dan 'Medium' ke dalam kelas 'Low', yang mengindikasikan keterbatasan KNN dalam membedakan batas antar kelas pada ruang fitur fingerprint Winnowing yang bersifat berdimensi tinggi dan jarang (sparse). Sebaliknya, SVM mampu

mempertahankan distribusi yang lebih seimbang dan konsisten, terutama pada kelas 'High', karena proses klasifikasinya tidak bergantung pada jarak antar seluruh data, melainkan pada support vector di sekitar batas kelas.

4.5.2 Analisis Hasil Berdasarkan Skenario Uji

Untuk memperoleh gambaran yang lebih komprehensif mengenai kinerja model, dilakukan pengujian dengan tiga skenario uji, yaitu:

1. Skenario 1 – Evaluasi Data Uji Eksternal Asli: n-gram = 5, window = 4.
2. Skenario 2 – Variasi Parameter n-gram dan Window: n-gram = 3, window = 3.
3. Skenario 3 – Simulasi Pengurangan 30% Isi Jawaban: meniru kondisi jawaban tidak lengkap atau adanya kesalahan ketik.

Hasil evaluasi pada ketiga skenario ditunjukkan pada Tabel 4.8.

Table 4.8 Perbandingan Performa Model pada Berbagai Skenario

Skenario	Model	Akurasi	F1-Score	Precision	Recall	Waktu Komputasi (s)
1	SVM dengan fitur hasil ekstraksi Winnowing	0.876	0.868430	0.891291	0.876	5.100589
1	KNN dengan fitur hasil ekstraksi Winnowing	0.707	0.686699	0.686783	0.707	0.099360
2	SVM dengan fitur hasil ekstraksi Winnowing	0.643	0.503285	0.413449	0.643	0.312290
2	KNN dengan fitur hasil	0.380	0.389606	0.428097	0.380	0.097886

	ekstraksi Winnowing					
3	SVM dengan fitur hasil ekstraksi Winnowing	0.846	0.823555	0.869654	0.846	3.515277
3	KNN dengan fitur hasil ekstraksi Winnowing	0.707	0.685312	0.687773	0.707	0.097212

Berdasarkan hasil di atas, model SVM dengan fitur hasil ekstraksi Winnowing pada skenario 1 menunjukkan performa terbaik dibandingkan seluruh skenario lainnya, baik dari sisi akurasi maupun F1-Score. Hal ini menunjukkan bahwa parameter n-gram dan window yang optimal ($n=5$, $window=4$) mampu menghasilkan fingerprint yang paling representatif untuk membedakan kategori skor esai.

Pada skenario 2, performa menurun signifikan untuk kedua model karena kombinasi n-gram yang lebih kecil ($n=3$) dan ukuran window yang lebih kecil menghasilkan fingerprint yang kurang informatif. Hal ini menyebabkan model kesulitan membedakan kelas dan meningkatkan kesalahan klasifikasi. Penurunan performa pada skenario 2 berdampak lebih signifikan pada KNN dibandingkan SVM. Hal ini menunjukkan bahwa KNN lebih sensitif terhadap perubahan kualitas fingerprint, karena mekanisme klasifikasinya bergantung pada kedekatan jarak antar vektor. Sebaliknya, SVM masih mampu mempertahankan pola pemisahan kelas meskipun kualitas fitur menurun, yang mengindikasikan bahwa model ini lebih adaptif terhadap variasi representasi fingerprint hasil Winnowing.

Skenario 3 menunjukkan bahwa pengurangan 30% isi jawaban menyebabkan sedikit penurunan akurasi pada model SVM (0,846), namun SVM tetap lebih unggul dibanding KNN yang akurasinya tetap 0,707. Hal ini menegaskan ketahanan SVM terhadap teks yang tidak lengkap atau adanya variasi konten, sedangkan KNN lebih sensitif terhadap kehilangan informasi karena prediksi bergantung pada jarak antar vektor dalam ruang fitur.

Berdasarkan hasil eksperimen yang ditunjukkan pada Tabel 4.8, model SVM secara konsisten menghasilkan nilai akurasi dan F1-Score yang lebih tinggi dibandingkan KNN pada seluruh skenario uji. Temuan ini diperoleh langsung dari hasil pengujian terhadap data uji asli, bukan semata-mata berdasarkan asumsi atau teori pada penelitian terdahulu. Secara keseluruhan, hasil ini menegaskan bahwa kombinasi SVM dengan fitur hasil ekstraksi Winnowing dengan n-gram 5 dan window 4 pada data uji asli adalah konfigurasi optimal untuk sistem penilaian esai otomatis berbasis teks.

4.5.3 Perbandingan dengan Penelitian Terdahulu

Berdasarkan hasil evaluasi performa model pada penelitian ini, kombinasi SVM dengan fitur hasil ekstraksi Winnowing menunjukkan akurasi tertinggi sebesar 87,6% dengan F1-Score 0,868 dan precision 0,891, sedangkan kombinasi KNN dengan fitur hasil ekstraksi Winnowing memiliki akurasi 70,7% dengan F1-Score 0,687. Performa SVM lebih

unggul secara konsisten dibandingkan KNN pada berbagai skenario uji, terutama pada skenario data uji eksternal asli.

Jika dibandingkan dengan penelitian terdahulu, terdapat beberapa poin penting:

1. SVM untuk klasifikasi teks

Penelitian oleh (Hidayat et al., 2024) menunjukkan SVM mampu mencapai akurasi 95% dalam analisis sentimen ulasan aplikasi SIREKAP, lebih tinggi dibanding Naive Bayes. Temuan dalam penelitian ini menunjukkan bahwa keunggulan SVM atas KNN bukan hanya didukung oleh literatur sebelumnya, tetapi juga tercermin langsung dari hasil eksperimen yang dilakukan. Pada seluruh skenario uji, SVM secara konsisten menghasilkan akurasi dan F1-Score yang lebih tinggi dibandingkan KNN, yang mengindikasikan bahwa karakteristik data fingerprint Winnowing lebih sesuai dengan mekanisme klasifikasi SVM dibandingkan pendekatan berbasis jarak seperti KNN.

2. Penggunaan Winnowing sebagai metode ekstraksi fitur:

Banyak penelitian sebelumnya menggunakan teknik NLP atau Cosine Similarity (Sihombing, 2022; Mufiid et al., 2021) atau metode Rabin-Karp (Tifli & Hamir, n.d.) untuk mendeteksi kemiripan teks. Namun, metode tersebut cenderung sensitif terhadap variasi teks dan kesalahan ketik. Penelitian ini menggabungkan SVM dengan fitur hasil ekstraksi Winnowing, KNN dengan fitur hasil ekstraksi Winnowing, sehingga ekstraksi

fitur lebih robust dan dapat menangkap kesamaan teks meskipun terjadi variasi kata atau sebagian jawaban hilang (skenario 3). Pendekatan ini memberikan efisiensi komputasi lebih baik dibanding teknik NLP berbasis deep learning seperti SBERT (Chamidah et al., 2022), yang membutuhkan dataset besar dan sumber daya komputasi tinggi.

3. Perbandingan dengan metode deep learning dan ANN:

(Yao & Jiao, 2023) melaporkan ANN dengan fitur keterbacaan dapat mencapai QWK 0,9724, namun waktu komputasi tinggi dan optimasi kompleks. Penelitian ini menunjukkan bahwa kombinasi SVM dengan fitur hasil ekstraksi Winnowing dapat memberikan performa yang kompetitif dengan waktu komputasi lebih rendah (5,1 detik untuk 1.000 data uji).

4. Robustness terhadap variasi data:

Skenario pengurangan 30% isi jawaban (skenario 3) menunjukkan SVM tetap mempertahankan akurasi 84,6%, sedangkan KNN stabil di 70,7%. Hal ini menunjukkan SVM dengan fitur hasil ekstraksi Winnowing lebih tahan terhadap noise atau jawaban tidak lengkap, sementara penelitian sebelumnya yang menggunakan Cosine Similarity atau Rabin-Karp menunjukkan sensitivitas tinggi terhadap kesalahan ketik atau jawaban parsial.

BAB 5 PENUTUP

5.1 Kesimpulan

Berdasarkan hasil pengujian yang dilakukan pada tiga skenario evaluasi, yaitu skenario parameter asli ($n\text{-gram} = 5$, $\text{window} = 4$), skenario variasi parameter ($n\text{-gram} = 3$, $\text{window} = 3$), serta skenario simulasi pengurangan 30% isi jawaban, maka kesimpulan penelitian ini dirumuskan berdasarkan temuan empiris dari hasil eksperimen sebagai berikut:

1. Kinerja model SVM dan KNN berbasis ekstraksi fitur Winnowing dalam sistem Automated Essay Scoring

Hasil eksperimen menunjukkan bahwa model Support Vector Machine (SVM) dan K-Nearest Neighbors (KNN) yang menggunakan ekstraksi fitur Winnowing secara empiris mampu melakukan klasifikasi esai otomatis pada seluruh skenario pengujian yang diterapkan. Representasi fitur berbasis fingerprint yang dihasilkan oleh algoritma Winnowing terbukti dapat diproses oleh kedua model, sehingga memungkinkan evaluasi kinerja model dalam sistem Automated Essay Scoring berbasis teks. Temuan ini diperoleh langsung dari hasil pengujian terhadap data uji eksternal dan bukan semata-mata didasarkan pada asumsi teoritis.

2. Perbandingan kinerja model SVM dan KNN berbasis fitur hasil ekstraksi Winnowing berdasarkan metrik evaluasi

Berdasarkan hasil pengujian pada seluruh skenario, kombinasi SVM dengan fitur hasil ekstraksi Winnowing secara konsisten menghasilkan performa yang lebih baik dibandingkan KNN, terutama pada metrik akurasi,

precision, recall, dan F1-Score. Pada skenario parameter asli, SVM mencapai akurasi tertinggi sebesar 0.876, sedangkan KNN hanya mencapai 0.707. Ketika parameter dipersempit, kedua model mengalami penurunan performa, namun SVM tetap menunjukkan akurasi yang lebih tinggi (0.643) dibandingkan KNN (0.380). Pada skenario simulasi jawaban tidak lengkap, SVM kembali mempertahankan kinerja yang stabil dengan akurasi 0.846, sementara KNN tidak menunjukkan peningkatan signifikan. Konsistensi hasil ini menunjukkan bahwa keunggulan SVM ditentukan berdasarkan hasil eksperimen yang dilakukan, bukan hanya merujuk pada temuan penelitian terdahulu.

3. Karakteristik perbedaan performa model SVM dan KNN dalam penilaian esai otomatis berbasis Winnowing

Hasil eksperimen memperlihatkan adanya perbedaan karakteristik performa antara model SVM dan KNN. Model SVM secara konsisten menunjukkan nilai precision dan F1-Score yang lebih tinggi pada seluruh skenario pengujian, yang mengindikasikan kemampuan generalisasi dan stabilitas klasifikasi yang lebih baik terhadap variasi data. Sebaliknya, model KNN memiliki keunggulan pada aspek efisiensi waktu komputasi dengan rata-rata waktu pemrosesan di bawah 0.1 detik. Temuan ini menunjukkan adanya trade-off antara tingkat ketepatan penilaian dan efisiensi waktu, di mana pemilihan model sebaiknya disesuaikan dengan kebutuhan sistem berdasarkan hasil pengujian empiris yang telah dilakukan dalam penelitian ini.

5.2 Saran

Secara keseluruhan, SVM lebih cocok digunakan pada aplikasi yang mengutamakan akurasi meskipun sedikit mengorbankan kecepatan, sementara KNN lebih efisien dalam aplikasi yang memiliki keterbatasan waktu. Penelitian ini juga membuka kemungkinan pengembangan selanjutnya dengan mengeksplorasi teknik ekstraksi fitur lainnya atau melakukan penyempurnaan parameter pada kedua model menggunakan grid search untuk mencari konfigurasi terbaik, yang dapat meningkatkan kinerja dan waktu komputasi model secara keseluruhan.



DAFTAR PUSTAKA

- Almi Yulistia Alwanda. (2022). ANALISIS PERBANDINGAN ALGORITMA SUPPORT VECTOR MACHINE, K-NEAREST NEIGHOR DAN NAÏVE BAYES DALAM PEMILIHAN KONSENTRASI MAHASISWA.
- Amri. (2024). KOMPARASI ALGORITMA NAÏVE BAYES, DECISION TREE, ARTIFICIAL NEURAL NETWORK, K- NEAREST NEIGHBOR, DAN SUPPORT VECTOR MACHINE UNTUK PREDIKSI TINGKAT KELULUSAN MAHASISWA.
- Bird, S., Klein, E., & Loper, E. (2009). *Natural Language Processing with Python*. O'Reilly Media.
- Burrows, S., Gurevych, I., & Stein, B. (2015). The eras and trends of automatic short answer grading. In *International Journal of Artificial Intelligence in Education* (Vol. 25, Issue 1, pp. 60–117). Springer New York LLC. <https://doi.org/10.1007/s40593-014-0026-8>
- Chamidah, N., Santoni, M. M., Irmanda, H. N., Astriratma, R., & Yulnelly, Y. (2022). Penilaian Esai Pendek Otomatis Berdasarkan Similaritas Semantik dengan SBERT. *Techno.COM Jurnal*, 21(4), 732–740. <https://doi.org/10.33633/tc.v21i4.6758>
- Contreras, J. O., Hilles, S., & Bakar, Z. A. (2021). Essay Question Generator based on Bloom's Taxonomy for Assessing Automated Essay Scoring System. 2021 2nd International Conference on Smart Computing and Electronic Enterprise (ICSCEE), 55–62. <https://doi.org/10.1109/ICSCEE50312.2021.9498166>
- Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1), 21–27. <https://doi.org/10.1109/TIT.1967.1053964>
- Cristianini, Nello, & Shawe-Taylor, J. (2001). An introduction to support vector machines and other kernel-based learning methods. Repr. *Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*, 22. <https://doi.org/10.1017/CBO9780511801389>
- Fitri, M., Widowati, S., & Riskiana, R. R. (n.d.). Implementasi Scenario Based Requirement Engineering Menggunakan Scenario-based Requirement Analysis Method Untuk Karyawan Bagian Program Dan Anggaran KPU Pusat.
- Fitri, R., Noor Asyikin, A., & Pengajar Jurusan Teknik Elektro Politeknik Negeri Banjarmasin Ringkasan, S. (2015). APLIKASI PENILAIAN UJIAN ESSAY OTOMATIS MENGGUNAKAN METODE COSINE SIMILARITY. 7(2), 54–105.
- Garbe, W. (2017). SymSpell: 1 million times faster spelling correction algorithm. <https://github.com/wolfgangarbe/SymSpell>
- Gomiasti, F., Wardo, W., Kartikadarma, E., Gondohanindijo, J., & Setiadi, D. R. I. M. (2024). Enhancing Lung Cancer Classification Effectiveness Through Hyperparameter-Tuned Support Vector Machine. *Journal of Computing Theories and Applications*, 2, 179–189. <https://doi.org/10.62411/jcta.10106>
- Goyal, S. (2021). Effective software defect prediction using support vector machines (SVMs). *International Journal of System Assurance Engineering and Management*, 13. <https://doi.org/10.1007/s13198-021-01326-1>

- Goyal, S. (2022). Handling Class-Imbalance with KNN (Neighbourhood) Under-Sampling for Software Defect Prediction. *Artificial Intelligence Review*, 55. <https://doi.org/10.1007/s10462-021-10044-w>
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques*. <https://doi.org/10.1016/C2009-0-61819-5>
- Hidayat, T., Nurchim, N., & Nurmalitasari, N. (2025a). Analisis Sentimen Ulasan Pengguna pada Aplikasi SIREKAP 2024 Menggunakan Machine Learning. *Smart Comp: Jurnalnya Orang Pintar Komputer*, 14(1). <https://doi.org/10.30591/smartcomp.v14i1.7021>
- Hidayat, T., Nurchim, N., & Nurmalitasari, N. (2025b). Analisis Sentimen Ulasan Pengguna pada Aplikasi SIREKAP 2024 Menggunakan Machine Learning. *Smart Comp: Jurnalnya Orang Pintar Komputer*, 14(1). <https://doi.org/10.30591/smartcomp.v14i1.7021>
- Ifenthaler, D. (2022). Automated Essay Scoring Systems. 1–15. https://doi.org/10.1007/978-981-19-0351-9_59-1
- Larose, D., & Larose, C. (2014). *Discovering Knowledge in Data: An Introduction to Data Mining: Second Edition*. <https://doi.org/10.1002/9781118874059>
- Mufiid, I., Lestanti, S., & Kholila, ma. (2021). APLIKASI PENILAIAN JAWABAN ESAI OTOMATIS MENGGUNAKAN METODE SYNONYM RECOGNITION DAN COSINE SIMILARITY BERBASIS WEB. In *Jurnal MNEMONIC* (Vol. 4, Issue 2).
- Narasimhulu, Y., Kolambkar, P., & China, V. V. (2024). Revisiting Winnow: A modified online feature selection algorithm for efficient binary classification. *Statistical Analysis and Data Mining*, 17(4). <https://doi.org/10.1002/sam.11707>
- Porter, M. F. (1980). An algorithm for suffix stripping. *Program*, 14(3), 130–137.
- Prabowo, H., & Setiawan, A. (2022). Analisis Kinerja Algoritma Winnowing pada Pendeteksian-Plagiarisme. *Jurnal Ilmiah Komputer Dan Informatika*, 10(1), 45–52.
- Pradeep, R., & Kowsalya, M. (2022). An Investigation of Several Models for Machine Learning Based Automated Essay Grading System. <https://doi.org/10.21203/rs.3.rs-1278674/v1>
- Ramli, M., Cokrowibowo, S., & Rustan, M. (2021). Uji Plagiarism pada Tugas Mahasiswa Menggunakan Algoritma Winnowing. *Journal of Applied Computer Science and Technology*, 2, 108–112. <https://doi.org/10.52158/jacost.v2i2.177>
- Schölkopf, B., & Šmola, A. J. (2001). Smola, A.: *Learning with Kernels - Support Vector Machines, Regularization, Optimization and Beyond*. MIT Press, Cambridge, MA. In *Journal of The American Statistical Association - JAMER STATIST ASSN* (Vol. 98).
- Shin, J., & Gierl, M. (2020). More efficient processes for creating automated essay scoring frameworks: A demonstration of two algorithms. *Language Testing*, 38, 026553222093783. <https://doi.org/10.1177/0265532220937830>
- Sihombing, D. O. (2022). Implementasi Natural Language Processing (NLP) dan Algoritma Cosine Similarity dalam Penilaian Ujian Esai Otomatis. *Jurnal Sistem Komputer Dan Informatika (JSON)*, 4(2), 396. <https://doi.org/10.30865/json.v4i2.5374>

- Tala, F. Z. (2009). A study of stemming effects on information retrieval in bahasa Indonesia.
- Team, S. (2016). Sastrawi: Indonesian stemmer. <https://github.com/sastrawi/sastrawi>
- Tifli, D., & Hamir, I. (n.d.). Penerapan Sistem Penilaian Otomatis pada Jawaban Tes Uraian menggunakan Algoritma Rabin Karp untuk Mata Pelajaran Informatika di MA Kanjeng Sepuh.
- Wild, F. (2016). Learning analytics in R with SNA, LSA, and MPIA. In Learning Analytics in R with SNA, LSA, and MPIA. <https://doi.org/10.1007/978-3-319-28791-1>
- Yao, L., & Jiao, H. (2023). Comparing Performance of Feature Extraction Methods and Machine Learning Models in Automatic Essay Scoring. 4(3). <https://doi.org/10.59863/dqiz8440>
- Yudhana, A., Iif Alfiatul Mukaromah, dan, Studi Teknik Elektro, P., Ahmad Dahlan, U., Dahlan, A., & Ji Soepomo, U. (n.d.). IMPLEMENTASI DETEKSI PLAGIARISME MENGGUNAKAN METODE N-GRAM DAN JACCARD SIMILARITY TERHADAP ALGORITMA WINNOWING. In TRANSMISI (Vol. 20, Issue 3). <https://ejournal.undip.ac.id/index.php/transmisi>
- Zupanc, K., & Bosnic, Z. (2017). Automated essay evaluation with semantic analysis. Knowledge-Based Systems, 120. <https://doi.org/10.1016/j.knosys.2017.01.006>

LAMPIRAN

Lampiran 1 Dataset

<https://www.kaggle.com/competitions/learning-agency-lab-automated-essay-scoring-2>

Lampiran 2 Source Code

https://colab.research.google.com/drive/1t_WEOoXcNtJjKeIV-X_rdz4jGzB113f?usp=sharing

