PREDIKSI MAHASISWA YANG BERPOTENSI MELAKUKAN PENGUNDURAN DIRI MENGGUNAKAN ALGORITMA *SUPPORT VECTOR MACHINE*

SKRIPSI

Program Studi Sarjana Informatika
Jurusan Informatika

Oleh:
SEPTIAN DWIRIZKIANTO
NIM D1041201014



FAKULTAS TEKNIK
UNIVERSITAS TANJUNGPURA
PONTIANAK
2025

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Septian Dwirizkianto

NIM : D1041201014

menyatakan bahwa dalam SKRIPSI yang berjudul "Prediksi Mahasiswa yang Berpotensi Melakukan Pengunduran Diri Menggunakan Algoritma *Support Vector Machine*" tidak terdapat karya yang pernah diajukan untuk memperoleh gelar sarjana di suatu perguruan tinggi manapun. Sepanjang pengetahuan Saya, tidak terdapat karya atau pendapat yang pernah ditulis atas diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam Daftar Rujukan.

Demikian pernyataan ini dibuat dengan sebenar-benarnya. Saya sanggup menerima konsekuensi akademis dan hukum di kemudian hari apabila pernyataan yang dibuat ini tidak benar.

Pontianak, 17 Juni 2025



Septian Dwirizkianto
NIM D1041201014

KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS TANJUNGPURA
FAKULTAS TEKNIK
Jalan Prof. Dr. H. Hadari Nawawi Pontianak 78124
Telepon (0561) 740186, WA: +6282152280907
Email : ft@untan.ac.id Website : http://teknik.untan.ac.id

HALAMAN PENGESAHAN

PREDIKSI MAHASISWA YANG BERPOTENSI MELAKUKAN PENGUNDURAN DIRI MENGGUNAKAN ALGORITMA *SUPPORT VECTOR MACHINE*

Program Studi Sarjana Informatika
Jurusan Informatika

Oleh:

Septian Dwirizkianto
NIM D1041201014

Telah dipertahankan di depan Penguji Skripsi pada tanggal 17 Juni 2025 dan diterima sebagai salah satu persyaratan untuk memperoleh gelar sarjana.

Susunan Penguji Skripsi:

Dosen Pembimbing Utama : Tursina, S.T., M.Cs.
NIP. 197801152002122003

Dosen Pembimbing Kedua : Niken Candraningrum, S.T., M.Cs.
NIP. 199309272022032010

Dosen Penguji Utama : Rina Septiriana, S.T., M.Cs.
NIP. 198709232020122001

Dosen Penguji Kedua : Ir. M. Azhar Irwansyah, S.T., M.Eng.
NIP. 198506062008121002

Pontianak, 17 Juni 2025

Dekan,

Dr. -Ing. Ir. Slamet Widodo, M.T., IPM.
NIP. 196712231992031002

Pembimbing Utama

Tursina, S.T., M.Cs.
NIP. 197801152002122003

HALAMAN PERSEMBAHAN

Alhamdulillahi Rabbil Aalamin sujud serta syukur kepada Allah SWT atas karunia-Nya yang telah memberikan kemudahan dan kelancaran sehingga skripsi ini dapat terselesaikan dengan baik.

Skripsi ini saya persembahkan kepada dua orang hebat dalam hidup saya yaitu kedua orang tua saya yang telah membuat segalanya menjadi mungkin sehingga saya bisa sampai pada tahap di mana skripsi ini akhirnya selesai. Terima kasih atas segala pengorbanan, nasihat dan doa terbaik yang tidak pernah berhenti diberikan kepada saya.

Terimakasih juga untuk saudara dan keluarga yang telah memberikan doa dan dukungan selama penyelesaian skripsi ini

Serta saya ucapkan terima kasih untuk seluruh dosen, teman, dan senior yang telah banyak membantu dan memberikan dukungan selama ini.

Terima kasih banyak untuk semuanya yang telah mendukung dan menyemangati dalam perjuangan ini

KATA PENGANTAR

Dengan memanjatkan puji dan syukur kehadirat Allah SWT atas berkah dan rahmat-Nya sehingga penulis dapat menyelesaikan penelitian dan penulisan skripsi dengan judul "Prediksi Mahasiswa yang Berpotensi Melakukan Pengunduran Diri Menggunakan Algoritma *Support Vector Machine* (Studi Kasus: Fakultas Teknik Universitas Tanjungpura)" sebagai salah satu syarat untuk memperoleh gelar Sarjana Program Studi Informatika, Fakultas Teknik, Universitas Tanjungpura.

Penulis menyadari bahwa skripsi ini tidak mungkin terselesaikan dengan baik tanpa bantuan, dukungan, dan bimbingan dari berbagai pihak selama penyusunan skripsi ini. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan rasa terima kasih yang sebesar - besarnya kepada Ibu Tursina, S.T., M.Cs. selaku dosen pembimbing utama dan Ibu Niken Candraningrum, S.T., M.Cs. selaku dosen pembimbing pendamping serta seluruh dosen Informatika yang telah memberikan ilmu dan pengetahuan selama penulis menempuh pendidikan.

Penulis menyadari masih banyak terdapat kekurangan dalam penyusunan skripsi ini, oleh karena itu dengan kerendahan hati penulis mengharapkan kritik dan saran demi kesempurnaan penyusunan skripsi ini. Penulis berharap penelitian ini dapat bermanfaat bagi perkembangan ilmu pengetahuan, bangsa dan negara.

Pontianak, 17 Juni 2025



Septian Dwirizkianto

NIM. D1041201014

ABSTRAK

Pengunduran diri mahasiswa merupakan isu krusial yang berdampak pada individu, institusi, dan kualitas pendidikan secara keseluruhan. Penelitian ini bertujuan untuk memprediksi mahasiswa yang berpotensi mengundurkan diri menggunakan algoritma Support Vector Machine (SVM). Data yang digunakan berasal dari mahasiswa Fakultas Teknik Universitas Tanjungpura, mencakup data akademik dan non-akademik mahasiswa dari angkatan 2014 hingga 2021. Metodologi penelitian mencakup pengumpulan data, preprocessing data, pengembangan model, pengujian, serta evaluasi kinerja model. Proses preprocessing mencakup pemilihan atribut, pembersihan data menggunakan Logical Imputation Listwise Deletion, transformasi data, ekstraksi fitur menggunakan Principal Component Analysis (PCA), serta normalisasi dengan metode Min-Max Normalization dan Standard Scaler. Model dikembangkan menggunakan algoritma SVM dan hyperparameter tuning, kemudian diuji menggunakan Mean Squared Error (MSE), Root Mean Square Error (RMSE), serta Confusion Matrix. Berdasarkan hasil evaluasi, model SVM menunjukkan performa yang baik dengan akurasi data testing sebesar 93,54%, MSE sebesar 0,07317, dan RMSE sebesar 0,27047. Selain itu, hasil prediksi menunjukkan tingkat kesesuaian yang tinggi dengan data aktual, yaitu sebesar 96,13%, di mana dari 83 mahasiswa yang benar-benar mengundurkan diri, sebanyak 74 mahasiswa berhasil terdeteksi melalui model, dan 9 mahasiswa tidak terdeteksi. Analisis lebih lanjut menunjukkan bahwa mahasiswa yang masuk melalui jalur mandiri, lulusan dari SMA, serta mereka yang mengalami penurunan Indeks Prestasi Semester (IPS) dari semester 1 hingga 4 memiliki kecenderungan lebih tinggi untuk mengundurkan diri. Temuan ini diharapkan dapat menjadi acuan bagi pihak institusi dalam menyusun strategi pencegahan dini untuk menekan angka pengunduran diri mahasiswa.

Kata Kunci: Pengunduran Diri Mahasiswa, Support Vector Machine (SVM), Prediksi Mahasiswa, Machine Learning, Data Mining.

ABSTRACT

Student resignation is a crucial issue that impacts individuals, institutions, and the overall quality of education. This research aims to predict students who have the potential to resign using the Support Vector Machine (SVM) algorithm. The data used comes from students of the Faculty of Engineering, Tanjungpura University, including academic and non-academic data of students from batch 2014 to 2021. The research methodology includes data collection, data preprocessing, model development, testing, and model performance evaluation. The preprocessing process includes attribute selection, data cleaning using Logical Imputation Listwise Deletion, data transformation, feature extraction using Principal Component Analysis (PCA), and normalization using the Min-Max Normalization and Standard Scaler methods. The model is developed using SVM algorithm and hyperparameter tuning, then tested using Mean Squared Error (MSE), Root Mean Square Error (RMSE), and Confusion Matrix. Based on the evaluation results, the SVM model shows good performance with testing data accuracy of 93.54%, MSE of 0.07317, and RMSE of 0.27047. In addition, the prediction results show a high level of conformity with the actual data, which is 96.13%, where out of 83 students who actually resigned, 74 students were successfully detected through the model, and 9 students were not detected. Further analysis shows that students who enter through independent pathways, graduates from high school, and those who experience a decrease in the Semester Achievement Index (IPS) from semester 1 to 4 have a higher tendency to resign. This finding is expected to be a reference for institutions in developing early prevention strategies to reduce student withdrawal rates.

Keywords: Student Resignation, Support Vector Machine (SVM), Student Prediction, Machine Learning, Data Mining.

DAFTAR ISI

HALAMAN PERNYATAAN ..... **ii**

HALAMAN PENGESAHAN ..... **iii**

HALAMAN PERSEMBAHAN ..... **iv**

KATA PENGANTAR ..... **v**

ABSTRAK ..... **vi**

ABSTRACT ..... **vii**

DAFTAR ISI ..... **viii**

DAFTAR TABEL ..... **xii**

DAFTAR GAMBAR ..... **xiii**

DAFTAR LAMPIRAN ..... **xv**

BAB I PENDAHULUAN **1**

1.1 Latar Belakang 1

1.2 Perumusan Masalah 2

1.3 Tujuan Penelitian 2

1.4 Pembatasan Masalah 2

1.5 Sistematika Penulisan 3

BAB II TINJAUAN PUSTAKA **4**

2.1 Studi Literatur 4

2.2 Data Mining 5

2.3 Klasifikasi 8

2.4 *Support Vector Machine* (SVM) 8

2.5 Confusion Matrix 12

2.6 Mean Square Error (MSE) 13

2.7 *Root Mean Square Error* (RMSE) 13

BAB III METODOLOGI PENELITIAN 15
3.1 Perangkat Penelitian 15
3.1.1 Perangkat Keras 15
3.1.2 Perangkat Lunak 15
3.2 Metodologi Penelitian 15
3.2.1 Studi Literatur 16
3.2.2 Pengumpulan Data 16
3.2.3 Preprocessing Data 19
3.2.3.1 Pemilihan Atribut 19
3.2.3.2 Pembersihan Data 20
3.2.3.3 Transformasi Data 20
3.2.3.4 Normalisasi Data 21
3.2.3.4.1 *Min-max Normalization* 21
3.2.3.4.2 *Standard Scaler* 22
3.2.3.5 Penyeimbangan Data 22
3.2.3.6 Ekstraksi Fitur 23
3.2.4 Pengembangan Model 23
3.2.4.1 Pembagian Data 23
3.2.4.2 Pemisahan Atribut Identitas 24
3.2.4.3 Pemisahan Fitur dan Target 24
3.2.4.4 *Hyperparameter Tuning* 24
3.2.4.5 Pelatihan 25
3.2.4.5.1 Algoritma *Support Vector Machine* (SVM) 25
3.2.4.6 Pengujian 27
3.2.4.6.1 Skenario Pengujian 27
3.2.5 Evaluasi 28

3.2.6 Prediksi 28
BAB IV HASIL DAN ANALISIS 30
4.1 Pengumpulan Data 30
4.2 Preprocessing 31
4.2.1 Pembersihan Data 31
4.2.2 Transformasi Data 38
4.2.3 Normalisasi Data 40
4.2.3.1 Min-max Normalization 40
4.2.3.2 *Standard Scaler* 41
4.2.4 Penyeimbangan Data 42
4.2.5 Ekstraksi Fitur 42
4.3 Pengembangan Model 43
4.3.1 Pembagian Data 43
4.3.2 Pemisahan Atribut Identitas 44
4.3.3 Pemisahan Fitur dan Target 45
4.3.4 Inisialisasi Model 45
4.4 Evaluasi 45
4.4.1 Skenario 1 45
4.4.2 Skenario 2 50
4.4.3 Skenario 3 51
4.4.4 Skenario 4 52
4.4.5 Perbandingan Hasil 53
4.5 Prediksi 54
4.5.1 Hasil Prediksi 54
4.5.2 Evaluasi Prediksi 56
BAB V KESIMPULAN DAN SARAN 62

5.1 Kesimpulan 62
5.2 Saran 62
DAFTAR PUSTAKA..... 63

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu 4
Tabel 2. 2 Confusion Matrix 12
Tabel 3. 1 Pemilihan Atribut 19
Tabel 3. 2 Pengkategorian Dataset 20
Tabel 3. 3 Pembagian Dataset 23
Tabel 3. 4 Konfigurasi Grid Search untuk Hyperparameter Tuning 25
Tabel 3. 5 Skenario Pengujian 28
Tabel 4. 1 Data Mahasiswa Mengundurkan Diri Berdasarkan Program Studi 31
Tabel 4. 2 Atribut Data 32
Tabel 4. 3 Data Mahasiswa Mengundurkan Diri Berdasarkan Program Studi Setelah Pembersihan Data 38
Tabel 4. 4 Pembagian Data Training dan Data Testing 44
Tabel 4. 5 Pemisahan Atribut Identitas 44
Tabel 4. 6 Hasil Evaluasi Skenario 1 Komposisi 60% : 40% 46
Tabel 4. 7 Hasil Evaluasi Skenario 1 Komposisi 70% : 30% 47
Tabel 4. 8 Hasil Evaluasi Skenario 1 Komposisi 80% : 20% 48
Tabel 4. 9 Hasil Evaluasi Skenario 1 Komposisi 90% : 10% 49
Tabel 4. 10 Hasil Evaluasi Skenario 2 50
Tabel 4. 11 Hasil Evaluasi Skenario 3 51
Tabel 4. 12 Hasil Evaluasi Skenario 4 52
Tabel 4. 13 Perbandingan Hasil Evaluasi 53
Tabel 4. 14 Hasil Evaluasi Terbaik 54
Tabel 4. 15 Rekapitulasi Data Aktual dan Prediksi Mahasiswa Mengundurkan Diri 59
Tabel 4. 16 Confusion Matrix Hasil Prediksi 60

DAFTAR GAMBAR

Gambar 2. 1 Tahapan Pada Data Mining 6
Gambar 2. 2 Konsep Hyperplane (Monika Parapat & Tanzil Furqon, 2018) 9
Gambar 2. 3 Hyperplane (Pratiwi & Setyawan, 2021) 10
Gambar 3. 1 Diagram Alir Penelitian 16
Gambar 3. 2 Dataframe Keseluruhan Data 17
Gambar 3. 3 Dataframe Mahasiswa Angkatan 2014 - 2020 18
Gambar 3. 4 Dataframe Mahasiswa Angkatan 2021 18
Gambar 3. 5 Diagram Alir Support Vector Machine 26
Gambar 3. 6 Alur Skenario Pengujian 27
Gambar 4. 1 Data Mahasiswa Mengundurkan Diri dan Tidak Mengundurkan Diri 30
Gambar 4. 2 Persentase Missing Value 33
Gambar 4. 3 Sampel Data Sebelum Proses Logical Imputation 34
Gambar 4. 4 Sampel Data Setelah Proses Logical Imputation 34
Gambar 4. 5 Persentase Missing Value 35
Gambar 4. 6 Sampel Data Hilang 36
Gambar 4. 7 Jumlah Missing Value 36
Gambar 4. 8 Data Mahasiswa Setelah Pembersihan Data 37
Gambar 4. 9 Dataframe Sebelum Transformasi Data 39
Gambar 4. 10 Dataframe Setelah Transformasi Data 40
Gambar 4. 11 Dataframe Setelah Normalisasi Data Menggunakan Min-max Normalization 41
Gambar 4. 12 Dataframe Setelah Normalisasi Data Menggunakan Standard Scaler 42
Gambar 4. 13 Dataframe Setelah Ekstraksi Fitur Menggunakan Principal Component Analysis (PCA) 43
Gambar 4. 14 Dataframe Mahasiswa Angkatan 2021 55
Gambar 4. 15 Dataframe Hasil Prediksi 55
Gambar 4. 16 Klasifikasi Hasil Prediksi Mahasiswa Angkatan 2021 56
Gambar 4. 17 Dataframe Aktual Mahasiswa Angkatan 2021 57

Gambar 4. 18 Data Aktual Mahasiswa Angkatan 2021 Berdasarkan Status Keluar 58
Gambar 4. 19 Deteksi Mahasiswa Mengundurkan Diri Berdasarkan Hasil Prediksi..... 59

DAFTAR LAMPIRAN

LAMPIRAN A DATASET MAHASISWA A1

LAMPIRAN B DATA AKTUAL MAHASISWA MENGUNDURKAN DIRI ANGKATAN 2021 ..... B1

LAMPIRAN C DATA HASIL PREDIKSI MAHASISWA MENGUNDURKAN DIRI ANGKATAN 2021 ..... C1

BAB I
PENDAHULUAN

1.1 Latar Belakang

Pendidikan Perguruan Tinggi di Indonesia terus mengalami peningkatan setiap tahunnya. Berdasarkan data dari Badan Pusat Statistik (BPS) dan Kemendikbudristek, jumlah mahasiswa terdaftar di Indonesia pada tahun 2020 mencapai 8,6 juta orang. Angka ini mengalami kenaikan pada tahun 2021 dengan total mahasiswa terdaftar mencapai 8,96 juta mahasiswa hingga pada tahun 2022 jumlah mahasiswa terdaftar di Indonesia mencapai 9,32 juta orang. Pada tahun akademik 2016/2017 hingga tahun akademik 2019/2020, rata – rata jumlah mahasiswa baru setiap tahun di seluruh perguruan tinggi di Indonesia mencapai 350 ribu orang. Namun, rata – rata lulusan pada tahun akademik 2016/2017 hingga tahun akademik 2019/2020 setiap tahun hanya 240 ribu orang. Berdasarkan data ini, sebanyak rata – rata 110 ribu mahasiswa memiliki status yang tidak diketahui. Status yang tidak diketahui tersebut dapat diartikan bahwa mahasiswa melakukan pengunduran diri, menempuh studi tidak tepat waktu, atau *drop out*.

Dalam dunia pendidikan tinggi, pengunduran diri mahasiswa merupakan salah satu isu kritis yang dihadapi oleh universitas di seluruh dunia, termasuk Universitas Tanjungpura. Pengunduran diri mahasiswa tidak hanya berdampak pada individu yang bersangkutan, melainkan juga pada reputasi dan kinerja akademik institusi. Fenomena ini seringkali mencerminkan berbagai tantangan yang dihadapi mahasiswa selama perjalanan akademis mereka, mulai dari faktor internal individu hingga kondisi eksternal yang berkaitan dengan lingkungan pendidikan.

Berdasarkan observasi dan wawancara terhadap beberapa mahasiswa aktif dari berbagai program studi di Fakultas Teknik Universitas Tanjungpura, diketahui banyak dari teman seangkatan mereka yang memutuskan untuk mengundurkan diri.

Upaya yang dilakukan untuk mengurangi jumlah mahasiswa yang melakukan pengunduran diri dengan cara melakukan analisis pola atau informasi yang didapatkan dari pusat data perguruan tinggi mengenai faktor - faktor yang memiliki pengaruh terhadap mahasiswa yang memilih mengundurkan diri dengan menggunakan *data mining* (Salam dkk., 2020).

Pada penelitian sebelumnya telah dilakukan penelitian untuk membandingkan algoritma SVM dan *Decision Tree* dalam melakukan prediksi mahasiswa yang aktif, menunjukan bahwa algoritma SVM mendapatkan hasil akurasi yang lebih baik yaitu sebesar 95% (Wiyono & Abidin, 2018). Pada penelitian sebelumnya juga telah dilakukan perbandingan beberapa algoritma salah satunya adalah ANN, SVM, dan *Naïve Bayes*, menunjukan bahwa SVM mendapatkan hasil akurasi yang lebih baik (Sinanto dkk., 2022).

Oleh karena itu, maka penelitian ini dilakukan untuk memprediksi mahasiswa yang berpotensi mengundurkan diri dengan melakukan pembuatan model prediksi menggunakan algoritma *Support Vector Machine* (SVM). Penelitian ini diharapkan dapat mengidentifikasi dan mengantisipasi mahasiswa yang berpotensi melakukan pengunduran diri.

1.2 Perumusan Masalah

Berdasarkan uraian latar belakang tersebut, maka perumusan masalah pada penelitian ini adalah sebagai berikut.

1. Bagaimana performa dari algoritma SVM dalam memprediksi mahasiswa yang mengundurkan diri?
2. Bagaimana analisis hasil prediksi algoritma SVM terhadap mahasiswa yang diprediksi berpotensi mengundurkan diri?

1.3 Tujuan Penelitian

Berdasarkan perumusan masalah tersebut, maka tujuan pada penelitian ini adalah sebagai berikut.

1. Mengetahui performa dari algoritma SVM dalam memprediksi mahasiswa yang mengundurkan diri.
2. Menganalisis hasil prediksi algoritma SVM terhadap mahasiswa yang diprediksi berpotensi mengundurkan diri.

1.4 Pembatasan Masalah

Pembatasan masalah berdasarkan perumusan masalah tersebut adalah sebagai berikut

1. Dataset yang digunakan merupakan data mahasiswa aktif dan data mahasiswa yang telah mengundurkan diri di tengah masa studi pada angkatan 2014 – 2021.

2. Dataset yang digunakan dalam melakukan prediksi mahasiswa yang berpotensi menggundurkan diri ditengah masa studi merupakan data mahasiswa aktif pada angkatan 2021.
3. Algoritma yang digunakan dalam melakukan prediksi mahasiswa yang berpotensi mengundurkan diri ditengah masa studi menggunakan *Support Vector Machine* (SVM).

1.5 Sistematika Penulisan

Sistematika penulisan skripsi pada penelitian ini akan terdiri dari lima bab dengan urutan sebagai berikut.

1. BAB I: PENDAHULUAN

   Bab pertama terdiri dari latar belakang penelitian, perumusan masalah, tujuan penelitian, pembatasan masalah, dan sistematika penulisan skripsi.
2. BAB II: TINJAUAN PUSTAKA

   Pada bab ini berisi landasan teori yang berkaitan dengan penelitian yang akan dilakukan.
3. BAB III: METODOLOGI PENELITIAN

   Pada bab ini berisi pembahasan mengenai data penelitian dan tahapan – tahapan penelitian.
4. BAB IV: HASIL DAN ANALISIS

   Pada bab ini berisi uraian hasil dan analisis secara rinci mengenai data penelitian, pelatihan, pengujian, evaluasi model, dan hasil prediksi.
5. BAB V: PENUTUP

   Pada bab ini berisi tentang kesimpulan dari penelitian yang telah dilakukan dan saran yang dapat digunakan untuk perbaikan dan pengembangan penelitian yang telah dilakukan.