

## IMPLEMENTASI ALGORITMA K-NEAREST NEIGHBORS PADA APLIKASI PREDIKSI KELAYAKAN CALON NASABAH

Ratna Rahmawati Rahayu <sup>1)</sup>, Satya Aryatama <sup>2)</sup>, Faiz Mubarak <sup>3)</sup>, Rafi Triamar Putra <sup>4)</sup>,  
Riski Tiara Ramadhan <sup>5)</sup>, Hendra Juli Ariansyah <sup>6)</sup>

<sup>1) 2) 3) 4) 5) 6)</sup> Sistem Informasi, Universitas Bani Saleh, Bekasi, Indonesia, 17113

Email : [ratna@ubs.ac.id](mailto:ratna@ubs.ac.id) <sup>1)</sup>, [satyaaryatama-m@ubs.ac.id](mailto:satyaaryatama-m@ubs.ac.id) <sup>2)</sup>, [faizmubarak-m@ubs.ac.id](mailto:faizmubarak-m@ubs.ac.id) <sup>3)</sup>, [rafitriamar-m@ubs.ac.id](mailto:rafitriamar-m@ubs.ac.id) <sup>4)</sup>,  
[riskitiara-m@ubs.ac.id](mailto:riskitiara-m@ubs.ac.id) <sup>5)</sup>, [hendrajuly-m@ubs.ac.id](mailto:hendrajuly-m@ubs.ac.id) <sup>6)</sup>

### ABSTRAK

Perkembangan teknologi informasi dalam sektor keuangan mendorong lembaga pembiayaan untuk memanfaatkan teknologi berbasis data dalam proses pengambilan keputusan, khususnya pada penentuan kelayakan calon nasabah. Proses seleksi yang dilakukan secara manual sering kali membutuhkan waktu yang lama dan berpotensi menimbulkan subjektivitas dalam penilaian. Penelitian ini bertujuan untuk mengimplementasikan algoritma *K-Nearest Neighbors* (K-NN) pada aplikasi prediksi kelayakan calon nasabah guna membantu proses klasifikasi secara lebih cepat dan objektif. Penelitian menggunakan atribut umur, pendapatan, status pekerjaan, lama kerja, dan status rumah sebagai parameter penilaian, sedangkan kelas keluaran berupa status kelayakan nasabah. Metode penelitian meliputi tahap pengumpulan data, preprocessing data, penerapan algoritma K-NN, serta pengujian performa model menggunakan metrik evaluasi *F1-Score*. Aplikasi dikembangkan menggunakan bahasa pemrograman Python untuk mempermudah proses pengolahan data dan implementasi sistem prediksi. Hasil pengujian menunjukkan bahwa algoritma KNN mampu melakukan klasifikasi kelayakan calon nasabah dengan nilai *F1-Score* sebesar 80%, sehingga metode ini dinilai cukup baik dalam mendukung proses pengambilan keputusan pada lembaga keuangan. Penelitian ini diharapkan dapat menjadi solusi alternatif dalam meningkatkan efektivitas dan akurasi proses seleksi calon nasabah.

Kata Kunci : *K-Nearest Neighbors*, prediksi kelayakan nasabah, klasifikasi, data mining, Python

### ABSTRACT

The development of information technology in the financial sector has encouraged financial institutions to utilize data-based technology in decision-making processes, particularly in determining the eligibility of prospective customers. Manual selection processes often require considerable time and may lead to subjectivity in assessment. This study aims to implement the *K-Nearest Neighbors* (K-NN) algorithm in an application for predicting the eligibility of prospective customers in order to support a faster and more objective classification process. The study uses several attributes, namely age, income, employment status, years of employment, and housing status as assessment parameters, while the output class is customer eligibility status. The research methodology includes data collection, data preprocessing, implementation of the K-NN algorithm, and model performance evaluation using the *F1-Score* metric. The application was developed using the Python programming language to facilitate data processing and prediction system implementation. The testing results indicate that the K-NN algorithm is capable of classifying prospective customer eligibility with an *F1-Score* of 80%, showing that the method performs adequately in supporting decision-making processes within financial institutions. This study is expected to provide an alternative solution for improving the effectiveness and accuracy of prospective customer selection processes.

Keywords: *K-Nearest Neighbors*, customer eligibility prediction, classification, data mining, Python

## 1. PENDAHULUAN

Kemajuan teknologi informasi yang berlangsung sangat cepat telah mendorong berbagai bidang, terutama sektor keuangan, untuk memanfaatkan data sebagai landasan dalam proses pengambilan keputusan. Salah satu keputusan krusial yang dihadapi lembaga keuangan adalah menentukan tingkat kelayakan calon nasabah. Ketidaktepatan dalam proses penilaian kelayakan dapat menimbulkan berbagai risiko, seperti terjadinya kredit bermasalah hingga kerugian finansial bagi perusahaan.

Meskipun demikian, pada implementasinya, proses evaluasi kelayakan calon nasabah masih kerap dilakukan secara manual dan bergantung pada penilaian subjektif. Kondisi tersebut menyebabkan hasil keputusan yang diperoleh belum mampu memberikan tingkat akurasi dan efektivitas yang optimal.

Untuk mengatasi hal tersebut, perlu dibuatkan suatu aplikasi sebagai rekomendasi penentuan kelayakan kredit dengan memanfaatkan data warehouse dan data mining menggunakan algoritma *K-Nearest Neighbors* (K-NN) yang mampu mengidentifikasi pola kelayakan nasabah dengan baik, sehingga dapat meminimalisir risiko kredit dan mengoptimalkan sistem penilaian pada lembaga keuangan (Purba, S., dan Situmorang, 2024).

Data mining merupakan salah satu pendekatan yang dapat digunakan untuk menggali informasi dan pola dari sekumpulan data. Salah satu algoritma klasifikasi yang sering digunakan dalam data mining adalah K-NN. Algoritma ini bekerja dengan cara membandingkan data baru dengan sejumlah data lama yang memiliki kemiripan paling dekat, kemudian menentukan kelas berdasarkan mayoritas tetangga terdekat tersebut.

Data mining atau penambangan data adalah kegiatan analisis terhadap kumpulan data yang sangat besar untuk mengidentifikasi pola, hubungan, atau wawasan baru yang

dapat membantu dalam membuat keputusan. Penambangan data tidak hanya berkaitan dengan pengambilan data, tetapi juga mencakup aspek pemahaman, pembersihan, dan modifikasi data sehingga hasil analisis dapat memberikan manfaat dalam konteks yang lebih luas (Saputri, 2025). Dalam konteks penelitian ini, data mining digunakan untuk mengolah data calon nasabah guna menemukan pola yang dapat digunakan sebagai dasar dalam merevisi kelayakan calon nasabah secara objektif.

Metode data mining yang tepat untuk digunakan dalam penelitian ini adalah klasifikasi. Karena klasifikasi adalah sebuah model yang mampu mendeteksi atau memprediksi kesamaan fitur dalam suatu kategori atau kelas. Teknik klasifikasi melakukan prediksi pada data yang memiliki atribut dalam bentuk label (Sholeh, Andayati, dan Rahmawati, 2022). Metode klasifikasi banyak digunakan dalam bidang keuangan, kesehatan, dan pendidikan karena kemampuannya dalam memberikan keputusan berbasis data. Dalam penelitian ini, klasifikasi digunakan untuk menentukan apakah calon nasabah termasuk dalam kategori layak, dipertimbangkan, dan tidak layak.

Algoritma K-NN merupakan sebuah metode klasifikasi yang memanfaatkan nilai K untuk mengidentifikasi kelas dari data baru yang perlu dikategorikan. Penentuan nilai K sangat berpengaruh terhadap hasil klasifikasi K-NN. Nilai K yang terlalu kecil dapat menyebabkan model sensitif terhadap *noise*, sedangkan nilai K yang terlalu besar dapat mengurangi ketepatan klasifikasi. Oleh karena itu, nilai K dipilih melalui pengujian untuk memperoleh akurasi terbaik. Metode ini beroperasi dengan menemukan data yang memiliki kedekatan tertinggi dengan data yang ingin diklasifikasikan. Aspek utama dari metode ini adalah melakukan pengelompokan pada objek baru dengan membandingkan data sasaran dengan data pelatihan yang sudah tersedia (Simbolon, D.S.J. dan Manik, 2024).

Karena itu, algoritma K-NN dinilai cocok untuk digunakan dalam prediksi kelayakan

calon nasabah karena memiliki konsep yang sederhana dan mudah diimplementasikan, serta mampu memberikan hasil prediksi berdasarkan kedekatan karakteristik data. Dengan memanfaatkan variabel umur, status pekerjaan, pendapatan, lama kerja, dan status rumah, algoritma K-NN diharapkan dapat membantu lembaga keuangan dalam melakukan klasifikasi kelayakan calon nasabah secara lebih objektif dan akurat.

Algoritma K-NN dengan seleksi fitur menggunakan korelasi Pearson untuk memprediksi kelayakan pengajuan pinjaman. Fitur dengan tingkat korelasi tertinggi digunakan dalam proses prediksi. Hasil pengujian menunjukkan kinerja terbaik pada rasio 80% data latih dan 20% data uji dengan nilai akurasi, presisi, dan recall sebesar 100%, sehingga metode KNN dinilai sangat efektif dalam membantu kreditur memprediksi kelayakan pinjaman (Hermawan dan Yamasari, 2022).

Penelitian dengan metode Algoritma K-NN digunakan untuk proses klasifikasi kelayakan kredit dengan 1.050 data responden. Hasil pengujian menunjukkan bahwa metode KNN mampu menghasilkan performa yang baik dengan nilai akurasi sebesar 0,85, presisi 0,87, recall 0,84, dan F1-score sebesar 0,85 (Winahyu dan Saputro, 2023).

Confusion Matrix merupakan sebuah perangkat yang berfungsi untuk mengevaluasi kinerja model pengklasifikasian dengan cara menilai seberapa efektif model tersebut dalam mengelompokkan data ke dalam kategori yang benar. Perangkat ini menyajikan suatu tabel yang menunjukkan hasil prediksi model dibandingkan dengan label yang sebenarnya, sehingga kita dapat mengamati jumlah prediksi yang tepat dan keliru untuk setiap kelas (Tatimu, Palilingan dan Rianto, 2025). Dalam penelitian ini, Confusion Matrix digunakan untuk mengevaluasi tingkat keberhasilan algoritma K-Nearest Neighbors (K-NN) dalam memprediksi kelayakan calon nasabah.

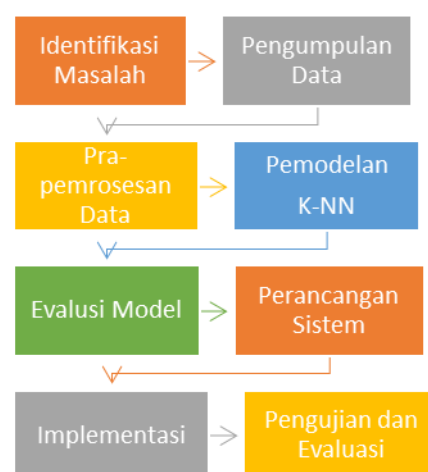
Hasil penerapan algoritma K-NN untuk mengklasifikasikan kelayakan pengajuan pinjaman menunjukkan bahwa implementasi metode ini efektif dalam mempercepat proses evaluasi yang sebelumnya dilakukan secara kesalahan manusia (*human error*) dalam pengambilan keputusan kredit (Listanto *et al.*, 2025).

Berdasarkan tinjauan penelitian terdahulu, dapat disimpulkan bahwa algoritma KNN memiliki performa yang baik untuk tugas klasifikasi dan layak digunakan dalam penelitian yang dilakukan dengan judul “Implementasi Algoritma *K-Nearest Neighbors* Pada Aplikasi Prediksi Kelayakan Calon Nasabah”, dengan harapan dapat memberikan gambaran penerapan metode data mining dalam mendukung pengambilan keputusan di sektor keuangan. Posisi penelitian ini dibandingkan dengan penelitian sebelumnya terletak pada perbedaan objek dan kombinasi atribut yang digunakan, yaitu umur, status pekerjaan, pendapatan, lama kerja, dan status rumah untuk memprediksi kelayakan calon nasabah.

## 2. METODE

### 2.1 Rancangan Penelitian

Penelitian ini dilakukan secara sistematis dengan tahapan sebagai berikut:



Gambar 1. Tahapan Penelitian

- a) **Identifikasi masalah**, mengidentifikasi kriteria kelayakan nasabah.
- b) **Pengumpulan data**, mengambil data calon nasabah.
- c) **Pra-pemrosesan data**, melakukan transformasi dan normalisasi data.
- d) **Pemodelan K-NN** mengimplementasikan masalah dan data ke dalam algoritma K-NN.
- e) **Evaluasi Model**: melakukan penilaian evaluasi kelayakan model K-NN.
- f) **Perancangan Sistem**, mendesain basis data, antarmuka dan proses penerapan model K-NN.
- g) **Implementasi**: mengodekan rancangan sistem ke dalam bahasa pemrograman.
- h) **Pengujian dan Evaluasi**, mengukur tingkat akurasi sistem

## 2.2 Pengumpulan Data

Data yang digunakan dalam penelitian ini bersumber dari database yang telah ditentukan sebelum penelitian dilakukan. Database tersebut berisi data historis calon nasabah yang disiapkan untuk kebutuhan penelitian dan pengujian algoritma *K-Nearest Neighbors* (K-NN). Data tersebut dibagi menjadi data latih dan data testing.

Tabel 1. Data Calon Nasabah

| Umur | Pendapatan | Status Pekerjaan | Lama kerja | Rumah    | Kelayakan       |
|------|------------|------------------|------------|----------|-----------------|
| 45   | 12500000   | Tetap            | 10         | Milik    | Layak           |
| 38   | 10000000   | Tetap            | 7          | Milik    | Layak           |
| 30   | 8000000    | Kontrak          | 4          | Sewa     | Dipertimbangkan |
| 29   | 5000000    | Freelance        | 2          | Keluarga | Dipertimbangkan |
| 50   | 15000000   | Tetap            | 20         | Milik    | Layak           |
| 28   | 4000000    | Freelance        | 1          | Sewa     | Tidak Layak     |
| 33   | 8000000    | Kontrak          | 3          | Keluarga | Dipertimbangkan |
| 41   | 9000000    | Tetap            | 6          | Sewa     | Layak           |
| 36   | 6500000    | Kontrak          | 4          | Sewa     | Dipertimbangkan |
| 27   | 3000000    | Freelance        | 1          | Sewa     | Tidak Layak     |
| ...  | ...        | ...              | ...        | ...      | ...             |

Data dalam database tidak diperoleh dari proses input pengguna secara langsung, melainkan merupakan data calon nasabah yang ada di gudang data.

Penggunaan database pada penelitian ini bertujuan untuk memudahkan pengelolaan data, proses normalisasi, serta perhitungan jarak antar data pada algoritma K-NN.

## 2.3 Pra-pemrosesan Data

Preprocessing data merupakan tahapan awal dalam proses data mining yang bertujuan untuk meningkatkan kualitas data sebelum dilakukan analisis lebih lanjut. Tahap ini dilakukan untuk memastikan bahwa data yang digunakan telah berada dalam kondisi yang sesuai dan siap diproses oleh algoritma data mining (Agung, Daniswara dan Nuryana, 2023).

Pada tahap pra-pemrosesan data dilakukan untuk mengubah data mentah menjadi format yang kompatibel dengan algoritma K-Nearest Neighbor (K-NN). Mengingat algoritma ini bekerja berdasarkan perhitungan jarak matematis (Euclidean Distance), seluruh data masukan wajib berbentuk numerik dan memiliki skala nilai yang seimbang. Algoritma K-NN tidak dapat memproses data yang berupa teks, yaitu Status Pekerjaan dan Status Rumah, yang dikonversi menjadi nilai numerik (0, 1, atau 2). Atribut Pendapatan untuk menyederhanakan digit angka atau penyesuaian skala tanpa menghilangkan variasi data dengan membagi 100.000.

### a) Atribut Status Pekerjaan:

- Freelance: Diubah menjadi nilai 0
- Kontrak: Diubah menjadi nilai 1
- Tetap: Diubah menjadi nilai 2

Pemberian nilai diurutkan berdasarkan tingkat stabilitas pekerjaan, dari yang paling rendah (0) hingga yang paling stabil (2).

### b) Atribut Status Rumah:

- Keluarga: Diubah menjadi nilai 0
- Sewa: Diubah menjadi nilai 1
- Milik: Diubah menjadi nilai 2

Pemberian nilai diurutkan berdasarkan tingkat kepemilikan aset, dari menumpang (0) hingga kepemilikan pribadi (2).

### c) Atribut Pendapatan:

Penyesuaian Nilai: Data nominal pendapatan dikonversi menjadi angka

yang lebih sederhana dengan cara membaginya dengan 100.000.

Transformasi ini dilakukan untuk memastikan seluruh variabel memiliki format numerik yang seragam dan sederhana sehingga dapat diproses menggunakan perhitungan jarak Euclidean dalam algoritma K-NN.

Tabel 2. Data Hasil Pre-pemrosesan

| Umur | Pendapatan | Status pekerjaan | Lama kerja | Rumah | Kelayakan       |
|------|------------|------------------|------------|-------|-----------------|
| 45   | 125        | 2                | 10         | 2     | Layak           |
| 38   | 100        | 2                | 7          | 2     | Layak           |
| 30   | 80         | 1                | 4          | 0     | Dipertimbangkan |
| 29   | 50         | 0                | 2          | 1     | Dipertimbangkan |
| 50   | 150        | 2                | 20         | 2     | Layak           |
| 28   | 40         | 0                | 1          | 0     | Tidak Layak     |
| 33   | 60         | 1                | 3          | 1     | Dipertimbangkan |
| 41   | 90         | 2                | 6          | 0     | Layak           |
| 36   | 65         | 1                | 4          | 0     | Dipertimbangkan |
| 27   | 30         | 0                | 1          | 0     | Tidak Layak     |
| ...  | ...        | ...              | ...        | ...   | ...             |

## 2.4 Pemodelan K-NN

Algoritma K-Nearest Neighbors (K-NN) merupakan metode klasifikasi yang menentukan label (*class*) dari suatu objek baru berdasarkan mayoritas kedekatan jarak dari kategori yang ada dalam K-NN (Rahayu dan Lidiawati, 2021). Salah satu pendekatan dalam algoritma K-NN adalah jarak Euclidean. Jarak Euclidean merupakan ukuran antara dua titik untuk menentukan posisi keduanya pada area, di mana panjang garis yang menghubungkan kedua titik tersebut dihitung. Sesuai namanya, ini mengukur jarak antar titik atau jarak garis lurus. Berikut adalah formula untuk jarak Euclidean yang ditunjukkan dalam persamaan 1 (Simbolon, D.S.J and Manik, 2024).

$$\text{dist}(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

Penerapan metode K-NN diawali dengan pengumpulan data calon nasabah yang telah memiliki label kelayakan. Data kemudian melalui tahap pra-pemrosesan, meliputi pembersihan dan normalisasi data.

$$\begin{aligned}
 &= \sqrt{(45 - 38)^2 + (125 - 65)^2 + (2 - 1)^2 + (10 - 7)^2 + (2 - 2)^2} \\
 &= \sqrt{49 + 3600 + 1 + 9 + 0} \\
 &= \sqrt{3659} \\
 &= 60,4897
 \end{aligned}$$

Selanjutnya, data dibagi menjadi data latih dan data uji.

Pada tahap klasifikasi, sistem menghitung tingkat kemiripan antara data calon nasabah uji dengan seluruh data latih. Sejumlah data latih terdekat dipilih sebagai tetangga. Status kelayakan calon nasabah ditentukan berdasarkan kelas yang paling dominan di antara tetangga tersebut. Hasil klasifikasi digunakan untuk mengevaluasi kinerja metode K-NN.

Langkah-langkah penerapan metode K-NN dalam penelitian ini adalah sebagai berikut:

- Menentukan nilai K sebagai jumlah tetangga terdekat. Penelitian ini menggunakan K = 5.
- Menyiapkan data latih dan data uji calon nasabah. Data latih digunakan sebagai acuan klasifikasi, sedangkan data uji digunakan untuk menguji kemampuan metode K-NN dalam menentukan kelayakan calon nasabah. Perbandingan data latih dan data uji adalah 80:20.
- Melakukan normalisasi data uji untuk menyamakan skala atribut.

Tabel 3. Normalisasi Data Uji

| Umur | Pendapatan | Status Pekerjaan | Lama Kerja | Rumah |
|------|------------|------------------|------------|-------|
| 38   | 65         | 1                | 7          | 2     |
| 32   | 60         | 1                | 5          | 1     |
| 25   | 25         | 0                | 1          | 0     |
| 43   | 110        | 2                | 12         | 2     |
| 34   | 70         | 1                | 4          | 1     |
| ...  | ...        | ...              | ...        | ...   |



- d) Menghitung jarak Euclidean antara data uji dan data latih. Perhitungan jarak Euclidean dilakukan menggunakan persamaan (1). Perhitungan data jarak data latih ke-1 dengan data uji ke-1.

Hasil perhitungan jarak antara data uji ke-1 dengan seluruh data latih terdapat di Tabel 4.

Tabel 4. Jarak Data Latih dan Uji

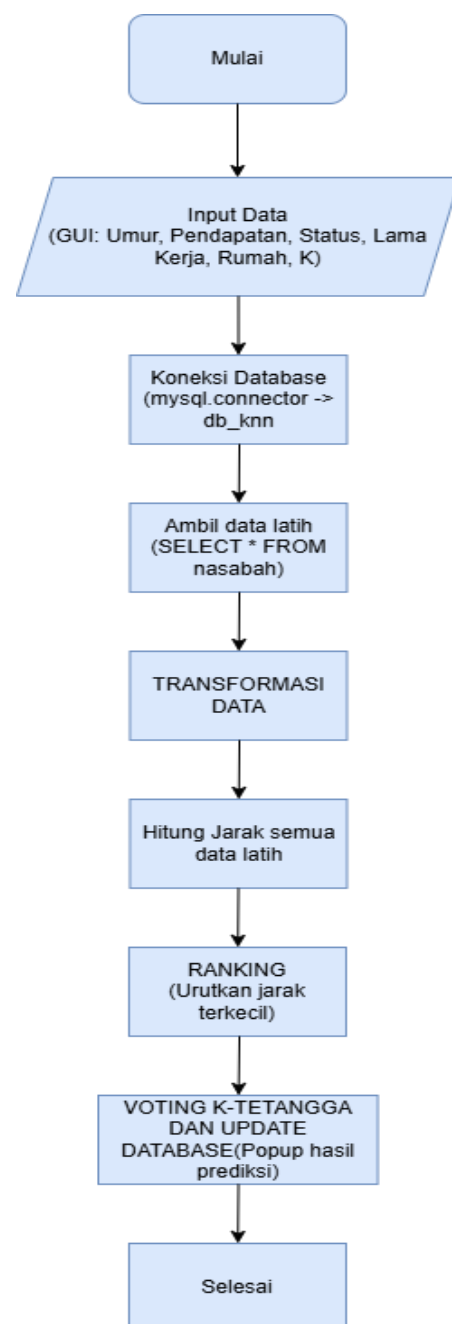
| Umur | Pendapatan | Status Pekerjaan | Lama Kerja | Rumah | Kelayakan       | Jarak   |
|------|------------|------------------|------------|-------|-----------------|---------|
| 45   | 125        | 2                | 10         | 2     | Layak           | 60,4897 |
| 38   | 100        | 2                | 7          | 2     | Layak           | 35,0143 |
| 30   | 80         | 1                | 4          | 0     | Dipertimbangkan | 17,2916 |
| 29   | 50         | 0                | 2          | 1     | Dipertimbangkan | 18,3303 |
| 50   | 150        | 2                | 20         | 2     | Layak           | 86,8274 |
| 28   | 40         | 0                | 1          | 0     | Tidak Layak     | 27,6225 |
| 33   | 60         | 1                | 3          | 1     | Dipertimbangkan | 8,3666  |
| 41   | 90         | 2                | 6          | 0     | Layak           | 25,2389 |
| 36   | 65         | 1                | 4          | 0     | Dipertimbangkan | 3,74166 |
| 27   | 30         | 0                | 1          | 0     | Tidak Layak     | 37,2022 |
| ...  | ...        | ...              | ...        | ...   | ...             | ...     |

- e) Memilih K tetangga terdekat berdasarkan kedekatan jarak atau data yang memiliki nilai jarak terkecil. Setelah nilai jarak diurutkan secara *ascending*, diperoleh 5 tetangga terdekat sebagai berikut:
- Dipertimbangkan: 5 Data (100%)
  - Layak: 0 Data (0%)
  - Tidak Layak: 0 Data (0%)
- f) Menentukan kelayakan calon nasabah berdasarkan kelas mayoritas untuk data uji ke-1 adalah “Dipertimbangkan”.

## 2.5 Evaluasi Model

Evaluasi kinerja atau *Confusion Matrix* merupakan sebuah perangkat yang berfungsi untuk mengevaluasi kinerja model pengklasifikasian dengan cara menilai seberapa efektif model tersebut dalam mengelompokkan data ke dalam kategori yang benar. Perangkat ini menyajikan suatu tabel yang menunjukkan hasil prediksi model dibandingkan dengan label yang sebenarnya, sehingga kita dapat mengamati jumlah prediksi yang tepat dan keliru untuk setiap

kelas (Tatimu, Palilingan dan Rianto, 2025). *Confusion Matrix* yang digunakan dalam penelitian ini adalah F1-Score untuk memastikan bahwa model memiliki sensitivitas yang baik dalam mengenali nasabah potensial serta memiliki keseimbangan performa (tidak bias), terutama jika terdapat ketimpangan jumlah data antarkelas. Hasil perhitungan menunjukkan nilai F1-Score yang mencapai 80%, hasil yang sangat positif dan menandakan bahwa pada data-data yang berhasil diprediksi, model memiliki tingkat presisi dan



Gambar 2. Flowchart Sistem

kepercayaan yang tinggi, sehingga keputusan yang dihasilkan tergolong solid.

## 2.6 Perancangan Sistem



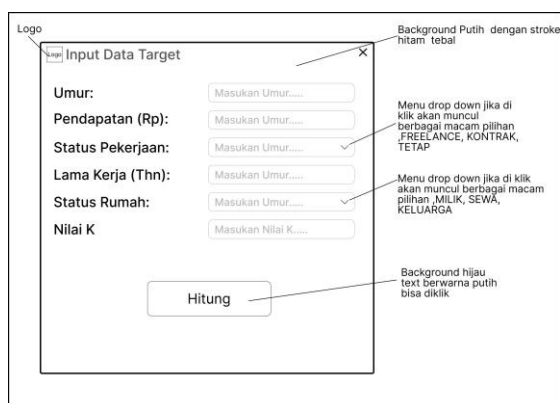
Gambar 4. Desain Halaman Hasil

Perancangan sistem bertujuan untuk menggambarkan alur kerja dan fungsi sistem yang akan dibangun.

Flowchart sistem menggambarkan urutan logika prosedur mulai dari input data nasabah baru, proses transformasi, hingga tahap akhir pemrosesan algoritma K-Nearest Neighbors (K-NN) yang menghasilkan prediksi kelayakan.

Antarmuka dirancang berbasis web/desktop sederhana menggunakan Python, yang terdiri dari:

- Halaman Input: Form untuk memasukkan data calon nasabah.



Gambar 3. Desain Halaman Input

- Halaman Hasil: Menampilkan nilai K yang digunakan, daftar tetangga terdekat, dan keputusan akhir.

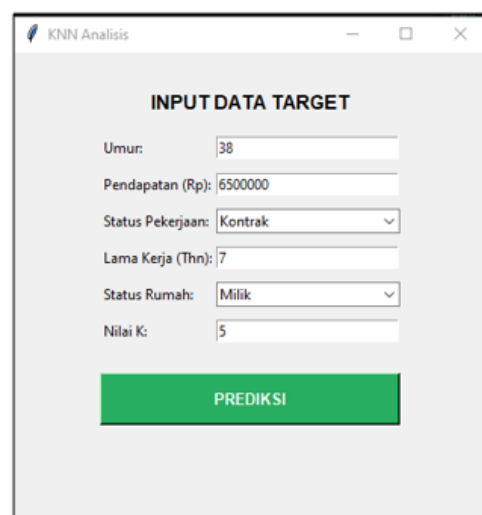
## 3 HASIL DAN PEMBAHASAN

### 3.1 Implementasi

Pada tahap ini, dilakukan implementasi antarmuka (*interface*) dari aplikasi prediksi kelayakan nasabah yang telah dibangun menggunakan bahasa pemrograman Python. Implementasi ini menampilkan halaman input data nasabah dan hasil prediksi yang diperoleh setelah sistem melakukan perhitungan menggunakan algoritma K-Nearest Neighbors (K-NN). Pengujian fungsional dilakukan dengan memasukkan sampel data uji untuk memastikan sistem mampu menerima input, memproses algoritma, dan menampilkan output keputusan serta memperbarui database.

Pengujian dilakukan menggunakan data sampel nasabah pertama sesuai dengan tabel Data Uji. Pada pengujian ini, pengguna memasukkan atribut data nasabah ke dalam form yang tersedia.

Setelah data dimasukkan, lalu di klik tombol PREDIKSI maka akan menampilkan status Kelayakan pengajuan kredit calon nasabah.



Gambar 5. Form Input Data

### 3.2 Pengujian dan Evaluasi

Dalam penelitian ini, metode pengujian yang diterapkan adalah *Black Box Testing* dan *Usability Testing*. Pemilihan metode ini dikarenakan *black box testing* memiliki sifat yang dinamis, tidak perlu diketahui proses yang terdapat dalam sistem aplikasi, yang diuji hanya masukan dan keluaran (Rahayu *et al.*, 2025). Setelah dilakukan pengujian sistem, hasilnya dapat dilihat pada Tabel 5 dan semua halaman yang dibuat tidak menampilkan kesalahan dan hasil sesuai dengan pemodelan K-NN dan tujuan dari penelitian.

Tabel 4. Hasil Pengujian Sistem

| Halaman | Fungsi                                                              | Hasil Eksekusi |
|---------|---------------------------------------------------------------------|----------------|
| Input   | Memasukkan data calon nasabah yang akan diprediksi status kelayakan | Berhasil       |
| Hasil   | Menampilkan hasil prediksi status kelayakan calon nasabah           | Berhasil       |

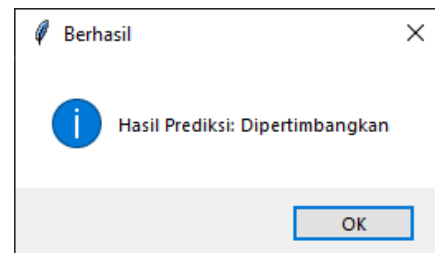
Dari pengujian *black box* yang hasilnya dapat dilihat pada Tabel 4, dapat dilihat bahwa semua pengujian yang dijalankan pada semua halaman berhasil dengan baik dan telah sesuai dengan tujuan penelitian.

## 4 PENUTUP

### 4.1. Kesimpulan

- Penerapan Algoritma: Algoritma K-Nearest Neighbors (K-NN) berhasil diterapkan untuk memprediksi kelayakan nasabah dengan menggunakan parameter input berupa umur, pendapatan, status pekerjaan, lama kerja, dan status rumah. Sistem dibangun dengan nilai parameter  $K=5$  yang dianggap optimal untuk jumlah data latih yang tersedia.

- Kinerja Sistem: Berdasarkan pengujian



Gambar 6. Tampilan Hasil

menggunakan Confusion Matrix F1-Score sebesar 80%.

- Efektivitas: Nilai F1-Score yang mencapai 80% menunjukkan bahwa sistem memiliki performa yang solid dan objektif dalam memberikan rekomendasi keputusan, meskipun diuji dengan jumlah sampel data yang terbatas.

### 4.2. Saran

Demi pengembangan sistem yang lebih baik di masa mendatang, penulis memberikan saran-saran sebagai berikut:

- Penambahan Data Latih: Disarankan untuk menambah jumlah data nasabah terdahulu (data latih) agar algoritma memiliki lebih banyak referensi pola, sehingga akurasi prediksi dapat meningkat.
- Penambahan Kriteria Penilaian: Sebaiknya ditambahkan variabel baru yang relevan dalam formulir input, seperti Jumlah Tanggungan atau Riwayat Kredit, agar penilaian kelayakan menjadi lebih detail dan akurat.
- Fitur Cetak Laporan: Sistem dapat dikembangkan dengan menambahkan fitur untuk mencetak hasil prediksi atau mengeksport rekapitulasi data ke dalam format laporan (PDF/Excel) untuk keperluan arsip administrasi.
- Pengembangan Platform: Aplikasi ini dapat dikembangkan lebih lanjut dari berbasis desktop menjadi berbasis web atau mobile agar lebih fleksibel dan mudah diakses oleh pengguna di berbagai perangkat.



## 5 DAFTAR PUSTAKA

- Agung, A., Daniswara, A. dan Nuryana, I.K.D. (2023) 'Data Preprocessing Pola Pada Penilaian Mahasiswa Program Profesi Guru', 05(01), pp. 97–100. Available at: <https://doi.org/https://doi.org/10.26740/jinacs.v5n01.p97-100>.
- Hermawan, F.F. dan Yamasari, Y. (2022) 'Implementasi K-Nearest Neighbor dengan Pemilihan Fitur pada Aplikasi Prediksi Kelayakan Pengajuan Pinjaman', *Journal of Informatics and Computer Science*, 03(04), pp. 411–424. Available at: <https://doi.org/https://doi.org/10.26740/jinacs.v3n04.p411-424>.
- Listanto, F. et al. (2025) 'IMPLEMENTASI ALGORITMA K-NEAREST NEIGHBOR UNTUK KLASIFIKASI PENGAJUAN PINJAMAN UANG DI BMT MU ' AMALAH SYARI ' AH TEBUIRENG', *Jurnal Ilmiah Inovasi Teknologi Informasi*, 9(2), pp. 61–68. Available at: <https://doi.org/https://doi.org/10.33752/inovate.v9i2.8872>.
- Purba, D.K., S, I.M.S. dan Situmorang, A. (2024) 'Penentuan Kelayakan Pemberian Kredit Pada Nasabah Cu Merdeka Berastagi Dengan Menggunakan Algoritma K-Nearest Neighbor (K-NN)', *Methosisfo*, 4(1), pp. 10–16.
- Rahayu, R.R. et al. (2025) 'Monitoring Pertumbuhan Fisik Balita Untuk Pendeteksian Dini Stunting Anak Berbasis Mobile', *Jurnal Teknik Informasi dan Komputer*, 8(1), pp. 368–379. Available at: <https://doi.org/10.37600/tekinkom.v8i1.2065>.
- Rahayu, R.R. dan Lidiawati, L. (2021) 'Implementasi Algoritma K-Nearest Neighbor Untuk Memprediksi Program Studi Bagi Calon Mahasiswa Baru', *Infotek: Jurnal Informatika dan Teknologi*, 4(2), pp. 131–141. Available at: <https://doi.org/10.29408/jit.v4i2.3546>.
- Saputri, E. (2025) 'Teknik dan aplikasi data mining di Indonesia: tinjauan literatur satu dekade ( 2015-2024 )', *Penerapan Teknologi Informasi dan Komunikasi*, 04(02), pp. 138–149. Available at: <https://doi.org/10.24246/itexplore.v4i2.2025.pp138-149>.
- Sholeh, M., Andayati, D. dan Rahmawati, R.Y. (2022) 'DATA MINING MODEL KLASIFIKASI MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBOR DENGAN NORMALISASI UNTUK PREDIKSI PENYAKIT DIABETES', *TelIKA*, 2, pp. 77–87. Available at: <https://doi.org/https://doi.org/10.36342/telika.v12i02.2911>.
- Simbolon, I.N., D.S.J, S.H. dan Manik, W.A. (2024) 'PREDIKSI KUALITAS AIR SUNGAI DI JAKARTA MENGGUNAKAN KNN YANG DIOPTIMALISASI DENGAN PSO', *Jurnal Informatika dan Teknik Elektro Terapan*, 12(2), pp. 1193–1203. Available at: <https://doi.org/https://doi.org/10.23960/jitet.v12i2.4191>.
- Tatimu, F., Palilingan, V.R. dan Rianto, I. (2025) 'Analisis Klasifikasi Kelulusan Mahasiswa Menggunakan Algoritma K-Nearest Neighbor di Fakultas Teknik Universitas Negeri Manado', *Journal of Education Method and Technology*, 3(1), pp. 117–123. Available at: <https://calamus.id/index.php/jemtech/article/view/111/73>.
- Winahyu, R.K.K. dan Saputro, V.A. (2023) 'Implementasi Algoritma Principal Component Analysis ( PCA ) dan K-Nearest Neighbor (KNN) Untuk Memprediksi Kelayakan Kredit Pengguna Smartphone di Indonesia Pada Masa Pandemi Covid-19', *Jurnal Darma Agung*, 31(3), pp. 1–10. Available at: <https://jurnal.universitasdarmaagung.ac>.

[id/jurnaluda/article/view/3207/3011](https://jurnaluda/article/view/3207/3011).

