

Penerapan Data Mining dalam Mendukung Sistem Penunjang Keputusan Penerima Beasiswa di Universitas: Literature Review

Muhammad Dafa Alviansyah*¹

¹Informatika, Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Sumatera Utara, Indonesia
Email: 1muhammad_dafa@students.usu.ac.id

Abstrak

Pengelolaan data yang besar untuk pengambilan keputusan dalam proses pemberian beasiswa di universitas menjadi tantangan tersendiri, terutama untuk memastikan bahwa keputusan yang diambil efektif dan efisien. Penelitian ini dilakukan untuk memberikan informasi terkait penerapan data mining dalam mendukung sistem penunjang keputusan (Decision Support System) sebagai solusi untuk membantu proses penentuan penerima beasiswa. Penelitian dilakukan dengan menganalisis artikel dan jurnal yang relevan melalui Google Scholar, IEEE Xplore, Springer, Science Direct dan database elektronik lainnya. Kata kunci yang digunakan meliputi “data mining”, “sistem penunjang keputusan” dan “beasiswa”. Hasil analisis yang telah dilakukan menunjukkan bahwa data mining dengan pemilihan algoritma yang sesuai dengan ukuran dataset akan berdampak pada performa algoritma tersebut. Penelitian ini menunjukkan bahwa penerapan data mining dengan algoritma artificial neural network (ANN) memiliki performa terbaik dibanding dengan algoritma pembandingnya dengan hasil diatas 79%. Dengan hasil tersebut penggunaan data mining disimpulkan dapat memberikan kemudahan dalam pengelolaan data yang besar serta dapat meningkatkan efisiensi dan efektivitas dalam pengambilan keputusan. Selain itu, teknologi ini juga berpotensi untuk diterapkan pada bidang lain seperti pertanian, kesehatan, dan keuangan untuk mendukung pengambilan keputusan yang lebih baik berdasarkan data yang dimiliki oleh setiap bidang ilmu.

Kata kunci: *Beasiswa, Data Mining, Sistem Penunjang Keputusan, Universitas*

Abstract

Managing large data for decision-making in the scholarship awarding process at universities is a challenge in itself, especially to ensure that the decisions taken are effective and efficient. This research was conducted to provide information related to the application of data mining in supporting a decision support system as a solution to help the process of determining scholarship recipients. The research was conducted by analyzing relevant articles and journals through Google Scholar, IEEE Xplore, Springer, Science Direct and other electronic databases. The keywords used include “data mining”, “decision support system” and “scholarship”. The results of the analysis that has been carried out show that data mining with the selection of algorithms that are suitable for the size of the dataset will have an impact on the performance of the algorithm. This research shows that the application of data mining with the artificial neural network (ANN) algorithm has the best performance compared to its comparison algorithm with results above 79%. With these results, the use of data mining is concluded to provide convenience in managing large data and can increase efficiency and effectiveness in decision making. In addition, this technology also has the potential to be applied to other fields such as agriculture, health, and finance to support better decision making based on data owned by each field of science.

Keywords: *Data Mining, Decision Support System, Scholarship, University*

1. PENDAHULUAN

Dengan perkembangan teknologi informasi yang sangat pesat, data merupakan salah satu hal yang penting di berbagai bidang. Salah satu bidang yang dimaksud yaitu bidang pendidikan. Di universitas, pengelolaan data mahasiswa yang akurat dan efisien dapat memudahkan pihak universitas untuk mengambil berbagai keputusan, terutama terkait pemberian beasiswa kepada mahasiswa berprestasi dan kriteria lain seperti dari segi ekonomi yang dimiliki setiap mahasiswa. Proses pemilihan penerima beasiswa sering melibatkan banyak kriteria yang kompleks, hal tersebut menjadi tantangan tersendiri

dalam hal kecepatan dan ketepatan pengambilan keputusan. Pengambilan keputusan secara konvensional sering melibatkan proses yang rumit dan rentan dalam mengidentifikasi dan berakibat pendistribusian dana beasiswa tidak tepat sasaran. Seorang ahli di sektor pendidikan bernama Jared Smith menekankan pentingnya transparansi di lingkungan pendidikan seperti sekolah dan sejenisnya untuk menumbuhkan kepercayaan [1]. Oleh sebab itu, salah satu teknologi yang dapat digunakan untuk membantu menyelesaikan permasalahan tersebut adalah dengan menggunakan sistem penunjang keputusan atau biasa disebut dengan *Decision Support System* menggunakan metode data mining [2]. Pada kajian literatur yang sudah ada seperti beasiswa pendidikan tinggi internasional beasiswa pendidikan tinggi internasional untuk siswa dari global south [3]. Pemberian beasiswa berdasarkan tempat tinggal [4]. Karena masih sedikit kajian yang membahas terkait data mining pada bidang pendidikan, maka jurnal ini akan membahas metode dalam data mining yang digunakan dalam kasus beasiswa.

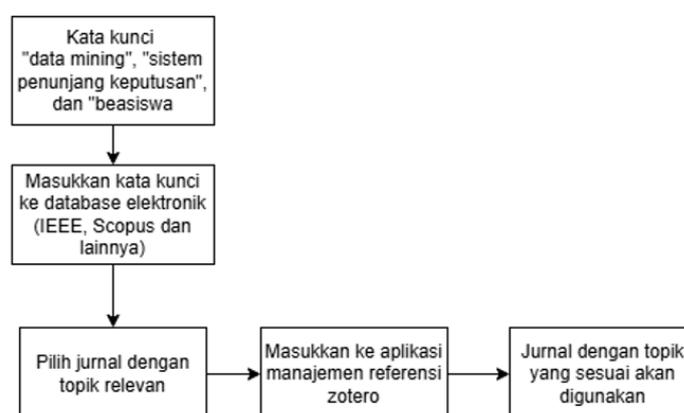
Data mining merupakan sebuah algoritma yang memungkinkan seorang analis dan *end user* untuk menyelesaikan masalah yang apabila dilakukan secara manual tidak akan menemukan solusi dari suatu permasalahan [5]. Data mining juga dapat dianggap sebagai suatu alat yang membantu menemukan pola, tren, atau informasi yang dapat digunakan dalam mengambil keputusan yang terbaik [6].

Sistem Penunjang Keputusan (*Decision Support System/DSS*) merupakan sistem berbasis komputer yang mendukung manusia dalam pengambilan Keputusan [7]. DSS bertujuan untuk menyediakan informasi, memberitahu dan memberikan prediksi terhadap alternatif yang memiliki nilai terbaik dari setiap kriteria yang ada [8]. Penerapan data mining dalam DSS memberikan akurasi yang baik dalam pengambilan keputusan dengan menggunakan teknik-teknik seperti klasifikasi dan prediksi.

Melalui kajian literatur ini, kita dapat memperoleh pemahaman terkait peran data mining dalam mendukung sistem pengambilan keputusan. Dalam jurnal ini, penulis akan memilih berbagai artikel ilmiah yang relevan dengan topik ini. Dengan melakukan analisis dari literatur yang ada, penulis akan memberikan gambaran terkait peran data mining dalam sistem penunjang keputusan. Penulis juga akan memberikan ringkasan literatur untuk memberikan pemahaman tentang data mining pada sistem penunjang keputusan.

2. METODE PENELITIAN

Pemilihan dan pengumpulan sumber literatur pada jurnal kali ini dilakukan secara sistematis dengan memilih artikel dan jurnal yang relevan dengan topik penelitian melalui *Google Scholar*, *Scopus*, *IEEE Xplore* dan situs database elektronik lainnya yang menyediakan akses ke publikasi artikel terkait. Seluruh jurnal yang didapatkan kemudian akan dikumpulkan pada aplikasi zotero, kemudian akan dilakukan penyaringan kembali sesuai dengan topik yang relevan. Apabila jurnal yang didapatkan sesuai dengan topik terkait maka akan dipertimbangkan untuk dijadikan sebagai referensi dari *literature review*. Proses pemilihan jurnal yang dilakukan juga disajikan dalam alur diagram berikut.



Gambar 1. Alur pemilihan jurnal

3. HASIL DAN PEMBAHASAN

3.1. Hasil

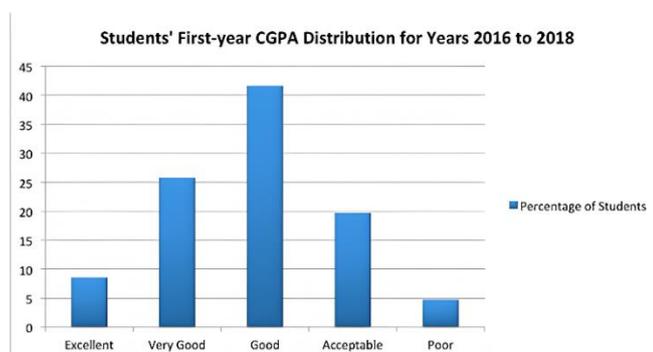
Sistem penunjang keputusan telah menjadi media yang efektif untuk membantu pemberi keputusan dalam mengambil keputusan, termasuk dalam pemberian beasiswa bagi mahasiswa berprestasi di lingkungan universitas. Penggunaan sistem penunjang keputusan dengan metode data mining dapat meningkatkan efisiensi waktu dalam pengambilan keputusan tanpa harus memfilter setiap data yang ada secara manual. Sistem penunjang keputusan juga bertujuan agar beasiswa yang diberikan sesuai dengan mahasiswa yang memenuhi kriteria seperti dari segi akademis maupun non-akademis.

3.2. Pembahasan

Universitas yang merupakan bagian dari sektor pendidikan yang berfungsi memfasilitasi mahasiswa yang ingin melanjutkan pendidikan tinggi di perguruan tinggi [9]. Beasiswa menjadi salah satu faktor yang sangat berpengaruh bagi mahasiswa untuk melanjutkan studi terutama bagi mahasiswa yang berprestasi namun kurang mampu dari segi finansial [10]. Universitas memiliki peran penting dalam menentukan penerima beasiswa yang disesuaikan dengan kriteria yang telah ditetapkan sesuai dengan kebutuhan dan potensi dari setiap mahasiswa.

Data dari setiap mahasiswa dari berbagai kriteria yang ditetapkan tentu sangat banyak. Untuk memudahkan memilih dan menentukan calon penerima beasiswa di lingkungan universitas, sistem penunjang keputusan dengan metode data mining memiliki peran penting. Peran penting yang dimaksud yaitu agar beasiswa yang diberikan bisa tepat sasaran kepada mahasiswa yang memiliki kriteria sesuai yaitu akademis dan non akademis yang baik.

Berdasarkan salah satu jurnal dari Hanan Abdullah Mengash yang berjudul “*Using Data Mining Techniques to Predict Student Performance to support Decision Making in University Admission Systems*”, penulis tersebut menggunakan bebarap algoritma data mining seperti *Artificial Neural Network* yang memiliki akurasi rata-rata diatas 79% [11]. Algoritma lain yang digunakan sebagai pembanding yaitu *Decision Trees*, *Support Vector Machines*, dan *Naïve Bayes*. Dataset yang digunakan setelah dilakukan normalisasi yaitu sebanyak 1430 data mahasiswa dari Angkatan 2016-2018 di *University Admission Systems* berdasarkan tiga kriteria yaitu Tes Bakat Umum (TBU), Tes Penerimaan Prestasi (TPR) dan Nilai Rata-rata Sekolah Menengah Atas(NRSMA). Untuk nilai TBU seperti verbal skill yang dimiliki mahasiswa dan TPR merupakan nilai matakuliah umum seperti bahasa inggris[12]. Pada Fig 1 ditunjukkan dataset pengkelompokan nilai mahasiswa dengan rentang 0% - 45% dari Angkatan 2016-2018 dengan pengkategorian nilai yaitu *Excellent*, *Very Good*, *Good*, *Acceptable* dan *Poor* [11].



Gambar 2 Dataset nilai mahasiswa [11]

Selanjutnya peneliti memasukkan nilai pada dataset pada Fig 1 ke setiap algoritma dan hasil dari setiap algoritma ditunjukkan pada Fig 2. Setiap model algoritma menggunakan validasi silang yang dimana 9 set data digunakan untuk melatih model sedangkan set data lainnya digunakan untuk pengujian. Proses ini diulang sebanyak 10 kali.

Kinerja dari setiap model data mining akan digunakan evaluasi matriks yang menggunakan konsep *True Positive Rate* (TP) apabila jumlah yang diprediksi dengan benar sebagai positif, *False Positive Rate* (FP) apabila jumlah yang diprediksi salah sebagai positif, *True Negative Rate* (TN) apabila jumlah yang diprediksi dengan benar sebagai negative, dan *False Negative Rate* (FN) jika jumlah yang diprediksi salah sebagai negative. Evaluasi matriks terdiri dari tiga bagian yaitu *accuracy* yang merupakan persentasi hasil yang diprediksi diukur menggunakan rumus berikut[12].

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (1)$$

Recall yaitu persentase positif yang diprediksi benar sebagai positif diukur menggunakan rumus berikut[12].

$$Recall = TP / (TP + FN) \quad (2)$$

Precision yang merupakan persentase pengamatan positif yang bernilai benar diukur menggunakan rumus berikut[12].

$$Precision = TP / (TP + FP) \quad (3)$$

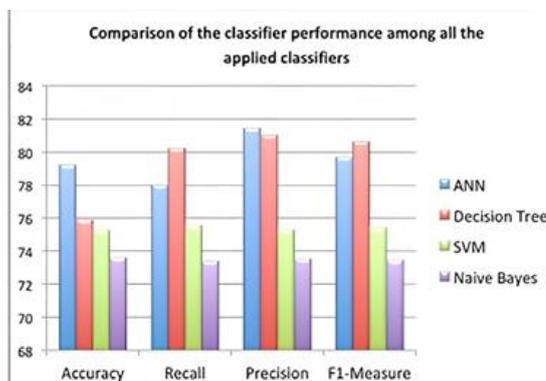
F1-Measure digunakan untuk mengukur keseimbangan antara *precision* dan *recall* dengan menggunakan rumus pengukuran sebagai berikut.

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (4)$$

Setelah tahap terakhir dilakukan menggunakan *F1-Measure* dan *training* data yang dilakukan sebanyak 10 kali. Hasil yang diperoleh yaitu algoritma ANN menempati urutan pertama dengan rata-rata akurasi diatas 79% menggunakan empat rumus pada evaluasi matriks [11]. Hasil data tersebut yang nanti dapat digunakan untuk akademisi dalam membuat keputusan seperti seberapa banyak mahasiswa yang mendapatkan beasiswa atau juga dapat digunakan untuk membuat keputusan terkait hal apa saja yang dapat diperbaiki agar performa mahasiswa bisa meningkat.

Tabel 1. Hasil akhir setelah dilakukan perhitungan menggunakan *F1-Measure* [12]

Classification Techniques	Accuracy (%)	Recall (%)	Precision (%)	F1-Measure (%)
ANN	79.22	78.03	81.44	79.70
Decision Tree	75.91	80.24	81.02	80.63
SVM	75.28	75.58	75.30	75.44
Naïve Bayes	73.61	73.38	73.54	73.46



Gambar 3. Hasil perbandingan antara algoritma [11]

Setiap algoritma pada data mining pasti memiliki kekurangan dan kelebihan yang dimiliki. Kelebihan dan kekurangan yang dimiliki oleh *Decision Trees*, *Support Vector Machines*, *Naïve Bayes*, dan *Artificial Neural Network* berdasarkan hasil yang didapatkan pada Fig 3 dan Fig 4 yaitu.

Tabel 2. Kelebihan dan kekurangan masing-masing algoritma

Algoritma	Kelebihan	Kekurangan
<i>Decision Trees</i>	<ul style="list-style-type: none"> - Mudah digunakan dan dipahami [13]. - Salah satu algoritma yang cepat dalam mengidentifikasi fitur baru [13]. 	<ul style="list-style-type: none"> - Rentan terhadap overfitting jika pohon terlalu dalam [13]. - Tidak efektif untuk data dengan variable yang memiliki banyak kategori tanpa melalui preprocessing [13]. - Tidak cocok untuk variabel yang bersifat berkelanjutan [13].
<i>Support Vector Machines</i>	<ul style="list-style-type: none"> - Efektif untuk dataset kecil hingga sedang [13]. - Cocok digunakan ketika terdapat pembagian yang jelas untuk setiap kelas-kelasnya [13]. 	<ul style="list-style-type: none"> - Tidak cocok untuk dataset besar karena memiliki waktu komputasi yang tinggi [13]. - Memiliki kinerja yang buru dalam kasus Dimana nilai numerik fitur pada setiap data lebih tinggi dibandingkan dengan sampel data pelatihan [13].
<i>Naïve Bayes</i>	<ul style="list-style-type: none"> - Cepat dan efisien bahkan untuk dataset yang besar [14]. - Sederhana dan mudah digunakan [14]. - Memiliki kinerja yang baik untuk masalah klasifikasi teks seperti spam [14]. 	<ul style="list-style-type: none"> - Performanya menurun jika ada korelasi yang kuat antara setiap fitur[14]. - Tidak cocok untuk dataset dengan jumlah data yang kecil [14]. - Mengasumsikan independensi antar fitur yang dimana jarang benar ketika dilakukan [14].
<i>Artificial Neural Network</i>	<ul style="list-style-type: none"> - Sangat fleksibel untuk menangani data kompleks seperti gambar, suara, dan teks [15]. - Dapat memodelkan hubungan non-linear dengan sangat baik [15]. - Memiliki kemampuan generalisasi yang baik jika bisa diatur dengan benar [15]. 	<ul style="list-style-type: none"> - Membutuhkan dataset yang besar agar menghasilkan performa yang optimal [15]. - Waktu pelatihan yang lama karena membutuhkan daya komputasi yang tinggi [15].

Berdasarkan jurnal yang telah dikaji maka dapat disimpulkan bahwa penggunaan sistem penunjang keputusan dengan metode data mining memiliki keefektifan yang berbeda tergantung dari algoritma data mining yang digunakan. Dengan mulai beralih dari mengelola data secara manual menjadi sistem melalui sistem penunjang keputusan, pengambilan keputusan bisa diambil secara tepat dan cepat, prediksi yang lebih akurat dan pengelolaan data yang lebih baik.

- a. Pengambilan keputusan yang tepat dan cepat
 Data mining dapat menemukan pola yang ada dalam data yang besar secara otomatis yang berguna untuk mempercepat proses pengambilan keputusan [16]. Dengan penggunaan sistem penunjang keputusan yang dipadukan dengan data mining, ketergantungan pada analisis manual bisa dikurangi yang dimana analisis manual memakan waktu yang lebih lama dan hasil yang tidak akurat.
- b. Prediksi yang lebih akurat
 Sistem penunjang keputusan dengan menggunakan data mining mengurangi tingkat kesalahan yang ditimbulkan oleh pengambil keputusan apabila dilakukan secara manusia. Sebaliknya dengan menggunakan data mining maka akurasi dari keputusan yang diambil bisa lebih akurat [17].
- c. Pengelolaan data yang lebih baik
 Sistem penunjang keputusan dengan pengelolaan data menggunakan data mining memberikan kemampuan untuk melakukan filter terhadap data yang besar secara efisien[18]. Selain itu penggunaan data mining juga diharapkan dapat menghasilkan laporan dalam bentuk data yang mendukung keputusan dengan akurasi yang lebih tinggi dibandingkan dilakukan secara manual.

4. KESIMPULAN DAN SARAN

4.1. Kesimpulan

Penggunaan sistem penunjang keputusan dengan metode data mining di bidang pendidikan khususnya di lingkungan universitas telah memberikan kemudahan terhadap pengelolaan data yang besar dan pengambilan keputusan bisa dilakukan secara efektif dan efisien. Pemilihan algoritma yang tepat sesuai dengan ukuran dataset yang dimiliki juga berpengaruh terhadap performa dari algoritma tersebut. Penggunaan *Artificial Neural Network* memiliki hasil yang lebih baik dari algoritma pembandingnya dengan hasil diatas 79% dikarenakan sesuai dengan ukuran dataset universitas yang tergolong ke dalam dataset yang besar. Dengan demikian penggunaan sistem penunjang keputusan menggunakan metode data mining diharapkan dapat memberikan kontribusi untuk meningkatkan efektifitas dan efisiensi dalam pengelolaan data dalam penentuan penerima beasiswa secara tepat sasaran sesuai kriteria yang ditetapkan oleh masing-masing universitas dan dengan jumlah data yang berbeda di setiap universitas.

4.2. Saran

Dengan perkembangan teknologi yang sangat cepat, sistem penunjang keputusan bisa digunakan untuk hal lainnya, seperti penentuan lulusan dengan pujian tertinggi (*cumlaude*). Sistem penunjang keputusan juga bisa digunakan di berbagai bidang seperti kesehatan, pertanian, dan keuangan. Diharapkan dengan perkembangan tersebut, setiap bidang yang ada bisa mulai menerapkan sistem penunjang keputusan dengan menggunakan data mining yang dimana algoritma yang digunakan dapat disesuaikan dengan ukuran dataset yang dimiliki. Untuk dataset yang berukuran besar seperti pada universitas bisa menggunakan *Artificial Neural Network* (ANN) dikarenakan algoritma ini juga fleksibel untuk berbagai jenis data seperti gambar suara dan teks. Namun apabila ingin diterapkan untuk bidang lainnya yang memiliki dataset lebih kecil bisa menggunakan *Support Vector Machine* atau *Naïve Bayes*.

DAFTAR PUSTAKA

- [1] T.-A. Nguyen-Hoang *et al.*, “Advancing Scholarship Management: A Blockchain-Enhanced Platform With Privacy-Secure Identities and AI-Driven Recommendations,” *IEEE Access*, vol. 12, pp. 168060–168090, 2024, doi: 10.1109/ACCESS.2024.3486078.
- [2] Y. Yun, D. Ma, and M. Yang, “Human–computer interaction-based Decision Support System with Applications in Data Mining,” *Future Generation Computer Systems*, vol. 114, pp. 285–289, Jan. 2021, doi: 10.1016/j.future.2020.07.048.
- [3] A. C. Campbell and E. Neff, “A Systematic Review of International Higher Education Scholarships for Students From the Global South,” *Review of Educational Research*, vol. 90, no. 6, pp. 824–861, Dec. 2020, doi: 10.3102/0034654320947783.
- [4] C. Anderson, “Local-level, place-based scholarships: a review of the literature,” *Educational Review*, vol. 73, no. 5, pp. 638–661, Sep. 2021, doi: 10.1080/00131911.2019.1619520.
- [5] P. E. Inc, “Mining, W. I. D,” in *Introduction to data mining*, New Jersey, 2006, pp. 2–12.
- [6] I. H. Sarker, “Data Science and Analytics: An Overview from Data-Driven Smart Computing, Decision-Making and Applications Perspective,” *SN COMPUT. SCI.*, vol. 2, no. 5, p. 377, Sep. 2021, doi: 10.1007/s42979-021-00765-8.
- [7] N. Aurelia, G. T. Murti, R. A. Putri, and R. M. Qodryanto, “Sistem Pendukung Keputusan, Decision Support System (DSS),” *INFORMATION SYSTEM FOR EDUCATORS AND PROFESSIONALS*, vol. 7, no. 1, 2022.
- [8] M. H. Natanael and D. Kusumaningsih, “PENERAPAN METODE WEIGHTED PRODUCT PADA SISTEM PENUNJANG KEPUTUSAN UNTUK PEMILIHAN ANGGOTA TERBAIK NAPOSO,” *Technologia*, vol. 12, no. 1, p. 41, Jan. 2021, doi: 10.31602/tji.v12i1.4181.

- [9] S. Mulyani, “Kebijakan Publik Di Bidang Pendidikan Tinggi Dalam Kaitannya Dengan Penerapan Rekognisi Pengalaman Lampau (RPL),” *I. Kom. dan Adm. Pub.*, vol. 11, no. 1, Jun. 2024, doi: 10.37676/professional.v11i1.6012.
- [10] R. Hegde, A. G. V, S. Madival, S. H. S, and S. U, “A Review on Data Mining and Machine Learning Methods for Student Scholarship Prediction,” in *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India: IEEE, Apr. 2021, pp. 923–927. doi: 10.1109/ICCMC51019.2021.9418376.
- [11] H. A. Mengash, “Using Data Mining Techniques to Predict Student Performance to Support Decision Making in University Admission Systems,” *IEEE Access*, vol. 8, pp. 55462–55470, 2020, doi: 10.1109/ACCESS.2020.2981905.
- [12] H. A. Mengash, “Using Data Mining Techniques to Predict Student Performance to Support Decision Making in University Admission Systems,” *IEEE Access*, vol. 8, pp. 55462–55470, 2020, doi: 10.1109/ACCESS.2020.2981905.
- [13] M. Bansal, A. Goyal, and A. Choudhary, “A comparative analysis of K-Nearest Neighbor, Genetic, Support Vector Machine, Decision Tree, and Long Short Term Memory algorithms in machine learning,” *Decision Analytics Journal*, vol. 3, p. 100071, Jun. 2022, doi: 10.1016/j.dajour.2022.100071.
- [14] Z. M. Fadhil, “Hybrid of K-means clustering and naive Bayes classifier for predicting performance of an employee,” *PEN*, vol. 9, no. 2, p. 799, Apr. 2021, doi: 10.21533/pen.v9i2.1898.
- [15] O. I. Abiodun *et al.*, “Comprehensive Review of Artificial Neural Network Applications to Pattern Recognition,” *IEEE Access*, vol. 7, pp. 158820–158846, 2019, doi: 10.1109/ACCESS.2019.2945545.
- [16] Y. Cheng, K. Chen, H. Sun, Y. Zhang, and F. Tao, “Data and knowledge mining with big data towards smart production,” *Journal of Industrial Information Integration*, vol. 9, pp. 1–13, Mar. 2018, doi: 10.1016/j.jii.2017.08.001.
- [17] F. Rolansa, Y. Yunita, and S. Suheri, “Sistem prediksi dan evaluasi prestasi akademik mahasiswa di Program Studi Teknik Informatika menggunakan data mining,” *JPIS*, vol. 9, no. 1, pp. 75–85, Jun. 2020, doi: 10.31571/saintek.v9i1.1696.
- [18] S. Maniyan, R. Ghousi, and A. Haeri, “Data mining-based decision support system for educational decision makers: Extracting rules to enhance academic efficiency,” *Computers and Education: Artificial Intelligence*, vol. 6, p. 100242, Jun. 2024, doi: 10.1016/j.caeai.2024.100242.

Halaman Ini Dikosongkan