IMPLEMENTASI DATA MINING PADA PROSES SELEKSI BEASISWA MENGGUNAKAN NAIVE BAYES DAN BACKWARD ELIMINATION

Irma Agustina[1], Gifthera Dwilestari[2], Ade Rizki Rinaldi[3]

[1]Program Studi Teknik Informatika, STMIK IKMI Cirebon
[2]Program Studi Sistem Informasi, STMIK IKMI Cirebon
[3]Program Studi Rekayasa Perangkat Lunak, STMIK IKMI Cirebon
[1,2,3] Jl. Perjuangan No 10. B, Kesambi, Kota Cirebon Jawa Barat 45131

Email: [1]agustinairma719@gmail.com , [2]ggdwilestari@gmail.com , [3]aderizki@ikmi.ac.id

ABSTRAK

Proses seleksi penerima beasiswa sering kali menghadapi tantangan dalam mengelola data yang kompleks dan memastikan keakuratan seleksi. Penelitian ini bertujuan mengoptimalkan algoritma *Naive Bayes* melalui teknik *Backward Elimination* untuk efisiensi proses seleksi. Dataset penelitian terdiri dari 1.042 data penerima beasiswa, mencakup variabel seperti Indeks Prestasi Kumulatif (IPK), penghasilan, jumlah tanggungan, dan status beasiswa. Penelitian dilakukan menggunakan platform *RapidMiner* versi 10.2 dengan tahapan meliputi *preprocessing*, transformasi data, pembagian data latih dan uji melalui *Split Data*. Teknik *Backward Elimination* diterapkan untuk menyederhanakan model dengan menghapus variabel yang kurang signifikan. Hasil penelitian menunjukkan bahwa penerapan *Naive Bayes* dengan teknik *Backward Elimination* menghasilkan tingkat akurasi sebesar 74,62%. Variabel utama yang paling berpengaruh adalah tanggungan orang tua dan penghasilan, yang secara signifikan memengaruhi keputusan seleksi. Selain itu, teknik ini juga berhasil mengurangi kompleksitas model, meningkatkan efisiensi proses analisis, dan meminimalkan waktu serta sumber daya yang dibutuhkan. Penelitian ini mendukung pengembangan sistem seleksi berbasis data yang lebih transparan dan efisien. Implementasi teknik *Backward Elimination* mempermudah interpretasi model. Dengan demikian, hasil ini diharapkan dapat menjadi landasan bagi pengembangan sistem seleksi beasiswa berbasis machine learning yang lebih efektif, serta membuka peluang untuk penelitian lanjutan yang berfokus pada optimalisasi algoritma dan seleksi fitur di berbagai sektor.

Kata Kunci: *Naive Bayes, Backward Elimination*, Seleksi Beasiswa, *Data Mining*, Akurasi Model

ABSTRACT

The scholarship recipient selection process often faces challenges in managing complex data and ensuring selection accuracy. This study aims to optimize the Naive Bayes algorithm through the Backward Elimination technique to efficiency of the selection process. The research dataset comprises 1,042 scholarship recipient records, including variables such as Grade Point Average (GPA), income, number of dependents, and scholarship status. The study was conducted using the RapidMiner platform version 10.2, involving stages such as preprocessing, data transformation, and the splitting of training and testing data through the Split Data operator. The Backward Elimination technique was applied to simplify the model by removing less significant variables. The study results indicate that implementing Naive Bayes with the Backward Elimination technique achieved an accuracy level of 74.62%. The key influencing variables were the number of dependents and parental income, which significantly affected the selection decisions. Additionally, this technique successfully reduced model complexity, improved the efficiency of the analysis process, and minimized the time and resources required. This research supports the development of a more transparent and efficient data-driven selection system. The implementation of the Backward Elimination technique simplifies model interpretation. Therefore, these findings are expected to serve as a foundation for the development of a more effective machine-learning-based scholarship selection system and to create opportunities for further research focused on optimizing algorithms and feature selection across various sectors.

Keywords: Naive Bayes, Backward Elimination, Scholarship Selection, Data Mining, Model Accuracy

1. PENDAHULUAN

Seleksi beasiswa merupakan komponen krusial dalam sistem pendidikan yang bertujuan memberikan akses kepada siswa yang membutuhkan dukungan finansial untuk melanjutkan studi ke jenjang yang lebih tinggi. Proses ini memerlukan pengolahan data yang tepat untuk menentukan siapa yang berhak menerima beasiswa, yang sering kali melibatkan analisis data yang besar dan kompleks[1]. Oleh karena itu, penerapan metode data mining menjadi solusi yang relevan untuk menangani tantangan ini, mengingat kemampuannya dalam mengolah data dengan efektif dan efisien.

Salah satu teknik yang sering digunakan dalam *data mining* adalah *Naive Bayes*, yang dapat melakukan klasifikasi dengan memanfaatkan probabilitas dari data yang tersedia. Metode ini berbasis pada *Teorema Bayes*, yang memungkinkan untuk menghitung probabilitas berdasarkan fitur yang ada, menjadikannya model yang sederhana namun kuat dalam memprediksi kategori suatu data[2]. Meskipun memiliki kemampuan yang baik, *Naive Bayes* dapat terpengaruh oleh adanya fitur-fitur yang tidak relevan, yang dapat menurunkan akurasi klasifikasi.

Untuk meningkatkan performa model, salah satu teknik yang dapat digunakan adalah *Backward Elimination*. Teknik ini bekerja dengan cara mengidentifikasi dan menghapus fitur yang kurang berpengaruh, sehingga hanya fitur yang paling relevan yang digunakan dalam model, meningkatkan akurasi dan efisiensi proses seleksi[3]. Kombinasi antara *Naive Bayes* dan *Backward Elimination* diharapkan dapat menghasilkan model klasifikasi yang lebih akurat dan efisien dalam konteks seleksi penerima beasiswa.

Penelitian ini bertujuan untuk menerapkan metode *Naive Bayes* yang dipadukan dengan teknik *Backward Elimination* dalam proses seleksi penerima beasiswa. Dengan harapan, penelitian ini dapat memberikan kontribusi dalam meningkatkan akurasi dan efisiensi sistem seleksi beasiswa di institusi pendidikan menggunakan pendekatan data mining yang lebih canggih.

2. TINJAUAN PUSTAKA

2.1. Implementasi

Implementasi merupakan proses penerapan suatu metode, konsep, atau rencana ke dalam tindakan nyata untuk mencapai tujuan yang telah ditetapkan. Dalam dunia penelitian, implementasi mengacu pada berbagai langkah praktis yang dilakukan untuk menerapkan teknik atau metode tertentu pada dataset yang sesuai. Proses ini biasanya mencakup tahapan mulai dari persiapan data, pemilihan algoritma, hingga evaluasi hasil yang diperoleh. Di bidang teknologi, implementasi sering kali berkaitan dengan penerapan algoritma dan pengembangan sistem berbasis perangkat lunak untuk memecahkan masalah tertentu. Dengan implementasi yang terencana dengan baik, solusi berbasis data atau teknologi dapat menghasilkan keputusan yang lebih akurat, efisien, dan relevan sesuai kebutuhan[4].

2.2. Data Mining

Data mining adalah proses mengidentifikasi pola, hubungan, atau informasi penting dari kumpulan data yang besar dengan memanfaatkan teknik analitik, metode statistik, serta algoritma berbasis kecerdasan buatan. Tahapan dalam proses ini biasanya dimulai dengan pengumpulan data yang relevan. Setelah itu, dilakukan pembersihan data atau *preprocessing* untuk menghilangkan nilai yang tidak konsisten atau hilang. Langkah berikutnya adalah transformasi data agar sesuai dengan kebutuhan analisis. Kemudian, data dianalisis menggunakan algoritma tertentu untuk menemukan pola yang signifikan. Hasil akhir dari proses ini diinterpretasikan untuk mendukung pengambilan keputusan yang lebih baik.

2.3. Naïve Bayes

Naive Bayes adalah algoritma *machine learning* berbasis probabilitas yang mengacu pada *Teorema Bayes* dan berasumsi bahwa setiap fitur dalam data bersifat independen satu sama lain. Algoritma ini sering digunakan untuk menyelesaikan masalah klasifikasi. Dalam konteks seleksi beasiswa, algoritma ini dapat membantu memprediksi kelayakan penerima beasiswa dengan menganalisis berbagai data sosial dan akademik calon penerima. Dengan memanfaatkan algoritma ini, lembaga pendidikan dapat membuat keputusan seleksi yang lebih berbasis data. Proses klasifikasi dilakukan dengan menghitung

peluang dari setiap kategori berdasarkan variabel yang tersedia. Hasil prediksi yang dihasilkan memungkinkan pengambilan keputusan yang lebih efisien dan akurat[5].

2.4. Backward elimination

Backward Elimination adalah teknik seleksi fitur dalam *data mining* yang berfungsi untuk menyederhanakan model prediksi dengan mengeliminasi variabel-variabel yang tidak signifikan. Proses ini diawali dengan memasukkan semua variabel ke dalam model. Selanjutnya, variabel yang kontribusinya dianggap tidak berpengaruh secara signifikan terhadap akurasi model dihapus secara bertahap. Langkah-langkah tersebut diulang hingga hanya variabel yang relevan dan memberikan pengaruh signifikan terhadap akurasi prediksi yang tersisa. Dengan metode ini, model menjadi lebih efisien dan mudah diinterpretasikan tanpa mengorbankan kualitas prediksi. *Backward Elimination* sering digunakan untuk meningkatkan performa model prediktif dengan fokus hanya pada fitur yang benar-benar relevan.

2.5. Seleksi Beasiswa

Seleksi beasiswa adalah proses evaluasi dan penentuan penerima beasiswa berdasarkan sejumlah kriteria yang telah ditetapkan. Biasanya, kriteria tersebut meliputi prestasi akademik yang menunjukkan kemampuan belajar calon penerima. Selain itu, kondisi finansial sering menjadi faktor penting agar beasiswa dapat diberikan kepada mereka yang benar-benar memerlukan dukungan ekonomi. Partisipasi dalam organisasi atau kegiatan ekstrakurikuler juga dapat dipertimbangkan sebagai indikator keterampilan kepemimpinan dan kontribusi sosial. Proses ini dirancang untuk memastikan bahwa bantuan pendidikan diberikan secara adil dan tepat sasaran. Dengan pendekatan seleksi yang terstruktur, lembaga pemberi beasiswa dapat mendukung pengembangan mahasiswa yang potensial secara maksimal[6].

2.6. Rapidminer

Sebuah perangkat lunak *data science* yang dirancang untuk mendukung pemodelan prediktif, analisis data, dan pembelajaran mesin *(machine learning)*. Platform ini menawarkan antarmuka visual berbasis *drag-and-drop*, sehingga memudahkan pengguna dalam mengelola seluruh proses analisis data. Proses tersebut mencakup tahap pra-pemrosesan data untuk memastikan kualitas data yang optimal, penggalian pola melalui teknik data mining, hingga pengujian dan evaluasi model untuk mendapatkan hasil yang akurat. Dengan fitur yang lengkap dan mudah digunakan, *RapidMiner* memungkinkan pengguna dari berbagai latar belakang untuk melakukan analisis data yang komprehensif tanpa memerlukan kemampuan pemrograman yang kompleks[7].

2.7. Referensi

Pada penelitian pertama yang berjudul "Implementasi Metode *Naïve Bayes* Dalam Memprediksi Kelulusan Mahasiswa Tepat Waktu Pada Program Studi Pendidikan Teknologi Informasi" menghasilkan model prediksi kelulusan mahasiswa tepat waktu dan tidak tepat waktu. Penelitian ini mempertimbangkan berbagai faktor pendukung kelulusan dan menunjukkan bahwa metode *Naïve Bayes* dapat digunakan untuk memprediksi kelulusan. Probabilitas kelulusan tepat waktu tercatat sebesar 11%, sementara tidak tepat waktu sebesar 89%, dengan akurasi mencapai 100% tanpa adanya error pada proses pelatihan dan pengujian[8].

Penelitian kedua yang berjudul "Klasifikasi Penerima Beasiswa Menggunakan Metode *Naïve Bayes* (Studi Kasus SMP Negeri 3 Selomerto)" menganalisis klasifikasi penerima beasiswa di SMP Negeri 3 Selomerto menggunakan algoritma *Naïve Bayes*. Dari 186 data siswa (150 data training dan 36 data testing), penelitian ini mencatat akurasi sebesar 91,67% baik melalui perhitungan manual maupun aplikasi *RapidMiner*. Hasil ini menunjukkan bahwa algoritma Naïve Bayes cocok digunakan untuk mempermudah proses klasifikasi penerima beasiswa, sehingga mempercepat dan meningkatkan akurasi pemberian beasiswa[9].

Penelitian ketiga berjudul "Penerapan Algoritma *Naïve Bayes* Berbasis *Backward Elimination* Untuk Prediksi Pemesanan Kamar Hotel" memanfaatkan algoritma *Naïve Bayes* yang mampu menghasilkan akurasi tinggi untuk memproses data dalam jumlah besar. Teknik backward elimination digunakan untuk memilih parameter yang optimal guna meningkatkan akurasi. Dengan 10.000 data pelanggan hotel, akurasi prediksi menggunakan algoritma *Naïve Bayes* tercatat sebesar 89,67%, sementara kombinasi

algoritma *Naïve Bayes* dengan backward elimination meningkatkan akurasi hingga 97,83%. Prediksi mencakup hasil *check-out*, pembatalan, dan *no-show*[10].

Penelitian keempat yang berjudul "Algoritma *Naïve Bayes* Dengan *Backward Elimination* Pada *Dataset Breast Cancer*" menerapkan algoritma *Naïve Bayes* dengan teknik *backward elimination* untuk mengoptimalkan akurasi, serta menggunakan validasi *split validation*. Hasil penelitian menunjukkan akurasi sebesar 77,14%, yang dianggap cukup baik untuk membantu dalam klasifikasi kanker payudara[11].

Penelitian kelima yang berjudul "Perbandingan Algoritma *C4.5* dan *Naive Bayes* dalam Klasifikasi Tingkat Kepuasan Mahasiswa Terhadap Pembelajaran Daring" Hasil perbandingan menunjukkan bahwa algoritma *Naïve Bayes* memiliki tingkat akurasi yang lebih tinggi dibandingkan dengan algoritma *C4.5*, dengan selisih sebesar 11,77%. Tingkat akurasi algoritma *C4.5* tercatat sebesar 58,82% berdasarkan uji validitas menggunakan metode *cross-validation*. Sebaliknya, algoritma *Naïve Bayes* menunjukkan tingkat akurasi yang lebih tinggi, yaitu 70,59%, dengan validitas diuji menggunakan *cross-validation*[12].

3. METODE PENELITIAN

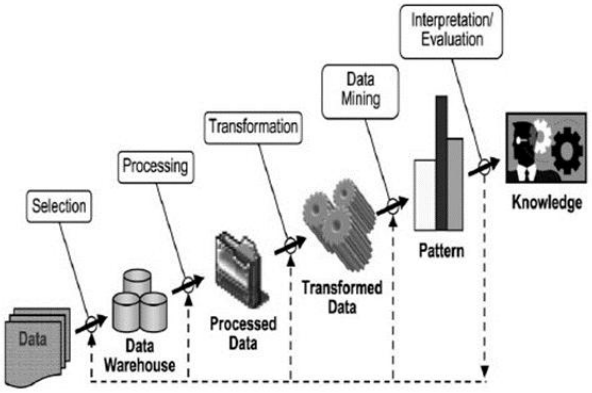
3.1 Sumber Data

Dalam penelitian ini, sumber data sekunder diambil dari *Kaggle*. Data yang digunakan mencakup informasi akademik dan sosial ekonomi, seperti IPK, jumlah tanggungan orang tua, serta faktor ekonomi lainnya, sehingga mendukung pengembangan model *Naïve Bayes* untuk meningkatkan efisiensi proses seleksi. Sumber data: https://www.*kaggle*.com/datasets/nirwanamy/beasiswa

3.2 Teknik Analisis Data

Penelitian ini mengunakan metode *knowledge discovery in database (KDD)* dengan pendekatan kuantitatif. Proses *Knowledge Discovery in Database (KDD)* adalah tahapan global dari penggalian data yang terdiri dari beberapa langkah / tahapan, dari pengolahan informasi, yang bertujuan untuk menggali dan menganalisis data yang sangat besar menjadi informasi yang berguna bagi pengetahuan[13].

Pada gambar 1 menunjukkan tahapan KDD (Knowledge Discovery in Databases) yang terdiri dari pemilihan data, prapemrosesan, transformasi, data mining, dan evaluasi pola untuk mendapatkan pengetahuan yang bermanfaat.



Gambar 1. Tahapan *KDD*

Tahapan proses *data mining* dimulai dari pemilihan data dari sumber data menuju data target, diikuti dengan tahap pra-pemrosesan untuk meningkatkan kualitas data, transformasi, proses *data mining*, serta interpretasi dan evaluasi yang diharapkan menghasilkan pengetahuan baru yang bermanfaat. Secara rinci, berikut adalah tahapan-tahapan tersebut:

a. Seleksi Data

Data untuk proses *data mining* dipilih dari data operasional yang ada dan disimpan dalam berkas terpisah dari basis data operasional utama. Pemilihan ini dilakukan sebelum memulai penggalian informasi dalam proses *Knowledge Discovery in Databases (KDD).*

b. Pra-pemrosesan/Pembersihan Data

Sebelum *data mining* dapat dilaksanakan, data perlu dibersihkan. Tahap ini mencakup penghapusan data duplikat, pengecekan konsistensi data, serta perbaikan kesalahan yang terdapat pada data[14].

c. Transformasi

Proses transformasi dilakukan pada data terpilih untuk menyesuaikannya dengan kebutuhan analisis dalam *data mining*. Transformasi data ini bersifat kreatif dan sangat bergantung pada informasi atau pola yang diinginkan dari basis data.

d. *Data mining*

Data mining merupakan tahap untuk mencari pola atau informasi menarik dari data yang sudah dipilih dengan menggunakan teknik atau metode tertentu. Metode atau algoritma yang digunakan beragam dan disesuaikan dengan tujuan dan proses *Knowledge Discovery in Databases (KDD)* yang diinginkan.

e. Interpretasi/Evaluasi

Pola atau informasi yang dihasilkan dari *data mining* perlu disajikan dengan cara yang mudah dipahami oleh pihak terkait. Tahap ini, yang disebut interpretasi, juga mencakup verifikasi bahwa pola atau informasi tersebut sesuai atau tidak bertentangan dengan fakta atau hipotesis yang ada sebelumnya[15].

4. PEMBAHASAN

Dataset yang akan diolah pada penelitian ini diambil dari *Kaggle*, dengan jumlah data sebanyak 1042 *record*. Data ini diperoleh dari dataset yang relevan dengan tema penerimaan beasiswa. Dataset tersebut mencakup variabel seperti jarak tempat tinggal ke kampus, IPK, tanggungan, penghasilan, dan status beasiswa. Variabel-variabel ini memberikan informasi yang penting terkait kriteria penerima beasiswa dan akan menjadi dasar dalam proses analisis untuk efisiensi seleksi penerima beasiswa.

4.1 Pengujian menggunakan Algoritma *Naïve Bayes* dengan *Rapidminer*

a. Data selection

Data ini akan dianalisis menggunakan algoritma Naïve Bayes, dengan teknik seleksi fitur *Backward elimination*, yang diterapkan melalui aplikasi *Rapidminer* versi 10.2.

b. *Preprocessing* atau *Cleaning*

Pada tahap *preprocessing*, dilakukan proses pembersihan data untuk mengatasi nilai-nilai yang hilang *(missing values).*

Tabel 1 memperlihatkan hasil akhir dari preprocessing, di mana data telah diproses secara cermat melalui verifikasi dan standardisasi untuk analisis yang optimal.

Tabel 1. Hasil *Preprocessing*

Name	*Type*	*Missing*	*Statistic*		
			Negative	*Positive*	*Values*
Status Beasiswa	Binominal	0	Terima	Tidak	Tidak (770), Terima (272)
Indeks Prestasi Kumulatif	Nominal	0	Sedang	Tinggi	Tinggi (1018), Sedang (24)
Penghasilan	Nominal	0	Rendah	Sedang	Sedang (418), Tinggi (357)
Tanggungan	Nominal	0	Banyak	Sedikit	Sedikit (552), Sedang (397

c. Transformation

Tahap transformasi ini dilakukan dengan tujuan untuk mengubah format atau struktur data menjadi bentuk yang lebih sesuai dan dapat diproses secara efisien dalam langkah pemodelan atau analisis data selanjutnya. Atribut yang tidak numerik dapat diubah menjadi numerik melalui proses nominal ke numerik. Data penerima beasiswa termasuk IPK, penghasilan, dan jumlah tanggungan.

Pada tabel 2 menunjukan hasil transformasi di mana variabel-variabel dengan nilai nominal telah direpresentasikan dalam bentuk numerik untuk memudahkan proses komputasi.

Tabel 2. Hasil Transformasi

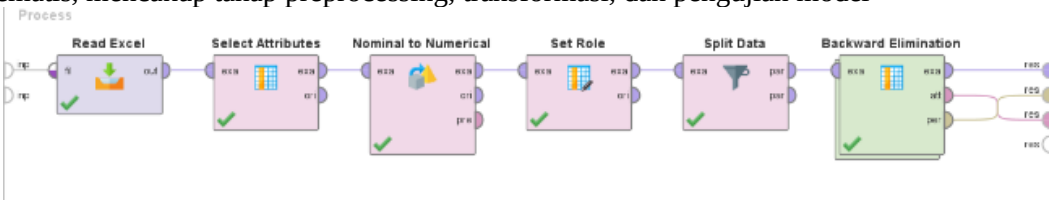
Row No.	Status Beasiswa	Indeks Prestasi Kumulatif	Penghasilan	Tanggungan
1	Terima	0	0	0
2	Tidak	0	0	1
3	Terima	0	0	0
4	Tidak	0	1	1
5	Tidak	0	0	1
6	Terima	0	2	0
7	Tidak	0	0	0

Gambar hasil transformasi *dataset dataset* telah siap untuk digunakan dalam algoritma prediktif, dengan seluruh atribut yang sudah terorganisasi dan nilai kategorikal yang telah diubah ke dalam format numerik.

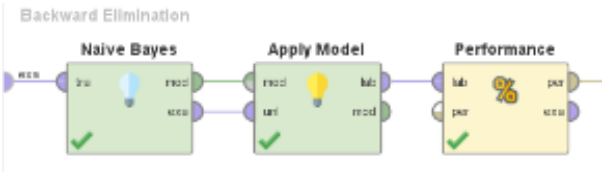
d. *Data mining*

Tahapan *data mining* ini menggunakan algoritma *Naïve Bayes* yang berperan dalam menyelesaikan masalah klasifikasi.

Gambar 2 dan 3 menampilkan operator-operator yang dirancang untuk mengolah data secara sistematis, mencakup tahap preprocessing, transformasi, dan pengujian model



Gambar 2. *Pre-Processing* dan *Backward Elimination*



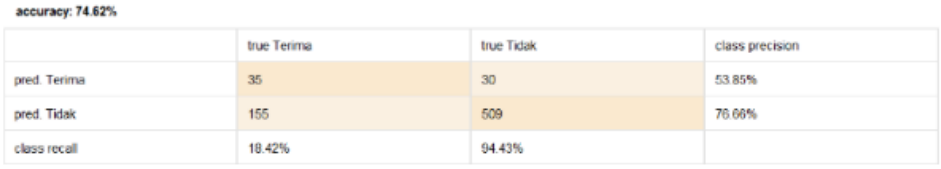
Gambar 3. *Naïve Bayes* dan *Apply Model*

Berdasarkan gambar di atas, *Operator Split Data* berfungsi untuk memisahkan data menjadi data latih dan data uji. Melalui proses ini, setiap subset data berperan baik sebagai data latih maupun data uji, sehingga menghasilkan evaluasi performa model yang lebih akurat dan tepercaya. Pada subproses pelatihan, *operator Naive bayes* digunakan untuk membangun model algoritma. Sementara itu, *operator Apply Model* diterapkan pada subproses pengujian guna menguji data yang dihasilkan dari model *Naive bayes*. *Operator Performance* digunakan untuk mengevaluasi performa model, secara otomatis memberikan daftar metrik kinerja yang relevan dengan tugas tertentu.

4.2 Hasil Tingkat Akurasi Yang Dicapai

Pada tahap ini, penerapan metode *Naive bayes* menggunakan fitur *Backward elimination* untuk klasifikasi menggunakan *Rapidminer*.

Pada gambar 4 dibawah menunjukan hasil akurasi model menggunakan algoritma *Naïve Bayes* menggunakan fitur *Backward Elimination*.



accuracy: 74.62%

	true Terima	true Tidak	class precision
pred. Terima	35	30	53.85%
pred. Tidak	155	509	76.66%
class recall	18.42%	94.43%	

Gambar 4. Hasil Akurasi Model *(Naïve Bayes* dan *Backward elimination)*

Berikut adalah hasil klasifikasi menggunakan algoritma *Naive bayes* dengan fitur *Backward elimination*:

a. Tingkat Akurasi: 74,62%.
b. Nilai *True Positive (TP)*: 35
c. Nilai *True Negative (TN)*: 509
d. Nilai *False Positive (FP)*: 30
e. Nilai *False Negative (FN)*:155

Pada Tabel 3, ditampilkan hasil dari metode backward elimination, di mana variabel-variabel yang tetap terpilih memiliki nilai *weight* sebesar 1.

Tabel 3. Hasil *Backwad Elimination*

Attribute	***Weight***
Penghasilan	1
Indeks Prestasi Kumulatif	0
Tanggungan	1

5. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Penerapan algoritma *Naive Bayes* dengan *Backward Elimination* menghasilkan tingkat akurasi 74,62% dalam seleksi penerima beasiswa. Proses ini menyederhanakan model dengan mempertahankan variabel yang relevan, seperti tanggungan orang tua dan penghasilan, sehingga proses seleksi menjadi lebih efisien.

5.2 Saran

Disarankan agar penelitian berikutnya menggunakan dataset yang lebih beragam dan mengeksplorasi algoritma lain seperti *Decision Tree* atau *Random Forest* untuk perbandingan performa model. Implementasi sistem seleksi berbasis web juga direkomendasikan untuk meningkatkan aksesibilitas dan penggunaan secara luas oleh lembaga pendidikan.

DAFTAR PUSTAKA

[1] M. A. W. Pratama, M. Fuad, Hazriani, and Yuyun, “Penentuan Status Penerima Bantuan Indonesia Pintar Pada Smkn 9 Bulukumba Dengan Metode Naive Bayes,” *Pros. Semin. Nas. Sist. Inf. dan Teknol.*, pp. 120–125, 2023.

[2] M. Hidayat, A. N. Fuadi, D. P. Utomo, E. D. Astuti, and D. Asmarajati, “Studi Komparasi Algoritma Naive Bayes Dan K-NN Untuk Klasifikasi Penerimaan Beasiswa Di MI Al-Islamiyah Karangsawah,” *J. Ilm. Tek. dan Ilmu Komput.*, vol. 2, no. 4, pp. 172–180, 2023.

[3] M. Arifin, “Naïve Bayes Algorithm Based On Backward Elimination For Predicting Cervical Cancer,” *Int. J. Innov. Sci. Res. Technol.*, vol. 7, no. 7, 2022, [Online]. Available: https://www.researchgate.net/publication/362221281

[4] E. Widodo, S. Fauziah, and A. A. Sulaeman, “Analisa Prediksi Hasil Produksi Popok Bayi Metode Naïve Bayes,” *Bull. Inf. Technol.*, vol. 4, no. 1, pp. 75–80, 2023.

[5] M. Afriansyah, J. Saputra, Y. Sa, V. Yoga, and P. Ardhana, “Optimasi Algoritma Naive Bayes Untuk Klasifikasi Buah Apel Berdasarkan Fitur Warna RGB,” *Bull. Comput. Sci. Res.*, vol. 3, no. 3, pp. 242–249, 2023, doi: 10.47065/bulletincsr.v3i3.251.

[6] A. Pebdika, R. Herdiana, and D. Solihudin, “Klasifikasi Menggunakan Metode Naive Bayes Untuk

Menentukan Calon Penerima PIP," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 452–458, 2023.
[7] Wahyuningsih, Budiman, and I. Umami, "Implementasi Algoritma Naive Bayes Untuk Menentukan Calon Penerima Beasiswa Di SMK YPM 14 Sumobito Jombang," *J. Teknol. Dan Sist. Inf. Bisnis*, vol. 4, no. 1, pp. 446–454, 2022.
[8] Darman and Z. Razilu, "Implementasi Metode Naïve Bayes Dalam Memprediksi Kelulusan Mahasiswa Tepat Waktu Pada Program Studi Pendidikan Teknologi Informasi," *J. Pendidik. Teknol. Inf.*, vol. 4, no. 3, pp. 915–927, 2024.
[9] A. Misbachudin Riyadi, H. Sibyan, I. Ahmad Ihsanuddin, and M. Alif Muwafiq Baihaqi, "Klasifikasi Penerima Beasiswa Menggunakan Metode Naïve Bayes (Studi Kasus SMP Negeri 3 Selomerto)," *J. Eng. Inform.*, vol. 1, no. 2, pp. 53–59, 2023, doi: 10.56854/jei.v1i2.61.
[10] H. Annur, "Penerapan Algoritma Naïve Bayes Berbasis Backward Elimination Untuk Prediksi Pemesanan Kamar Hotel," *J. Ilm. Ilmu Komput. Banthayo Lo Komput.*, vol. 1, no. 1, pp. 1–5, 2022, doi: 10.37195/balok.v1i1.99.
[11] R. A. Anggraini, "Algoritma Naïve Bayes Dengan Backward Elimination Pada Dataset Breast Cancer," *J. Kaji. Ilm.*, vol. 23, no. 1, pp. 87–94, 2024, doi: 10.31599/tzsb5v61.
[12] Farmawati and Narti, "Perbandingan Algoritma C4.5 dan Naive Bayes Dalam Klasifikasi Tingkat Kepuasan Mahasiswa Terhadap Pembelajaran Daring," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 4, no. 1, pp. 1–12, 2022, doi: 10.35746/jtim.v4i1.196.
[13] D. H. D. Herudin, A. Faqih, and A. Bahtiar, "Klasifikasi Penerimaan Peserta Didik Baru Menggunakan Algoritma Naïve Bayes Dengan Smote Pada Smkn 1 Jamblang," *JURSIMA (Jurnal Sist. Inf. dan Manajemen)*, vol. 10, no. 2, 2022.
[14] I. Loelianto, M. S. S. Thayf, and H. Angriani, "Implementasi Teori Naive Bayes Dalam Klasifikasi Calon Mahasiswa Baru Stmik Kharisma Makassar," *SINTECH (Science Inf. Technol. J.*, vol. 3, no. 2, pp. 110–117, 2020, doi: 10.31598/sintechjournal.v3i2.651.
[15] A. Widhiantoyo, B. N. Sari, and D. Yusuf, "Penerapan Algoritma Naïve Bayes Dengan Backward Elimination Untuk Prediksi Waktu Tunggu Alumni Mendapatkan Pekerjaan," *JIKO (Jurnal Inform. dan Komputer)*, vol. 4, no. 3, pp. 145–151, 2021, doi: 10.33387/jiko.v4i3.3272.