

Prediksi Kelayakan Mahasiswa sebagai Penerima Beasiswa Bank Indonesia pada Tahap Seleksi Administrasi di Universitas Nurul Jadid Menggunakan Algoritma K Nearest Neighbor

Uky Oktavia Risti Permatasari¹, Wali Jafar Shudiq², Moh. Jasri³

Prodi Teknik Informatika¹, Prodi Rekayasa Perangkat Lunak², Prodi Sistem Informasi³
Fakultas Teknik Universitas Nurul Jadid Probolinggo

Article Info

Article history:

Diterima 28 Maret 2024

Revisi 12 April 2024

Diterbitkan 25 April 2024

Keywords:

Beasiswa

Administrasi

K Nearest Neighbor (KNN)

Python

ABSTRAK

Beasiswa ialah jenis bantuan keuangan yang diberikan kepada mahasiswa untuk membantu mereka membayar biaya pendidikan. Tahap seleksi administrasi merupakan langkah awal dalam menilai kelayakan mahasiswa. Proses seleksi administrasi jika dilakukan secara manual dapat memakan waktu dan sumber daya, serta keputusan manual dapat rentan terhadap subjektivitas, dan perbedaan penilaian antar panitia seleksi. Oleh karena itu dibutuhkan Prediksi yang akurat untuk dapat mengoptimalkan penggunaan sumber daya, mengetahui faktor utama dan faktor pendukung untuk membantu pihak terkait dalam menentukan kelayakan mahasiswa untuk dinyatakan lolos seleksi administrasi secara lebih mendalam. Tujuan penelitian ini ialah meminimalkan adanya pengaruh keputusan yang bersifat subjektivitas serta meminimalisir adanya human error. Penelitian ini mengusulkan Prediksi Kelayakan Mahasiswa Algoritma K Nearest Neighbor (KNN). Perhitungan jarak yang digunakan dalam penelitian ini ialah Euclidean distance yang dimana digunakan untuk mengukur seberapa mirip data yang akan di prediksi dan data latih yang ada. Implementasi algoritma ini menggunakan python di google colab. Dataset yang digunakan dalam penelitian ini ialah sebanyak 350 record data, dengan membagi 75% sebagai data training, dan 25% sebagai data testing. Hasil dari penelitian ini menunjukkan bahwa Algoritma K Nearest Neighbor (KNN) mampu menjadi model prediksi kelayakan yang baik, ditunjukkan dengan nilai akurasi sebesar 93%.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Wali Ja'far Shudiq,

Universitas Nurul Jadid, Karanganyar Paiton, Probolinggo 67291, Indonesia

Email: wali.jafar@unuja.ac.id

1. PENDAHULUAN

Beasiswa ialah jenis bantuan keuangan yang diberikan kepada mahasiswa untuk membantu mereka membayar biaya pendidikan. Ada banyak sumber yang dapat memberikan beasiswa, seperti pemerintah, lembaga swasta, atau organisasi. Tujuan utama dari beasiswa ialah untuk memberi kesempatan pendidikan kepada mereka yang memiliki potensi akademik atau prestasi di bidang lain, tetapi memiliki keterbatasan

keuangan. Salah satu Program Beasiswa yang tersebar dalam perguruan tinggi negeri maupun perguruan tinggi swasta ialah beasiswa Bank Indonesia (BI). Beasiswa Bank Indonesia adalah beasiswa yang dikeluarkan oleh Bank Indonesia dalam bentuk pengabdian Bank Indonesia kepada dunia pendidikan[1].

Universitas Nurul Jadid merupakan salah satu perguruan tinggi swasta yang mendapatkan kesempatan sebagai kampus penerima beasiswa Bank Indonesia Malang. Kuota mahasiswa penerima beasiswa Bank Indonesia ialah sebanyak 50 mahasiswa. Ada beberapa tahap dalam proses seleksi beasiswa, diantaranya ialah seleksi administrasi oleh Pihak Universitas dan seleksi wawancara oleh pihak Bank Indonesia. Selain bantuan finansial, Bank Indonesia juga memberikan fasilitas kepada mahasiswa dalam meningkatkan prestasi, dan program kerja di masyarakat. Para penerima beasiswa nantinya akan menjadi satu komunitas yaitu komunitas GenBI. GenBI sendiri merupakan singkatan dari Generasi Baru Bank Indonesia, yang dimana didalamnya ialah mahasiswa penerima beasiswa dibawah naungan Bank Indonesia.

Tahap seleksi administrasi merupakan langkah awal dalam menilai kelayakan mahasiswa. Proses seleksi administrasi jika dilakukan secara manual dapat memakan waktu dan sumber daya, serta keputusan manual dapat rentan terhadap subjektivitas, dan perbedaan penilaian antar panitia seleksi. Oleh karena itu dibutuhkan Prediksi yang akurat untuk dapat mengoptimalkan penggunaan sumber daya, mengurangi beban kerja panitia seleksi, mengetahui faktor utama dan faktor pendukung untuk membantu pihak terkait dalam menentukan kelayakan mahasiswa untuk dinyatakan lolos seleksi administrasi secara lebih mendalam. Penggunaan data historis sebelumnya juga dapat digunakan untuk menemukan dan menentukan pola terkait dalam menentukan kelayakan mahasiswa. Prediksi yang akurat akan membantu memastikan bahwa calon penerima beasiswa ialah mahasiswa yang benar-benar layak dan memenuhi semua kriteria yang telah ditentukan.

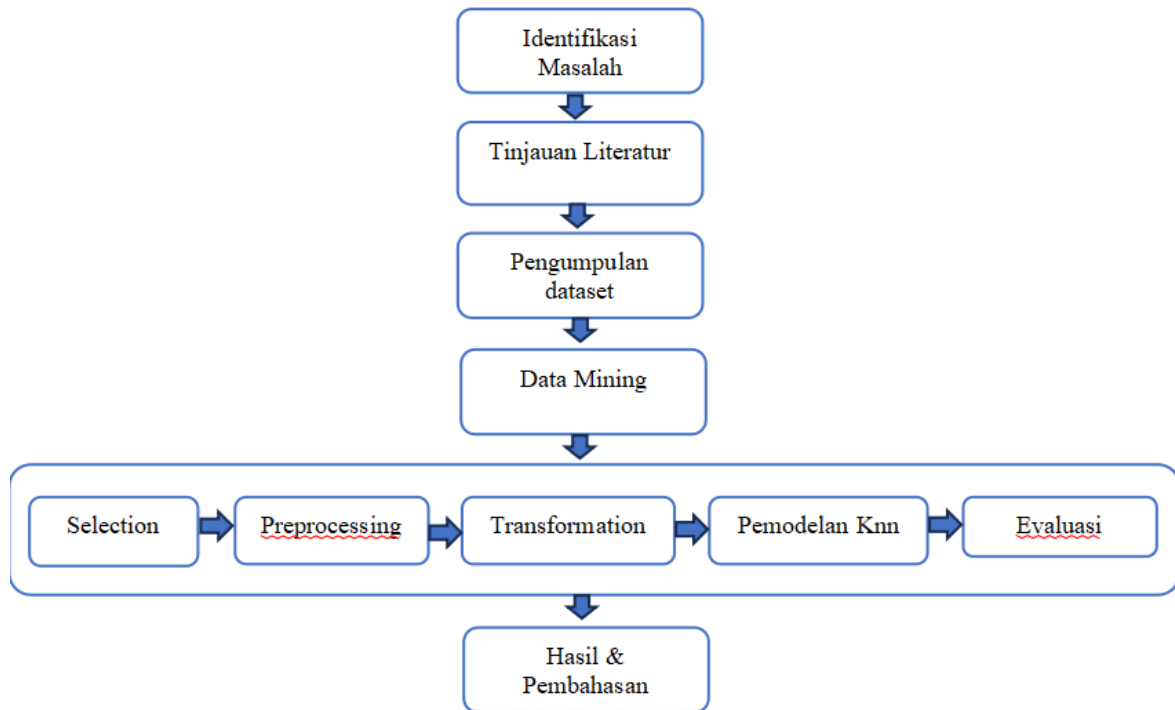
Penelitian ini mengusulkan Prediksi Kelayakan Mahasiswa sebagai Penerima Beasiswa menggunakan Algoritma *K Nearest Neighbor* (KNN). Penelitian ini menggunakan teknik data mining. Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari database yang besar [2]. Algoritma *K Nearest Neighbor* (KNN) merupakan salah satu metode machine learning yang banyak digunakan dalam tugas klasifikasi[3]. Beberapa peneliti terdahulu juga menggunakan metode *K Nearest Neighbor* (KNN) dalam bidang kesehatan seperti Pendeteksian Penggunaan Masker Untuk Pencegahan Penyebaran Covid-19 menggunakan algoritma *K Nearest Neighbor* [4]. Penerapan metode *K Nearest Neighbor* dalam sistem pendukung keputusan penerimaan KIP (Kartu Indonesia Pintar)[5]. Dalam beberapa tahun terakhir, peneliti lain telah meneliti penelitian yang serupa. Oleh karena itu, penelitian ini dapat mengutip beberapa penelitian sebagai referensi.

Penelitian yang dilakukan oleh Khatib Sulaiman & Hendri tahun 2021 [6] dengan topik “Implementasi Data Mining Dengan Metode C4.5 Untuk Prediksi Mahasiswa Penerima Beasiswa” Penelitian ini dilakukan dengan tujuan untuk mengetahui cara melakukan prediksi atau peramalan dari sejumlah mahasiswa yang akan mendapatkan beasiswa. Penelitian ini menggunakan 100 record data dan tingkat akurasi yang didapat pada penelitian ini ialah 80% dengan menerapkan metode C4.5 sebagai model prediksi mahasiswa penerima beasiswa. Penelitian selanjutnya yang dilakukan oleh Fauzi firdaus tahun 2023 [7] dengan topik “Analisis Algoritma C4.5, Naïve Bayes dan *K-Nearest Neighbor* Untuk Menentukan Penerimaan Beasiswa”. Penelitian ini bertujuan untuk menentukan klasifikasi penerimaan beasiswa di SMK Indra Bangsa dengan data yang digunakan sejumlah 56 record yang akan dibagi menjadi 70% data training dan 30% data testing. Hasil yang didapatkan ialah Tingkat akurasi yang diperoleh algoritma C4.5 dan naïve bayes sebesar 82.35% sedangkan algoritma *K-Nearest Neighbor* memperoleh nilai akurasi 88.24%. Penelitian ini menunjukkan bahwa Algoritma *K-Nearest Neighbor* memberikan nilai akurasi yang lebih baik dibandingkan dengan algoritma C4.5 dan naïve bayes.

Adapun perbedaan penelitian ini dengan penelitian sebelumnya adalah fokus penelitian ini dalam bidang Pendidikan tepatnya memprediksi kelayakan mahasiswa sebagai penerima beasiswa menggunakan algoritma *K Nearest Neighbor* (KNN). Data historis penerimaan beasiswa pada tahun sebelumnya akan menjadi dataset yang akan diolah dengan algoritma *K Nearest Neighbor* (KNN) untuk mengenali pola terkait hubungan antar atribut. Perhitungan jarak yang digunakan dalam penelitian ini ialah Euclidean distance yang dimana digunakan untuk mengukur seberapa mirip data yang akan di prediksi dan data latih yang ada. Implementasi algoritma ini menggunakan python di google colab. Python dipilih karena memiliki penulisan sintaksis yang mudah, selain itu python juga memiliki library yang lengkap dan memiliki dukungan komunitas yang kuat karena python bersifat open source[8].

2. METODE

Metode penelitian merupakan serangkaian alur tahapan dalam mengumpulkan, menganalisis serta untuk mencapai tujuan penelitian.



Gambar 1. Tahapan Metode Penelitian

a. Identifikasi Masalah

Banyaknya mahasiswa pendaftar dalam program Beasiswa bank Indonesia tentunya membuat proses seleksi cukup memakan waktu, dengan kuota yang ditentukan Bank Indonesia (BI) sebanyak 50 mahasiswa terpilih. Dengan proses seleksi yang masih dilakukan secara manual dapat memakan waktu dan sumber daya, serta keputusan manual dapat rentan terhadap subjektivitas, dan perbedaan penilaian antar panitia seleksi. Oleh karena itu dibutuhkan Prediksi yang akurat untuk dapat mengoptimalkan penggunaan sumber daya, meminimalkan adanya Keputusan yang bersifat subjektivitas, mengetahui faktor utama dan faktor pendukung untuk membantu pihak terkait dalam menentukan kelayakan mahasiswa untuk bisa dinyatakan lolos seleksi administrasi secara lebih mendalam

b. Tinjauan Literatur

Beberapa peneliti sebelumnya telah meneliti judul yang selaras menggunakan beberapa algoritma yang berbeda seperti menggunakan algoritma C4.5, *Naïve Bayes*, *K-Nearest Neighbor* maupun melakukan perbandingan antara beberapa algoritma klasifikasi. Hasil tinjauan kemudian dibuat kedalam sebuah ringkasan untuk memudahkan dalam proses banding [9]. Beberapa penelitian terkait menggunakan aplikasi rapid miner dalam pengolahan data serta masih ada yang melakukan pengolahan data secara manual. Oleh karena itu pembeda penelitian ini dengan penelitian sebelumnya ialah dataset yang berbeda serta pengolahan data yang digunakan ialah python di google colab dengan menggunakan algoritma *K-Nearest Neighbor*, karena algoritma tersebut banyak digunakan dan terbukti dapat menghasilkan prediksi yang baik.

c. Pengumpulan dataset

Dalam sebuah penelitian diperlukan adanya dataset yang nantinya akan diolah untuk mencapai tujuan tertentu. Peneliti mengumpulkan data melakukan wawancara dengan pihak terkait untuk mengetahui variabel yang tepat dan dapat diuji [10]. Dataset yang digunakan dalam penelitian ini ialah data pendaftar beasiswa Bank Indonesia (BI) pada Tahun 2022 dan 2023. Sumber dataset ini diperoleh dari pihak terkait yang bertanggung jawab dalam proses seleksi administrasi beasiswa Bank Indonesia (BI) di Universitas Nurul Jadid.

d. Data Mining

1. Selection

Tahapan ini merupakan tahapan persiapan dalam pemilihan data. Data yang telah didapatkan dari observasi dan wawancara akan dilakukan proses pemilahan atribut untuk selanjutnya digunakan untuk preprocessing[11] Dataset ini mencakup informasi terkait kelayakan mahasiswa.

2. Preprocessing Data

Preprocessing data dilakukan untuk memastikan dan memilah kesesuaian data untuk dapat diolah lebih lanjut. Dari data yang diperoleh, dilakukan cleaning agar terdeteksi data ganda, data hilang dan data outlier [12]. Pemilihan fitur dalam tahapan ini untuk mengidentifikasi variabel variabel yang relevan dalam prediksi kelayakan mahasiswa sebagai penerima beasiswa Bank Indonesia.

3. Transformation

Setelah melakukan preprocessing data kemudian masuk ke tahap transformasi, dimana data akan dipilah dengan kebersesuaian inliner atau outliner[13]. Transformasi data dilakukan untuk mempersiapkan data yang akan diolah.

4. Pemodelan *K-Nearest Neighbor*

Pada tahap pemodelan dataset yang sudah melalui proses pembersihan data, akan diolah menggunakan pemodelan Algoritma *K-Nearest Neighbor* yang merupakan salah satu metode dalam machine learning untuk digunakan dalam proses pengklasifikasian[14]. Pemodelan algoritma *K-Nearest Neighbor* menggunakan python di google colab, dengan perhitungan jarak yang digunakan ialah Euclidean distance. Hasil dari pemodelan ini harapannya dapat memberikan gambaran yang akurat dan dapat diandalkan terkait dengan potensi kelayakan mahasiswa untuk menerima beasiswa

5. Evaluasi

Evaluasi dilakukan untuk mengevaluasi performa model prediksi yang telah dilakukan, dengan menampilkan nilai confusion matrix dan F1 score untuk dapat mengukur sejauh mana model dapat memprediksi secara akurat. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya [15].

3. HASIL DAN PEMBAHASAN

3.1. Data selection

Dataset yang digunakan dalam penelitian ini ialah data mahasiswa pendaftar Beasiswa Bank Indonesia pada tahun 2022-2023. Total data yang digunakan ialah 383 record dengan beberapa atribut yang ditunjukkan pada tabel dibawah ini. Beberapa atribut diatas, akan dipilih variabel yang dapat memprediksi kategori mahasiswa yang dapat dinyatakan lolos atau tidak lolos pada tahap seleksi administrasi. Variabel yang akan digunakan ialah Indek Prestasi Kumulatif (IPK), Sktm, dan juga Prestasi. Keterlibatan pihak Universitas dalam penyediaan dataset tentunya menjamin keakuratan data yang akan diolah untuk didapatkan hasil prediksi.

Tabel 1. Dataset Penelitian

No	NIM	NAMA	PRODI	JK	IPK	SKTM	PRESTASI	Keterangan
1.	19104000	Ira Febrihana	Ekonomi Syariah	P	3.94	Ada	4	Lulus
2.	2010700062	Nadhifah Shofiyah	PAI	P	3.93	Tidak Ada	4	Lulus
3.	2010700070	Husein	PAI	P	3.93	Tidak Ada	2	Tidak Lulus
4.	2010900003	Agus Salim	(MPI)	L	3.90	Ada	3	Lulus
5.	1910800019	Fadilatur Rohma	PBA	P	3.90	Tidak Ada	3	Lulus
...
...
383.	2142400079	Nurul kamaliah	Ekonomi	P	3.67	Ada	3	Tidak Lulus

3.2. Preprocessing

Tahapan ini melakukan pembersihan pada data yang ada, seperti data duplikat dan data kosong. Data yang mulanya berjumlah 383 record menjadi 350 data melalui tahapan preprocessing, yaitu dengan menghapus data duplikat yang berjumlah 33 record. Penghapusan data duplikat dilakukan secara manual di microsoft excel. Data duplikat ini dapat terjadi karena data pendaftar pada tahun 2023 berisi mahasiswa yang sama dan pernah mendaftar pada tahun sebelumnya atau pada Tahun 2022. Pada tahap preprocessing ini juga dilakukan *label encoding* pada variabel Surat Keterangan Tidak Mampu (SKTM), yaitu mengubah kategori kelas dalam bentuk kategorikal menjadi numerik.

Keterangan : **Tidak Ada** (1), **Ada** (2) Seperti yang ditunjukkan pada tabel dibawah.

SKTM	PRESTASI	Keterangan
Ada	4	Lulus
Tidak	4	Lulus
Ada	2	Tidak Lulus
Ada	3	Lulus
Tidak	3	Lulus
Ada
...
...
Ada	3	Tidak Lulus

SKTM	PRESTASI	Keterangan
2	4	Lulus
1	4	Lulus
1	2	Tidak Lulus
2	3	Lulus
1	3	Lulus
...
...
2	3	Tidak Lulus

Gambar 2. Data sebelum dan sesudah dilakukan label encoding

Pada tahap preprocessing, penentuan data yang akan diuji sebanyak 3 variabel, yaitu Indeks Prestasi Kumulatif (ipk), Surat Keterangan Tidak Mampu (sktm), dan Prestasi. Seperti yang ditunjukkan pada gambar dibawah ini

```

[1] import numpy as np
import matplotlib.pyplot as plt
import pandas as pd

dataset = pd.read_csv('/content/DatasetPenelitianFix.csv')

[3] dataset.head()

[4] x = dataset.iloc[:,[5,6,7]].values
y = dataset.iloc[:, -1].values

print(x)

[[3.98 1.  4. ]
 [3.97 2.  3. ]
 [3.95 1.  4. ]
 ...
 [3.27 1.  4. ]
 [3.63 1.  4. ]
 [3.79 2.  2. ]]

```

Gambar 3. Tahap Preprocessing

3.3. Transformation

Tranformasi data yang efektif dapat mempengaruhi analisis atau hasil pada pengolahan data. Pada penelitian ini dilakukan tahap normalisasi untuk mengubah nilai variabel, agar nilai antar variabel seragam dan tidak terjadi perbedaan skala terlalu jauh. Tahapan Transformation ditunjukkan dengan gambar dibawah ini.

```

from sklearn.preprocessing import StandardScaler
sc = StandardScaler()
x_train = sc.fit_transform(x_train)
x_test = sc.transform(x_test)

[11] print(x_train)

print(x_test)

[[-0.24350766 -1.05494229 -0.587811 ]
 [-0.11578192 -1.05494229 -1.47290573]
 [ 0.01194382  0.94791916 -0.587811 ]
 [ 0.2673953  -1.05494229  1.18237845]
 [ 0.01194382 -1.05494229 -0.587811 ]
 [ 0.84216113  0.94791916 -0.587811 ]
 [ 0.906024    0.94791916  1.18237845]
 [ 0.13966956  0.94791916 -1.47290573]
 [-1.52076507  0.94791916 -0.587811 ]
 [ 0.71443539 -1.05494229  1.18237845]
 [ 0.39512104  0.94791916  1.18237845]
 [-2.03166803 -1.05494229  0.29728373]
 [ 0.77829826 -1.05494229 -1.47290573]
 [ 0.39512104 -1.05494229  0.29728373]
 [ 0.33125817  0.94791916  1.18237845]
 [ 0.77829826  0.94791916  2.06747318]
 [ 0.84216113 -1.05494229 -1.47290573]
 [ 0.13966956 -0.23704016  0.29728373]

```

Gambar 4. Tahap Transformation (Normalisasi data)

Setelah melakukan normalisasi data selanjutnya ialah membagi data training dan juga data testing untuk dapat diolah lebih lanjut. Dari 350 data, akan dibagi menjadi 75% sebagai data training data 25% data testing menggunakan split data seperti yang ditunjukkan oleh gambar dibawah ini.

```

Membagi data Training dan Testing

[7] from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(x,y, test_size= 0.25, random_state = 0)

[8] len(x_train)

262

len(x_test)

88

```

Gambar 5. Pembagian data training dan data testing

3.4 Pemodelan K-Nearest Neighbor

Pada tahap pemodelan dataset yang sudah melalui proses pembersihan data, kemudian akan diolah menggunakan pemodelan Algoritma *K-Nearest Neighbor* (knn). Perhitungan jarak yang digunakan dalam penelitian ini ialah *Euclidean distance* yang dimana digunakan untuk mengukur seberapa mirip data yang akan di prediksi dan data latih yang ada. Implementasi algoritma ini menggunakan python di google colab. Nilai parameter atau K tetangga yang digunakan pada penelitian ini ialah K=5.

```

[14] from sklearn.neighbors import KNeighborsClassifier
      classifier = KNeighborsClassifier(n_neighbors=5, metric = 'euclidean', p=2)
      classifier.fit(x_train, y_train)

      KNeighborsClassifier
      KNeighborsClassifier(metric='euclidean')

[15] y_pred = classifier.predict(x_test)

[16] from sklearn.metrics import classification_report
      print(classification_report(y_test,y_pred))

```

	precision	recall	f1-score	support
Lulus	0.85	0.92	0.88	25
Tidak Lulus	0.97	0.94	0.95	63
accuracy			0.93	88
macro avg	0.91	0.93	0.92	88
weighted avg	0.93	0.93	0.93	88

Gambar 6. Pemodelan K-Nearest Neighbor (knn)

Seperti hasil yang ditunjukkan pada gambar diatas, bahwa akurasi yang didapat ialah sebesar 93%. Hal ini menjelaskan bahwa Algoritma *K-Nearest Neighbor* mampu menjadi model prediksi yang baik dalam prediksi kelayakan mahasiswa sebagai penerima beasiswa.

3.5 Evaluasi

Evaluasi dilakukan untuk mengevaluasi performa model prediksi yang telah dilakukan, dengan menampilkan nilai confusion matrix untuk dapat mengukur sejauh mana model dapat memprediksi secara akurat. Confusion matrix dapat dilihat pada gambar dibawah ini

```

[17] from sklearn.metrics import confusion_matrix, accuracy_score
      cm = confusion_matrix(y_test, y_pred)
      print(cm)
      accuracy_score(y_test, y_pred)

```

```

[[23  2]
 [ 4 59]]
0.9318181818181818

```

Gambar 6. Confusion Matrix

Hasil yang didapat dalam confusion matrix diatas menjelaskan bahwa terdapat 23 data yang termasuk kedalam True Positive (TP), 59 data di klasifikasikan menjadi True Negatif (TN), 4 data diklasifikasikan sebagai false Negative (FN) dan 2 data diklasifikasikan sebagai false Positif (FP).

4. KESIMPULAN

Setelah melakukan pengolahan data dengan jumlah dataset 350 record yang diambil dari data pendaftar Beasiswa Bank Indonesia di Universitas Nurul Jadid Tahun 2022-2023, kemudian di implementasikan menggunakan python di google colab dengan pemodelan Algoritma *K-Nearest Neighbor* maka dihasilkan hasil akurasi sebesar 93% yang dimana hasil ini membuktikan bahwa algoritma mampu memberikan nilai akurasi yang baik dan mampu menjadi model prediksi dalam kelayakan mahasiswa sebagai penerima beasiswa Bank Indonesia. Hasil dari penelitian ini diharapkan dapat membantu pihak terkait dalam proses penyeleksian mahasiswa dalam tahap administrasi. Dengan adanya nilai akurasi yang baik dan akurat tentunya akan menghasilkan keputusan yang objektif, sehingga mahasiswa yang dinyatakan lolos seleksi administrasi ialah mahasiswa yang memang benar benar layak dengan tetap mempertimbangkan kinerja akademis mahasiswa.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih banyak kepada semua pihak terkait yang turut mendukung, dan membantu dalam proses penyelesaian penelitian ini. Terimakasih kepada Dosen pembimbing yang turut memberikan arahan dan membimbing penulis sehingga dapat menyelesaikan setiap proses penelitian. Terimakasih kepada orangtua, keluarga maupun teman yang terus mendukung penulis sehingga akhirnya dapat menyelesaikan penelitian ini.

REFERENSI

- [1] A. Mardiana, D. Abdurahman, and P. Putriani, "SISTEM PENDUKUNG KEPUTUSAN PENERIMAAN BEASISWA BANK INDONESIA STUDI KASUS UNIVERSITAS MAJALENGKA," *INFOTECH journal*, vol. 8, no. 1, pp. 13–21, Jan. 2022, doi: 10.31949/infotech.v8i1.1664.
- [2] N. Afriani Manihuruk, M. Zarlis, E. Irawan, and H. Satria Tambunan, "Penerapan Data Mining Dalam Mengelompokkan Calon Penerima Beasiswa Dengan Menggunakan Algoritma K-Means," vol. 4, no. 1, 2020, doi: 10.30865/komik.v4i1.2575.
- [3] "tik-52-9-hal-penerapan-algoritma-k-nearest-neighbor-untuk-prediksi-pengelompokkan-tingkat-risiko-penyebaran-covid-19-jawa-barat".
- [4] A. Saleh, "Pendeteksian Penggunaan Masker Untuk Pencegahan Penyebaran Covid-19 Menggunakan Algoritma K-nearest neighbor." [Online]. Available: <https://www.kaggle.com/datasets/jamesnogra/face->
- [5] D. Desa *et al.*, "Arifin, Shudiq, Maghfiroh-Penerapan Metode Knn (K-Nearest Neighbor) Dalam Sistem Pendukung Keputusan Pen-erimaan KIP (Kartu Indonesia Pintar) PENERAPAN METODE KNN (K-NEAREST NEIGHBOR) DALAM SISTEM PENDUKUNG KEPUTUSAN PENERIMAAN KIP (KARTU INDONESIA PINTAR) DI DESA PANDEAN BERBASIS WEB DAN MYSQL."
- [6] J. Khatib Sulaiman and H. Hendri, "Implementasi Data Mining Dengan Metode C4.5 Untuk Prediksi Mahasiswa Penerima Beasiswa," *Indonesian Journal of Computer Science Attribution-ShareAlike*, vol. 4, no. 2, pp. 2021–312.
- [7] J. Penelitian Ilmu Komputer and M. Fauzi Firdaus, "ANALISIS ALGORITMA C 4.5, NAÏVE BAYES DAN K-NEAREST NEIGHBOR UNTUK MENENTUKAN PENERIMAAN BEASISWA," 2023. [Online]. Available: <https://mypublikasi.com/index.php/JUPIK/156>
- [8] M. Riziq sirfatullah Alfarizi, M. Zidan Al-farish, M. Taufiqurrahman, G. Ardiansah, and M. Elgar, "PENGUNAAN PYTHON SEBAGAI BAHASA PEMROGRAMAN UNTUK MACHINE LEARNING DAN DEEP LEARNING," 2023.
- [9] L. Setiyani, M. Wahidin, D. Awaludin, and S. Purwani, "Analisis Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Data Mining Naïve Bayes : Systematic Review," *Faktor Exacta*, vol. 13, no. 1, p. 35, Jun. 2020, doi: 10.30998/faktorexacta.v13i1.5548.
- [10] Y. Andini, J. Tata Hardinata, Y. Pranayama Purba, and P. A. Studi Sistem Informasi STIKOM Tunas Bangsa JJend Sudirman Blok No, "PENERAPAN DATA MINING TERHADAP TATA LETAK BUKU DI PERPUSTAKAAN SINTONG BINGEI PEMATANGSIANTAR MENGGUNAKAN METODE APRIORI," 2022. [Online]. Available: <http://ejournal.stmik-time.ac.id>
- [11] F. Alghifari and D. Juardi, "Fauzan Alghifari Penerapan Data Mining Pada Penerapan Data Mining Pada Penjualan Makanan Dan Minuman Menggunakan Metode Algoritma Naïve Bayes."
- [12] A. Al Arif, M. Firdaus, Y. Maruhawa, S. AMIK Riau, and J. Purwodadi Panam, "SENTIMAS: Seminar Nasional Penelitian dan Pengabdian Masyarakat Comparison of Data Mining Methods for Prediction of Rainfall with C4.5, Naïve Bayes, and KNN Algorithm Perbandingan Metode Data Mining untuk

- Prediksi Curah Hujan dengan Algoritma C4.5, Naïve Bayes, dan KNN.” [Online]. Available: <https://journal.irpi.or.id/index.php/sentimas>
- [13] M. Muslim, S. Windarti, D. Manajemen, I. Kesehatan, S. Akbidyo, and J. Parangtritis, “Visualisasi Data Mining Untuk Skiring Digital Covid-19 Pada Instansi,” *Jurnal Teknik Informatika Kaputama (JTIK)*, vol. 7, no. 1, 2023.
- [14] M. N. Muttaqin and I. Kharisudin, “Analisis Sentimen Pada Ulasan Aplikasi Gojek Menggunakan Metode Support Vector Machine dan K Nearest Neighbor,” *UNNES Journal of Mathematics*, vol. 10, no. 2, pp. 22–27, 2021, [Online]. Available: <http://journal.unnes.ac.id/sju/index.php/ujm>
- [15] R. Bahtiar, “Implementasi Data Mining Untuk Prediksi Penjualan Kusen Terlaris Menggunakan Metode K-Nearest Neighbor.” [Online]. Available: <https://jurnal.publikasitecno.id/index.php/jim203>