

Implementation of K-NN Algorithm to classify the Scholarship Recipients of Aceh Carong at Universitas Malikussaleh

Riski Yanti^{1*}, Sujacka Retno², Balqis Yafis³

^{1,2} Universitas Malikussaleh, Indonesia

³ National Yang Ming Chiao Tung University, Taiwan

*Corresponding Author Email: [*riski.190170002@mhs.unimal.ac.id](mailto:riski.190170002@mhs.unimal.ac.id)

ABSTRAK

Received: 29 December 2023
Revised: 31 December 2023
Accepted: 31 December 2023
Available online: 1 January 2024

Kata Kunci:

K-Nearest Neighbors (KNN), Data Mining, klasifikasi, penerimaan beasiswa, Aceh Carong

Dalam upaya meningkatkan efisiensi proses seleksi beasiswa, penelitian ini bertujuan untuk mengimplementasikan algoritma K-Nearest Neighbors (K-NN) dalam klasifikasi penerimaan beasiswa. Metode penelitian melibatkan pengumpulan data penerimaan beasiswa dari beberapa tahun sebelumnya berdasarkan kriteria yang telah ditentukan seperti Pekerjaan Ayah, Pekerjaan Ibu, Pendapatan Orang Tua, Jumlah Orang Tua Bekerja, Pendidikan Terakhir Ayah, Dan Pendidikan Terakhir Ibu. Selanjutnya, algoritma K-NN diterapkan untuk mengklasifikasikan calon penerima beasiswa berdasarkan kesamaan profil mereka dengan mahasiswa yang telah menerima beasiswa sebelumnya. Hasil penelitian ini menunjukkan bahwa implementasi algoritma K-NN dalam klasifikasi penerimaan beasiswa di Universitas Malikussaleh Aceh Carong dapat meningkatkan akurasi seleksi. Hasil eksperimen nilai akurasi yang diperoleh menunjukkan bahwa menggunakan algoritma K-Nearest Neighbors dengan pendekatan Euclidean Distance dan nilai $K = 3$ menghasilkan tingkat keakuratan algoritma sebesar 87,55%. Dengan demikian, algoritma K-NN dapat menjadi metode yang berguna bagi penyeleksi beasiswa untuk mendukung pengambilan keputusan yang lebih tepat dan objektif.

ABSTRACT

Keywords:

K-Nearest Neighbors (KNN), Data Mining, classification, scholarship recipients, Aceh Carong

In an effort to increase the efficiency of the scholarship selection process, this research aims to implement the K-Nearest Neighbors (K-NN) algorithm in the classification of scholarship recipients. The research method involves collecting data on scholarship receipts from several previous years based on predetermined criteria such as father's job, mother's job, parent's income, number of parents working, father's last education, and mother's last education. Next, the K-NN algorithm is applied to classify prospective scholarship recipients based on the similarity of their profiles to students who have received previous scholarships. The results of this research indicate that the implementation of the K-NN algorithm in the classification of scholarship admissions at Malikussaleh Aceh Carong University can increase selection accuracy. The experimental results of the accuracy values obtained show that using the K-Nearest Neighbors algorithm with the Euclidean Distance approach and a value of $K = 3$ produces an algorithm accuracy level of 87.55%. Thus, the K-NN algorithm can be a useful method for scholarship selectors to support more precise and objective decision making.

1. INTRODUCTION

Dengan adanya era teknologi 4.0 sudah sangat mempermudah pekerjaan manusia, misalnya adanya teknologi data mining dimana teknologi ini termasuk teknologi yang memiliki banyak data. Data mining adalah suatu teknik yang dapat membantu menentukan suatu fenomena untuk suatu objek yang telah dipelajari dengan lebih mendalam memahami atau tidak memahami suatu kelas tertentu untuk objek tersebut melalui penggunaan big data. Data mining dapat membantu populasi tertentu lebih memahami lingkungan mereka. Dimana big data juga merupakan subset dari data mining.

Beasiswa adalah suatu bentuk dukungan yang diberikan kepada setiap sekolah atau instansi. Tujuannya adalah untuk membantu mengurangi beban finansial siswa selama masa sekolah, khususnya dalam hal biaya atau keuangan. Program ini berdasarkan ketentuan pasal 4 ayat (1) UU PPh/2000 yang

mendefinisikan penghasilan sebagai tambahan dana keuangan yang diterima atau diperoleh dari sumber baik dari dalam negeri maupun dari luar negeri, yang digunakan untuk kebutuhan atau aset yang termasuk dalam kewajiban membayar pajak.

Universitas Malikussaleh menyediakan beasiswa bagi mahasiswa berprestasi dan kurang mampu secara ekonomi yang bersumber dari beberapa jenis beasiswa, salah satunya yaitu beasiswa Aceh Carong. Beasiswa Aceh Carong adalah salah satu jenis beasiswa yang diterima oleh mahasiswa universitas untuk memberikan mereka motivasi yang lebih menarik untuk belajar. Namun, karena pemberian beasiswa tidak tepat dan banyak mahasiswa yang tidak dapat memanfaatkannya, mahasiswa tersebut akan kehilangan motivasi untuk terus belajar. Oleh sebab itu, proses seleksi untuk mendapatkan beasiswa harus benar-benar dilaksanakan secara cermat agar hasil yang didapatkan efektif dan efisien

serta tepat pada sasaran, Hanya Mahasiswa yang memenuhi syarat yang berhak menerima beasiswa tersebut seperti yang dipersyaratkan oleh pemerintah Aceh. Saat ini proses seleksi Beasiswa Aceh carong di universitas malikussaleh masih dilakukan secara manual sehingga dalam pelaksanaan membutuhkan waktu yang lama serta tingkat efektifitas yang rendah.

Untuk memudahkan pencarian mahasiswa dalam proses perekrutan beasiswa ini, diperlukan data mining menggunakan Algoritma *K-Nearest Neighbours (K-NN)* pada mahasiswa yang telah dipetakan sebelumnya untuk mengidentifikasi mahasiswa yang layak menerima beasiswa tersebut.

Penelitian yang dilakukan oleh Sri Widaningsih, Sonia Yusuf dengan judul “Penerapan Data Mining Untuk Memprediksi siswa Berprestasi Dengan Menggunakan Algoritma *K Nearest Neighbor*” dimana tujuan penelitian ini adalah Aplikasi Data Mining ini ditujukan kepada bagian tata usaha untuk penentuan atau prediksi siswa berprestasi. Sistem ini memudahkan petugas dalam proses pendataan siswa. Aplikasi data mining ini memudahkan proses perhitungan nilai kedekatan antara siswa yang akan diprediksi dengan siswa yang sebelumnya. Aplikasi ini mampu menyajikan informasi mengenai kemampuan siswa yang nantinya bisa dijadikan sebagai acuan dalam menentukan siswa berprestasi yang sesuai dengan kemampuan serta prestasi akademik dan nonakademiknya.

2. RESEARCH METHODS

Prinsip dasar dari KNN adalah melakukan pengelompokan atau klasifikasi pada data baru yang tidak diketahui kelasnya berdasarkan jaraknya ke beberapa tetangga terdekat. Tetangga terdekat dalam hal ini merujuk pada objek latihan yang memiliki tingkat kemiripan terbesar atau tingkat ketidakmiripan terkecil dengan data baru. Beberapa cara digunakan untuk pengukuran jarak antara data baru (*data testing*) dengan data lama (*data training*), antara Lain Manhattan Distance (*City Block Distance*) dan Euclidean Distance. Pengukuran jarak yang paling sering digunakan adalah *euclidean distance*. Jarak euclidean distance didefinisikan seperti pada persamaan 1.

K-Nearest Neighbors adalah sebuah algoritma atau metode yang bertujuan untuk mengklasifikasikan suatu data berdasarkan pembelajaran yang dikenal sebagai train data sets. Algoritma KNN bekerja dengan cara mencari k tetangga terdekat dari data yang akan diklasifikasikan. Di sini, nilai k merupakan jumlah tetangga dekat yang akan dipertimbangkan pada proses pengklasifikasi. K-Nearest Neighbors (KNN) merupakan metode yang berguna dalam mengatasi masalah klasifikasi. Prinsip dasar dari KNN adalah melakukan pengelompokan atau klasifikasi pada data baru yang tidak diketahui kelasnya berdasarkan jaraknya ke beberapa tetangga terdekat. Tetangga terdekat dalam hal ini merujuk pada objek latihan yang memiliki tingkat kemiripan terbesar atau tingkat ketidakmiripan terkecil dengan data baru. Terdapat beberapa metode yang digunakan agar dapat dilakukan pengukuran jarak antara data baru (*data testing*) dengan data lama (*data training*), termasuk di antaranya *Manhattan Distance (City Block Distance)* dan *Euclidean Distance*. Di antara kedua metode tersebut, Euclidean Distance adalah yang paling umum digunakan dalam pengukuran jarak.

Perancangan sistem yakni rangkaian tindakan yang mendetail tentang cara sistem akan berfungsi. Tujuannya

untuk menciptakan suatu perangkat lunak disesuaikan dengan kebutuhan dari pengguna.

2.1 Skema Sistem

Skema sistem pada implementasi K-Nearest Neighbors untuk klasifikasi penerimaan beasiswa berprestasi pada Universitas Malikussaleh. Dimana dapat dilihat pada gambar dibawah ini:



Gambar 1. Skema Sistem

3. RESULT AND DISCUSSION

3.1 Perhitungan Manual Algoritma K-Nearest Neighbors

Membangun sistem yang dapat mengklasifikasi kelayakan Penerima Beasiswa Aceh Carong dengan algoritma *K-Nearest Neighbors*. Mengakses database untuk menampilkan tabel hasil perhitungan serta mengakses data untuk menampilkan hasil pengklasifikasian data berupa klasifikasi kelayakan yang didapatkan berdasarkan kriteria yang telah ditetapkan sebagai berikut:

Tabel 1. Tabel kriteria

Kode kriteria	Kriteria
X1	Pekerjaan Ayah
X2	Pekerjaan Ibu
X3	Pendapatan Orang Tua
X4	Jumlah Orang Tua Bekerja

X5	Pendidikan Terakhir Ayah
X6	Pendidikan Terakhir Ibu

SD Sederajat	2
SLTP / SMP Sederajat	3
SLTA / SMA Sederajat	4
D3 / D4 / S1	5

Selanjutnya menentukan nilai bobot dari setiap kriteria. Table 1 menunjukkan kriteria apa saja yang akan dibobotkan, dan dapat dilihat pada table di bawah ini:

Tabel 2. Pekerjaan Ayah

Value	Bobot
Almarhum	1
Tidak Bekerja	2
buruh	3
Wiraswasta/wirausaha	4
karyawan	5
PNS	6

Tabel 3. Pekerjaan Ibu

Value	Bobot
Almarhum	1
Tidak Bekerja	2
buruh	3
Wiraswasta/wirausaha	4
karyawan	5
PNS	6

Tabel 4. Pendapatan Orang Tua

Value	Bobot
< Rp.500.000	1
Rp.500.000 s.d. Rp.1.500.000	2
Rp.1.500.000 s.d. Rp.2.500.000	3
Rp.2.500.000 s.d. Rp.4.000.000	4
> Rp.4.000.000	5

Tabel 5. Jumlah Orang Tua Bekerja

Value	Bobot
1	1
2	2

Tabel 6. Pendidikan Terakhir Ayah

Value	Bobot
Tidak Sekolah	1
SD Sederajat	2
SLTP / SMP Sederajat	3
SLTA / SMA Sederajat	4
D3 / D4 / S1	5

Tabel 7. Pendidikan Terakhir Ibu

Value	Bobot
Tidak Sekolah	1

Tabel 8. Normalisasi

Value	Bobot
Layak	1
Tidak Layak	2

Tujuan dari algoritma K-Nearest Neighbors adalah untuk mengklasifikasikan objek baru berdasarkan data latih yang ada. Dalam proses ini, perhitungan jarak antara dua titik, yaitu titik data latih dan titik data uji, dapat menggunakan rumus Euclidean Distance seperti yang dijelaskan berikut ini:

$$d(x, y) = \sqrt{(x_i - y_i)^2 + (x_j - y_j)^2 + \dots + (x_n - y_n)^2} \quad (1)$$

Keterangan:

x_i = Sampel data

y_i = Data testing

i = Variabel/Kriteria data

$d(x,y)$ = Jarak

Setelah mendapat hasil dari perhitungan menggunakan algoritma k-nearest neighbors kemudian menghitung tingkat akurasi algoritma dengan menggunakan Confusion Matrix yaitu mencari akurasi, presisi dan recallnya.

Adapun rumus yang digunakan untuk confusion matrix adalah sebagai berikut:

1. Akurasi (*Accuracy*):

Akurasi mengukur seberapa sering model benar dalam melakukan klasifikasi.

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

2. Presisi (*Precision*):

Presisi mengukur seberapa banyak yang benar dari prediksi positif yang dibuat oleh model.

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

3. Recall (*Sensitivity atau True Positive Rate*):

Recall mengukur seberapa banyak data positif yang berhasil diidentifikasi oleh model.

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

Dimana didapat nilai prediksi $k=5$ yaitu dapat dilihat pada tabel dibawah ini:

Tabel 9. Prediteded class $K=5$

Kelas prediksi	Layak	Tidak Layak
Layak	35	7
Tidak Layak	3	30

Setelah mendapat hasil prediksi kemudian dihitung menggunakan confusion matrix dan kemudian didapat hasil sebagai berikut:

Tabel 10. Nilai akurasi, presisi, dan recall

Akurasi	presisi	recall	Nilai k
85%	86%	88%	K2
85%	89%	83%	K3
85%	89%	83%	K4
92%	83%	83%	K5

3.2 Implementasi Sistem

Perancangan sistem melibatkan serangkaian tindakan yang secara rinci menggambarkan cara pengimplementasian sistem merupakan proses yang dilaksanakan guna mewujudkan terselesaikannya perancangan sistem yang disetujui meliputi menguji, menginstal, dan mengawali dengan sistem yang terbaru atau yang sedang disempurnakan:

1. Halaman login



Gambar 2. Halaman Login

2. Halaman menu utama admin



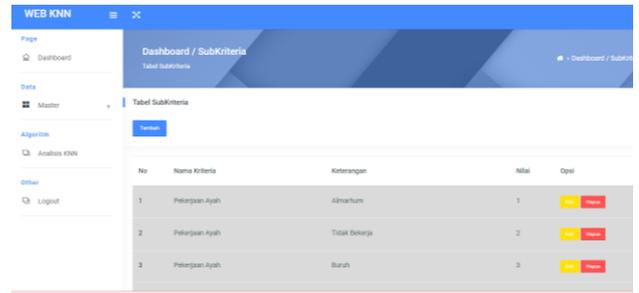
Gambar 3. Halaman menu utama admin

3. Halaman data kriteria



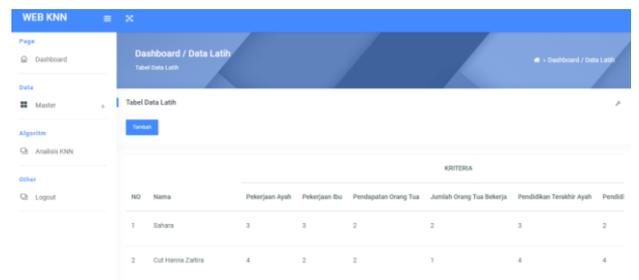
Gambar 4. Halaman data kriteria

4. Halaman data subkriteria



Gambar 5. Halaman data subkriteria

5. Halaman data latihan



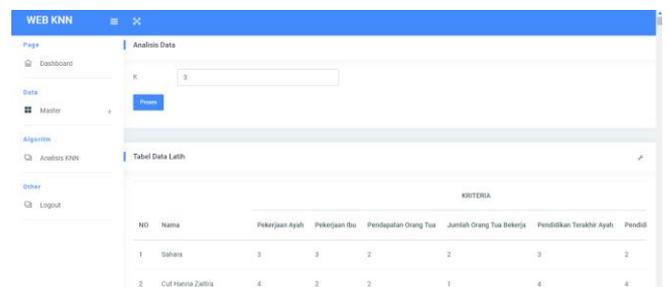
Gambar 6. Halaman data latihan

6. Halaman data uji



Gambar 7. Halaman data uji

7. Halaman Perhitungan Algoritma K-Nearest Neighbors



Gambar 8. Halaman Perhitungan Algoritma KNN

4. CONCLUSION

Sistem klasifikasi kelayakan penerima bantuan beasiswa Aceh Carong ini dihitung menggunakan algoritma K-Nearest Neighbors telah berhasil melakukan pengklasifikasian berupa layak dan tidak layak dengan menggunakan 6 (enam) kriteria yaitu pekerjaan ayah, pekerjaan ibu, jumlah pendapatan orang tua, jumlah orang tua bekerja, pendidikan terakhir ayah dan pendidikan terakhir ibu. Dengan menggunakan algoritma K-Nearest Neighbors dengan pendekatan menggunakan rumus Euclidean Distance dimana nilai $K = 3$ didapatkan hasil dari jumlah data latih 150 dan data uji 75 di dapatkan output klasifikasi berupa 39 data layak dan 36 data tidak layak dalam penerimaan beasiswa Aceh Carong Pada Universitas Malikussaleh.

Dalam penerapannya sistem klasifikasi kelayakan penerima bantuan beasiswa Aceh Carong maka diperoleh hasil nilai akurasi dimana menunjukkan bahwa menggunakan algoritma K-Nearest Neighbors dengan pendekatan menggunakan rumus Euclidean Distance dimana nilai $K = 5$ dan dengan perhitungan Confusion Matrix menghasilkan klasifikasi K-Nearest Neighbors nilai k terendah yaitu $k=1$ sebesar 84% dan $k=2$ sebesar 85%, $k=3$ sebesar 85%, $k=4$ sebesar 85% dan nilai k tertinggi dan yang paling baik yaitu $k=5$ sebesar 86%. Kesimpulan penelitian ini yaitu didapatkan hasil k terbaik yaitu $k=5$ dengan tingkat akurasi 86%.

REFERENCES

- [1] Arifin, Z. (2019). Penerapan Metode Knn (K-Nearest Neighbor) Dalam Sistem Pendukung Keputusan Penerimaan Kip (Kartu Indonesia Pintar) Di Desa Pandean Berbasis Web Dan Mysql. *NJCA (Nusantara Journal of Computers and Its Applications)*, 4(1). <https://doi.org/10.36564/njca.v4i1.101>
- [2] Cholil, S. R., Handayani, T., Prathivi, R., & Ardianita, T. (2021). *IJCIT (Indonesian Journal on Computer and Information Technology) Implementasi Algoritma Klasifikasi K-Nearest Neighbor (KNN) Untuk Klasifikasi Seleksi Penerima Beasiswa. IJCIT (Indonesian Journal on Computer and Information Technology)*, 6(2), 118–127.
- [3] Jogiyanto. (2018). Desain Algoritma Operasi Perkalian Matriks Menggunakan Metode Flowchart. *Jurnal Teknik Komputer Amik Bsi*, 1(1), 144–151.
- [4] Khasanah, F. N., & Rofiah, S. (2019). Metode Simple Additive Weighting Dalam Menentukan Rekomendasi Penerima Beasiswa. 6(1), 65–74.
- [5] 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>
- [6] Putriani, J. D., & Hudaidah, H. (2021). Penerapan Pendidikan Indonesia Di Era Revolusi Industri 4.0. *Edukatif: Jurnal Ilmu Pendidikan*, 3(3), 830–838. <https://edukatif.org/index.php/edukatif/article/view/407>
- [7] Razi, A. (2022). Klasifikasi Penerima Beasiswa Aceh Carong (Aceh Pintar) Di Universitas Malikussaleh Menggunakan Algoritma Knn (K-Nearest Neighbors). *Jurnal Tika*, 7(1), 79–84. <https://doi.org/10.51179/tika.v7i1.1116>
- [8] Retno, S., & Hasdyna, N. (2022). Profile Matching in Government Scholarship Acceptance System for Student in Aceh Utara. *Journal of Informatics and Telecommunication Engineering*, 5(2), 268–275. <https://doi.org/10.31289/jite.v5i2.6031>
- [9] Sulistyono, D., & Winiarti, S. (2015). Penentuan Beasiswa Siswa Kurang Mampu. *Jurnal Informatika*, 9(1), 965–974.
- [10] Sumiah, A., & Mirantika, N. (2020). Perbandingan Metode K-Nearest Neighbor dan Naive Bayes untuk Rekomendasi Penentuan Mahasiswa Penerima Beasiswa pada Universitas Kuningan. *Buffer Informatika*, 6(1), 1–10.
- [11] Widaningsih, S., Yusuf, S., Informatika, J. T., Teknik, F., & Suryakencana, U. (2022). Penerapan Data Mining Untuk Memprediksi Siswa Berprestasi Dengan Menggunakan Algoritma K Nearest Neighbor. 9(3), 2598–2611. Yuli Mardi. (2019). Data Mining: Klasifikasi Menggunakan Algoritma C4.5 Data mining merupakan bagian dari tahapan proses Knowledge Discovery in Database (KDD). *Jurnal Edik Informatika. Jurnal Edik Informatika*, 2(2), 213–219