

Optimasi Backward Elimination untuk Klasifikasi Kepuasan Pelanggan Menggunakan Algoritme k-Nearest Neighbor (k-NN) dan Naïve Bayes



Yunitasari¹
Hopi Siti Hopipah²
Rini Mayasari³

Universitas Singaperbangsa Karawang^{1,2,3}

HS. Ronggo Waluyo, Telukjambe Timur, Karawang, Jawa Barat, Indonesia

E-mail: yunitasari.yunita17007@student.unsika.ac.id¹, hopi.siti17022

Hopipah@student.unsika.ac.id², rini.mayasari@staff.unsika.ac.id³

Notifikasi Penulis
26 Januari 2021
Akhir Revisi
05 Juli 2021
Terbit
01 Agustus 2021

Yunitasari, Hopipah, H. S., & Mayasari, R. (2021). Optimasi Optimasi Backward Elimination untuk Klasifikasi Kepuasan Pelanggan Menggunakan Algoritme k-nearest neighbor (k-NN) and Naive Bayes. *Technomedia Journal*, 6(1 Agustus).

https://doi.org/10.33050/tmj.v6i1_Agustus.1531

ABSTRAK

Mempertahankan kepuasan pelanggan merupakan sebuah tantangan besar bagi perusahaan. Salah satu upaya yang dapat dilakukan adalah memberikan pelayanan terbaik terhadap pelanggan berdasarkan aspek yang paling berpengaruh. Pada penelitian ini dilakukan optimasi fitur Backward Elimination pada klasifikasi kepuasan pelanggan dengan algoritme k-NN dan Naïve Bayes. Penggunaan fitur Backward Elimination bertujuan meningkatkan akurasi dan mengurangi jumlah atribut yang kurang berpengaruh. Hasilnya, dapat diketahui bahwa pemodelan terbaik tanpa Backward Elimination adalah algoritme Naïve Bayes dengan akurasi 99.04% dan nilai AUC mencapai 1. Sedangkan penerapan Backward Elimination bekerja lebih optimal pada algoritme k-NN dengan peningkatan sebesar 33.74% menjadi 97.28% dengan AUC 0.996. Hal ini menunjukkan bahwa kinerja fitur Backward Elimination efektif dalam optimasi klasifikasi kepuasan pelanggan dan dapat mengurangi atribut yang kurang berpengaruh.

Kata kunci : Backward Elimination; Klasifikasi Pelanggan; k-Nearest Neighbor; Naïve Bayes; Feature Selection.

ABSTRACT

Maintaining customer satisfaction is a big challenge for companies. One effort that can be done is to provide the best service to customers based on the most influential aspects. In this study, the optimization of the Backward Elimination feature in the classification of customer satisfaction using the k-NN and Naïve Bayes algorithm. The use of the Backward Elimination feature aims to increase accuracy and reduce the number of less influential attributes. As a result, it can be seen that the best modeling without Backward Elimination is the Naïve Bayes algorithm with an accuracy of 99.04% and an AUC value of 1. While the application of Backward Elimination works more optimally on the k-NN algorithm with an increase of 33.74% to 97.28% with AUC 0.996. This shows that the performance of the Backward Elimination feature is effective in optimizing the classification of customer satisfaction and can reduce the less influential attributes.

Keywords: Backward Elimination; Klasifikasi Pelanggan; k-Nearest Neighbor; Naïve Bayes; Feature Selection.

PENDAHULUAN

Dalam era bisnis saat ini, setiap perusahaan pasti mengalami kondisi persaingan yang sangat ketat salah satunya yaitu pada perusahaan penerbangan baik penerbangan berjadwal maupun penerbangan tidak berjadwal. Kepuasan pelanggan menjadi aspek yang sangat penting untuk diperhatikan dalam rangka meningkatkan kinerja perusahaan. Perusahaan perlu memperhatikan hal apa saja yang dianggap penting oleh pelanggan agar kepuasan para pelanggan dapat tercapai [1]

Dengan demikian perusahaan perlu melakukan penilaian untuk mengetahui faktor apa saja yang akan mempengaruhi kepuasan pelanggannya dan bagaimana cara untuk memenuhi harapan pelanggannya sehingga loyalitas pelanggan akan tercipta apabila pelanggan merasakan kepuasannya dalam menerima manfaat jasa perusahaan tersebut. Akan tetapi, mempertahankan pelanggan yang loyal menjadi salah satu tantangan bagi sebuah perusahaan [2]. Jumlah pelanggan sangat mempengaruhi profit perusahaan, maka kepuasan pelanggan penting untuk diperhatikan. Kepuasan dan loyalitas pelanggan saling berhubungan, kepuasan pelanggan dapat mempengaruhi peningkatan dan penurunan loyalitas pelanggan [3]. Upaya dalam mempertahankan dan meningkatkan loyalitas pelanggan salah satunya adalah dengan menganalisis data klasifikasi kepuasan pelanggan. Hasil tersebut dapat dijadikan acuan perusahaan untuk menentukan langkah selanjutnya dalam memberikan pelayanan terbaik bagi pelanggan.

Algoritme *k*-NN merupakan salah satu algoritme *supervised learning* yang dapat digunakan dalam mengklasifikasikan data baru berdasarkan mayoritas tetangga terdekat [4]. Penelitian sebelumnya dilakukan analisis algoritme *k-Nearest Neighbor (k-NN)* dalam menentukan nilai akurasi kepuasan pelanggan, hasilnya metode klasifikasi dengan algoritme *k-NN* ini memiliki nilai akurasi sebesar 83.33%. Berdasarkan hasil tersebut, disimpulkan bahwa pelanggannya merasa puas terhadap pelayanan yang diberikan perusahaan [5].

Akurasi algoritme *k-NN* dipengaruhi oleh tingkat relevansi atribut atau fitur yang digunakan [6]. Bahkan, sebagian besar riset membahas bagaimana meningkatkan performa klasifikasi dengan melakukan pemberian bobot terhadap fitur pada data. *Feature selection* dapat digunakan dalam pengurangan dimensi dengan mengidentifikasi fitur yang berpengaruh, dan mengurangi atribut lain yang kurang relevan [7]. Salah satu jenis *feature selection* adalah *Backward Elimination*. *Backward Elimination* memiliki kemampuan untuk mendapatkan beberapa atribut yang semula memiliki tingkat klasifikasi rendah secara individu namun menghasilkan akurasi yang tinggi ketika digabungkan dengan atribut lain. *Backward Elimination* bekerja dengan melakukan seleksi atribut yang memiliki kedekatan setiap kombinasi dengan target melalui pemanfaatan regresi statistik [8]. Fitur ini biasanya digunakan untuk meningkatkan efektivitas dan akurasi algoritma klasifikasi [9].

Penelitian melakukan penambahan fitur *Backward Elimination* pada klasifikasi *client telemarketing* menggunakan algoritme *k-NN* [10]. Hasilnya, percobaan dengan algoritme *k-NN* memiliki akurasi tertinggi sebesar 88.25% dengan nilai *k* 11 dan AUC sebesar 0.793. Sedangkan hasil penerapan algoritme *k-NN* dan *Backward Elimination* mendapatkan akurasi tertinggi 89.43% dengan nilai *k* 9 dan nilai AUC sebesar 0.80, yang termasuk kategori *Good Classification* [11]. Artinya, akurasi meningkat sebesar 1,18% dibandingkan hanya dengan penggunaan algoritme *k-NN* saja [12].

Di sisi lain, algoritme *Naïve Bayes* disebut sebagai salah satu algoritme terbaik dalam klasifikasi dataset yang memiliki tipe data campuran [13]. *Naïve Bayes* dapat menganalisis

variabel yang paling mempengaruhi dalam bentuk Penelitian menunjukkan bahwa *Naïve Bayes* mampu mengklasifikasikan pengujian data dengan benar, yang menghasilkan akurasi rata-rata sebesar 86,16% dan nilai *error* 13,84% dari hasil pengujiannya [14].

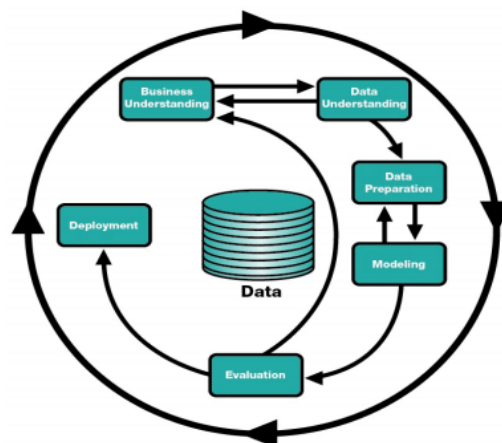
PERMASALAHAN

Dari paparan tersebut diatas maka permasalahan pada penelitian ini yaitu ingin mengetahui faktor apa saja yang dapat meningkatkan kepuasan pelanggan pada dataset pelanggan pada sebuah perusahaan yang bergerak di bidang jasa transportasi udara. Terdapat 23 atribut yang tersedia berdasarkan tingkat kepuasan pelanggan terhadap pelayanan yang diberikan maskapai. Fitur *Backward Elimination* akan diterapkan sebagai perbandingan akurasi kepuasan pelanggan dengan menggunakan algoritme *k*-NN dan algoritme *Naïve Bayes*. Oleh karena itu, berdasarkan latar belakang diatas dilakukan penelitian untuk klasifikasi kepuasan pelanggan dengan optimasi fitur *Backward Elimination* pada algoritme *k*-NN dan *Naïve Bayes*, sehingga diharapkan dapat mengetahui atribut yang berpengaruh terhadap peningkatan pelayanan pelanggan.

METODOLOGI PENELITIAN

METODOLOGI PENELITIAN

Penelitian ini menggunakan tahapan *Cross-Industry Standard Process Model for Data Mining* (CRISP-DM) dalam penerapan algoritme *k*-NN dan *Naïve Bayes* dengan *Backward Elimination* pada klasifikasi kepuasan pelanggan. CRISP-DM bertujuan untuk memberikan pedoman kepada praktisi untuk melakukan penambangan data pada set data besar (Plotnikova et al., 2020). Adapun tahapannya dapat dilihat pada Gambar 1.



Gambar 1. Tahap CRISP-DM
(Sumber: media.neliti.com)

2.1 *Business Understanding*

Tahap ini dilakukan untuk mendapatkan pemahaman tujuan berdasarkan perspektif bisnis untuk kemudian diubah menjadi definisi masalah *data mining*. Setelah itu ditentukan bagaimana solusi yang tepat dalam menangani permasalahan tersebut [15]. Adapun masalah penelitian ini adalah bagaimana klasifikasi kepuasan pelanggan pada salah satu maskapai penerbangan berdasarkan beberapa aspek pelayanan yang telah diberikan. Solusi yang ditawarkan yaitu penerapan *data mining* menggunakan algoritme *k*-NN dan *Naïve Bayes* untuk mengetahui efektifitas fitur *Backward Elimination* dalam pengurangan atribut dan peningkatan akurasi.

2.2 Data Understanding

Setelah melakukan pengumpulan data, tahap ini melakukan proses pemahaman dan identifikasi kualitas data, lalu menemukan pengetahuan awal untuk dapat membentuk hipotesis. Hipotesis awal peneliti berdasarkan data yang tersedia dan penelitian sebelumnya yang dilakukan fitur *Backward Elimination* dapat mengurangi atribut yang kurang berpengaruh sehingga meningkatkan performa akurasi dibandingkan dengan hanya penggunaan algoritme *k-NN* [16].

2.3 Data Preparation

Data preparation merupakan tahap persiapan data yang cenderung dilakukan berulang kali hingga data benar-benar siap untuk selanjutnya dilakukan pemodelan. Tahap ini terdiri dari *data cleaning*, *data selection*, dan penentuan label pada atribut data (Plotnikova et al., 2020).

2.4 Modelling

Pada tahap ini dilakukan pemilihan dan penerapan teknik pemodelan berdasarkan parameter yang tersedia. Biasanya, beberapa teknik digunakan untuk menangani masalah data mining yang serupa. Pada penelitian ini teknik klasifikasi dipilih untuk membandingkan penerapan algoritme *k-NN* dengan dan tanpa *Backward Elimination*, serta algoritme *Naïve Bayes* dengan dan tanpa *Backward Elimination*[17].

2.5 Evaluation

Evaluasi dilakukan terhadap hasil penerapan model apakah sudah memenuhi tujuan penelitian atau belum, untuk selanjutnya dilakukan penentuan keputusan mengenai penggunaan hasil data mining [18]. Tahap evaluasi merupakan tahap analisis terhadap hasil penerapan model berdasarkan nilai tingkat akurasi, *recall*, *precision*, nilai AUC dan jumlah pengurangan atributnya. *Recall* dan *Precision* diambil berdasarkan tingkat keberhasilan sistem dan ketepatan antara informasi yang diminta oleh user dengan jawaban yang diberikan oleh system [19].

2.6 Deployment

Tahapan penyebaran dilakukan untuk merepresentasikan hasil pengetahuan dan informasi yang didapatkan ke dalam bentuk yang mudah dipahami pengguna [20].

HASIL DAN PEMBAHASAN

3.1 Business Understanding

Dengan memanfaatkan sumber data yang ada, dapat dianalisis dan diprediksi menggunakan teknik data mining yang tujuan bisnisnya adalah membuat klasifikasi kepuasan pelanggan berdasarkan optimalisasi *Backward Elimination* untuk prediksi kepuasan pelanggan terhadap pelayanan perusahaan.

3.2 Data Understanding

Dataset yang digunakan merupakan kumpulan detail pelanggan salah satu maskapai penerbangan. Tujuan utama dari kumpulan data ini adalah untuk mengetahui aspek mana yang harus ditekankan dari layanan yang ditawarkan perusahaan, sehingga dapat meningkatkan kepuasan dan loyalitas pelanggan. *Dataset* yang digunakan adalah data publik dari www.kaggle.com, yang terdiri dari 23 atribut dan berjumlah 129.880 data. Adapun keterangan mengenai atribut data terdapat pada Tabel 1.

Tabel 1. Atribut Dataset

No	Nama Atribut	Keterangan	Nilai Referensi
1	<i>Satisfaction</i>	Kepuasan Pelanggan	Puas, Tidak puas
2	<i>Gender</i>	Jenis Kelamin	Perempuan, Laki-laki
3	<i>Customer Type</i>	Tipe Pelanggan	Loyal, Tidak loyal
4	<i>Age</i>	Usia	Min. 7, Maks. 85 tahun
5	<i>Type of Travel</i>	Jenis Perjalanan	Bisnis, Pribadi
6	<i>Class</i>	Kelas	Bisnis, Bisnis Plus, Ekonomi
7	<i>Flight Distance</i>	Jarak Penerbangan	Min. 50, Maks. 6951
8	<i>Seat comfort</i>	Kenyamanan Kursi	Nilai kepuasan dari 0-5
9	<i>Departure/Arrival time convenient</i>	Waktu Keberangkatan/Kedatangan	Nilai kepuasan dari 0-5
10	<i>Food and drink</i>	Makanan dan Minuman	Nilai kepuasan dari 0-5
11	<i>Gate location</i>	Lokasi <i>Gate</i>	Nilai kepuasan dari 0-5
12	<i>Inflight wifi service</i>	Layanan <i>wifi</i> dalam pesawat	Nilai kepuasan dari 0-5
13	<i>Inflight entertainment</i>	Hiburan dalam Pesawat	Nilai kepuasan dari 0-5
14	<i>Online support</i>	Dukungan <i>Online</i>	Nilai kepuasan dari 0-5
15	<i>Ease of Online booking</i>	Kemudahan pemesanan <i>online</i>	Nilai kepuasan dari 0-5
16	<i>On-board service</i>	Layanan <i>on-board</i>	Nilai kepuasan dari 0-5
17	<i>Leg room service</i>	Layanan <i>leg room</i>	Nilai kepuasan dari 0-5
18	<i>Baggage handling</i>	Penanganan Bagasi	Nilai kepuasan dari 0-5
19	<i>Checkin service</i>	Layanan <i>Checkin</i>	Nilai kepuasan dari 0-5
20	<i>Cleanliness</i>	Kebersihan	Nilai kepuasan dari 0-5
21	<i>Online boarding</i>	<i>Boarding Online</i>	Nilai kepuasan dari 0-5
22	<i>Departure Delay in Minutes</i>	Keterlambatan Keberangkatan (Menit)	Min. 0, Maks. 1592
23	<i>Arrival Delay in Minutes</i>	Keterlambatan Kedatangan (Menit)	Min. 0, Maks. 1585

(Sumber : www.kaggle.com)

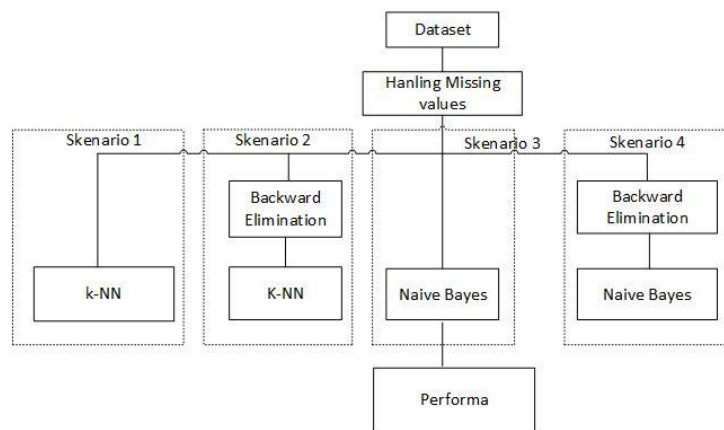
3.3 Data Preparation

Dari total 129.880 data, dilakukan pengambilan sampel secara acak menjadi 5000 data. Atribut *satisfaction* dipilih sebagai label pada data. Kemudian, dilakukan juga penanganan *missing value* menggunakan nilai rata – rata (*mean*) sebagai pengganti pada data yang bernilai kosong.

satisfaction	Gender	Customer Type	Age	Type of Travel	Class	Flight Distance	Arrival Delay in Minutes ↑
dissatisfied	Male	Loyal Customer	58	Personal Travel	Eco Plus	1726	?
dissatisfied	Male	Loyal Customer	42	Personal Travel	Eco	3408	?
satisfied	Female	Loyal Customer	70	Personal Travel	Eco	636	?
dissatisfied	Male	Loyal Customer	59	Personal Travel	Eco Plus	2426	?
dissatisfied	Male	Loyal Customer	39	Personal Travel	Eco	1787	?
satisfied	Female	Loyal Customer	23	Personal Travel	Eco	2906	?
dissatisfied	Male	Loyal Customer	57	Personal Travel	Eco	2778	?
satisfied	Female	Loyal Customer	64	Personal Travel	Eco	593	?
satisfied	Female	Loyal Customer	55	Personal Travel	Business	480	?
dissatisfied	Male	Loyal Customer	18	Personal Travel	Eco	2754	?
satisfied	Female	Loyal Customer	43	Personal Travel	Eco	1212	?
satisfied	Female	Loyal Customer	29	Personal Travel	Eco	2726	?
dissatisfied	Male	Loyal Customer	46	Personal Travel	Eco	1680	?

Gambar 2. Missing Value pada dataset

3.4 Modelling



Gambar 2. Diagram Skenario Pengujian

Pemodelan dilakukan menggunakan algoritme *k*-NN dan *Naive Bayes* dengan menerapkan *Backward Elimination* untuk mengurangi atribut yang kurang relevan serta meningkatkan akurasi model klasifikasi kepuasan pelanggan. Dalam mengimplementasikan pemodelan ini dilakukan 4 skenario pengujian untuk mengetahui model *data mining* terbaik. Setiap skenario diuji menggunakan 10 *k-fold Cross Validation*, untuk dilakukan perbandingan berdasarkan nilai *recall*, akurasi, *precision*, nilai AUC, dan jumlah atribut.

3.5 Evaluation

Dari hasil penerapan klasifikasi *data mining* pada data pelanggan maskapai penerbangan, diperoleh perbandingan pada skenario pengujian yang telah dilakukan seperti pada Tabel 2.

Tabel 2. Hasil Perbandingan 4 Skenario Pengujian

	Skenario 1	Skenario 2	Skenario 3	Skenario 4
Algoritme	<i>k</i> -NN	<i>k</i> -NN + BE	NB	NB + BE
Akurasi (%)	63.54	97.28	98.92	99.04
Precision (%)	57.46	95.80	97.80	97.85
Recall (%)	55.56	97.94	99.72	99.95
AUC	0.689	0.996	1	1
Jumlah Atribut	23	11	23	22

Keterangan :

BE = *Backward Elimination*
NB = *Naïve Bayes*

Berdasarkan Tabel 2, hasil pengujian menggunakan *Cross Validation* dengan nilai *k-fold* 10 ditemukan bahwa skenario 4 yaitu algoritme *Naïve Bayes* dengan *Backward Elimination* memiliki akurasi tertinggi sebesar 99.04%, dengan nilai AUC 1 yang masuk dalam kategori *excellent classification*. Sedangkan akurasi terendah terdapat pada skenario 1 yaitu algoritme *k-NN* tanpa *Backward Elimination* dengan nilai akurasi 63.54% dan nilai AUC 0.689 yang masuk kategori *poor classification*. Adapun klasifikasi nilai AUC tertera pada Tabel 3.

Tabel 3. Klasifikasi Nilai AUC

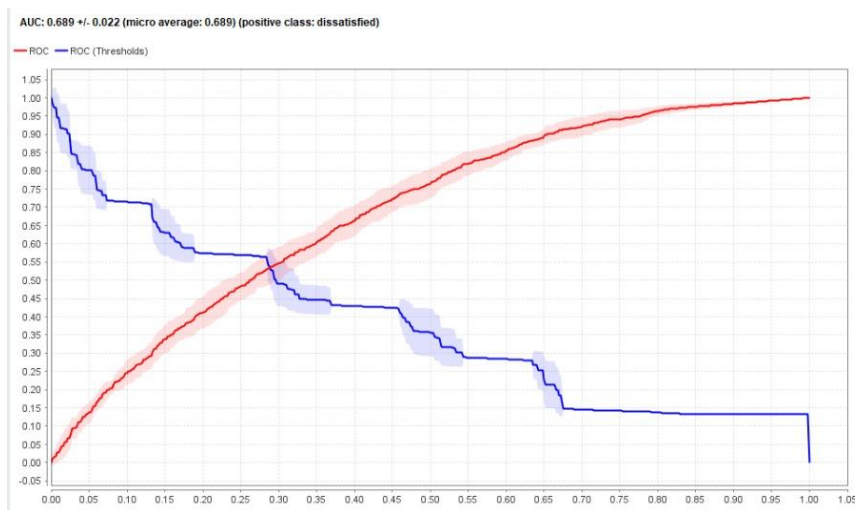
Nilai AUC	Klasifikasi
0.90 - 1.00	<i>Excellent Classification</i>
0.80 - 0.90	<i>Good Classification</i>
0.70 - 0.80	<i>Fair Classification</i>
0.60 - 0.70	<i>Poor Classification</i>
0.50 - 0.60	<i>Failure</i>

(Sumber : Amri, 2020)

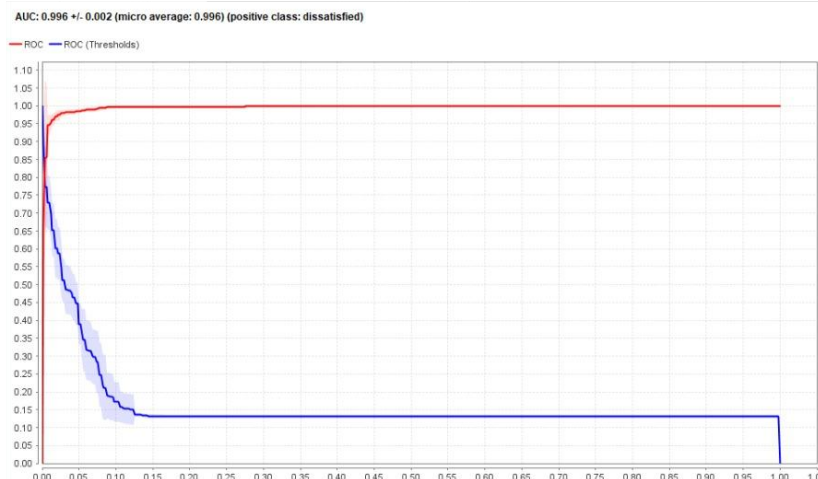
Selain meningkatnya akurasi, nilai *precision* dan *recall* dari masing-masing skenario juga meningkat. Artinya, tingkat keberhasilan sistem dan ketepatan informasi lebih baik dengan penerapan fitur *Backward Selection*.

Perbandingan Skenario 1 dan Skenario 2

Grafik perbandingan skenario 1 dan skenario 2 yaitu penerapan algoritme *k-NN* tanpa dan dengan *Backward Elimination* disajikan pada Gambar 3 dan Gambar 4.



Gambar 3. Grafik AUC Skenario 1



Gambar 4. Grafik AUC Skenario 2

Berdasarkan hasil pengujian, nilai AUC yang terdapat pada Gambar 3 grafik skenario 1 yaitu penerapan *k*-NN tanpa *Backward Elimination* menghasilkan nilai sebesar 0.689. Nilai tersebut meningkat menjadi 0.996 dengan penambahan fitur *Backward Elimination* seperti yang terdapat pada gambar 4 grafik skenario 2, artinya nilai AUC meningkat sebesar 0.307. Selain itu, didapatkan sebanyak 13 atribut yang paling relevan dengan data. Jumlah atribut berkurang cukup banyak sehingga selain meningkatkan nilai akurasi yang sangat signifikan, fitur *Backward Elimination* paling efektif dalam pengurangan jumlah atribut data menggunakan algoritme *k*-NN. Hasil perbandingan terdapat pada Tabel 4.

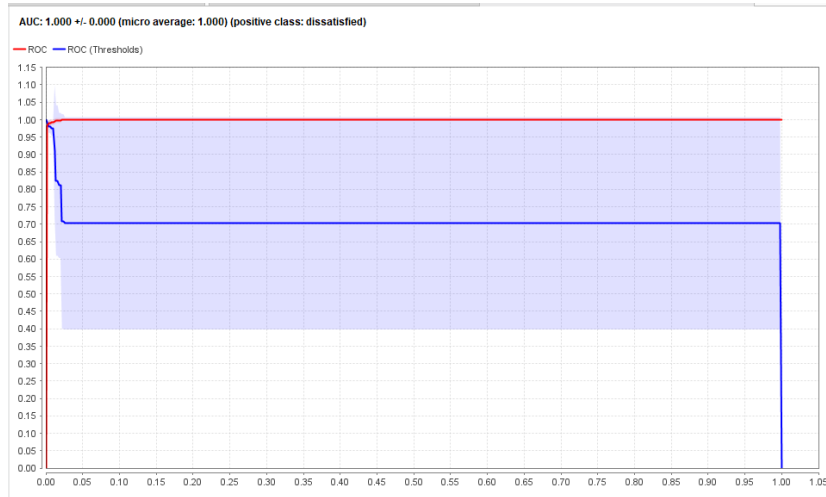
Tabel 4. Perbandingan Hasil Skenario 1 dan Skenario 2

Atribut Setelah Pengurangan	Selisih			AUC
	Akurasi	Precision	Recall	
<i>Satisfaction, Gender, Customer Type, Type of Travel, Class, Seat Comfort, Gate Location, Inflight Wifi Service, Online Support, Ease of Online Booking, Online Boarding</i>	33.74 %	38.34 %	42.38 %	0.307

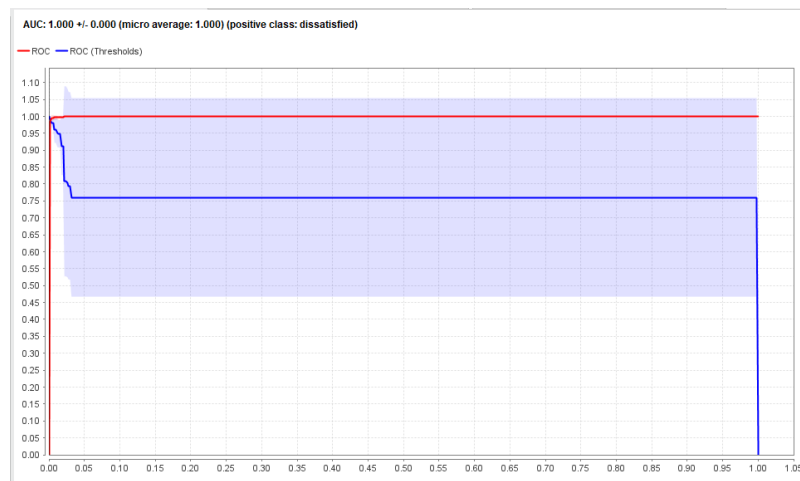
Tabel 4 menunjukkan kenaikan akurasi sebesar 33.74%, dari yang semula 63.54% menjadi 97.28%, dapat disimpulkan bahwa fitur *Backward Elimination* sangat berpengaruh dalam meningkatkan nilai akurasi. Tidak hanya itu, nilai *precision* dan *recall* pun meningkat dengan masing-masing peningkatan *precision* sebesar 38.34% dan *recall* meningkat 42.38%.

Perbandingan Skenario 3 dan Skenario 4

Grafik perbandingan skenario 3 dan skenario 4 yaitu penerapan algoritme *Naïve Bayes* tanpa dan dengan *Backward Elimination* disajikan pada Gambar 5 dan Gambar 6.



Gambar 5. Grafik AUC Skenario 3



Gambar 6. Grafik AUC Skenario 4

Pada perbandingan skenario 3 dan skenario 4 yaitu penerapan algoritme *Naïve Bayes* tanpa dan dengan *Backward Elimination*, dapat menunjukkan bahwa fitur *Backward Elimination* dapat meningkatkan nilai akurasi sebesar 0.12%, *precision* meningkat 0.05%, serta *recall* meningkat sebesar 0.23%. Meskipun peningkatannya tidak sebesar pada algoritma *k-NN*, namun cukup berpengaruh dalam menambah nilai akurasi, *precision*, dan *recall*. Jumlah atribut yang berkurang hanya 1, yang berarti penggunaan *Backward Elimination* pada algoritme *Naïve Bayes* tidak terlalu signifikan. Adapun perbandingannya terdapat pada Tabel 5.

Tabel 5. Perbandingan Hasil Skenario 3 dan Skenario 4

Atribut Setelah Pengurangan	Selisih			
	Akurasi	Precision	Recall	AUC
<i>Satisfaction, Gender, Customer Type, Age, Type of Travel, Flight Distance, Seat comfort, Departure/Arrival time convenient, Food and</i>	0.12 %	0.05 %	0.23 %	0

*drink, Gate location,
Inflight wifi service,
Inflight entertainment,
Online support, Ease of
Online booking, On-board
service, Leg room service,
Baggage handling, Checkin
service, Cleanliness,
Online boarding,
Departure Delay in
Minutes, Arrival Delay in
Minutes*

Berdasarkan hasil penelitian, pada Tabel 5 menunjukkan selisih dari hasil skenario 3 dan 4 yang telah dilakukan, dapat dibandingkan bahwa terdapat selisih dari 24 atribut menjadi 23 atribut setelah dilakukan pengujian pada algoritma *Naïve Bayes* dengan menerapkan fitur *Backward Elimination*.

Dari hasil seluruh pengujian yang dilakukan dapat dikatakan bahwa kinerja *Backward Elimination* dapat meningkatkan akurasi terhadap kedua algoritme tersebut. Fitur *Backward Elimination* ternyata dapat bekerja lebih optimal pada algoritme *k-NN* daripada algoritme *Naïve Bayes*. Seperti yang tertera pada Tabel 4, algoritme *k-NN* dapat meningkatkan akurasi sebesar 33,74% dengan $k=7$ sedangkan pada Tabel 5 tertera bahwa algoritme *Naïve Bayes* akurasi hanya meningkat sebesar 0,12%.

3.6 Deployment

Berdasarkan hasil penerapan *Backward Elimination* yang paling optimal, terdapat 11 atribut yang berpengaruh. Oleh karena itu, rekomendasi untuk aspek yang perlu ditekankan diantaranya kenyamanan kursi, lokasi *gate*, layanan *wifi* dalam pesawat, dukungan *online*, kemudahan pemesanan *online*, serta *boarding online*.

KESIMPULAN

Hasil penelitian pada klasifikasi kepuasan pelanggan dengan algoritme *k-NN* dan *Naïve Bayes* menggunakan seleksi fitur *Backward Elimination* menghasilkan akurasi yang cukup signifikan. Akurasi pada algoritme *k-NN* dengan nilai $k=7$ mendapatkan nilai akurasi sebesar 97.28% sedangkan algoritme *Naïve Bayes* menghasilkan akurasi sebesar 99.04%. Penggunaan *Backward Elimination* dalam klasifikasi kepuasan pelanggan bekerja dengan optimal. Kinerja fitur tersebut telah meningkatkan akurasi sebesar 33.74% pada algoritme *k-NN* dan 0,12 % terhadap algoritme *Naïve Bayes*.

Dengan menggunakan seleksi fitur *Backward Elimination*, performa akurasi dan nilai AUC yang dihasilkan lebih baik dibandingkan tanpa penggunaan fitur tersebut. Selain itu, *Backward Elimination* juga bekerja dengan baik dalam pengurangan atribut yang kurang relevan. Hal tersebut menunjukkan bahwa penerapan Fitur *Backward Elimination* lebih optimal terhadap algoritme *k-NN*, dibandingkan algoritme *Naïve Bayes*.

SARAN

Dari penelitian yang penulis lakukan maka diharapkan untuk penelitian selanjutnya dilakukan perbandingan nilai k pada algoritme *k-NN* agar dapat mengetahui perbedaan nilai akurasi, serta melakukan pembobotan dari setiap atribut dengan seleksi fitur yang lain.

DAFTAR PUSTAKA

- [1] A. Firatmadi, "Pengaruh Kualitas Pelayanan dan Persepsi Harga Terhadap Kepuasan Pelanggan Serta Dampaknya Terhadap Loyalitas Pelanggan," *Journal of Business Studies*, vol. 2, no. 2, pp. 80–105, 2017.
- [2] M. G. Pradana and P. H. Saputro, "KOMPARASI METODE NAÏVE BAYES DAN C4.5 DALAM KLASIFIKASI LOYALITAS PELANGGAN TERHADAP LAYANAN PERUSAHAAN," *Indonesian Journal of Business Intelligence (IJUBI)*, vol. 3, no. 1, pp. 20–24, 2020.
- [3] A. Maheswari and N. M. A. Aksari, "Peran Kepuasan Pelanggan Memediasi Kualitas Layanan Terhadap Loyalitas Pelanggan Pada Pt. Airasia Indonesia," *E-Jurnal Ekonomi Dan Bisnis Universitas Udayana*, vol. 3, p. 315, 2019.
- [4] R. T. Prasetyo, "SELEKSI FITUR DAN OPTIMASI PARAMETER k-NN BERBASIS ALGORITMA GENETIKA PADA DATASET MEDIS," *Jurnal Responsif: Riset Sains & Informatika*, vol. 2, no. 2, pp. 213–221, 2020.
- [5] M. Fansyuri, "ANALISA ALGORITMA KLASIFIKASI K-NEAREST NEIGHBOR DALAM MENENTUKAN NILAI AKURASI TERHADAP KEPUASAN PELANGGAN (STUDY KASUS PT. TRIGATRA KOMUNIKATAMA)," *Jurnal Ilmiah Humanika*, vol. 3, no. 1, pp. 29–33, 2020.
- [6] M. A. Ghani, "ANALISIS PERFORMANSI QUALITY OF SERVICE (QOS) PADA JARINGAN MULTI PROTOCOL LABEL SWITCHING DENGAN METODE INTSERV," *Jurnal Manajemen Informatika*, vol. 9, no. 2, 2019.
- [7] F. Ma'arif and T. Arifin, "Optimasi Fitur Menggunakan Backward Elimination Dan Algoritma SVM Untuk Klasifikasi Kanker Payudara," *Jurnal Informatika*, vol. 4, no. 1, 2017.
- [8] I. W. Gamadarenda and I. Waspada, "Implementasi Data Mining untuk Deteksi Penyakit Ginjal Kronis (PGK) menggunakan K-Nearest Neighbor (KNN) dengan Backward Elimination," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, vol. 7, no. 2, 2020.
- [9] I. Herliawan, M. Iqbal, W. Gata, A. Rifai, and J. J. Purnama, "CLASSIFICATION OF LIVER DISEASE BY APPLYING RANDOM FOREST ALGORITHM AND BACKWARD ELIMINATION," *JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer)*, vol. 6, no. 1, pp. 89–94, 2020.
- [10] M. Sadikin and F. Alfiandi, "Comparative Study of Classification Method on Customer Candidate Data to Predict its Potential Risk.," *International Journal of Electrical & Computer Engineering (2088-8708)*, vol. 8, no. 6, 2018.
- [11] B. E. Sibarani, "Smart Farmer Sebagai Optimalisasi Digital Platform Dalam Pemasaran Produk Pertanian Pada Masa Pandemi Covid-19," *Technomedia Journal*, vol. 6, no. 01 Agustus, 2021.
- [12] P. P. P. Pangestu and R. Yusuf, "Implementasi Metode QINQ Pada Jaringan Metro Ethernet Untuk Memaksimalkan Penggunaan VLAN Menggunakan Teknologi GPON Studi Kasus: PT. Telkom Indonesia," *Technomedia Journal*, vol. 6, no. 1 Agustus, 2021.
- [13] K. Nalakhudin, M. Imron, and M. A. W. Prasetyo, "Pemanfaatan Notifikasi Telegram Untuk Monitoring Perangkat CCTV Rumah Sakit Orthopaedi Purwokerto," *Technomedia Journal*, vol. 6, no. 01 Agustus, 2021.
- [14] R. Rosyid and M. A. W. Prasetyo, "Robot Peraga 12 Gerakan Pengaturan Lalu Lintas Berbasis Arduino Mega 2560," *Technomedia Journal*, vol. 5, no. 2, pp. 193–205, 2021.
- [15] B. Basri and A. Qashlim, "Relay Kontrol Menggunakan Google Firebase dan Node MCU pada Sistem Smart Home," *Technomedia Journal*, vol. 6, no. 01 Agustus, 2021.

- [16] I. B. A. Peling, I. N. Arnawan, I. P. A. Arthawan, and I. G. N. Janardana, "Implementation of Data Mining To Predict Period of Students Study Using Naive Bayes Algorithm," *Int. J. Eng. Emerg. Technol*, vol. 2, no. 1, p. 53, 2017.
- [17] V. Plotnikova, M. Dumas, and F. Milani, "Adaptations of data mining methodologies: a systematic literature review," *PeerJ Computer Science*, vol. 6, p. e267, 2020.
- [18] S. Amri, "Perbandingan Kerangka Model Klasifikasi untuk Pemilihan Metode Kontrasepsi dengan Pendekatan CRIPS-DM," *Information Science and Library*, vol. 1, no. 1, pp. 14–23, 2020.
- [19] F. Ruan, L. Hou, T. Zhang, and H. Li, "A modified backward elimination approach for the rapid classification of Chinese ceramics using laser-induced breakdown spectroscopy and chemometrics," *Journal of Analytical Atomic Spectrometry*, vol. 35, no. 3, pp. 518–525, 2020.
- [20] E. Musvida, "Recall dan Precision pada Sistem Temu Kembali Informasi Bidang Ilmu Fikih di Perpustakaan UIN Ar-Raniry Banda Aceh." UIN Ar-Raniry Banda Aceh, 2017.