sudo
Jurnal Teknik Informatika
https://jurnal.ilmubersama.com/index.php/sudo

Machine Learning

Analisis Algoritma K-Nearest Neighbors dengan Teknik SMOTE untuk Klasifikasi Kepuasan Pelanggan pada Layanan Laundry

Dea Syahfira Hasibuan[*], *Sriani*

Fakultas Sains dan Teknologi, Ilmu Komputer, Universitas Islam Negeri Sumatera Utara, Medan, Indonesia

INFORMASI ARTIKEL

Diterima Redaksi: 01 September 2026
Revisi Akhir: 22 September 2026
Diterbitkan *Online*: 29 September 2026

KATA KUNCI

K-Nearest Neighbors
Kepuasan Pelanggan
Klasifikasi
Laundry
SMOTE

KORESPONDENSI (*)

Phone: +62 852-6091-5899
E-mail: deasyahfirahsb06@gmail.com

A B S T R A K

Zayn Laundry menghadapi persaingan bisnis laundry yang semakin ketat di Kota Medan sehingga diperlukan pengukuran kepuasan pelanggan secara objektif. Penelitian ini bertujuan menerapkan algoritma K-Nearest Neighbors (KNN) untuk mengklasifikasikan tingkat kepuasan pelanggan, mengetahui tingkat kepuasan berdasarkan data kuesioner, dan mengetahui akurasi model yang dihasilkan. Data penelitian diperoleh dari 250 pelanggan melalui kuesioner yang terdiri dari 15 pernyataan dengan skala Likert 1–4 yang mencakup lima variabel, yaitu kualitas pelayanan, harga, kemudahan akses, pengalaman pelanggan, dan kualitas hasil layanan. Data melalui tahap pra-pemrosesan, uji validitas dan reliabilitas, kemudian dikelompokkan menjadi tiga kategori, yaitu Tidak Puas, Cukup Puas, dan Puas dengan kategori Puas sebagai kategori mayoritas. Ketidakseimbangan kelas ditangani menggunakan Synthetic Minority Over-sampling Technique (SMOTE). Klasifikasi dilakukan menggunakan KNN dengan Euclidean Distance, sedangkan nilai K terbaik ditentukan melalui Stratified K-Fold Cross Validation. Hasil penelitian menunjukkan K=3 sebagai nilai terbaik. Model KNN sebelum SMOTE menghasilkan akurasi sebesar 96%, precision sebesar 0,96%, recall sebesar 0,96%, dan F1-score sebesar 0,95%. Setelah menggunakan SMOTE, model memperoleh accuracy sebesar 98%, precision, recall, dan F1-score masing-masing sebesar 0,98%. Hasil tersebut menunjukkan KNN dengan SMOTE dapat meningkatkan akurasi model klasifikasi.

PENDAHULUAN

Laundry merupakan salah satu jasa penting yang mendukung gaya hidup modern dan masyarakat yang ada diperkotaan. Tingginya tingkat kesibukan masyarakat diperkotaan untuk menjalankan aktivitas sehari-hari menyebabkan terbatasnya waktu mereka dalam mengurus kebutuhan rumah tangga seperti, mencuci pakaian, sehingga hal tersebut mendorong meningkatnya kebutuhan terhadap jasa laundry. Dari adanya keterbatasan waktu yang mereka alami dapat menjadikannya potensi bisnis strategis untuk mengembangkan perekonomian lokal. Oleh karena itu, Layanan Laundry merupakan salah satu sektor jasa yang patut ditingkatkan kualitasnya dan dikembangkan secara berkelanjutan [1].

Di era bisnis saat ini, setiap pelaku bisnis pasti mengalami persaingan yang sangat ketat, salah satunya adalah bisnis Laundry. Kepuasan pelanggan menjadi faktor kunci dalam mempertahankan keberlangsungan bisnis. Kepuasan pelanggan memiliki peran yang penting dalam menciptakan loyalitas pelanggan yang berdampak pada peningkatan keuntungan usaha. Oleh karena itu, pelaku usaha perlu memahami faktor-faktor yang mempengaruhi kepuasan pelanggan agar dapat memberikan layanan yang sesuai dengan yang diharapan pelanggan [2].

Pengelola Zayn Laundry menghadapi kesulitan utama dalam menjaga loyalitas pelanggan di tengah persaingan ketat sesama penyedia Layanan Laundry di Medan. Hal ini dikarenakan belum adanya data yang memadai dan jelas untuk mengukur tingkat kepuasan pelanggan secara akurat. Dan penilaian kepuasan saat ini masih bersifat subjektif berdasarkan

https://doi.org/10.56211/sudo.v5i3.2014 Attribution-ShareAlike 4.0 International Some rights reserved

pengamatan pribadi pengelola tanpa dukungan data kuesioner, sehingga belum diketahui apakah sebagian besar pelanggan termasuk dalam kategori puas, cukup puas atau tidak puas.

Menurut data dari Asosiasi perusahaan laundry Indonesia, industri laundry nasional tumbuh sebesar 15% per tahun, dengan permintaan untuk Layanan Laundry professional yang semakin meningkat. Kondisi ini menuntut pelaku usaha untuk mengelola bisnis secara lebih terstruktur dan berbasis data agar mampu meningkatkan kualitas layanan serta mempertahankan kepercayaan konsumen [3].

Salah satu metode yang dapat digunakan untuk mengatasi permasalahan tersebut adalah data mining dengan algoritma K-Nearest Neighbors (KNN). Algoritma KNN merupakan metode klasifikasi yang mengelompokkan data berdasarkan kedekatan jarak antar data, sehingga dapat digunakan untuk mengklasifikasikan tingkat kepuasan pelanggan secara akurat berdasarkan data kuesioner. Namun, pada data kepuasan pelanggan umumnya ditemukan kondisi imbalanced data, yaitu jumlah pelanggan pada kategori puas jauh lebih banyak dibandingkan kategori cukup puas dan tidak puas. Kondisi ini dapat menyebabkan algoritma KNN bias dalam memprediksi kelas mayoritas dan kurang akurat dalam mengenali kelas minoritas. Oleh karena itu, diperlukan teknik Synthetic Minority Oversampling Technique (SMOTE) untuk menyeimbangkan distribusi data sebelum proses klasifikasi dilakukan, sehingga akurasi model KNN dapat lebih optimal.

Analisis kepuasan yang telah diteliti oleh [4] Dalam penelitiannya yang berjudul "Klasifikasi Kepuasan Masyarakat terhadap Kinerja Petugas Sensus menggunakan K-Nearest Neighbours". Penelitian ini menyimpulkan bahwa tingkat kepuasan pelanggan terbagi ke dalam lima kategori kepuasan: Sangat Puas, Puas, Cukup, Tidak Puas, dan Sangat Tidak Puas. Model KNN yang diterapkan menggunakan Python menghasilkan akurasi sebesar 75%.

Analisis kepuasan pelanggan yang telah diteliti oleh [5] Dalam penelitiannya yang berjudul "Implementasi Metode K-Nearest Neighbor (K-NN) untuk Analisis Sentimen Kepuasan Pengguna Aplikasi Teknologi Finansial FLIP". Penelitian ini menyimpulkan bahwa algoritma K-Nearest Neighbor (KNN) mampu mengklasifikasikan sentimen pengguna ke dalam tiga kategori, yaitu positif, negatif, dan netral, serta menghasilkan tingkat akurasi klasifikasi sebesar 76,68%.

Berdasarkan kedua penelitian tersebut, ditemukan bahwa gap dari penelitian ini adalah belum adanya penelitian yang menerapkan algoritma K-Nearest Neighbors (KNN) yang dikombinasikan dengan teknik SMOTE untuk mengklasifikasikan tingkat kepuasan pelanggan pada sektor jasa laundry. [4] telah menerapkan KNN untuk klasifikasi kepuasan, namun tanpa penanganan imbalanced data menggunakan SMOTE sehingga akurasi model berpotensi tidak optimal. [5] juga telah membuktikan kemampuan KNN dalam klasifikasi, namun pada objek yang berbeda dan tanpa mempertimbangkan ketidakseimbangan data. Selain itu, penilaian kepuasan pelanggan di Zayn Laundry masih bersifat subjektif dan belum menggunakan metode analisis data yang terstruktur, sehingga belum menghasilkan penilaian yang akurat.

Berdasarkan uraian permasalahan tersebut maka penelitian ini diberi judul "Analisis Algoritma K-Nearest Neighbors Dengan Teknik SMOTE Untuk Klasifikasi Kepuasan Pelanggan Pada Layanan Laundry". Penelitian ini bertujuan untuk mengetahui tingkat kepuasan pelanggan pada Zayn Laundry berdasarkan data kuesioner, serta menerapkan algoritma K-Nearest Neighbors (KNN) dalam mengklasifikasikan pelanggan ke dalam kategori puas, cukup puas dan tidak puas dengan akurasi yang optimal menggunakan SMOTE.

TINJAUAN PUSTAKA

Laundry

Gaya hidup modern saat ini menuntut agar seseorang dapat mengatur waktunya se-efesien dan se-efektif mungkin didalam kehidupan sehari-hari. Perubahan tersebut hidup agar lebih praktis dan lebih cepat. Perubahan ini menyebabkan kebutuhan pribadi yang di tuntut menjadi lebih praktis, misalnya dalam mencuci pakaian dan menyetrika. Dengan adanya tuntutan tersebut akhirnya memunculkan jasa yang menawarkan kemudahan dalam mencuci dan menyetrika pakaian atau yang disebut juga dengan laundry [6].

Skala Likert

Skala Likert adalah skala pengukuran yang digunakan untuk mengukur sikap, pendapat, atau persepsi seseorang terhadap suatu objek atau pernyataan. Skala ini terdiri atas beberapa pilihan jawaban yang menunjukkan tingkat persetujuan responden, seperti sangat setuju hingga sangat tidak setuju [7].

Data Mining

Data Mining merupakan salah satu teknik untuk menemukan, mencari, atau menggali informasi atau pengetahuan baru dari sekumpulan dat a yang sangat besar, dengan integrasi atau penggabungan dengan disiplin ilmu lain seperti statistika, kecerdasan buatan, serta machine learning, menjadikan Data Mining sebagai salah satu alat bantu untuk menganalisa data yang kemudian menghasilkan informasi yang berguna [8].

Klasifikasi

Klasifikasi merupakan salah satu teknik dalam data mining untuk mengelompokkan data atau pengamatan yang memiliki kesamaan ke target variabel kategori. Tujuan utama klasifikasi adalah untuk mengidentifikasi pola atau hubungan antara fitur-fitur data yang ada dan target variabel kategori. Hasil dari proses klasifikasi adalah model yang dapat digunakan untuk memprediksi kategori dari data yang belum pernah dilihat sebelumnya [9].

K-Nearest Neighbor (KNN)

Algoritma K-Nearest Neighbors (KNN) merupakan salah satu metode dari Supervised Learning yang digunakan untuk proses klasifikasi. Cara kerja algoritma ini yaitu dengan menghitung jarak antara data baru dengan data yang sudah ada pada dataset, kemudian menentukan kelas berdasarkan mayoritas nilai K dari tetangga terdekat [10].

Untuk menentukan tetangga, KNN menghitung jarak antara data baru dengan setiap data pada dataset. Metode yang sering digunakan dalam menghitung jarak adalah Euclidean Distance. Euclidean Distance adalah ukuran jarak paling terkenal dan paling banyak digunakan dalam data mining, terutama di KNN [11].

$$d(x,y) = \sqrt{x_1 - y_1^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2} \quad (1)$$

Dimana d adalah Menyatakan Jarak, i adalah menyatakan Variabel data, x adalah menyatakan data testing, y adalah menyatakan data training, n adalah Jumlah data.

Data Tak Seimbang (Imbalanced Data)

Data tidak seimbang merupakan suatu keadaan dimana distribusi kelas data tidak seimbang, jumlah kelas data yang satu lebih sedikit atau lebih banyak dibanding dengan jumlah kelas data lainnya. Kelompok kelas data yang lebih sedikit dikenal dengan kelompok minoritas, kelompok kelas data yang lainnya disebut dengan kelompok mayoritas [12].

Ketidakseimbangan data dapat membuat performa model menjadi bias dan berkinerja rendah dalam mendeteksi kelas minoritas. Bias ini disebabkan karena model cenderung lebih mengenal pola mayoritas yang memiliki subset data dominan dan tidak sensitif terhadap kelas minoritas yang memiliki data kurang representatif [13].

SMOTE

SMOTE (*Synthetic Minority Over-sampling Technique*) adalah metode *oversampling* yang menangani ketidak seimbangan kelas pada *dataset* dengan menciptakan sampel sintetis dalam kelas minoritas melalui interpolasi. Ketidakseimbangan kelas terjadi saat frekuensi kelas minoritas jauh lebih rendah. Ini dapat menyebabkan bias dalam prediksi model. SMOTE mengatasi ini dengan menciptakan sampel sintetis dalam kelas minoritas [14].

SMOTE adalah teknik *oversampling* canggih di mana algoritma menciptakan contoh sintetis dari kelas minoritas, bukan hanya menduplikasi contoh yang sudah ada. Ini dilakukan dengan membuat contoh baru dengan interpolasi antara dua contoh kelas minoritas yang sudah ada. SMOTE membantu menghindari masalah duplikasi yang ada pada *oversampling* sederhana [15].

Evaluasi model

Evaluasi adalah tahap akhir dalam proses KDD, tahap ini merupakan visualisasi data dan menghitung performa model. Dan pada tahap ini menggunakan Confusion Matrix.

Confusion Matrix adalah berfungsi untuk menilai output suatu metode klasifikasi dengan menyajikan hasil penelitian model melalui tabel matrik. Proses penilaian dengan penerapan *Confusion Matrix* memperoleh sejumlah nilai *Matrix* penting, seperti *accuracy*, *precision*, *recall*, F1-*Score*.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

$$\text{Precission} = \frac{TP}{TP+FP} \quad (3)$$

$$\text{Reccal} = \frac{TP}{TP+FN} \quad (4)$$

$$\text{F1-Score} = 2 \text{ x } \frac{Precision\ X\ Reccal}{Precision + Reccal} \quad (5)$$

Dimana TP adalah jumlah data yang bernilai positif dan diprediksi benar sebagai positif. FP adalah jumlah data yang bernilai negative tetapi diprediksi sebagai positif. FN adalah jumlah data yang bernilai positif tetapi diprediksi sebagai negatif. TN adalah jumlah data yang bernilai negatif dan diprediksi benar sebagai negatif.

METODOLOGI

Penelitian ini dilaksanakan di Zayn Laundry, yang berlokasi di Jl. Suka Maju Indah, Tj. Selamat, Kec. Sunggal, Kabupaten Deli Serdang, Sumatera Utara. Penelitian ini merupakan penelitian kuantitatif dengan pendekatan data mining, yang bertujuan untuk menganalisis tingkat kepuasan pelanggan pada layanan Zayn Laundry serta mengkalsifikasikan data kepuasan pelanggan dengan menggunakan algoritma K-Nearest Neighbors (KNN) dengan teknik SMOTE untuk menangani ketidakseimbangan data. teknik pengumpulan data yang diterapkan adalah menggunakan studi literatur, wawancara, dan juga kuesioner.

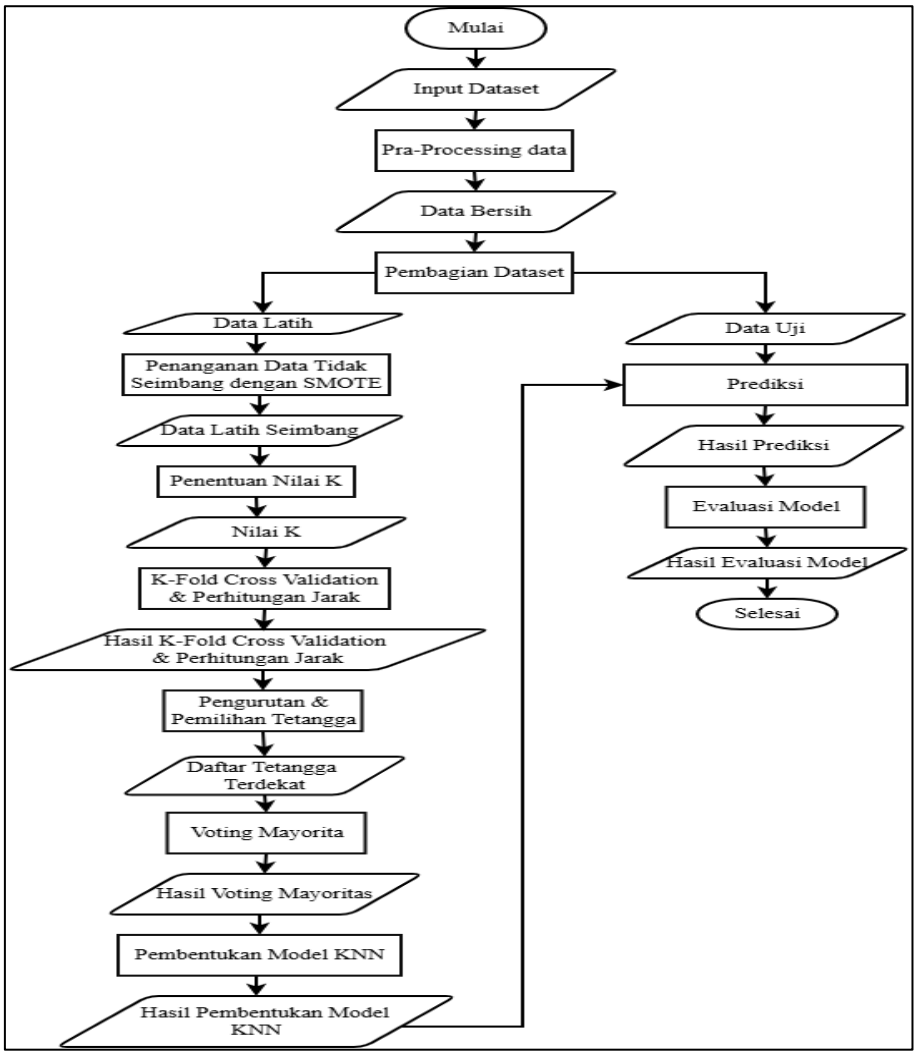
Teknik pengumpulan data yang diterapkan dalam penelitian ini adalah melalui kuesioner yang didistribusikan kepada pelanggan Zayn Laundry untuk memperoleh data mengenai tingkat kepuasan pelanggan. Kuesioner dalam penelitian ini terdiri atas 5 variabel, yaitu kualitas pelayanan, harga, kemudahan akses, pengalaman pelanggan, dan kualitas hasil layanan, dengan masing-masing variabel terdiri atas 3 pernyataan, sehingga total terdapat 15 pernyataan. Data yang diperoleh dari kuesioner selanjutnya digunakan sebagai data penelitian untuk proses klasifikasi menggunakan algoritma K-Nearest Neighbors (KNN). Penelitian ini menggunakan rentang skor total kuesioner untuk mengetahui tingkat kepuasan pelanggan yang dikategorikan menjadi tiga kelas, yaitu Tidak Puas, Cukup Puas, dan Puas. Berikut adalah rentang skor yang digunakan:

Tabel 1. Rentang Kategori

Rentang Skor	Kategori
46-60	Puas
31-45	Cukup Puas
15-30	Tidak Puas

Sebelum masuk pada model KNN, penelitian ini melakuka tahapan pra-pemrosesan data terlebih dulu yang dimulai dari melakukan pembersihan data, dimana data yang tidak dibutuhkan akan dihapus. Pengkodean data, tahap ini untuk mengubah data yang masih berbentuk kategori atau teks ke dalam bentuk angka. Dan pembagian data, dimana data dibagi menjadi data latih dan data uji.

 https://doi.org/10.56211/sudo.v5i3.2014



Gambar 1. Tahapan algoritma *K-Nearest Neighbors* dengan teknik SMOTE

Pemodelan dimulai dari input dataset yang kemudian melalui tahap pra-pemrosesan untuk menyiapkan data sebelum dilakukan klasifikasi. Setelah diperoleh data bersih, dataset dibagi menjadi data latih dan data uji. Data latih selanjutnya diproses menggunakan SMOTE untuk menangani ketidakseimbangan kelas sehingga diperoleh data latih yang lebih seimbang. Setelah itu dilakukan pengujian beberapa nilai K menggunakan K-Fold Cross Validation. Pada setiap fold, data training digunakan untuk membangun model, sedangkan data validasi digunakan untuk mengukur accuracy. Nilai accuracy dari setiap fold kemudian dirata-ratakan untuk setiap nilai K. Nilai K dengan rata-rata accuracy tertinggi dipilih sebagai nilai K terbaik.

Setelah nilai K terbaik diperoleh, jarak antar data dihitung menggunakan Euclidean Distance, kemudian diurutkan untuk menentukan tetangga terdekat. Kelas ditentukan berdasarkan voting mayoritas dari K tetangga. Model KNN selanjutnya digunakan untuk memprediksi data uji. Hasil prediksi dibandingkan dengan label sebenarnya dan dievaluasi menggunakan confusion matrix, accuracy, precision, recall, dan F1-score untuk mengetahui kinerja klasifikasi.

HASIL DAN PEMBAHASAN

Data yang digunakan adalah data primer yang diperoleh dari penyebaran kuesioner dengan Google Form yang respondennya ditujukan kepada pelanggan laundry. Kuesioner ini terdiri dari 15 butir pernyataan dan juga pada kuesioner ini pernyataan diukur menggunakan Skala Likert yaitu Sangat Tidak Setuju = 1, Tidak Setuju = 2, Setuju = 3, dan Sangat Setuju = 4. Data yang akan digunakan sebanyak 250 data. Berikut adalah dataset yang diperoleh setelah melakukan pra-pemrosesan data:

Tabel 1. Dataset setelah pra-pemrosesan

No	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	…	X15	Label
58	3	4	3	3	4	3	4	4	3	4	…	3	2
59	3	3	2	4	4	3	3	4	4	3	…	3	2
60	4	3	4	4	3	4	4	3	3	4	…	4	2
61	2	3	4	4	4	3	3	3	2	4	…	4	2
62	4	3	4	3	4	4	4	4	4	4	…	4	2
63	3	4	3	3	4	3	4	4	3	4	…	3	2
64	4	4	4	4	4	4	4	4	4	4	…	4	2
65	4	3	3	4	4	3	4	4	3	4	…	4	2
66	4	3	4	4	3	4	3	3	4	3	…	4	2
57	3	4	4	3	4	4	3	4	3	3	…	4	2

Penentuan Nilai K Terbaik

Berdasarkan hasil pengujian nilai K terbaik, dimana yang diuji nilai K = 1 sampai K = 10 menggunakan 5-fold cross validation. Hasil pengujian menunjukkan bahwa nilai akurasi yang diperoleh berada pada nilai K = 3 dengan rata-rata accuracy 0.9750.

Penerapan SMOTE

Tabel berikut menampilkan hasil untuk data sebelum SMOTE dan setelah SMOTE. diterapkan hanya pada data latih dengan parameter k_neighbors=3 dan random_state=12. Hasil dari proses SMOTE berupa data fitur dan label yang telah diseimbangkan. Data tersebut selanjutnya digunakan sebagai data latih pada proses klasifikasi menggunakan nilai K terbaik yang telah diperoleh sebelumnya, sedangkan data uji tetap menggunakan data asli.

Tabel 2. Distribusi Kelas Sebelum dan Sesuadah SMOTE

Label	Sebelum SMOTE	Setelah SMOTE
Puas	176	176
Cukup Puas	18	176
Tidak Puas	6	176
Total	200	528

Pembentukan Model KNN

Setelah diperoleh nilai K terbaik dari proses dan data latih diseimbangkan menggunakan SMOTE, selanjutnya model dilatih ulang menggunakan seluruh data latih yang telah melalui proses SMOTE untuk model klasifikasi akhir. Model KNN menggunakan nilai K terbaik yang diperoleh sebelumnya dengan Euclidean Distance sebagai metode pengukuran jarak. Hasil dari tahap ini digunakan untuk melakukan klasifikasi terhadap data uji.

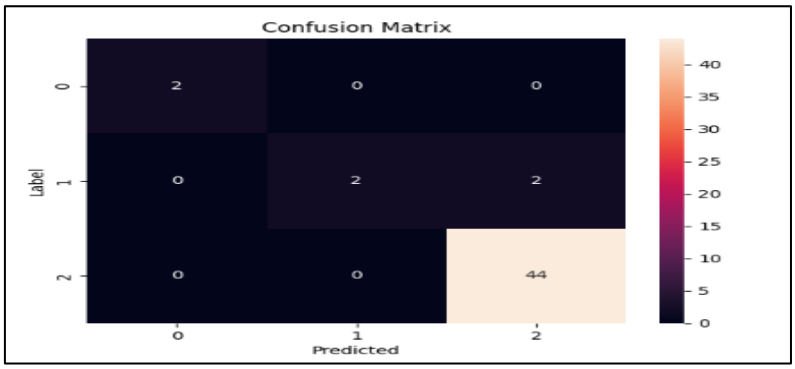
Prediksi Label

Setelah proses pembentukan model KNN selesai, kemudian melakukan prediksi terhadap data uji. Prediksi dilakukan menggunakan fungsi predict() dengan memasukkan X_test sebagai data yang akan diklasifikasikan. Hasil prediksi disimpan pada variabel y_pred, yang berisi label kelas hasil klasifikasi untuk setiap data uji. Berdasarkan hasil prediksi, data uji diklasifikasikan ke dalam kelas 0, 1, atau 2 sesuai dengan karakteristik dan kedekatannya terhadap data latih.

.

Confusion Matrix Sebelum SMOTE

Gambar berikut merupakan hasil *confusion matrix* sebelum SMOTE. Berdasarkan hasil Confusion Matrix sebelum SMOTE pada data uji, diperoleh matriks klasifikasi dengan tiga kelas, yaitu kelas 0, 1, dan 2. Hasil menunjukkan bahwa dari 2 data kelas 0, seluruhnya berhasil diklasifikasikan dengan benar sebagai kelas 0. Pada kelas 1, terdapat 4 data, dengan 2 data berhasil diklasifikasikan dengan benar sebagai kelas 1, sedangkan 2 data diklasifikasikan sebagai kelas 2.



Gambar 1. Hasil Confusion Matrix Sebelum SMOTE

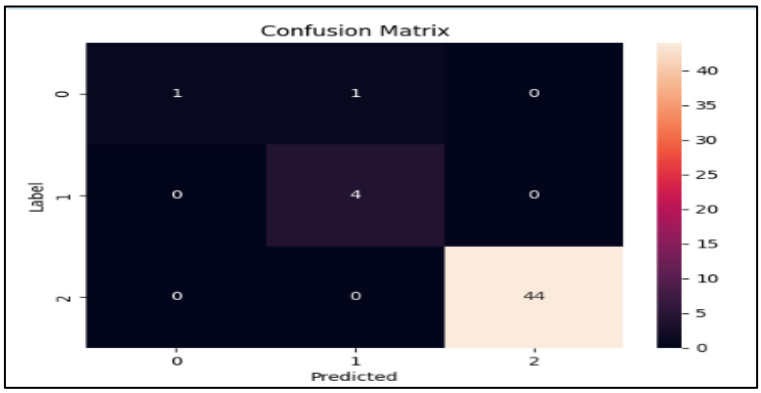
$$Accuracy = \frac{TP_0 + TP_1 + TP_2}{Total\ Data}$$

$$\frac{2 + 2 + 44}{50} = \frac{48}{50} = 0{,}96\%$$

Berdasarkan confusion matrix sebelum SMOTE yang diperoleh dari hasil pengujian, algoritma *K-Nearest Neighbors* (KNN), model KNN menghasilkan nilai accuracy sebesar 96%.

Confusion Matrix Setelah SMOTE

Gambar berikut merupakan hasil *confusion matrix* setelah SMOTE. Pada Label 0, terdapat satu data yang berhasil diprediksi dengan benar dan satu data yang diprediksi sebagai Label 1. Pada Label 1, seluruh data berhasil diklasifikasikan dengan tepat. Sementara itu, pada Label 2, seluruh data juga berhasil diklasifikasikan dengan benar. Dengan demikian, kesalahan klasifikasi hanya terjadi pada satu data Label 0 yang diprediksi sebagai Label 1.



Gambar 2. Hasil Confusion Matrix Setelah SMOTE

$$Accuracy = \frac{TP_0 + TP_1 + TP_2}{Total\ Data}$$

$$\frac{1 + 4 + 44}{50} = \frac{49}{50} = 0{,}98\%$$

Berdasarkan confusion matrix setelah SMOTE yang diperoleh dari hasil pengujian, algoritma *K-Nearest Neighbors* (KNN), model KNN menghasilkan nilai accuracy sebesar 98%.

Perbandingan KNN Sebelum dan Sesudah SMOTE

Berdasarkan hasil perbandingan, penerapan SMOTE meningkatkan kinerja model KNN, yang terlihat dari peningkatan akurasi dari 96% sebelum SMOTE, menjadi 98% Setelah SMOTE. Peningkatan akurasi dari 96% menjadi 98% setelah penerapan SMOTE menunjukkan adanya perubahan kemampuan model dalam mengenali kelas minoritas. sebelum SMOTE, dari 4 data kelas 1 terdapat 2 data yang diklasifikasikan sebagai kelas 2 (Puas). Setelah SMOTE, seluruh data kelas 1 berhasil diklasifikasikan dengan benar. Hal ini menunjukkan bahwa penambahan data sintetis pada kelas minoritas

membantu memberikan representasi data yang lebih baik sehingga model KNN lebih mampu mengenali pola kelas Cukup Puas. Sementara itu, pada kelas 0 (Tidak Puas), sebelum SMOTE seluruh data berhasil diklasifikasikan dengan benar, sedangkan setelah SMOTE terdapat 1 data yang diprediksi sebagai kelas 1. Dengan demikian, SMOTE tidak meningkatkan seluruh kelas secara bersamaan, tetapi perbaikan pada kelas 1 memberikan pengaruh terhadap peningkatan hasil klasifikasi secara keseluruhan.

Penggunaan nilai K=3 juga mendukung proses klasifikasi karena kelas ditentukan berdasarkan tiga tetangga terdekat menggunakan Euclidean Distance. Hasil Cross Validation menunjukkan bahwa K=3 memperoleh rata-rata accuracy tertinggi sebesar 97,5%, sehingga digunakan sebagai nilai K pada model akhir. Hasil tersebut menunjukkan bahwa kedekatan antar data cukup mendukung proses klasifikasi pada penelitian ini.

KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan hasil penelitian dapat diambil kesimpulan bahwa algoritma K-Nearest Neighbors (KNN) berhasil diterapkan dalam mengklasifikasikan tingkat kepuasan pelanggan Laundry ke dalam tiga kategori, yaitu Tidak Puas, Cukup Puas, dan Puas. Penerapan dilakukan melalui tahapan pra-pemrosesan data, pembagian data menjadi 80% data latih (200 data) dan 20% data uji (50 data), penanganan ketidakseimbangan kelas menggunakan SMOTE, hingga penentuan nilai K terbaik melalui Stratified K-Fold Cross Validation. Hasil pengujian menunjukkan nilai K optimal diperoleh pada K = 3.

Berdasarkan hasil pengolahan data kuesioner, tingkat kepuasan pelanggan Zayn Laundry terbagi ke dalam tiga kategori, yaitu Puas, Cukup Puas, dan Tidak Puas. Hasil menunjukkan bahwa mayoritas pelanggan berada pada kategori Puas.

Hasil evaluasi model K-Nearest Neighbors (KNN) menunjukkan bahwa penerapan SMOTE meningkatkan akurasi model klasifikasi. Dimana sebelum menggunakan SMOTE, model memperoleh accuracy sebesar 96%, precision sebesar 0,96%, recall sebesar 0,96%, dan F1-score sebesar 0,95%. Setelah menggunakan SMOTE, model memperoleh accuracy sebesar 98%, precision sebesar 0,98%, recall sebesar 0,98%, dan F1-score sebesar 0,98% pada data uji.

Saran

Penelitian selanjutnya disarankan untuk membandingkan algoritma K-Nearest Neighbors (KNN) dengan algoritma klasifikasi lain, yaitu Random Forest, atau Support Vector Machine (SVM), guna mengetahui algoritma mana yang memberikan tingkat akurasi terbaik untuk kasus klasifikasi kepuasan pelanggan, terutama dalam menangani kelas minoritas.

DAFTAR PUSTAKA

[1] M. Siringoringo, E. O. P. Damanik, dan J. W. Pangihutan, “Analisis motivasi kerja untuk meningkatkan kinerja karyawan pada fresh laundry pematangsiantar,” *Manaj. J. Ekon. USI*, vol. 5, no. 1, hal. 59–64, 2023.

[2] M. F. Arif dan A. H. Hasugian, “Penentuan Tingkat Kepuasan Pelanggan untuk Menentukan Pengembangan Bisnis Batik menggunakan Algoritma Naive Bayes Determining Customer Satisfaction Level to Determine Batik Business Development u sing Naive Bayes Algorithm,” *Sist. J. Sist. Inf.*, vol. 13, hal. 1662–1678, 2024.

[3] R. Putra, “Shaka Laundry Kabupaten Lahat Tesis Untuk memenuhi sebagian persyaratan Mencapai derajat Sarjana S2 Program Magister Manajemen Disusun Oleh : Shaka Laundry Kabupaten Lahat,” Universitas Islam Negeri Sultan Agung Semarang, 2025.

[4] M. F. Auji, W. Fuadi, dan A. Razi, “Klasifikasi Kepuasan Masyarakat terhadap Kinerja Petugas Sensus menggunakan K-Nearest Neighbours,” *Algoritma*, hal. 205–214, 2025, doi: 10.33364/algoritma/v.22-2.2502.

[5] S. Rahayu, Y. MZ, J. E. Bororing, dan R. Hadiyat, “Implementasi Metode K-Nearest Neighbor (K-NN) untuk Analisis Sentimen Kepuasan Pengguna Aplikasi Teknologi Finansial FLIP,” vol. 6, no. 1, hal. 98–106, 2022, doi: 10.29408/edumatic.v6i1.5433.

[6] S. Khaerunnisa dan S. Nursyamsiah, “Pengembangan Layanan Ramah Lingkungan pada Perusahaan Laundry IronMom di Cirebon Jawa Barat,” *Innov. J. Soc. Sci. Res.*, vol. 3, hal. 8249–8258, 2023.

[7] H. Kurniawan, *Pengantar Praktis Penyusunan Instrumen Penelitian*. Yogyakarta: Deepublish, 2022. [Daring]. Tersedia pada: https://books.google.co.id/books?id=-bM-EQAAQBAJ

[8] C. Hironima, B. Sihotang, M. Syaifuddin, dan I. Mariami, “Penerapan Data Mining Dalam Penjualan Alat Kerja

Cleaning Service Menggunakan Metode Apriori," *STMIK Triguna Dharma*, vol. 2, hal. 249–259, 2023.
[9] B. J. Z. Abidin, *Klasifikasi Dengan Artificial Neural Network di Perguruan Tinggi*. Sukabumi: Naureen Digital Education, 2023. [Daring]. Tersedia pada: https://books.google.co.id/books?id=STTCEQAAQBAJ
[10] Rahmaddeni, D. Wulandari, S. Aziz, S. Adrianto, dan F. Pratiwi, *Teori Dan Implementasi Machine Learning Menggunakan Python*. Sumatera Barat: Serasi Media Teknologi, 2025.
[11] A. M. Hasibuan, "Classification of Customer Loan Eligibility Using The K-Nearest Neighbor (K-NN) Algorithm Klasifikasi Kelayakan Peminjaman Nasabah Menggunakan Algoritma K-Nearest Neighbor (K-NN)," *MALCOM*, vol. 4, no. 10, hal. 1525–1532, 2024.
[12] S. Junaidi, I. S. Beno, M. Farkhan, dan I. K. D. G. Supartha, *Buku Ajar Machine Learning*. Jambi: PT. Sonpedia Publishing Indonesia, 2024.
[13] A. N. Putro dan M. A. Muslim, "Penanganan Ketidakseimbangan Data Ekstrim pada Sistem Prediksi," *Tecno.COM*, vol. 24, no. 4, hal. 1359–1367, 2025.
[14] A. S. Firmansyah, A. Aziz, dan M. Ahsan, "OPTIMASI K-NEAREST NEIGHBOR MENGGUNAKAN ALGORITMA SMOTE UNTUK MENGATASI IMBALANCE CLASS PADA KLASIFIKASI ANALISIS SENTIMEN," *JATI*, vol. 7, no. 6, 2023.
[15] Rianto dan P. I. Santosa, *Data Preparation untuk Machine Learning & Deep Learning*. Yogyakarta: Penerbit Andi, 2025. [Daring]. Tersedia pada: https://books.google.co.id/books?id=Y5U9EQAAQBAJ

NOMENKLATUR

d = Jarak Antara Data
i = Variabel Data
x = Data Testing
y = Data Araining
n = Jumlah Atribut
TP = True Positif
TN = True Negatif
FP = False Positif
FN = False Negatif