

Klasifikasi Calon Mahasiswa Penerima Beasiswa Baznas Kota Pariaman Menggunakan Metode *K-Nearest Neighbor*

Bestia Dhelfiariani¹, Dewi Murni²

^{1,2}Departemen Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Padang (UNP)

Article Info

Article history:

Received October 27, 2023
Revised November 03, 2023
Accepted June 20, 2024

Keywords:

Baznas
Classification
K-Nearest Neighbor

Kata Kunci:

Baznas
Klasifikasi
K-Nearest Neighbor

ABSTRACT

Scholarship is a financial aid provided to students and scholars who face economic constraints. One of the organizations that offers scholarships is the National Amil Zakat Agency (Baznas). The selection process for Baznas scholarship recipients requires a data classification system to assist in the decision-making process. This research aims to analyze the classification results and determine the most optimal level of accuracy. The method used in this research is the k-nearest neighbor. The classification results using $k=22$ on 41 testing data obtained 37 accurately classified data and 4 misclassified data. Out of the 37 accurately classified data, 17 are students receiving Baznas scholarships, and 20 are students not receiving Baznas scholarships, with the most optimal accuracy rate of 90% (excellent classification).

ABSTRAK

Beasiswa merupakan bantuan biaya pendidikan yang diberikan kepada siswa maupun mahasiswa yang terkendala secara ekonomi. Salah satu pihak yang memberikan beasiswa adalah Badan Amil Zakat Nasional (Baznas). Proses penyeleksian penerima beasiswa Baznas membutuhkan sistem pengklasifikasian data untuk membantu proses pengambilan keputusan penerima beasiswa Baznas. Penelitian ini bertujuan menganalisis hasil klasifikasi dan mengetahui tingkat akurasi paling optimal. Metode yang digunakan pada penelitian ini yaitu *k-nearest neighbor*. Hasil klasifikasi menggunakan $k = 22$ pada 41 data *testing* diperoleh 37 data tepat klasifikasi dan terdapat 4 data tidak tepat klasifikasi. Dari 37 data tepat klasifikasi terdapat 17 data mahasiswa menerima beasiswa dan 20 data mahasiswa tidak menerima beasiswa Baznas dengan tingkat akurasi paling optimal 90% (*excellent classification*).

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



Penulis pertama

(Bestia Dhelfiariani)

Departemen Matematika, Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Negeri Padang, Jl.Prof.Dr. Hamka, Air Tawar barat, Padang Utara, Padang, 25171 Padang, Sumatera Barat
Email: bestdhlfm18@gmail.com

1. PENDAHULUAN

Beasiswa merupakan pemberian dana kepada siswa dalam bentuk pemberian gratis atau pinjaman tanpa bunga, untuk mendukung mereka dalam menyelesaikan pendidikannya. Tingginya biaya pendidikan, terutama pada jenjang perguruan tinggi, telah menjadi perhatian utama dalam upaya meningkatkan kualitas pendidikan dalam suatu negara. Permasalahan ini tercermin dalam data Badan Pusat Statistik (BPS) Kota Pariaman tahun 2022 menunjukkan bahwa sebanyak 22,66% dari penduduk yang menamatkan pendidikan Sekolah Menengah Atas, hanya 13,15% yang melanjutkan pendidikan ke Perguruan Tinggi (BPS, 2023). Satu-satunya lembaga resmi yang ditunjuk oleh pemerintah berdasarkan keputusan Presiden No. 8 tahun 2001 adalah Badan Amil Zakat Nasional (Baznas), yang bertugas mengumpulkan dan mendistribusikan zakat, infak, dan sedekah (ZIS) dalam skala nasional. Baznas berbagi tugas dengan pemerintah untuk mengawasi administrasi zakat sesuai dengan prinsip-prinsip hukum Islam yaitu keandalan, manfaat, keadilan, kejelasan hukum, integrasi, dan akuntabilitas [1]. Salah satu hibah yang membantu masalah pengeluaran biaya pendidikan yang tinggi adalah beasiswa Baznas.

Berdasarkan permasalahan pendidikan yang telah diidentifikasi yaitu rendahnya jumlah penduduk Kota Pariaman yang melanjutkan pendidikan ke Perguruan Tinggi, maka langkah selanjutnya yang perlu diambil adalah memastikan bahwa program beasiswa yang dikelola oleh Baznas Kota Pariaman benar-benar efektif dalam menjangkau mereka yang perlu bantuan. Proses seleksi beasiswa Baznas harus memenuhi persyaratan sebagai kelengkapan administrasi beasiswa tersebut. Persyaratan beasiswa yang cukup banyak mengakibatkan pihak Baznas kesulitan dalam melakukan proses seleksi beasiswa agar penerimaan beasiswa Baznas tersebut bisa tersalurkan tepat sasaran. Untuk mengurangi kesalahan ini, diperlukan pengklasifikasian dalam membantu pengambilan keputusan mahasiswa mana yang berhak mendapatkan bantuan beasiswa dengan menggunakan metode klasifikasi.

Klasifikasi adalah bentuk analisis data yang membantu pengguna dalam menetapkan kelas untuk sampel yang ingin dikategorikan. KNN ini merupakan metode dalam pembelajaran terawasi yang berupaya untuk mengidentifikasi hubungan antara atribut yang diberikan sebagai input dan atribut target yang diinginkan. Tujuan utama dari klasifikasi adalah untuk meningkatkan validitas dan keandalan hasil yang diperoleh dari data [2].

Metode *k-nearest neighbor* (KNN), salah satu teknik klasifikasi data dasar yang menggunakan perhitungan jarak terpendek sebagai metrik untuk mengkategorikan contoh baru berdasarkan tingkat kemiripannya, merupakan teknik klasifikasi yang digunakan dalam penelitian ini [3]. Salah satu teknik klasifikasi data yang sederhana untuk diterapkan pada dataset kecil adalah pendekatan KNN, namun metode KNN memiliki kekurangan dan menjadi tidak efisien ketika memproses dataset yang besar dan rumit [4].

Penelitian sebelumnya mengenai klasifikasi penerimaan beasiswa oleh Jaman [5], Zerlinda [6], dan Pratama [7] dengan memakai metode *k-nearest neighbor*. Selanjutnya, Rachma [8] juga melakukan studi tentang klasifikasi tingkat kedalaman kemiskinan dengan memakai metode *k-nearest neighbor*. Data indeks kedalaman kemiskinan pada penelitian tersebut diklasifikasikan menggunakan pendekatan KNN dan menghasilkan nilai akurasi terbaik sebesar 76,67%.

2. METODE PENELITIAN

Salah satu nilai parameter k merepresentasikan jumlah tetangga terdekat dalam pendekatan *k-nearest neighbor*. Metode ini mencari kelompok k objek pada data *training* yang paling kecil dengan objek dalam data baru atau data *testing* [9]. Penentuan variabel yang dipilih berdampak pada akurasi dari hasil klasifikasi. Data *testing* diproses menggunakan pendekatan *k-nearest neighbor* yang mengambil jarak terdekat kemudian digunakan untuk membandingkannya dengan data *training*. Untuk mengukur tingkat



akurasi data pada metode KNN digunakan *confusion matrix* yang melibatkan tingkat akurasi yang lebih baik [10].

Jenis penelitian ini merupakan penelitian terapan menggunakan data sekunder yang secara tidak langsung dikumpulkan atau melalui perantara dan didokumentasikan oleh pihak lain. Sumber data yang digunakan sebanyak 410 pendaftar beasiswa Baznas pada tahun 2022 dengan 369 di antaranya dibagi menjadi *training* dan 41 menjadi *testing*. Terdapat beberapa proses klasifikasi data penelitian ini yang mengacu pada [11]:

2.1 Menentukan Nilai Parameter K .

K-nearest neighbor digunakan dalam penelitian ini untuk mengklasifikasikan data dengan terlebih dahulu mencari nilai parameter k . Pengambilan nilai k yang besar secara umum akan mengurangi dampak *noise* pada klasifikasi, tetapi juga akan menyebabkan batas antara setiap klasifikasi menjadi lebih kabur [12].

Pertimbangan lain terdapat pada penggunaan angka k ganjil atau genap. Terdapat kemungkinan bahwa suara dari kedua klasifikasi akan cocok untuk k genap dengan jumlah klasifikasi genap. Akan lebih sederhana jika menggunakan k ganjil dengan jumlah kategori yang genap akan memudahkan karena dipastikan kedua kelas yang menerima suara yang sama.

2.2 Melakukan Normalisasi Data.

Salah satu prosedur pada transformasi data yang digunakan dalam data mining menggunakan normalisasi data yang menskalakan atribut numerik dalam rentang yang lebih kecil [13]. Proses normalisasi data melibatkan penskalaan properti data agar sesuai dengan interval [0,1].

$$x_{i'} = \frac{x_i - x_{min}}{x_{max} - x_{min}}. \quad (1)$$

Dengan:

$x_{i'}$ = data normalisasi ke- i

x_i = data actual ke- i

x_{max} = nilai maksimal dari data x

x_{min} = nilai minimal dari data x

2.3 Melakukan Data *Splitting*

Data *splitting* merupakan proses pemisahan data memerlukan pemisahan dataset menjadi dua bagian diantaranya data *training* dan *testing* yang dipilih secara random. Tujuan pada data *training* adalah untuk mengajarkan sistem untuk mengidentifikasi karakteristik yang dimiliki oleh mahasiswa yang mendapatkan beasiswa dan yang tidak mendapatkan beasiswa. Data *testing* digunakan mengevaluasi kinerja model klasifikasi yang dikembangkan dengan membandingkan kelas sebenarnya pada setiap set data *testing* dengan temuan hasil klasifikasi [14].

2.4 Menghitung Jarak Menggunakan Rumus Jarak *Euclidean*.

Pengukuran dua titik dalam satu dimensi pada jarak *Euclidean* menampilkan hasil yang mirip dengan perhitungan Pythagoras [15].

$$d(x, y) = \sqrt{\sum_{i=1}^p (x_i - y_i)^2}. \quad (2)$$

Dimana:

$d(x, y)$ = jarak antara objek x ke objek y

x_i = data *training*

y_i = data *testing*

p = banyaknya variabel yang diamati n

i = atribut individu antara 1 sampai dengan n

Hasil perhitungan jarak *Euclidean* selanjutnya diurutkan, dimulai dari jarak terdekat berdasarkan nilai yang paling kecil ke paling besar, kemudian ditentukan jarak terdekat sebanyak nilai k . Selanjutnya untuk

mendapatkan hasil klasifikasi, maka peneliti mengambil jumlah kelas maksimum menjadi kelas pada setiap objek data *testing*.

Setelah itu, peneliti menetapkan kelas terbanyak sebagai kelas dari masing-masing objek data *testing* untuk mendapatkan hasil dari klasifikasi.

2.5 Menghitung Evaluasi Akurasi Menggunakan *Confusion Matrix*.

Perhitungan evaluasi akurasi menggunakan tabel khusus yang disebut *confusion matrix* yang digunakan untuk menunjukkan seberapa baik kinerja algoritma. Setiap baris dalam matriks menggambarkan suatu contoh dalam kelas sebenarnya, sedangkan setiap kolom menggambarkan contoh dalam kelas prediksi, atau sebaliknya [16].

2.5.1 Akurasi

Akurasi merupakan presentase jumlah data yang dilakukan pada klasifikasi atau prediksi secara benar.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% . \quad (3)$$

2.5.2 Presisi

Presisi adalah kemampuan untuk mengulangi pengukuran dalam keadaan yang sama dan memperoleh hasil yang sama setiap kali.

$$Precision = \frac{TP}{TP + FP} \times 100\% . \quad (4)$$

2.5.3 Recall

Recall merupakan nilai yang dapat mengukur hasil berapa presentase data yang terklasifikasikan dengan benar.

$$Recall = \frac{TP}{TP + FN} \times 100\% . \quad (5)$$

Dimana:

True Positive (TP) = Jumlah kelas positif yang diklasifikasikan sebagai positif.

True Negative (TN) = Jumlah kelas negatif yang diklasifikasikan sebagai negatif.

False Negative (FN) = Jumlah kelas positif yang diklasifikasikan sebagai negatif.

False Positive (FP) = Jumlah kelas negatif yang diklasifikasikan sebagai positif.

Berdasarkan hasil pengujian akurasi, untuk mengekspresikan *confusion matrix* maka terdapat beberapa tingkat nilai diagnosa yaitu [17]:

Tabel 1. Tingkat Nilai Diagnosa

Nilai Akurasi	Keterangan
0,50 – 0,60	<i>Failure</i>
0,60 – 0,70	<i>Poor Classification</i>
0,70 – 0,80	<i>Fair Classification</i>
0,80 – 0,90	<i>Good Classification</i>
0,90 – 1,00	<i>Excellent Classification</i>

Berdasarkan Tabel 1 menunjukkan bahwa hasil perhitungan nilai akurasi pada *confusion matrix* akan menunjukkan hasil akurasi paling optimal apabila nilai akurasi berkisar pada 0,90 – 1,0.

2.6 Menentukan *K* Optimal

Pemilihan *k* dalam metode *k-nearest neighbor* memiliki peran utama dalam menentukan akurasi klasifikasi. Metode KNN akan menghasilkan tingkat akurasi yang optimal apabila nilai *k* yang dipilih memberikan hasil akurasi yang tinggi [18]. Penentuan nilai akurasi pada *confusion matrix* dapat dihitung menggunakan persamaan (3) dengan tingkat nilai diagnosa terlihat pada Tabel 1.



3. HASIL DAN PEMBAHASAN

3.1 Statistik Deskriptif

Pada penelitian ini terdapat variabel *dependen* (terikat) yang merupakan status penerimaan beasiswa (0: tidak diterima dan 1: diterima) serta variabel *independen* (bebas) yang terdiri dari IPK (X_1), masa studi (X_2), Perguruan Tinggi (X_3), pendapatan orang tua (X_4), dan jumlah tanggungan (X_5). Setelah selesainya pengumpulan data maka tahapan selanjutnya yaitu mendeskripsikan data pada masing-masing variabel dan mencari karakteristik data, dalam menentukan karakteristik data akan dicari nilai rata-rata, nilai maksimum dan minimum dengan menggunakan *software Google Colabs* dengan perolehan data yaitu:

Tabel 2. Analisis Deskriptif Variabel *Independen* (Bebas)

Variabel	N	Mean	Std, Deviasi	Minimum	Maximum
X_1	410	3,25	0,13	2,79	3,70
X_2	410	4,59	2,14	1,00	7,00
X_3	410	0,34	0,47	0,00	1,00
X_4	410	1.043.927	534.060,6	400.000	2.750.000
X_5	410	3,15	0,97	2,00	6,00

Berdasarkan Tabel 2, hasil nilai rata-rata X_1 sebesar 3,25 dengan nilai maksimum 3,70 dan minimum 2,79. Serta X_4 memiliki rentang data yang besar dengan nilai maksimum sebesar 2.750.000 dan minimum 1.043.927. Dikarenakan rentang data variabel yang sangat besar antara data akan mempengaruhi hasil pengolahan data sehingga harus dilakukan normalisasi data terlebih dahulu sebelum mengolah data lebih lanjut.

3.2 Normalisasi Data

Dalam studi ini, normalisasi data pada variabel *independen* (variabel bebas) untuk data *training* dan *testing* dengan rentang nilai [0,1].

Tabel 3. Normalisasi Data *Training*

No	X_1	X_2	X_3	X_4	X_5	Y
1	0,462	0,333	0,000	0,617	0,250	0
2	0,626	0,333	0,000	0,574	0,000	0
3	0,462	0,333	0,000	0,213	0,000	0
4	0,560	0,667	0,000	0,213	0,500	1
5	0,473	1,000	0,000	0,149	0,500	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮
369	0,341	0,000	0,000	0,255	0,250	0

Tabel 3 menampilkan hasil normalisasi data pada data *training* menggunakan Persamaan (1) dengan memberikan bobot yang sama pada setiap variabel *independen* (bebas).

Tabel 4. Normalisasi Data *Testing*

No	X_1	X_2	X_3	X_4	X_5	Y
1	0,343	0,333	0,000	0,250	0,500	1
2	0,164	0,000	0,000	0,375	0,250	0
3	0,328	0,000	1,000	0,100	0,250	1
4	0,060	0,333	0,000	0,650	0,250	0
5	0,328	1,000	0,000	0,125	0,250	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮
41	1,000	1,000	0,000	0,088	0,500	1

Tabel 4 menampilkan hasil normalisasi data pada data *testing* menggunakan Persamaan (1) dengan memberikan bobot yang sama pada setiap variabel *independen* (bebas).

3.3 Perhitungan Jarak *Euclidean*

Jarak *Euclidean* pada data *training* dan *testing* dihitung dengan menggunakan Persamaan (1). Salah satu perhitungan pada jarak *Euclidean* antara data *testing* pertama dengan data *training* ke-1 hingga data *training* ke-369 sebagai berikut.

$$d(1,1) = \sqrt{(0,46 - 0,34)^2 + (0,33 - 0,33)^2 + (0,00 - 0,00)^2 + (0,62 - 0,25)^2 + (0,25 - 0,50)^2} = 0,460$$

$$d(2,1) = \sqrt{(0,63 - 0,34)^2 + (0,33 - 0,33)^2 + (0,00 - 0,00)^2 + (0,57 - 0,25)^2 + (0,00 - 0,50)^2} = 0,660$$

$$d(3,1) = \sqrt{(0,46 - 0,34)^2 + (0,33 - 0,33)^2 + (0,00 - 0,00)^2 + (0,21 - 0,25)^2 + (0,00 - 0,50)^2} = 0,515$$

⋮

$$d(369,1) = \sqrt{(0,34 - 0,34)^2 + (0,00 - 0,33)^2 + (0,00 - 0,00)^2 + (0,26 - 0,25)^2 + (0,25 - 0,50)^2} = 0,417.$$

Hasil perhitungan jarak *Euclidean* yang dihitung menggunakan Persamaan (2) bertujuan untuk menentukan sejauh mana data baru (data *testing*) mendekati data yang telah ada sebelumnya (data *training*).

3.4 Mengurutkan Hasil Perhitungan Jarak

Hasil perhitungan jarak untuk menentukan urutan terkecil hingga terbesar yang ditampilkan pada Tabel 4.

Tabel 5. Ranking Jarak *Euclidean* antar Data *Training* dan Data *Testing*

Data <i>Testing</i>	Data <i>Training</i> (1-369)	Ranking
1	0,460	49
	0,660	122
	0,515	75
	⋮	⋮
	0,417	30
2	0,508	34
	0,653	66
	0,537	46
	⋮	⋮
	0,213	1
⋮	⋮	⋮
40	1,077	131
	1,108	142
	1,117	149
	⋮	⋮
	1,153	171
41	1,038	194
	1,035	192
	1,000	178
	⋮	⋮
	1,235	264

Hasil perhitungan jarak *Euclidean* pada Tabel 5 yang sudah diurutkan berdasarkan dari jarak terdekat dari masing-masing data *testing* dan *training*.

3.5 Confusion Matrix

Berdasarkan hasil klasifikasi ditentukan nilai $k = 1$, kemudian dilakukan evaluasi hasil klasifikasi metode KNN menggunakan *confusion matrix*. Berikut hasil perhitungan *confusion matrix* pada data *testing* sebagai berikut [19]:

Tabel 6. *Confusion Matrix*

Nilai Prediksi	Nilai Asli	
	Diterima	Tidak Diterima
Diterima	17	2
Tidak Diterima	4	18

Hasil perhitungan *confusion matrix* pada Tabel 6 menunjukkan bahwa dari 41 data *testing* terdapat 17 mahasiswa diklasifikasikan diterima dan 18 mahasiswa diklasifikasikan tidak diterima. Terdapat 2 mahasiswa tidak diterima tetapi diklasifikasikan diterima dan 4 mahasiswa tidak diterima tetapi diklasifikasikan diterima. Hasil perhitungan akurasi menggunakan Persamaan (3) menghasilkan tingkat akurasi sebesar 0,85. Selanjutnya, untuk mencari hasil perhitungan *presisi* dan *recall* menggunakan *software Google Colabs*.

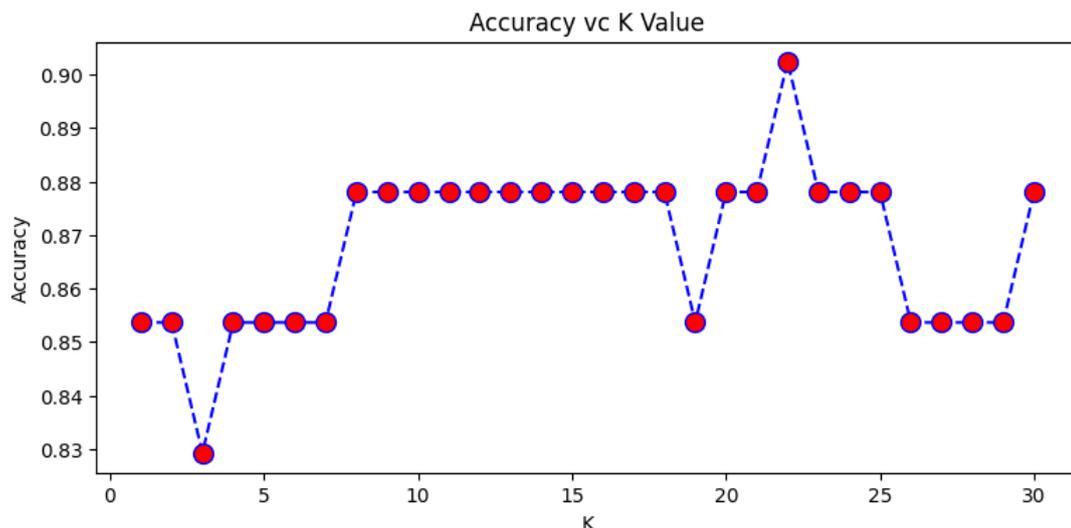
Tabel 7. Tingkat Akurasi, Presisi dan, Recall

	Akurasi	Presisi	Recall
Diterima	0,85	0,89	0,81
Tidak Diterima		0,82	0,90

Berdasarkan Tabel 7 menunjukkan tingkat *presisi* dan *recall* yang dihasilkan menggunakan rumus pada Persamaan (4) dan (5). *Presisi* “Diterima” ini mengukur sejauh mana dari semua prediksi positif yang benar-benar positif dalam kategori “Diterima”, sedangkan *recall* “Diterima” ini mengukur sejauh mana model mampu mengidentifikasi semua kasus positif yang sebenarnya dalam kategori “Diterima”

3.6 Menentukan K Optimal

Untuk menentukan nilai k optimal, dapat dilakukan dengan melihat nilai *akurasi*, *presisi* dan *recall* dengan bantuan *software Google Colabs*. Adapun nilai akurasi untuk nilai $k = 1, 2, \dots, 30$ dengan hasil percobaan yang dapat ditemukan pada Gambar 1.



Gambar 1. Output Percobaan Model K Terbaik

Berdasarkan Gambar 1, diketahui bahwa nilai *akurasi* yang paling rendah sebesar 0,83 terdapat pada $k = 3$ dan *akurasi* paling tinggi dihasilkan sebesar 0,90, sehingga dapat disimpulkan bahwa klasifikasi status mahasiswa yang menerima beasiswa Baznas pada metode *k-nearest neighbor* akan membuat akurasi yang optimal apabila menggunakan $k = 22$.

4. KESIMPULAN

Hasil studi yang dicobakan menggunakan metode KNN digunakan dalam menentukan klasifikasi calon mahasiswa penerima beasiswa Baznas pada 41 data *testing* dengan $k = 22$ diperoleh 37 data tepat klasifikasi dan terdapat 4 data tidak tepat klasifikasi. Pada 37 data tepat klasifikasi terdapat 17 data mahasiswa menerima dan 20 data mahasiswa tidak menerima beasiswa Baznas. Hasil klasifikasi diperoleh kesimpulan nilai akurasi paling optimal sebesar 90,0% dengan nilai $k = 22$.

REFERENSI

- [1] (Anggelia, O. (2020). Analisis Pengelolaan Program Beasiswa Satu Keluarga Satu Sarjana (Skss) Di Badan Amil Zakat Nasional (Baznas) Kabupaten Dharmasraya Dalam Mensejahterakan Pendidikan Mahasiswa Dharmasraya.
- [2] Pandie, E. S. Y. (2012). *Sistem Informasi Pengambilan Keputusan Pengajuan Kredit Dengan Algoritma K-Nearest Neighbour (Studi Kasus: Koperasi Simpan Pinjam)*.
- [3] Khairi, A., Fais Ghozali, A., & Darul Nur Hidayah, A. (n.d.). *Implementasi K-Nearest Neighbor (KNN) Untuk Klasifikasi Masyarakat Pra Sejahtera Desa Sapikerep Kecamatan Sukapura*. 2(3), 319–323.
- [4] 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.
- [5] Jaman, J. H., & Fahlevi, S. A. (2019). Klasifikasi Calon Mahasiswa Bidikmisi dengan Algoritma K-Nearest Neighbor. *Annual Research Seminar (ARS)*, 5(1), 1–5.
- [6] Zerlinda, H. N., Slamet, I., & Zukhronah, E. (2019). 15. Klasifikasi Calon Penerima Bidikmisi Dengan Menggunakan Algoritma K-Nearest Neighbor. *Seminar Nasional Penelitian Pendidikan Matematika (SNP2M)*, 88–93.
- [7] Pratama, A., Ma'ruf, F. A., Iin, & Rinaldi, A. R. (2021). Klasifikasi Penerima Beasiswa Dengan Menggunakan Algoritma K Nearest Neighbor. *Jurnal Data Science & Informatika*, 1(1), 11–15.
- [8] Rachma, C. A. (2022). *Implementasi Algoritma K-Nearest Neighbor Dalam Penentuan Klasifikasi Tingkat Kedalaman Kemiskinan Provinsi Jawa Timur*.
- [9] Leidiyana, H. (2013). Penerapan Algoritma K-Nearest Neighbor Untuk Penentuan Resiko Kredit Kepemilikan Kendaraan Bermotor. In *Jurnal Penelitian Ilmu Komputer, System Embedded & Logic* (Vol. 1, Issue 1).
- [10] Sumarlin. (2015). Implementasi Algoritma K-Nearest Neighbor Sebagai Pendukung Keputusan Klasifikasi Penerima Beasiswa. *Prosiding Semmau*, 16–23.
- [11] Afifanto, R. (2022). *Algoritma K-Nearest Neighbour(K-NN) Dengan Euclidean dan Manhattan Distance untuk Klasifikasi Tingkat Kesejahteraan di Kabupaten Temanggung*.
- [12] Ramaulidyah, F. N. (2020). *Perbandingan Metode Klasifikasi Naive Bayes Dank-Nearest Neighbor Pada Data Status Pembayaran Pajak Pertambahan Nilai di Kantor Pelayanan Pajak Pratama Samarinda Ulu*.
- [13] Imron, M. A. (2020). *Peningkatan Akurasi Algoritmak-Nearestneighbormenggunakan Normalisasi Z-Score danparticle Swarm Optimizationuntuk Prediksi Customer Churn*.
- [14] Azizah, N., Firdaus, M. R., Suyaningsih, R., Indrayatna, F., & Padjadjaran, U. (2023). *Penerapan Algoritma Klasifikasi K-Nearest Neighbor pada Penyakit Diabetes*. <http://prosiding.snsa.statistics.unpad.ac.id>.
- [15] Mustofa, Z., & Suasana, I. S. (2018). Algoritma Clustering K-Medoids Pada E-Government Bidang Information and Communication Technology dalam Penentuan Status EDGI. *Jurnal Teknologi Informasi Dan Komunikasi*, 9(1), 1–10.
- [16] Baskoro, Sriyanto, & Rini, L. S. (2021). Prediksi Penerima Beasiswa dengan Menggunakan Teknik Data Mining di Universitas Muhammadiyah Pringsewu. *Seminar Nasional Hasil Penelitian Dan Pengabdian Masyarakat 2021*, 87–94.
- [17] Rosandy, T. (2016). Perbandingan Metode Naive Bayes Classifier Dengan Metode Decision Tree (C4.5) Untuk Menganalisa Kelancaran Pembiayaan (Study Kasus : KSPPS / BMT AL-FADHILA). *Jurnal TIM Darmajaya*, 02(01), 52–62.
- [18] Hidayatul Nufus, A. (2023). Klasifikasi Masyarakat Penerima BPNT Program Sembako 2021 di Kelurahan Tiakar dengan Menggunakan Metode KNN Classifier. *Journal Of Mathematics UNP*, 8(2), 51–61.
- [19] Rahmadanti, S. (2020). *Implementasi Metode K-Nearest Neighbor dalam Menentukan Kualitas Massa Batuan*.