PENERAPAN TEKNIK FEATURE SELECTION DALAM MENINGKATKAN AKURASI PREDIKSI CHURN MENGGUNAKAN MACHINE LEARNING

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana

Program Studi Informatika

disusun oleh

HERNANDA GIRI PRAMUDITA

19.11.2761

Kepada

FAKULTAS ILMU KOMPUTER

UNIVERSITAS AMIKOM YOGYAKARTA

YOGYAKARTA

2025

PENERAPAN TEKNIK FEATURE SELECTION DALAM MENINGKATKAN AKURASI PREDIKSI CHURN MENGGUNAKAN MACHINE LEARNING

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana

Program Studi Informatika

disusun oleh

HERNANDA GIRI PRAMUDITA

19.11.2761

Kepada

FAKULTAS ILMU KOMPUTER

UNIVERSITAS AMIKOM YOGYAKARTA

YOGYAKARTA

2025

HALAMAN PERSETUJUAN

HALAMAN PERSETUJUAN

SKRIPSI

PENERAPAN TEKNIK FEATURE SELECTION DALAM MENINGKATKAN AKURASI PREDIKSI CHURN MENGGUNAKAN MACHINE LEARNING

yang disusun dan diajukan oleh

Hernanda Giri Pramudita

19.11.2761

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 25 Ferbruari 2025

Dosen Pembimbing,

Mulia Sulistiyono, S.Kom., M.Kom.
NIK. 190302248

HALAMAN PENGESAHAN

HALAMAN PENGESAHAN

SKRIPSI

PENERAPAN TEKNIK FEATURE SELECTION DALAM MENINGKATKAN AKURASI PREDIKSI CHURN MENGGUNAKAN MACHINE LEARNING

yang disusun dan diajukan oleh

Hernanda Giri Pramudita

19.11.2761

Telah dipertahankan di depan Dewan Penguji
pada tanggal 25 Februari 2025

Susunan Dewan Penguji

Nama Penguji	Tanda Tangan
Ahmad Sa`di, S.Kom, M.Eng. **NIK. 190302459**	
Majid Rahardi, S.Kom., M.Eng. **NIK. 190302393**	
Mulia Sulistiyono, S.Kom., M.Kom. **NIK. 190302248**	

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 25 Februari 2025

DEKAN FAKULTAS ILMU KOMPUTER

Hanif Al Fatta,S.Kom., M.Kom., Ph.D.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Hernanda Giri Pramudita
NIM : 19.11.2761

Menyatakan bahwa Skripsi dengan judul berikut:

Penerapan Teknik Feature Selection Dalam Meningkatkan Akurasi Prediksi Churn Menggunakan Machine Learning

Dosen Pembimbing : Mulia Sulistiyono, S.Kom., 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, <tanggal lulus ujian skripsi>

Yang Menyatakan,

Hernanda Giri Pramudita

HALAMAN PERSEMBAHAN

Dengan penuh rasa syukur dan ketulusan, skripsi ini saya persembahkan kepada:

1. **Allah SWT**

   Atas limpahan rahmat, nikmat, dan karunia-Nya yang selalu menyertai setiap langkah dalam menyelesaikan skripsi ini.

2. **Kedua Orang Tua Tercinta**

   Terima kasih atas doa, kasih sayang, dukungan, dan pengorbanan yang tiada henti. Semangat dan cinta kalian adalah kekuatan terbesar dalam hidup saya.

3. **Saudara dan Keluarga Besar**

   Terima kasih atas dorongan, motivasi, dan perhatian yang selalu kalian berikan di setiap perjalanan hidup saya.

4. **Dosen Pembimbing**

   Terima kasih atas bimbingan, ilmu, dan arahan yang telah membantu saya menyelesaikan penelitian ini dengan baik.

5. **Teman-teman Tersayang**

   Terima kasih atas dukungan, semangat, dan kebersamaan yang telah menjadi penguat dalam menyelesaikan perjalanan ini.

6. **Almamater Tercinta**

   Sebagai wujud rasa cinta dan pengabdian atas semua ilmu dan pengalaman yang telah saya dapatkan.

Semoga skripsi ini dapat memberikan manfaat dan menjadi langkah awal untuk perjalanan yang lebih baik di masa depan.

KATA PENGANTAR

Puji syukur ke hadirat Allah SWT atas rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul *"Penerapan Teknik Feature Selection dalam Meningkatkan Akurasi Prediksi Churn Menggunakan Machine Learning"*. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Informatika di Universitas Amikom Yogyakarta.

Dalam penyusunan skripsi ini, penulis mendapatkan banyak bantuan, bimbingan, serta dukungan dari berbagai pihak. Oleh karena itu, dengan segala kerendahan hati, penulis ingin menyampaikan terima kasih yang sebesar-besarnya kepada:

1. **Prof. Dr. M. Suyanto, M.M.**, selaku Rektor Universitas Amikom Yogyakarta, yang telah memberikan kesempatan bagi penulis untuk menempuh pendidikan di institusi ini.
2. **Hanif Al Fatta,S.Kom., M.Kom., Ph.D.**, selaku Dekan Fakultas Ilmu Komputer, yang telah memberikan fasilitas serta dukungan akademik dalam penyusunan skripsi ini.
3. **Windha Mega Pradnya Dhuhita, M.Kom.**, selaku Ketua Program Studi Informatika, yang telah memberikan arahan dan dukungan selama perkuliahan.
4. **Pak Mulia Sulistiyono, S.Kom., M.Kom**, selaku dosen pembimbing, yang dengan penuh kesabaran memberikan bimbingan, kritik, dan saran yang sangat berharga bagi penyelesaian skripsi ini.
5. **Mulia Sulistiyono, S.Kom., M.Kom., Majid Rahardi, S.Kom., M.Eng., Ahmad Sa`di, S.Kom, M.Eng**, yang telah memberikan masukan serta kritik yang konstruktif dalam penyempurnaan penelitian ini.
6. Orang tua dan keluarga tercinta yang selalu memberikan doa, semangat, serta dukungan moral dan materiil selama proses penyusunan skripsi ini.

7. Sahabat dan rekan-rekan seperjuangan di Universitas Amikom Yogyakarta yang telah memberikan motivasi, dukungan, serta kebersamaan yang tak ternilai.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan. Oleh karena itu, penulis terbuka terhadap kritik dan saran yang membangun demi perbaikan di masa mendatang. Semoga skripsi ini dapat memberikan manfaat bagi perkembangan ilmu pengetahuan, khususnya dalam bidang analisis data dan machine learning, serta menjadi referensi bagi penelitian selanjutnya.

Akhir kata, penulis berharap agar penelitian ini dapat memberikan kontribusi yang berarti bagi akademisi maupun praktisi dalam memahami dan menerapkan teknik feature selection dalam meningkatkan akurasi prediksi churn.

Yogyakarta, 25 Februari 2025

Penulis

DAFTAR ISI

HALAMAN JUDUL i
HALAMAN PERSETUJUAN ii
HALAMAN PENGESAHAN iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI iv
HALAMAN PERSEMBAHAN v
KATA PENGANTAR vi
DAFTAR ISI viii
DAFTAR TABEL x
DAFTAR GAMBAR xi
DAFTAR LAMPIRAN xii
DAFTAR LAMBANG DAN SINGKATAN xiii
INTISARI xiv
ABSTRACT xv
BAB I PENDAHULUAN 1
1.1 Latar Belakang 1
1.2 Rumusan Masalah 2
1.3 Batasan Masalah 3
1.4 Tujuan Penelitian 3
1.5 Manfaat Penelitian 3
1.6 Sistematika Penulisan 4
BAB II TINJAUAN PUSTAKA 5
2.1 Studi Literatur 5
2.8 Dasar Teori 10

BAB III METODE PENELITIAN 17
3.1 Alur Penelitian 17
3.2 Alat dan Bahan 23
BAB IV HASIL DAN PEMBAHASAN 24
4.5.2. Akurasi Model *Naïve Bayes* 29
4.5.4. Akurasi Model *Random Forest* 31
4.5.6. Akurasi Model *Support Vector Machine (SVM)* 32
BAB V PENUTUP 34
5.1 Kesimpulan 34
5.2 Saran 34
REFERENSI 36
LAMPIRAN 38

DAFTAR TABEL

Tabel 2.1. Keaslian Penelitian 7
Tabel 2.2. Hasil Perbandingan 32

DAFTAR GAMBAR

Gamber 2.1. Alur Metode Filter 11
Gamber 2.2. Alur Metode Wrapper 12
Gambar 2.3. Alur metode embedded 12
Gambar 3.1 Alur Penelitian 17
Gambar 4.1 Dataset WA_Fn-UseC_-Telco-Customer-Churn.csv 24
Gambar 4.2 Pengolahan data 24
Gambar 4.3 Memuat dataset 24
Gambar 4.4 Melihat Struktur dataset 25
Gambar 4.5 Menghapus NaN Value 25
Gambar 4.6 Mengubah data 26
Gambar 4.7 Penerapan Feature Selection 26
Gambar 4.8 *Naïve Bayes* 27
Gambar 4.9 *Random Forest* 27
Gambar 4.10 SVM 28
Gambar 4.11 Report AUC Klasifikasi *Naïve Bayes* 28
Gambar 4.12 Curva ROC Klasifikasi *Naïve Bayes* 29
Gambar 4.13 Report AUC *Random Forest* 29
Gambar 4.14 Curva ROC Klasifikasi *Random Forest* 30
Gambar 4.15 Report AUC SVM 30
Gambar 4.13 Curva ROC SVM 31

DAFTAR LAMPIRAN

DAFTAR LAMBANG DAN SINGKATAN

SVM	Support Vector Machines
M	mean
Xi	sample value
N	number of samples
O	nilai Observasi (pengamatan)
E	nilai Expected (harapan)
Y	Class sought
Yj	Y syb-class searched.

INTISARI

Churn atau perpindahan pelanggan merupakan tantangan besar bagi perusahaan di berbagai industri, karena dapat berdampak signifikan terhadap pendapatan dan pertumbuhan bisnis. Oleh karena itu, prediksi churn menjadi aspek penting dalam strategi retensi pelanggan. Penelitian ini bertujuan untuk meningkatkan akurasi prediksi churn dengan menerapkan teknik *feature selection* dalam model *machine learning*.

Dalam penelitian ini, digunakan beberapa algoritma klasifikasi, yaitu *Naïve Bayes*, *Support Vector Machine (SVM)*, dan *Random Forest*. Teknik *feature selection* diterapkan untuk mengidentifikasi atribut yang paling berpengaruh dalam memprediksi churn, sehingga dapat meningkatkan kinerja model. Evaluasi dilakukan dengan membandingkan akurasi prediksi sebelum dan sesudah penerapan *feature selection* pada ketiga algoritma tersebut.

Hasil penelitian menunjukkan bahwa teknik *feature selection* secara tidak signifikan meningkatkan akurasi prediksi churn. Model *SVM* dengan *feature selection* memberikan hasil terbaik dengan mendapatkan *score* 78,8% dibandingkan dengan *Naïve Bayes* yang mendapatkan *score* 77% dan *Random Forest* dengan *score* 78,3%. Dengan demikian, penelitian ini memberikan wawasan bahwa efektivitas teknik *feature selection* dapat bervariasi tergantung pada algoritma yang digunakan serta karakteristik data yang dianalisis. Temuan ini dapat menjadi pertimbangan bagi perusahaan dalam mengoptimalkan model prediksi churn untuk strategi retensi pelanggan yang lebih efektif.

Kata kunci: *Feature Selection, Machine Learning, Churn Prediction, Naïve Bayes, Support Vector Machine, Random Forest*

ABSTRACT

Churn or customer turnover is a significant challenge for companies across various industries, as it can have a substantial impact on revenue and business growth. Therefore, churn prediction has become a crucial aspect of customer retention strategies. This study aims to improve the accuracy of churn prediction by applying feature selection techniques in machine learning models.

In this research, several classification algorithms were used, namely Naïve Bayes, Support Vector Machine (SVM), and Random Forest. Feature selection techniques were applied to identify the most influential attributes in predicting churn, thereby enhancing model performance. Evaluation was conducted by comparing the prediction accuracy before and after the application of feature selection across the three algorithms.

The results of the study indicate that feature selection techniques did not significantly improve churn prediction accuracy. The SVM model with feature selection yielded the best results compared with score 78,8% to Naïve Bayes 77% and Random Forest 78,3%. Thus, this research provides insights that the effectiveness of feature selection techniques can vary depending on the algorithm used and the characteristics of the data analyzed. These findings can serve as a consideration for companies in optimizing churn prediction models for more effective customer retention strategies.

Keyword**:** *Feature Selection, Machine Learning, Churn Prediction, Naïve Bayes, Support Vector Machine, Random Forest*