PERBANDINGAN METODE SUPPORT VECTOR MACHINE DAN LOGISTIC REGRESSION UNTUK KLASIFIKASI KATEGORI TUGAS AKHIR
(Studi Kasus: Prodi Sistem Informasi Untan)

DIMAS AQIL SALSABIL
H1101221057

SKRIPSI

PROGRAM STUDI SISTEM INFORMASI
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS TANJUNGPURA
PONTIANAK
2026

PERBANDINGAN METODE SUPPORT VECTOR MACHINE DAN LOGISTIC REGRESSION UNTUK KLASIFIKASI KATEGORI TUGAS AKHIR
(Studi Kasus: Prodi Sistem Informasi Untan)

DIMAS AQIL SALSABIL
H1101221057

Skripsi
Sebagai salah satu syarat untuk memperoleh gelar
Sarjana Komputer pada Program Studi Sistem Informasi

PROGRAM STUDI SISTEM INFORMASI
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS TANJUNGPURA
PONTIANAK
2026

LEMBAR PENGESAHAN

Judul Tugas Akhir	: Perbandingan Metode Support Vector Machine dan Logistic Regression untuk Klasifikasi Kategori Tugas Akhir (Studi Kasus: Prodi Sistem Informasi Untan)
Nama Mahasiswa	: Dimas Aqil Salsabil
NIM	: H1101221057
Jurusan/Program Studi	: Sistem Informasi
Tanggal Lulus	: 29 Juli 2026
SK Pembimbing	: 4024/DST/UN22.8/DT.00.09/2026/27 November 2025
SK Penguji	: 4764/DST/UN22.8/DT.00.09/2026/24 Juli 2026

Dosen Pembimbing

Pembimbing I	Pembimbing II
Dian Prawira, S.T., M.Eng.	Ilhamsyah, S.Si., M.Cs.
NIP. 198411132015041001	NIP. 198405102012121001

Dosen Penguji

Ketua Penguji	Anggota Penguji
Nurul Mutiah, ST., MT.	Renny Puspita Sari, ST., MT.
NIP. 198711182015042002	NIP. 198704182015042001
Pimpinan Sidang (merangkap anggota penguji)	Sekretaris Sidang (merangkap anggota penguji)
Dian Prawira, S.T., M.Eng.	Ilhamsyah, S.Si., M.Cs.
NIP. 198411132015041001	NIP. 198405102012121001

Mengesahkan
Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Tanjungpura

Prof. Dr. Gusrizal, S.Si., M.Si.

PERNYATAAN INTEGRITAS AKADEMIK

Saya yang bertanda tangan di bawah ini:

Nama	: Dimas Aqil Salsabil
NIM	: H1101221057
Program Studi/ Jurusan	: Sistem Informasi
Fakultas	: Matematika dan Ilmu Pengetahuan Alam

dengan ini menyatakan bahwa dokumen ilmiah Tugas Akhir yang disajikan ini tidak mengandung unsur pelanggaran integritas akademik sesuai Peraturan Menteri Pendidikan, Kebudayaan, Riset, Dan Teknologi Republik Indonesia Nomor 39 Tahun 2021. Apabila di kemudian hari dokumen ilmiah Tugas Akhir ini mengandung unsur pelanggaran integritas akademik sesuai ketentuan perundangan tersebut, maka saya bersedia menerima sanksi akademik dan/atau sanksi hukum yang berlaku.

Demikian pernyataan ini untuk digunakan sebagaimana mestinya.

Pontianak, 15 Juni 2026

Dimas Aqil Salsabil

H1101221057

Perbandingan Metode Support Vector Machine dan Logistic Regression untuk Klasifikasi Kategori Tugas Akhir (Studi Kasus: Prodi Sistem Informasi Untan)

Abstrak

Program Studi Sistem Informasi Universitas Tanjungpura menghasilkan dokumen tugas akhir setiap tahunnya yang berasal dari berbagai bidang peminatan dengan karakteristik penelitian yang beragam. Informasi mengenai bidang peminatan tersebut tercermin pada judul dan abstrak, namun belum tersedia model klasifikasi yang mampu mengklasifikasikan dokumen tugas akhir ke dalam kategori bidang peminatan secara otomatis. Kondisi tersebut mendorong perlunya pendekatan klasifikasi teks berbasis *machine learning* untuk membantu menentukan kategori bidang peminatan secara lebih efisien dan konsisten.. *Support Vector Machine* (SVM) dan *Logistic Regression* merupakan dua algoritma yang umum digunakan dalam klasifikasi teks, namun performa keduanya tidak dapat digeneralisasikan pada setiap kasus karena dipengaruhi oleh karakteristik dataset, representasi fitur, dan distribusi kelas. Oleh karena itu, penelitian ini bertujuan membandingkan performa *Support Vector Machine* (SVM) dan *Logistic Regression* dalam mengklasifikasikan kategori bidang peminatan tugas akhir menggunakan data judul dan abstrak mahasiswa Program Studi Sistem Informasi Universitas Tanjungpura. Tahapan penelitian meliputi pengumpulan data, pelabelan kategori bidang peminatan, *preprocessing* teks, pembobotan *Term Frequency–Inverse Document Frequency* (TF-IDF), pelatihan model, serta evaluasi model menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Berdasarkan hasil evaluasi akhir menggunakan pembagian data *train-test split* rasio 70:30, kedua model tersebut terbukti menghasilkan tingkat keberhasilan yang identik. Keduanya sama-sama memperoleh *accuracy* sebesar 93,02%, *precision* sebesar 93,36%, *recall* sebesar 93,02%, dan *F1-score* sebesar 92,47%. Berdasarkan hasil evaluasi tersebut, model *Support Vector Machine* maupun *Logistic Regression* menunjukkan tingkat yang setara dan baik pada dataset yang digunakan. Hasil penelitian ini diharapkan dapat menjadi referensi dalam pemilihan metode klasifikasi kategori tugas akhir pada Program Studi Sistem Informasi Universitas Tanjungpura.

Kata Kunci: Klasifikasi Kategori Tugas Akhir, *Machine Learning*, *Support Vector Machine*, *Logistic Regression*, TF-IDF.

Comparison of Support Vector Machine and Logistic Regression Methods for Final Project Category Classification (Case Study: Information Systems Study Program, Universitas Tanjungpura)

Abstract

The Information Systems Study Program at Universitas Tanjungpura produces final project documents annually that originate from various fields of interest with diverse research characteristics. Information regarding these fields of interest is reflected in the titles and abstracts; however, there is currently no classification model available capable of automatically classifying final project documents into their respective field of interest categories. This condition drives the need for a machine learning-based text classification approach to help determine the field of interest categories more efficiently and consistently. Support Vector Machine (SVM) and Logistic Regression are two commonly used algorithms in text classification, but their performance cannot be generalized across all cases as it is influenced by dataset characteristics, feature representation, and class distribution. Therefore, this research aims to compare the performance of Support Vector Machine (SVM) and Logistic Regression in classifying the final project's field of interest categories using the title and abstract data of students from the Information Systems Study Program at Universitas Tanjungpura. The research stages include data collection, labeling of field of interest categories, text preprocessing, Term Frequency–Inverse Document Frequency (TF-IDF) weighting, model training, and model evaluation using accuracy, precision, recall, and F1-score metrics. Based on the final evaluation results using a 70:30 train-test split ratio, both models were proven to produce identical success rates. Both achieved an accuracy of 93.02%, a precision of 93.36%, a recall of 93.02%, and an F1-score of 92.47%. Based on these evaluation results, both the Support Vector Machine and Logistic Regression models demonstrated an equivalent and good level of performance on the dataset used. The results of this study are expected to serve as a reference in selecting the classification method for final project categories at the Information Systems Study Program, Universitas Tanjungpura.

Keywords*: Final Project Classification, Machine Learning, Support Vector Machine, Logistic Regression, TF-IDF.*

PRAKATA

Segala puji dan syukur penulis panjatkan dengan penuh kerendahan hati ke hadirat Allah Subhanahu wa Ta'ala atas rahmat dan hidayah-Nya penulis mendapatkan kemudahan dan kelancaran dalam penyusunan skripsi penulis yang berjudul, "Perbandingan Metode *Support Vector Machine* dan *Logistic Regression* untuk Klasifikasi Kategori Tugas Akhir (Studi Kasus: Prodi Sistem Informasi Untan)." Penulis begitu bersyukur dapat menyelesaikan skripsi ini, dan penulis berharap skripsi ini dapat memberikan manfaat dan menjadi referensi yang membantu bagi pembaca. Penulis juga menyampaikan rasa terima kasih yang sebesar-besarnya kepada pihak-pihak yang telah mendukung dan membantu kelangsungan proses belajar serta selama penyusunan skripsi ini, terutama kepada:

1. Orang tua dan keluarga yang senantiasa selalu memberikan dukungan baik secara materil maupun moril yang telah menjadi fondasi utama dalam setiap langkah penulis.
2. Bapak Dian Prawira, S.T., M.Eng., selaku dosen pembimbing akademik sekaligus dosen pembimbing I, atas waktu, bimbingan, serta arahan yang telah diberikan selama proses penyusunan tugas akhir maupun selama bimbingan akademik.
3. Bapak Ilhamsyah, S.Si., M.Cs., selaku dosen pembimbing II, atas waktu, bimbingan, serta arahan yang telah diberikan selama proses penyusunan tugas akhir.
4. Ibu Nurul Mutiah, ST., MT., selaku dosen penguji I, atas banyak saran dan masukan yang telah diberikan selama proses penyusunan tugas akhir.
5. Ibu Renny Puspita Sari, ST., MT., selaku dosen penguji II, atas banyak saran dan masukan yang telah diberikan selama proses penyusunan tugas akhir.
6. Bapak dn Ibu, serta staf Program Studi Sistem Informasi yang telah memberikan ilmu pengetahuan, dukungan, dan bantuan selama masa perkuliahan.

7. Teman-teman penulis pada grup *profile*, yang telah memberikan semangat, bantuan, serta menjadi tempat berbagi cerita selama proses awal masuk perkuliahan maupun penyusunan tugas akhir.
8. Teman-teman Program Studi Sistem Informasi Angkatan 2022 (*Matrix*) atas kebersamaan, bantuan, dan semangat saling mendukung selama masa perkuliahan hingga penyusunan tugas akhir.

Pontianak, 1 Juli 2026

Dimas Aqil Salsabil
H1101221057

DAFTAR ISI

Halaman

LEMBAR PENGESAHAN ii
PERNYATAAN INTEGRITAS AKADEMIK iii
Abstrak iv
Abstract v
PRAKATA vi
DAFTAR ISI viii
DAFTAR GAMBAR xi
DAFTAR TABEL xii
DAFTAR PERSAMAAN xiv
BAB I PENDAHULUAN 1
1.1 Latar Belakang 1
1.2 Rumusan Masalah 3
1.3 Tujuan Penelitian 3
1.4 Batasan Masalah 3
1.5 Manfaat Penelitian 4
1.6 Sistematika Penulisan 5
BAB II LANDASAN TEORI 7
2.1 Landasan Teori 7
2.1.1 Prodi Sistem Informasi Untan 7
2.1.2 Tugas Akhir 8
2.1.3 *Machine learning* 8
2.1.4 *Text mining* 9
2.1.5 Klasifikasi Teks 9
2.1.6 *Term Frequency-Inverse Document Frequency* (TF-IDF) 9
2.1.7 *Support Vector Machine* (SVM) 10
2.1.8 *Logistic Regression* 12
2.1.9 *Confusion Matrix* 13
2.1.10 *Google Colab* 14

2.1.11 *Python* 14
2.2 Penelitian Terdahulu 15
BAB III METODOLOGI PENELITIAN 17
3.1 Bahan Penelitian 17
3.2 Alat Penelitian 17
3.3 Kerangka Kerja Penelitian 17
3.3.1 *Environment* 19
3.3.2 *IS Research* 19
3.3.3 *Knowledge Base* 20
3.4 Metode Penelitian 21
3.5 Data Penelitian 25
3.5.1 Objek Penelitian 25
3.5.2 Populasi dan Sampel 25
3.5.3 Jenis dan Sumber Data 26
BAB IV HASIL DAN PEMBAHASAN 27
4.1 Pengumpulan Dokumen Tugas Akhir 27
4.2 Ekstraksi Judul dan Abstrak 27
4.3 Penyusunan Dataset Penelitian 31
4.4 *Labelling* 33
4.5 Data *Preprocessing* 42
4.5.1 Normalisasi Teks dan Angka 43
4.5.2 Pembersihan Data 45
4.5.3 *Stopword removal* 47
4.5.4 *Stemming* 49
4.6 *Encoding Label* 51
4.7 *Train-test Split* 52
4.8 *Stratified K-Fold Cross Validation* 52
4.9 Ekstraksi Fitur TF-IDF 55
4.10 *Hyperparameter Tuning* 57
4.10.1 Hasil *Hyperparameter Tuning* SVM 58
4.10.2 Hasil *Hyperparameter Tuning Logistic Regression* 59

4.11 Pelatihan dan Evaluasi Model 60
4.11.1 *Support Vector Machine* (SVM) 61
4.11.2 *Logistic Regression* 65
BAB V PENUTUP 69
5.1 Kesimpulan 69
5.2 Saran 69
DAFTAR PUSTAKA 71
LAMPIRAN 74

DAFTAR GAMBAR

Halaman

Gambar 2. 1 Struktur Organisasi Jurusan Sistem Informasi Untan 7
Gambar 3. 1 Kerangka Kerja Penelitian 18
Gambar 3. 2 *Flowchart* Penelitian 22
Gambar 4. 1 *Confusion Matrix* SVM 63
Gambar 4. 2 *Grafik Metrik Evaluasi* SVM 64
Gambar 4. 3 *Confusion Matrix Logistic Regression* 66
Gambar 4. 4 *Grafik Metrik Evaluasi Logistic Regression* 67

DAFTAR TABEL

Halaman

Tabel 2. 1 *Confusion Matrix* 13
Tabel 2. 2 Penelitian Terdahulu 15
Tabel 4. 1 Struktur Data Judul dan Abstrak Tugas Akhir 28
Tabel 4. 2 Data Dokumen Penelitian Mahasiswa 29
Tabel 4. 3 Dataset Teks Sebelum Penggabungan 32
Tabel 4. 4 Dataset Teks Hasil Penggabungan 32
Tabel 4. 5 *Mapping* Kategori Bidang Peminatan ke Label Klasifikasi 35
Tabel 4. 6 Hasil Pelabelan Dataset Teks 42
Tabel 4. 7 Hasil Normalisasi Teks dan Angka 43
Tabel 4. 8 Hasil Teks Setelah Normalisasi dan Penghapusan URL 45
Tabel 4. 9 Hasil Teks Setelah Penghapusan Simbol dan Normalisasi Spasi 47
Tabel 4. 10 Hasil Teks Setelah *Stopword Removal* 48
Tabel 4. 11 Hasil Teks Setelah *Stemming* 50
Tabel 4. 12 Pembagian Data *Stratified 5-Fold CV* 54
Tabel 4. 13 Pembagian Data *Stratified 10-Fold CV* 54
Tabel 4. 14 Perbandingan Skenario Hasil *Tuning* Terbaik SVM 59
Tabel 4. 15 Perbandingan Skenario Hasil *Tuning* Terbaik *Logistic Regression* 59

DAFTAR LISTING CODE

Halaman

Listing Code 4.1 *Load Data* dan Penggabungan Teks 31
Listing Code 4. 2 Normalisasi Teks dan Angka 43
Listing Code 4. 3 Menghapus URL 45
Listing Code 4. 4 Penghapusan Simbol dan Normalisasi Spasi 46
Listing Code 4. 5 Proses *Stopword Removal* Menggunakan Pustaka Sastrawi 47
Listing Code 4. 6 *Stemming* 49
Listing Code 4. 7 *Encoding Label* 51
Listing Code 4. 8 *Train-test Split* 52
Listing Code 4. 9 *Stratified 5-Fold Cross Validation* 53
Listing Code 4. 10 *Stratified 10-Fold Cross Validation* 54
Listing Code 4. 11 Ekstraksi Fitur TF-IDF 56
Listing Code 4. 12 *Hyperparameter Tuning TF-IDF* dan *SVM* 58
Listing Code 4. 13 Model *Support Vector Machine* 62
Listing Code 4. 14 Model *Logistic Regression* 65

DAFTAR PERSAMAAN

Halaman

Persamaan 2.1 TF-IDF 11
Persamaan 2.2 *Hyperplane* SVM 12
Persamaan 2.3 Fungsi *Sigmoid* 13
Persamaan 2.4 *Accuracy* 15
Persamaan 2.5 *Precision* 15
Persamaan 2.6 *Recall* 15
Persamaan 2.7 *F1-Score* 15

BAB I
PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi berlangsung sangat pesat di berbagai aspek kehidupan, memberikan pengaruh besar terhadap cara manusia berkomunikasi, bekerja, serta memandang dan memahami dunia. Berbagai aspek kehidupan manusia semakin dipengaruhi oleh kemajuan teknologi informasi (Muttaqin et al., 2021). Penerapan Teknologi Informasi dan Komunikasi (TIK) dalam pendidikan berpengaruh secara signifikan, berfungsi sebagai media pembelajaran yang komprehensif untuk meningkatkan kualitas pendidikan sekaligus memperluas akses bagi masyarakat (Huraerah et al., 2021).

Program Studi Sistem Informasi di Universitas Tanjungpura (UNTAN) merupakan salah satu program sarjana yang berada di bawah naungan Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA). Kampus Universitas Tanjungpura terletak di Jalan Prof. Dr. H. Hadari Nawawi, Pontianak, Kalimantan Barat. Sejak resmi berdiri pada tahun 2014, program studi ini telah berkomitmen untuk menghasilkan lulusan yang memiliki kemampuan di bidang teknologi informasi, bisnis, dan manajemen. Tugas akhir adalah karya tulis ilmiah berupa laporan hasil penelitian yang membahas suatu masalah pada bidang ilmu tertentu secara sistematis dengan menggunakan kaidah-kaidah yang berlaku (Fitria et al., 2024). Sebagai bagian dari standar akademik, mahasiswa tingkat akhir Program Studi Sistem Informasi Universitas Tanjungpura harus menyelesaikan tugas akhir sebagai salah satu syarat kelulusan.

Berdasarkan data dokumen tugas akhir Program Studi Sistem Informasi Universitas Tanjungpura yang diperoleh dari repositori Google Classroom Tugas Akhir Sisfo, terdapat 122 dokumen pada periode 2014–2017 dan 164 dokumen pada periode

2018–2021, sehingga total data yang tersedia berjumlah 286 dokumen. Dokumen tersebut memiliki karakteristik penelitian yang beragam, baik dari sisi topik, metode, maupun fokus penelitian. Penentuan kategori bidang peminatan tidak selalu mudah dilakukan secara manual, terutama pada dokumen yang memiliki karakteristik lintas bidang atau menggunakan lebih dari satu pendekatan penelitian. Oleh karena itu, diperlukan suatu pendekatan klasifikasi berbasis *machine learning* yang mampu menentukan kategori bidang peminatan secara lebih efisien, konsisten, dan akurat.

Salah satu pendekatan yang dapat digunakan untuk mengatasi permasalahan tersebut adalah *text mining*, khususnya melalui teknik klasifikasi teks berbasis *machine learning* (Wijaya et al., 2024). Dengan memanfaatkan informasi yang terdapat pada judul dan abstrak, algoritma *machine learning* dapat mempelajari pola dari setiap kategori sehingga dokumen dapat diklasifikasikan secara otomatis ke dalam kategori peminatan.

Hasil penelitian Jahan et al. (2024) menunjukkan bahwa *Support Vector Machine* (SVM) dan *Logistic Regression* memiliki performa yang baik dalam klasifikasi teks, dibandingkan algoritma *machine learning* lain yang diuji. Namun, performa kedua metode tersebut tidak dapat digeneralisasikan pada semua kasus karena dipengaruhi oleh karakteristik dataset, representasi fitur, distribusi kelas, serta domain permasalahan yang digunakan. Oleh karena itu, diperlukan studi komparatif untuk mengevaluasi kedua metode pada klasifikasi kategori tugas akhir Program Studi Sistem Informasi Universitas Tanjungpura

Berdasarkan permasalahan tersebut, penelitian ini dilakukan untuk menganalisis dan membandingkan performa metode *Support Vector Machine* (SVM) dan *Logistic Regression* dalam menentukan kategori pada dokumen tugas akhir berdasarkan judul dan abstrak menggunakan representasi fitur *Term Frequency–Inverse Document Frequency* (TF-IDF). Hasil penelitian diharapkan dapat menjadi referensi dalam pemilihan metode yang sesuai untuk klasifikasi dokumen tugas akhir serta menjadi landasan bagi penerapan model klasifikasi otomatis pada sistem akademik di penelitian selanjutnya.

1.2 Rumusan Masalah

Berdasarkan latar belakang permasalahan diatas, maka didapatlah beberapa rumusan masalah yang digunakan sebagai landasan penelitian yaitu:

1. Bagaimana cara melakukan kategorisasi tugas akhir mahasiswa pada Program Studi Sistem Informasi Universitas Tanjungpura?
2. Bagaimana performa metode yang digunakan dalam melakukan kategorisasi tugas akhir mahasiswa pada Program Studi Sistem Informasi Universitas Tanjungpura?

1.3 Tujuan Penelitian

Adapun tujuan dari adanya penelitian ini antara lain:

1. Membangun model klasifikasi untuk menentukan kategori tugas akhir mahasiswa berdasarkan data judul dan abstrak pada Program Studi Sistem Informasi Universitas Tanjungpura.
2. Menganalisis dan membandingkan performa metode *Support Vector Machine* (SVM) dan *Logistic Regression* dalam melakukan kategorisasi tugas akhir mahasiswa pada Program Studi Sistem Informasi Universitas Tanjungpura.

1.4 Batasan Masalah

Adapun batas masalah dari penelitian antara lain:

1. Data yang digunakan dalam penelitian ini berupa dokumen tugas akhir mahasiswa Program Studi Sistem Informasi Universitas Tanjungpura yang diperoleh dari repositori *Google Classroom* Tugas Akhir Sisfo.
2. Dataset yang digunakan pada tahap pelatihan model berupa gabungan judul dan abstrak dari masing-masing tugas akhir, sehingga tidak mencakup isi lengkap dokumen seperti bab pendahuluan, pembahasan, kesimpulan, maupun lampiran.

3. Label yang digunakan dibatasi pada bidang peminatan Program Studi Sistem Informasi Universitas Tanjungpura, meliputi: IT Governance, Business Intelligence dan Data Analytics, dan Software Engineering. Setiap dokumen diklasifikasikan ke dalam satu kategori peminatan utama (*single-label classification*).
4. Model *Machine Learning* yang digunakan dalam penelitian ini difokuskan pada perbandingan metode *Support Vector Machine* (SVM) dan *Logistic Regression*.

1.5 Manfaat Penelitian

Adapun manfaat dengan adanya penelitian ini antara lain:

1. Bagi Peneliti

   Penelitian ini menjadi sarana bagi peneliti untuk mengembangkan wawasan, keterampilan, dan pengalaman dalam menerapkan model *machine learning* untuk klasifikasi tugas akhir secara otomatis. Selain itu, penelitian ini memberikan kesempatan bagi peneliti untuk mendalami penerapan metode *Support Vector Machine* (SVM) dan *Logistic Regression* dalam pengolahan data teks, khususnya pada klasifikasi dokumen berdasarkan kategori bidang peminatan.
2. Bagi Bidang Ilmu (Sistem Informasi)

   Penelitian ini diharapkan dapat memberikan kontribusi terhadap pengembangan ilmu Sistem Informasi, khususnya pada bidang *text mining* dan *machine learning* untuk klasifikasi teks. Hasil penelitian dapat menjadi referensi dalam pemilihan metode klasifikasi yang sesuai untuk dokumen akademik serta menjadi acuan bagi penelitian selanjutnya yang berkaitan dengan klasifikasi kategori tugas akhir.
3. Bagi Pengguna

   Penelitian ini diharapkan dapat memberikan informasi mengenai performa metode *Support Vector Machine* (SVM) dan *Logistic Regression* dalam

mengklasifikasikan kategori tugas akhir berdasarkan judul dan abstrak. Hasil penelitian dapat menjadi bahan pertimbangan bagi Program Studi Sistem Informasi Universitas Tanjungpura dalam menentukan metode klasifikasi yang sesuai di masa mendatang dilakukan pengembangan mekanisme atau fitur klasifikasi kategori tugas akhir.

1.6 Sistematika Penulisan

Penyusunan laporan tugas akhir ini disusun pada pedoman penulisan ilmiah Program Studi Sistem Informasi. Laporan disusun secara terstruktur ke dalam lima bab utama yang saling berhubungan untuk mendukung alur penelitian secara menyeluruh. Adapun sistematika penulisannya dijelaskan sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi penjelasan awal mengenai penelitian yang dilakukan. Pembahasan mencakup latar belakang penelitian, rumusan masalah, tujuan penelitian, ruang lingkup atau batasan penelitian, manfaat yang diharapkan, serta gambaran umum mengenai susunan laporan tugas akhir.

BAB II LANDASAN TEORI

Bab ini berisi pembahasan mengenai teori-teori yang mendasari penelitian, mencakup kajian pustaka, referensi ilmiah, serta penelitian terdahulu yang relevan sebagai acuan dalam mendukung proses analisis dalam penelitian ini.

BAB III METODOLOGI PENELITIAN

Bab III Metodologi Penelitian menjelaskan secara rinci terkait langkah-langkah yang digunakan pada penelitian ini mulai dari bahan penelitian, alat penelitian, kerangka kerja penelitian, serta data penelitian untuk mencapai tujuan dari penelitian ini.

BAB IV HASIL DAN PEMBAHASAN

Bab ini membahas hasil penelitian yang telah dilakukan, meliputi proses pengumpulan dan persiapan dataset, *preprocessing* data, ekstraksi fitur menggunakan

Term Frequency–Inverse Document Frequency (TF-IDF), pelatihan model *Support Vector Machine* (SVM) dan *Logistic Regression*, evaluasi performa model menggunakan metrik yang telah ditentukan, serta analisis perbandingan hasil klasifikasi kedua metode.

BAB V PENUTUP

Bab ini memuat kesimpulan dari hasil penelitian yang telah dilakukan serta saran untuk pengembangan penelitian dan sistem pada tahap selanjutnya.