

Jurnal JTik (Jurnal Teknologi Informasi dan Komunikasi)

DOI: <https://doi.org/10.35870/jtik.v9i3.3848>

Penerapan Prediksi untuk Klasifikasi Penerima Beasiswa Berprestasi pada SMK Islam Pemalang Berdasarkan Algoritma *K-Nearest Neighbor*

Agung Yuliyanto Nugroho ¹, Tundo ^{2*}, Riolandi Akbar ³

¹ Program Studi Informatika, Universitas Cendekia Mitra Indonesia, Yogyakarta, Provinsi Yogyakarta, Indonesia.

^{2*} Program Studi Teknik Informatika, Fakultas Teknik, Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika, Kota Jakarta Timur, Daerah Khusus Ibukota Jakarta, Indonesia.

³ Program Studi Pendidikan Agama Islam, STIT Al Quraniyah Manna, Kabupaten Bengkulu Selatan, Provinsi Bengkulu, Indonesia.

article info

Article history:

Received 10 January 2025

Received in revised form

10 February 2025

Accepted 10 March 2025

Available online July 2025.

Keywords:

Predicting; Scholarship; K-Nearest Neighbor; Algorithm; Classification.

Kata Kunci:

Prediksi; Beasiswa; K-Nearest Neighbor; Algoritma; Klasifikasi.

abstract

This research aims to help Pemalang Islamic SMK in identifying outstanding students and predicting potential scholarship recipients, by utilizing the algorithm, K-Nearest Neighbor (K-NN) in determining students who have the potential to receive scholarships. This research used 100 student data involving attributes such as report card grades, academic achievement, parental responsibilities, parental salary, and participation in organizations. Meanwhile, the testing process is carried out by adding 6 data on potential scholarship recipients to be predicted. The data is then processed and normalized before being applied to the K-NN algorithm. The K-NN steps involve determining the K parameter (number of nearest neighbors), calculating the Euclidean distance, sorting the distance results, and selecting the majority category as a prediction for the new object class. The research results show that the application of the K-NN algorithm with K=3 is successful in providing predictions of outstanding students by considering relevant attributes. This process is carried out with the help of JAVA programming to calculate and analyze data. The research conclusion shows that the K-NN algorithm can be used as an effective prediction tool for classification to determine students who excel and are worthy of receiving scholarships. This research contributes to increasing efficiency and accuracy in the selection of outstanding scholarship recipients in the school environment with an accuracy of 83.33%.

abstrak

Penelitian ini bertujuan untuk membantu SMK Islam Pemalang dalam mengidentifikasi siswa berprestasi dan memprediksi potensi penerima beasiswa, dengan memanfaatkan algoritma K-Nearest Neighbor (K-NN) dalam menentukan siswa yang mempunyai potensi menerima beasiswa. Penelitian ini menggunakan 100 data siswa yang melibatkan atribut seperti nilai rapor, prestasi akademik, tanggungan orang tua, gaji orang tua, dan keikutsertaan dalam organisasi. Sedangkan proses pengujian dilakukan dengan menambahkan 6 data calon penerima beasiswa untuk diprediksi. Data tersebut kemudian diolah dan dinormalisasi sebelum diaplikasikan pada algoritma K-NN. Langkah-langkah K-NN melibatkan penentuan parameter K (jumlah tetangga terdekat), perhitungan jarak Euclidean, pengurutan hasil jarak, dan pemilihan kategori mayoritas sebagai prediksi kelas objek baru. Hasil penelitian menunjukkan bahwa penerapan algoritma K-NN dengan K=3 berhasil memberikan prediksi siswa berprestasi dengan mempertimbangkan atribut yang relevan. Proses ini dilakukan dengan bantuan pemrograman JAVA untuk menghitung dan menganalisis data. Kesimpulan penelitian menunjukkan bahwa algoritma K-NN dapat digunakan sebagai alat prediksi untuk klasifikasi yang efektif untuk menentukan siswa berprestasi dan layak menerima beasiswa. Penelitian ini memberikan kontribusi dalam meningkatkan efisiensi dan akurasi dalam seleksi penerima beasiswa berprestasi di lingkungan sekolah dengan keakuratan 83,33%.

Corresponding Author. Email: asna8mujahid@gmail.com ^{2}.



ACM Computing Classification System (CCS)

EBSCOhost

Communication and Mass Media Complete (CMMC)

Copyright 2025 by the authors of this article. Published by Lembaga Otonom Lembaga Informasi dan Riset Indonesia (KITA INFO dan RISET). This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

1. Pendahuluan

Perkembangan teknologi, terutama dalam bidang komputer, telah mengalami kemajuan yang signifikan dalam beberapa tahun terakhir. Pada masa lalu, manusia umumnya menggunakan alat tulis tradisional seperti pena atau pensil untuk menulis. Namun, saat ini, proses menulis manual dapat digantikan dengan menggunakan komputer (Tundo & Mahardika, 2023). Cukup dengan menekan tombol pada papan ketik, huruf atau angka yang diinginkan akan muncul secara langsung dilayar. Pentingnya teknologi komputer juga tercermin dalam dunia pendidikan, terutama di lembaga-lembaga seperti sekolah. Komputer telah menjadi alat yang mempermudah kinerja setiap guru dan staf yang bertugas di sekolah. Penggunaan komputer juga memiliki dampak positif dalam proses penerimaan beasiswa. Beasiswa merupakan bantuan keuangan yang diberikan kepada perorangan dengan tujuan mendukung kelangsungan pendidikan yang sedang dijalani (Tundo, Wijonarko, & Raffiudin, 2023). Di setiap lembaga pendidikan, terutama Sekolah Menengah Atas/Kejuruan, terdapat dana bantuan sekolah yang ditawarkan kepada siswa-siswi, terutama yang memiliki kondisi keuangan kurang mampu, namun secara akademik memiliki prestasi yang baik. Hal ini bertujuan untuk memastikan bahwa setiap siswa memiliki akses yang adil terhadap pendidikan dan mendukung kelangsungan proses pembelajaran.

SMK Islam Pemalang yang terletak di Jalan Yos Sudarso No. 41 Widuri Kecamatan Taman, Kabupaten Pemalang memiliki tujuan untuk meningkatkan prestasi akademis siswanya. Oleh karena itu, diperlukan suatu sistem yang dapat membantu memprediksi siswa yang berpotensi berprestasi agar dapat mendapatkan perhatian dan pembinaan lebih lanjut. Dalam pemberian beasiswa kepada siswa berprestasi, terdapat keterbatasan anggaran yang membuat tidak dapat mencakup semua penerima beasiswa. Selain itu, seringkali terjadi ketidaktepatan sasaran dalam pemberian beasiswa. Oleh karena itu, dipilihlah pendekatan yang memfokuskan pada siswa-siswa yang benar-benar layak untuk menerima beasiswa tersebut. Meskipun banyak siswa yang memenuhi persyaratan dan mengajukan permohonan, pihak sekolah membutuhkan waktu yang cukup untuk secara teliti

menentukan siswa-siswa yang paling pantas menerima beasiswa (Purwanto & Nugroho, 2023). Penelitian ini dilakukan dengan maksud membantu pihak sekolah dalam mengkategorikan siswa berprestasi di lingkungan sekolah, dengan tujuan untuk memprediksi dengan lebih tepat siswa-siswa yang layak menerima beasiswa berprestasi (Hasanah, Suprpti, Rahaningsih, & Ali, 2022). Algoritma K-NN dipilih karena kemampuannya dalam melakukan prediksi berdasarkan kemiripan data. Penelitian ini akan fokus pada penerapan algoritma K-NN untuk memprediksi siswa berprestasi berdasarkan data yang telah dikumpulkan (Dinata, Akbar, & Hasdyna, 2020). Algoritma K-NN dianggap mempunyai kesederhanaan dalam pengolahan data *training* dan data *testing* dalam jumlah yang sangat besar (Homepage, Cholil, Handayani, Prathivi, & Ardianita, 2021), (Widaningsih & Yusuf, 2022). Sebagai penguatan terhadap penelitian ini, berikut beberapa penelitian serupa terkait penentuan beasiswa. Penelitian oleh (Pratama, Ali Ma, & Rizki Rinaldi, 2021) terkait prediksi untuk klasifikasi beasiswa berdasarkan peningkatan prestasi akademik (PPA) dan bantuan belajar mahasiswa (BBM) dilakukan dengan menggunakan variabel-variabel yang telah ditetapkan melalui penerapan algoritma K-Nearest Neighbor. Pengujian performa algoritma dilakukan dengan metode *Cross Validation*, *Confusion Matrix*, dan kurva *Receiver Operating Characteristic* (ROC).

Hasil pengujian menunjukkan bahwa akurasi untuk klasifikasi beasiswa peningkatan prestasi akademik mencapai 88,33%, dengan nilai *Area Under the Curve* (AUC) sebesar 0,925 dari total 227 rekaman data dalam dataset. Sementara itu, akurasi untuk klasifikasi beasiswa bantuan belajar mahasiswa mencapai 90%, dengan nilai AUC sebesar 0,937 dari total 183 rekaman data dalam dataset. Selain itu, akurasi untuk klasifikasi gabungan beasiswa peningkatan prestasi akademik dan bantuan belajar mahasiswa mencapai 85,56%, dengan nilai AUC sebesar 0,958. Karena nilai AUC berada dalam rentang 0,9 hingga 1, hasil metode yang diterapkan dianggap sangat baik (*excellent*) dalam melakukan klasifikasi dan prediksi terkait penerimaan beasiswa. Sehingga algoritma K-NN yang digunakan terbukti efektif dalam membedakan antara penerima beasiswa peningkatan prestasi akademik, penerima beasiswa bantuan belajar mahasiswa, dan gabungan keduanya.

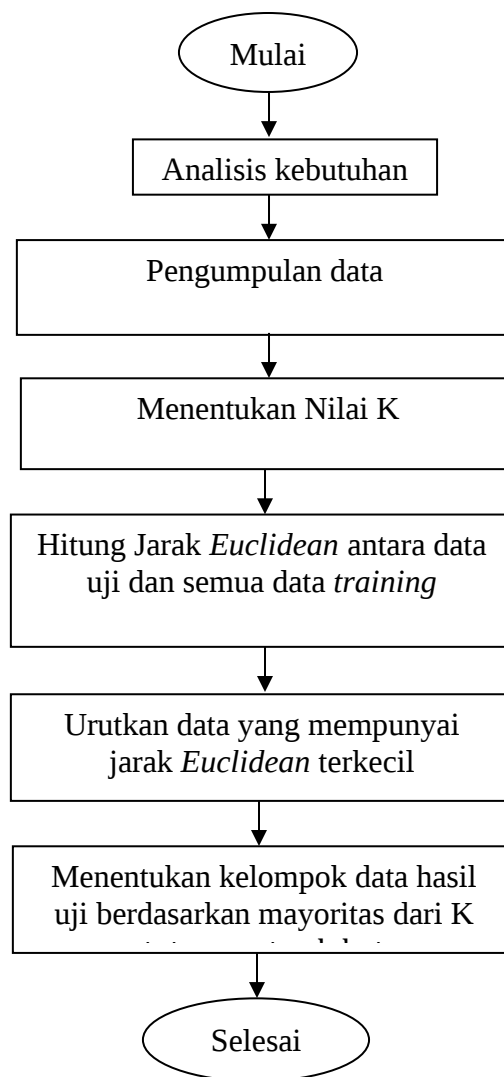
Sedangkan menurut (Karepesina & Zahrotun, 2023), beasiswa yang diberikan melibatkan berbagai jenis, termasuk beasiswa peningkatan prestasi akademik (PPA) yang diberikan kepada mahasiswa berprestasi. Seiring dengan peningkatan jumlah mahasiswa yang mengajukan beasiswa, kebutuhan akan metode klasifikasi yang dapat membantu menentukan siapa yang berhak menerima beasiswa PPA dari Universitas Sebelas Maret menjadi semakin penting. Hasil penelitian ini menyoroiti keakuratan dan tingkat kesalahan dari algoritma K-NN dan algoritma C 4.5 dalam menentukan penerima beasiswa PPA. Algoritma K-NN menunjukkan tingkat akurasi sebesar 87.7% dengan tingkat kesalahan 12.3%. Sementara itu, pada algoritma C 4.5, akurasi yang diperoleh adalah sebesar 85.3%, dengan tingkat kesalahan sebesar 14.7%. Temuan ini menggambarkan bahwa algoritma K-NN memiliki tingkat akurasi yang lebih tinggi dibandingkan dengan algoritma C4.5 dalam konteks klasifikasi penerima beasiswa PPA.

Informasi ini dapat memberikan panduan yang berharga dalam proses pengambilan keputusan terkait penerimaan beasiswa, memungkinkan penyelenggara beasiswa untuk membuat keputusan yang lebih tepat dan efisien. Berdasarkan pertimbangan informasi yang telah diuraikan sebelumnya, peneliti tertarik untuk mengadakan penelitian yang berkaitan dengan prediksi penerima beasiswa berprestasi. Penelitian ini akan menerapkan metode K-NN dengan menggunakan data yang diperoleh dari SMK Islam Pemalang. Penelitian ini akan memanfaatkan lima atribut sebagai variabel penentu, yaitu nilai raport, prestasi akademik, tanggungan orang tua, gaji orang tua, dan keikutsertaan dalam organisasi. Dengan menggunakan data dan metode ini, penelitian bertujuan untuk memprediksi dengan lebih tepat siapa saja yang memiliki potensi untuk menjadi penerima beasiswa berprestasi di lingkungan SMK Islam Pemalang.

2. Metodologi Penelitian

Tahapan metode yang dilakukan dalam penelitian secara garis besar terlihat pada Gambar 1 dengan melalui 6 tahapan yang kemudian hasil dari algoritma

K-NN dianalisa untuk mengetahui tingkat keakuratan, *precision*, dan *recall* dengan menggunakan *Confussion Matrix*.



Gambar 1. Alur Penelitian

Berdasarkan Gambar 1 tentang Alur Penelitian dapat dijelaskan sebagai berikut:

- 1) Analisis Kebutuhan
Analisis Kebutuhan dilakukan untuk menentukan atribut yang harus ada untuk memprediksi penerima beasiswa yaitu berupa Nilai Raport, Prestasi Akademik, Tanggungan Orang Tua, Gaji Orang Tua dan Keikutsertaan Organisasi.
- 2) Pengumpulan Data
Pengumpulan data dilakukan setelah memutuskan atribut apa yang menjadi kualifikasi penentuan penerimaan beasiswa, kemudian dilakukan pengumpulan data dengan menginputkan data

Nilai Raport, Prestasi Akademik, Tanggungan Orang Tua, Gaji Orang Tua dan Keikutsertaan Organisasi kedalam tabel.

3) Menentukan Nilai K.

Menentukan Jumlah nilai tetangga paling dekat untuk memperoleh klasifikasi.

4) Hitung Jarak *Euclidean*

Jarak euclidean yaitu jarak yang dihitung sebagai teknik pencarian tetangga terdekat dan berfungsi untuk menguji ukuran yang bisa digunakan sebagai intepretasi kedekatan jarak antara dua objek.

5) Urutkan data yang mempunyai jarak *Euclidean* terkecil.

Mengurutkan objek-objek data ke dalam kelompok yang mempunyai jarak terkecil.

6) Menentukan kelompok data hasil uji berdasarkan mayoritas dari K.

Menentukan Output yang telah dihasilkan berdasarkan urutan jarak dimana output yang dihasilkan adalah jumlahnya yang lebih banyak.

Pengumpulan Data

Data yang digunakan dalam penelitian ini berjumlah 80 data set, adapun atribut yang mencakup nilai raport, Prestasi akademik, tanggungan orang tua, Gaji orang tua dan keikutsertaan Organisasi. Disajikan pada tabel 1 dibawah ini:

Tabel 1. Atribut Data Siswa SMK Karya Mandiri Nusawungu

Nama	Nilai Raport	Prestasi Akademik	Tanggungan Orang Tua	Gaji Orang Tua	Keikutsaertaan Organisasi	Prediksi
Achmad Syah	78	1	1	505.000	1	Tidak Dapat
Ade Januar	80	1	1	555.000	1	Dapat
Kurniawan	79	1	1	950.000	1	Tidak Dapat
Afan Mardian	80	1	1	500.000	1	Dapat
Ahbe Saiful	77	2	2	500.000	1	Dapat
Alif Andre	77	1	1	700.000	1	Tidak Dapat
Andika R	78	1	1	500.000	1	Tidak Dapat
Angga R	79	1	1	1.100.000	1	Tidak Dapat
Arif Wahyu	80	1	1	500.000	1	Dapat
.....						
.....						
Aris Priyanto	77	1	1	500.000	1	Tidak Dapat

Pemrosesan Data

Dalam tahap ini, data yang telah dikumpulkan melalui berbagai sumber diolah dan dinormalisasi untuk memastikan konsistensi dan relevansi sebagai input bagi algoritma *K-Nearest Neighbor* (K-NN). Proses normalisasi bertujuan untuk menghilangkan perbedaan skala antara variabel yang dapat mempengaruhi hasil prediksi. Setelah data siap, dilakukan uji coba dengan menambahkan enam data baru yang akan diprediksi menggunakan algoritma K-

NN. Tujuan dari uji coba ini adalah untuk menentukan kelayakan siswa-siswa tersebut dalam menerima Beasiswa Berprestasi berdasarkan kriteria yang telah ditetapkan. Hasil dari pengolahan dan prediksi data calon penerima beasiswa disajikan secara rinci pada Tabel 2, yang memuat informasi penting mengenai atribut yang digunakan dalam penilaian. Proses ini memastikan bahwa keputusan yang diambil bersifat objektif dan didasarkan pada analisis data yang akurat.

Tabel 2. Data Siswa Calon Penerima

Nama	Nilai Raport	Prestasi Akademik	Tanggungan Orang Tua	Gaji Orang Tua	Keikutsertaan Organisasi	Prediksi
Dzikry	81	1	1	1.000.000	1	?
Erwin J	77	1	2	600.000	1	?
Evan Z	81	1	2	1.400.000	1	?
Fikri A	78	1	1	750.000	1	?
Gilang P	79	1	1	850.000	1	?
Iklil Abi	78	2	1	1.000.000	3	?

Algoritma *K-Nearest Neighbor* (K-NN)

Algoritma K-NN adalah algoritma yang menentukan nilai jarak pada pengujian data testing dengan data training berdasarkan nilai terkecil dari nilai ketetanggaan terdekat. Algoritma k tetangga terdekat (*K-Nearest Neighbour algorithm*, K-NN) adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut (Nuraeni, Syam, Wajdi, Firmansyah, & Malkan, 2023). Tujuannya adalah untuk mengklasifikasi objek baru berdasarkan atribut dan training samples. Algoritma K-NN menggunakan klasifikasi ketetanggaan sebagai nilai prediksi dari sampel uji yang baru. Langkah-langkah dalam algoritma K-NN:

- 1) Menentukan parameter K (jumlah tetangga paling dekat). Nilai K dipilih berdasarkan ketentuan apabila klasifikasi genap maka ambil nilai k ganjil dan nilai K > 1. Dan dalam kasus ini sudah ditentukan bahwa nilai K=3.
- 2) Menghitung kuadrat jarak euclidean objek terhadap data training yang diberikan.
- 3) Mengurutkan hasil jarak euclidean dari yang terkecil.
- 4) Mengumpulkan kategori (Klasifikasi Nearest Neighbor berdasarkan K).

Dengan menggunakan kategori nearest neighbor yang paling mayoritas maka dapat diklasifikasikan kategori objek baru.

Jarak Euclidean

Jarak euclidean yaitu jarak yang dihitung sebagai teknik pencarian tetangga terdekat dan berfungsi untuk menguji ukuran yang bisa digunakan sebagai interpretasi kedekatan jarak antara dua objek. Euclidean distance adalah perhitungan untuk mengukur jarak dua titik dalam euclidean space yang

mempelajari hubungan antara sudut dan jarak (Miftahuddin, Umaroh, & Karim, 2020). Dalam matematika euclidean distance d. Jarak euclidean didefinisikan sebagai berikut (Pahrudin & Harianto, 2022).

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Keterangan:

d = Jarak Euclidean

x1, y1 = Variabel Data Testing

x2, y2 = Variabel Data Training

Pengujian Confusion Matrix

Confusion matrix digunakan untuk pengujian dalam memperkirakan objek yang bernilai *True* dan *False* (Deden Martia Nanda, Tacbir Hendro Pudjiantoro, 2022). Urutan pengujian dilakukan dalam bentuk tabulasi confusion matrix dimana *class* yang diprediksi akan ditampilkan pada bagian atas matriks dan *class* yang diamati pada bagian kiri (Nurfauzan & Fatimah, 2022). Setiap sel berisi angka yang menunjukkan berapa banyak kasus yang sebenarnya dari kelas yang diamati untuk diprediksi. Ilustrasi tersebut terlihat pada Tabel 3.

Tabel 3. Confusion Matrix

Nilai Prediksi	Nilai Aktual	
	TP	TN
	FP	FN

Keterangan:

TP = *True Positive*

TN = *True Negative*

FP = *False Positive*

FN = *False Negative*

Adapun rumus perhitungan *confusion matrix* jika dituliskan tampak terlihat sebagai berikut.

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

Accuracy berguna untuk mengukur kinerja sebuah metode.

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

Precision berguna untuk mengukur tingkat ketepatan antara informasi yang diminta oleh pengguna dengan jawaban yang diberikan oleh sistem.

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

Recall berguna untuk mengukur tingkat keberhasilan sistem dalam menemukan kembali sebuah informasi.

3. Hasil dan Pembahasan

Hasil

Jenis dan Sumber Data

Data yang digunakan dalam penelitian adalah data siswa SMK Karya mandiri Nusawungu Kelas XI yang berjumlah 80 data training dan 6 data Uji, dengan Prediksi Dapat dan Tidak Dapat. Adapun atribut data Training disimbolkan dengan Nilai Raport (V_1), Prestasi Akademik (W_1), Tanggungan Orang Tua (X_1), Gaji Orang Tua (Y_1), dan Keikutsertaan Organisasi (Z_1). Dan data Uji disimbolkan dengan Nilai Raport (V_2), Prestasi Akademik (W_2), Tanggungan Orang Tua (X_2), Gaji Orang Tua (Y_2), dan Keikutsertaan Organisasi (Z_2).

Metode Analisa

Setelah data dikumpulkan proses selanjutnya dilakukan sebagai berikut:

- 1) Menghitung jarak Euclidean pada data training pertama dan data Uji pertama, dimana $V_1=78$, $W_1=1$, $X_1=1$, $Y_1=505.000$, $Z_1=1$ dan $V_2=81$, $W_2=1$, $X_2=1$, $Y_2=1.000.000$, $Z_2=1$.

$$d = \sqrt{(v_2 - v_1)^2 + (w_2 - w_1)^2 + \dots + (z_2 - z_1)^2}$$

$$d = \sqrt{(81 - 78)^2 + (1 - 1)^2 + \dots + (1 - 1)^2}$$

$$d = \sqrt{(3)^2 + (0)^2 + (0)^2 + (495.000)^2 + (0)^2}$$

$$d = 3 + 495.000$$

$$d = 498.000$$

Perhitungan ini dilakukan kepada seluruh data training dengan data uji pertama.

- 2) Menghitung jarak Euclidean pada data training pertama dan data Uji ke 2, dimana $V_1=78$, $W_1=1$, $X_1=1$, $Y_1=505.000$, $Z_1=1$ dan $V_2=77$, $W_2=1$, $X_2=2$, $Y_2=600.000$, $Z_2=1$.

$$d = \sqrt{(v_2 - v_1)^2 + (w_2 - w_1)^2 + \dots + (z_2 - z_1)^2}$$

$$d = \sqrt{(77 - 78)^2 + (1 - 1)^2 + \dots + (1 - 1)^2}$$

$$d = \sqrt{(-1)^2 + (0)^2 + (1)^2 + (95.000)^2 + (0)^2}$$

$$d = 1 + 1 + 95.000$$

$$d = 97.000$$

Perhitungan ini dilakukan kepada seluruh data training dengan data uji ke 2, kemudian perhitungan ini dilakukan keseluruhan data training dengan data uji ke 2, begitu seterusnya sampai perhitungan data uji ke 6. Setelah Proses perhitungan dari data uji ke data training selesai, kemudian mengurutkan hasil jarak euclidean dari yang terkecil ke yang terbesar. Dengan parameter K (jumlah tetangga paling dekat) = 3, sehingga menghasilkan data seperti tampak pada Tabel 4 dibawah ini.

Tabel 4. Hasil Pengujian Algoritma KNN dengan K=3

Nama	Nilai Raport	Prestasi Akademik	Tanggungan Orang Tua	Gaji Orang Tua	Keikutsertaan Organisasi	Prediksi
Dzikry	81	1	1	1.000.000	1	Dapat
Erwin J	77	1	2	600.000	1	Dapat
Evan Z	81	1	2	1.400.000	1	Tidak Dapat
Fikri A	78	1	1	750.000	1	Tidak Dapat
Gilang P	79	1	1	850.000	1	Tidak Dapat
Iklil Abi	78	2	1	1.000.000	3	Dapat

Hasil prediksi pada Tabel 4 dilakukan pengujian performa dengan menggunakan metode *confussion Matrix* dengan menghitung nilai *accuracy*, *precision*, dan *recall* untuk memvisualisasikan kinerja dari algoritma KNN dengan K=3 yang digunakan. Berikut Pengujian *confussion matrix* yang dihasilkan tampak pada Tabel 5.

Tabel 5. Pengujian performa dengan Confussion Matrix

Actual	Predicted		Total
	Dapat	Tidak Dapat	
Dapat	TP = 3	FN = 0	3
Tidak Dapat	FP = 1	TN = 2	3
Total	4	2	6

Setelah nilai TP, FP, FN, dan TN diketahui, maka dapat dihitung nilai *Accuracy*, *Precision*, dan *Recall*. Berdasarkan persamaan (2), (3), dan (4):

Accuracy = 83,33%

Precision = 75 %

Recall = 100 %

Berdasarkan perhitungan di atas maka dapat disimpulkan bahwa dengan menguji hasil algoritma KNN dengan K=3 dari pengujian banyaknya data sesungguhnya dan data prediksi memperoleh hasil performa terbaik dengan nilai *accuracy* sebesar 83,33%, *precision* 75% dan *recall* 100% untuk prediksi penerima beasiswa berprestasi.

Pembahasan

penerapan algoritma K-Nearest Neighbor (K-NN) dalam memprediksi penerima beasiswa berprestasi di SMK Islam Pemalang. Algoritma K-NN dipilih karena kemampuannya yang telah terbukti dalam berbagai penelitian, seperti yang dilakukan oleh Deden Martia Nanda *et al.* (2022) dalam memprediksi curah hujan di Bandung dan oleh Dinata *et al.* (2020) untuk klasifikasi transportasi bus. Penerapan K-NN dalam penelitian ini melibatkan penggunaan atribut seperti nilai rapor, prestasi akademik, tanggungan orang tua, gaji orang tua, dan keikutsertaan organisasi. Proses normalisasi dilakukan untuk memastikan data yang digunakan konsisten, sebagaimana disarankan oleh Hasanah *et al.* (2022) dalam konteks penentuan buku berdasarkan peminatan. Algoritma K-NN bekerja dengan

menghitung jarak Euclidean antara data uji dan data pelatihan, seperti yang dijelaskan oleh Miftahuddin *et al.* (2020) yang membandingkan metode perhitungan jarak dalam metode lain. Dalam penelitian ini, penggunaan nilai K=3 dipilih untuk menghindari ambiguitas dalam klasifikasi, sejalan dengan temuan Pratama *et al.* (2021) yang menunjukkan efektivitas K-NN dalam klasifikasi penerima beasiswa. Hasil dari penelitian ini menunjukkan akurasi sebesar 83,33%, *precision* 75%, dan *recall* 100%, yang mengindikasikan bahwa K-NN merupakan alat yang efektif untuk klasifikasi ini.

Temuan ini sejalan dengan penelitian oleh Karepesina dan Zahrotun (2023), yang juga menemukan bahwa K-NN memiliki tingkat akurasi yang tinggi dalam penentuan penerima beasiswa. Selain itu, penelitian ini juga menggarisbawahi pentingnya pemilihan parameter yang tepat dalam algoritma K-NN, seperti yang diuraikan oleh Purwanto dan Nugroho (2023) dalam analisis mereka tentang kinerja algoritma K-NN dibandingkan dengan algoritma lain. Dengan memanfaatkan algoritma K-NN, penelitian ini berhasil meningkatkan efisiensi dan akurasi dalam seleksi penerima beasiswa berprestasi, memberikan kontribusi yang signifikan dalam konteks pendidikan di SMK Islam Pemalang.

4. Kesimpulan dan Saran

Berdasarkan penelitian yang dilakukan, ditarik kesimpulan bahwa penerapan Algoritma KNN dengan K=3 telah menghasilkan data calon penerima Beasiswa Berprestasi dengan mempertimbangkan atribut berupa Nilai Rapor, Prestasi Akademik, Tanggungan Orang Tua, Gaji Orang Tua dan Keikutsertaan Organisasi. Dari jumlah 6 data uji yang telah dihitung maka menghasilkan prediksi Penerima Beasiswa Prestasi untuk siswa SMK Karya Mandiri Nusawungu tingkat akurasi sebesar 83,33%. Penelitian ini fokus pada penerapan algoritma K-NN dengan K=3 yang dipengaruhi oleh Nilai Rapor, Prestasi Akademik, Tanggungan Orang Tua, Gaji Orang Tua, dan Keikutsaertaan Organisasi beserta implementasi sederhana dengan menggunakan pemrograman JAVA dalam menentukan hasil prediksi penerima beasiswa berprestasi. Sehingga saran untuk peneliti selanjutnya terhadap penelitian serupa dapat

menambahkan kriteria yang mempengaruhi terkait penerima beasiswa berprestasi serta melakukan perbandingan dengan nilai K yang beranekaragam disertai perbandingan algoritma yang lain untuk mengetahui tingkat keakuratan yang paling efektif dalam menentukan prediksi penerima beasiswa berprestasi.

5. Daftar Pustaka

- Nanda, D. M., Pudjiantoro, T. H., & Sabrina, P. N. (2022, April). Metode K-Nearest Neighbor (KNN) Dalam Memprediksi Curah Hujan di Kota Bandung. In *Prosiding Seminar Nasional Teknik Elektro, Sistem Informasi, dan Teknik Informatika (SNESTIK)* (Vol. 1, No. 1, pp. 387-393).
<https://doi.org/10.31284/p.snestik.2022.2750>.
- Dinata, R. K., Akbar, H., & Hasdyna, N. (2020). Algoritma K-Nearest Neighbor dengan Euclidean Distance dan Manhattan Distance untuk Klasifikasi Transportasi Bus. *ILKOM Jurnal Ilmiah*, 12(2), 104-111.
<https://doi.org/10.33096/ilkom.v12i2.539.104-111>.
- Hasanah, F., Suprapti, T., Rahaningsih, N., & Ali, I. (2022). Implementasi algoritma K-Nearest Neighbor dalam menentukan buku berdasarkan peminatan. *Jurnal Accounting Information System (AIMS)*, 5(1), 102-111.
<https://doi.org/10.32627/aims.v5i1.467>.
- Cholil, S. R., Handayani, T., Prathivi, R., & Ardianita, T. (2021). Implementasi algoritma klasifikasi k-nearest neighbor (knn) untuk klasifikasi seleksi penerima beasiswa. *IJCIT (Indonesian Journal on Computer and Information Technology)*, 6(2), 118-127.
- Karepesina, F., & Zahrotun, L. (2023). Penerapan Data Mining Untuk Penentuan Penerima Beasiswa Dengan Metode K-Nearest Neighbor (K-NN). *Techno (Jurnal Fak. Tek. Univ. Muhammadiyah Purwokerto)*, 24(1), 1.
- Miftahuddin, Y., Umaroh, S., & Karim, F. R. (2020). Perbandingan metode perhitungan jarak Euclidean, Haversine, dan Manhattan dalam penentuan posisi karyawan. *Jurnal Tekno Insentif*, 14(2), 69-77.
<https://doi.org/10.36787/jti.v14i2.270>.
- Nuraeni, S., Syam, S. P. A., Wajdi, M. F., Firmansyah, B., & Malkan, M. (2023). Implementasi metode K-NN untuk menentukan jurusan siswa di SMAN 02 Manokwari. *G-Tech: Jurnal Teknologi Terapan*, 7(1), 89-95.
<https://doi.org/10.33379/gtech.v7i1.1905>.
- Nurfauzan, D., & Fatimah, T. (2022, September). Implementasi algoritma K-Nearest Neighbors Regression dalam memprediksi harga saham. In *Prosiding Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFIT)* (Vol. 1, No. 1, pp. 576-584).
- Pahrudin, P., & Harianto, K. (2022). Penerapan algoritma K-Nearest Neighbor untuk klasifikasi warga penerima bantuan sosial. *Building of Informatics, Technology and Science (BITS)*, 4(3), 1241-1245.
<https://doi.org/10.47065/bits.v4i3.2276>.
- Pratama, A., Ma'ruf, F. A., & Rinaldi, A. R. (2021). Klasifikasi Penerima Beasiswa Dengan Menggunakan Algoritma K Nearest Neighbor. *Jurnal Data Science & Informatika*, 1(1), 11-15.
- Purwanto, A., & Nugroho, H. W. (2023). Analisa perbandingan kinerja algoritma C4.5 dan algoritma K-Nearest Neighbors untuk klasifikasi penerima beasiswa. *Jurnal Teknoinfo*, 17(1), 236.
<https://doi.org/10.33365/jti.v17i1.2370>.
- Tundo, T., & Mahardika, F. (2023). Fuzzy Inference System Tsukamoto-Decision Tree C 4.5 in Predicting the Amount of Roof Tile Production in Kebumen. *JTAM (Jurnal Teori dan Aplikasi Matematika)*, 7(2), 533-544.
- Tundo, T., Wijonarko, P., & Raffiudin, M. (2023). The WASPAS method in determining BSM

recipients objectively. *IJID (International Journal on Informatics for Development)*, 12(1), 338–349. <https://doi.org/10.1442401/ijid.2023.4089>.

Widaningsih, S. (2022). Penerapan Data Mining untuk Memprediksi Siswa Berprestasi dengan Menggunakan Algoritma K Nearest Neighbor. *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, 9(3), 2598-2611.