

Implementation of Data Mining for Interpretation of KSE Scholarship Applicant Number Data using Naive Bayes Algorithm

Riyan Hidayah Purnama, Ali Ikhwan*

Program Studi Fakultas Sains dan Teknologi, Universitas Islam Negeri Sumatera Utara, Medan, Indonesia

Email: ¹riyanhidayah391@gmail.com, ^{2,*} ali_ikhwan@uinsu.ac.id

Correspondence Author Email: ali_ikhwan@uinsu.ac.id

Submitted: **20/05/2024**; Accepted: **30/06/2024**; Published: **30/06/2024**

Abstract—The purpose of this study is to interpret the large number of KSE scholarship applicant data, which is expected to provide a positive contribution in developing the KSE scholarship branding strategy, optimizing resource allocation and increasing the attractiveness of companies to allocate their CSR funds to the Karya Salemba Empat Foundation using data analysis techniques. The problem currently experienced is that the Karya Salemba Empat Foundation has been selecting KSE scholarship recipients manually, which results in the decision-making process not being able to be carried out quickly, accurately and efficiently. As one way to improve data accuracy, a method or computational model is needed in the form of a machine learning algorithm using the Naive Bayes method. With this Naive Bayes method, it is very appropriate to use to produce Knowledge. This study shows how the Karya Salemba Empat Foundation can utilize data to increase its value. From the results of the test carried out using 4,492 rows of data and 6 data variables and the pattern accuracy of 92% with an error margin of 8%, it shows that the naive bayes method is almost perfect in processing its data. The results of this study are expected to provide in-depth insight into how the application of data science can help the Karya Salemba Empat Foundation increase its appeal and strategy.

Keywords: Data Science; Naive Bayes; KSE Scholarship

1. INTRODUCTION

Yayasan Karya Salemba Empat is a scholarship institution founded by 8 alumni students of FE UI in 1995, currently KSE has grown and partnered with 3 state universities in North Sumatra and has provided scholarships to 15,000 children of the nation. In the era of the Industrial Revolution 4.0, every organization or institution uses information technology to improve its performance, so that it becomes more efficient and effective. The impact of the use of this information technology is the creation of a large amount of data that has not been fully utilized [1]. The stack of data can be processed into knowledge, thereby facilitating decision-making to formulate future policies [2]. One of the solutions is to utilize data science techniques or the science of data [3].

Data science is the combination of computer science, statistics, and business knowledge to process large amounts of data into knowledge, which allows the identification of certain patterns or trends.[4] This is certainly very valuable for interpreting large amounts of data to facilitate decision making [5]. Naive Bayes is a classification method in Machine Learning that can predict [6]. This method was chosen because this algorithm has a fairly high level of accuracy [7]. The Salemba Empat Karya Foundation is quite well organized, but it would be even better if it was processed with a data processing system such as data science which will be used to determine the quality pattern of KSE scholarship branding to increase the attractiveness of companies as donors [8]. In this study, the Naive Bayes algorithm is used to classify KSE scholarship applicant data. Naive Bayes is one of the classification methods that uses simple probability. Naive Bayes can classify data by calculating the probability of a class for each existing attribute[9]. The Naive Bayes Classifier method involves two stages in the classification process, namely the training stage and the classification stage. In the training stage, data samples are processed to create the best possible representation of the data. This method is the best method of classification by calculating the probability of each category[10]. In an effort to achieve this goal, the classification process forms a model that is able to differentiate data into different classes based on certain rules or functions [11].

The ongoing state of information science use is as yet restricted in the field of information examination, so the expectation for what's in store is that the usage of information science can work on the nature of the Salemba Empat Karya Establishment in taking future procedures [12]. In view of the aftereffects of the Joint Gathering Conversation Discussion with KSE staff, the ongoing issue is that the Karya Salemba Empat Establishment is as yet choosing KSE grant beneficiaries physically, and that implies that the dynamic cycle can't be completed rapidly, precisely and productively [13]. Aside from that, dealing with the information of students who register is easier, so it's important to use specific procedures or strategies to make the process easier.[14]So in this review, the scientist utilized Guileless Bayes examination to group and address the quantity of KSE grant candidates.

In this way, the Salemba Empat Work Foundation can be more responsive to the needs of Indonesian students who excel and have the potential to continue their education [15]. The Naive Bayes Classification Method used to classify scholarship recipients produced good accuracy, with a feasibility value of 5.04919, according to Said. Iskandar et al.'s "Naive Bayes Classifier Method in Determining Bidikmisi Scholarship Recipients at Medan State University," which the researcher used to support this study. 2. The highest accuracy of 79% was achieved during testing using the 80:20 comparison of training and testing data. Of the 318 students who participated, 250 were found to be eligible and 68 were found to be ineligible [16]. Antony Anwari Rahman and Agus Suryanto named "Execution of Grant

Beneficiary Determination Data Framework Utilizing Innocent Bayes Classifier Technique" The exploration strategy utilized is Innovative work (Research and development). The gadget is made with Microsoft Visual Studio 2013 system. The plausibility test was completed with tests 1 and 2. The consequences of the review from test 1 were 96.56%, and test 2 was 90.33%, remembered for the truly possible classification. The investigation discovered that the framework utilizing the Gullible Bayes Classifier calculation delivered high precision. When selecting potential scholarship recipients, it is helpful to have a system that is both accurate and effective. Various benefits integrate the progression of data about the execution of data mining, explicitly the Blameless Bayes Classifier calculation[17]"Implementation of the Naive Bayes Algorithm for Classification of Scholarship Recipients (Case Study of Hamzanwadi University)" by Nurhidayawati et al. From the implementation and testing of the Naive Bayes algorithm data mining technique for classifying the graduation of Hamzanwadi University student files that registered or proposed themselves to get a Bidikmisi scholarship in 2021, the highest accuracy value was obtained at 91 While the AUC was 0.996%, an excellent diagnosis level for classification was achieved. The Naive Bayes algorithm is more accurate and very good at analyzing and classifying Bidikmisi scholarship recipients at Hamzanwadi University Selong, as evidenced by the obtained accuracy and AUC values [18]. Application of Naive Bayesian Classifier in Ppa Scholarship Selection by Cici Alfiani Pradika Dita et al. This study uses the Naive Bayes Classifier to classify Budi Darma University scholarship recipients. In order to decide whether students who apply for the scholarship program will be accepted or rejected, the Naive Bayes Classifier method was chosen. The classifier technique is to characterize text in light of the greatest likelihood set as another report classification. From the consequences of this review, it was tracked down that the worth of "Acknowledge" is higher than "Not" which is 0.0351 contrasted with 0, so it very well may be inferred that the understudy named Riska Ramadhani got the 2019 PPA grant[19]. The researcher intends to conduct research on the advantages of data science analysis in interpreting KSE scholarship applicants to increase the attractiveness of businesses. In order to extract useful information from the number of KSE scholarship applicants, the researcher will investigate a variety of data science methods, such as statistical analysis and data visualization [20]. With this methodology, it is trusted that the consequences of this study can give a more top to bottom perspective on the data on KSE grant candidates [21], with this methodology, it is trusted that the consequences of this study can give a more top to bottom perspective on the data on KSE grant candidates[22].

2. RESEARCH METHODOLOGY

2.1 Research Stages

This piece makes use of descriptive qualitative research, which explains data in accordance with the findings of extensive, broad, and in-depth research [23]. The names of applicants for KSE scholarships will serve as the research data, and statistics will be used as a tool for calculation and testing. Starting with determining the problem, identifying the problem, collecting data, and managing data in a structured manner, this process is related to the problem being studied to produce conclusions. The phases of this examination allude to Information Disclosure in Information, which should be visible in the accompanying picture:

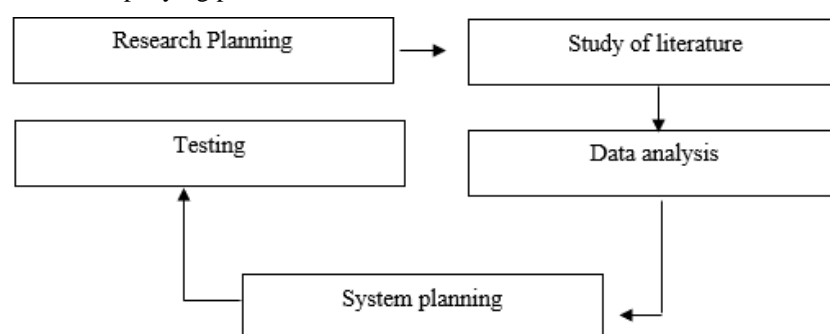


Figure 1. Research Flow

2.2 System planning

By and large, arranging is characterized as an action that looks at what will be finished, how to make it happen and when to do it. In various research textbooks, the term "research design" refers to all of the steps involved in planning and carrying out research [24]. In order to classify the number of KSE scholarship applicants, the Naive Bayes Classifier method was used in this study.

2.3 Data collection

An essential step in the research process, data collection aims to gather relevant and accurate data in order to provide answers to the research's questions. Various methods, including interviews, observations, and literature reviews, can be used to collect data.[25] In order to find information and read literature studies to find solutions to problems by

tracing previously written sources, researchers held a Forum Group Discussion with KSE staff. In the examination cycle, writing studies are vital on the grounds that they assist specialists with grasping the ebb and flow framework, recognize framework needs, and explain the components to be concentrated by perusing and figuring out a few past examinations, whether as diaries, articles, proposals, or propositions that are connected with the momentum research. The KSE scholarship applicant dataset is the one that was used. This study used data from the registration website database and FGDs with KSE staff to create the dataset. Thus, there were 4492 rows of data, each containing six variables. These rows were then divided into training data (80%) and testing data (80%).

2.4 System planning

The process of figuring out what data and processes a newsystem needs is called system design. The objective is to address the issues of framework clients and present a reasonable picture and complete plan. The first step in the system design process is to create a flowchart that follows this research method. The flowchart will be utilized as an aide in completing each stage or cycle in the characterization framework by applying the Gullible Bayes Classifier technique, including the accompanying:

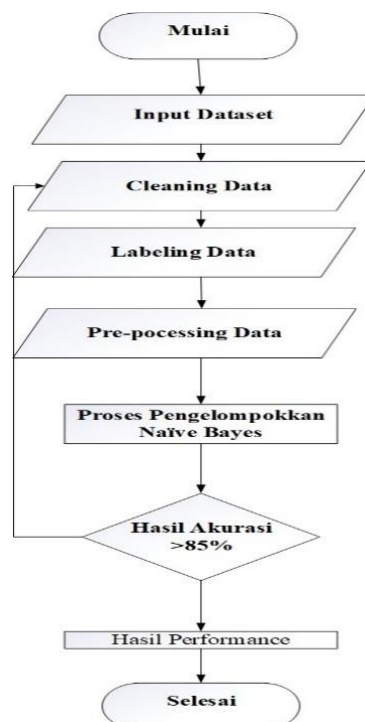


Figure 2. Flowchart of the Naïve Bayes Algorithm

The Naive Bayes algorithm will be used to interpret KSE scholarship applicant data from three state universities in the North Sumatra region in this study. Gullible Bayes is a technique used to foresee future open doors in light of past experience. When used on databases with a lot of data, this method typically performs well and quickly [26]. The information in this study were acquired from a succeed organized data set. Data collection, data cleaning, data labeling, data pre-processing, application of the Naive Bayes algorithm, performance evaluation, and conclusion or result drawing are all parts of the research methodology.

2.6 Testing

This testing is a process that is done to make sure the built data mining system follows the plan for the built system. This testing includes a few stages, for example, data choice, information cleaning, information change, and understanding of results.[27]Testing the data that has been successfully collected is the next step. This review involves Python language and Google Collaboratory as a cloud to run it and utilizations Rapidminer programming as an instrument to test precision and show information representation.

3. RESULT AND DISCUSSION

3.1 Data analysis

Information examination is a deliberate cycle for portraying, deciphering and handling information so significant ends can be drawn. This is a basic move toward transforming crude information into helpful data. An accuracy table was created with the help of the Rapidminer tool and the Python programming language in this study, with Google Collab serving as the terminal.

3.1.1 Data Collection

According to the dissemination table, the KSE grant candidate information is spread across 3 state colleges in North Sumatra. Information Assortment This examination involves crude information from the KSE grant data set in 2023. This information comprises of 4491 lines of information and 6 information factors.

Table 1. KSE Registrant Data Table

No	Nama	NPM	Jenis Kelamin	Fakultas	Complete %	Universitas
0	RIKA RISKY	2202090	Perempuan	Syariah dan Hukum	98	Universitas Islam Negeri Sumatera Utara
1	KAMILLA GANING ADZANI	0502202092	Perempuan	Ekonomi dan Bisnis Islam	96	Universitas Islam Negeri Sumatera Utara
2	NADILA JELITA	0506211009	Perempuan	Ekonomi dan Bisnis Islam	95	Universitas Islam Negeri Sumatera Utara
3	M. FAROUHU WAROYHAN	0203212049	Laki-laki	Syariah dan Hukum	91	Universitas Islam Negeri Sumatera Utara
4	ATIKA LINGGA	0704191067	Perempuan	Sains dan Teknologi	100	Universitas Islam Negeri Sumatera Utara
....
4487	YOHANES RICHARD SANJAYA PARAPAT	220100126	Laki-laki	Kedokteran	96	Universitas Sumatera Utara
4488	ANUGERAH MALIKA LEODZARLY PANE	210805079	Perempuan	Matematika dan IPA	57	Universitas Sumatera Utara
4489	AFWAN HAFIZAHULLAH	231201003	Laki-laki	Kehutanan	70	Universitas Sumatera Utara
4490	DITA SHABRINA	221000094	Perempuan	Kesehatan Masyarakat	57	Universitas Sumatera Utara
4491	MUHAMMAD NURWANA	230708048	Laki-laki	Ilmu Budaya	54	Universitas Sumatera Utara

Based on the table image above, the data consists of the variables name, npm, gender, faculty, %complete, and university.

3.1.2 Data Cleaning

After the data collection stage, the next step is to carry out Data Cleaning or cleaning using the Python language in Google Colab. The goal is to eliminate data duplication, check data consistency, and correct errors such as misprints. This is done so that the data is ready to be processed in the data mining process. At this stage, namely cleaning the data on Gender and NPM data which previously contained missing data which must be cleaned for the next analysis process using Python, namely.

Table 2. Pembersihan Data

Variabel	Missing Values
Nama	0
NPM	0
Jenis Kelamin	0
Fakultas	0
% Complete	0
Universitas	0

3.1.3 Data Labeling

In this process the data is labeled to determine the classification of KSE scholarship applicants. At this stage, university data is used as the y-axis and the data is given a label as a benchmark in classifying KSE scholarship applicant data.



Table 3. Data Labeling

Universitas	Angka
UINSU	0
UNIMED	1
USU	2

3.1.4 Preprocessing Data

In data preprocessing, not all the data in the database is used, only the data that is relevant for the analysis is retrieved. Data processing involves datasets that have gone through a preprocessing process, including cleaning, integration, transformation and reduction processes. This makes the dataset ready to be used in the training data and testing data stages.

Table 4. Preprocessing Data

Nama	NPM	Jenis Kelamin	Fakultas	% Complete	Universitas
1962	3700	0	22	3	0
2258	558	0	28	3	0
1016	1013	1	13	26	0
...
2707	3420	1	0	24	1
3272	3485	0	13	1	1
354	3705	1	22	22	1
2733	2241	1	27	1	1
1769	3547	0	22	0	1
3662	2996	1	17	12	1
...
4132	2985	0	3	24	2
2305	3131	1	20	25	2
489	4093	0	27	3	2

In processing this data, the data type in string form must be changed to numeric so that machine learning can perform calculations automatically and create testing and training data, where in this study 80% of the training data was used and 20% of the test data.

3.1.5 Naïve Bayes

Naïve Bayes is a machine learning method for classifying data. At this stage the training data and testing data are combined and grouped, the data has been processed using Python programming which will produce a data modeling dataset.

	Nama	NPM	Jenis Kelamin	Fakultas	% Complete	Universitas	target
157	4,117	59	1	1	96	0	0
4,246	2,721	2,222	1	27	54	2	2
1,020	3,445	1,136	1	25	80	0	0
1,119	2,682	310	1	15	95	0	0
731	3,995	858	1	4	96	0	0

Figure 3. Modeling Dataset

3.1.6 Performance

After everything is connected, the next step is to see the accuracy and performance results of the data using the Naive Bayes model. You can see the results after starting the process by clicking the "Run" button. Once the calculations are complete, you can see the results of the Naive Bayes process using RapidMiner by examining the output of the "Apply Model" process and the results of the performance evaluation. The image below is the accuracy result of Naïve Bayes, it can be seen that the result is 92%, which shows that the Naïve Bayes method is almost perfect in processing the data. Most of the results are accurate, the better and more precise the data produced. To measure the level of accuracy of the Naive Bayes Classifier method, it is calculated using the equation.

	precision	recall	f1-score	support
0	0.97	1	0.98	269
1	0.86	0.9	0.88	286
2	0.94	0.88	0.91	343
accuracy	0.92	0.92	0.92	0.92
macro avg	0.92	0.93	0.92	898
weighted avg	0.92	0.92	0.92	898

Figure 4. Accuracy results

This bar column image is a classification of data produced by Naïve Bayes where there are 29 faculties from 3 universities and 899 students in North Sumatra. As a result, the engineering faculty has the highest number of KSE scholarship applicants at number 95 and is followed by the faculty of economics and Islamic business at number 76 and followed by other faculties

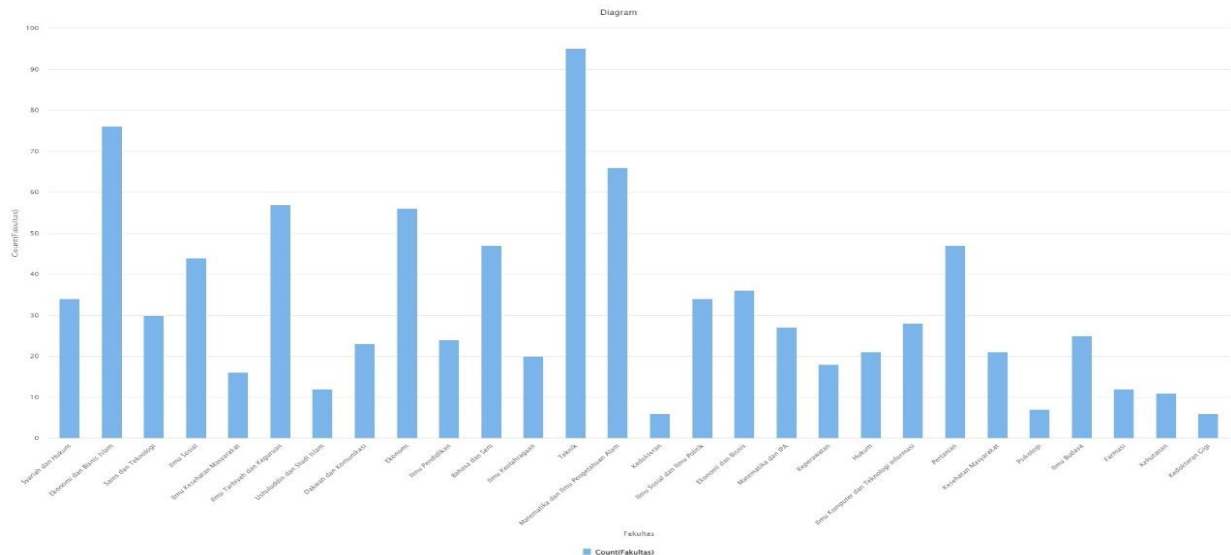


Figure 5. Faculty Column Bar Graph

As seen in Figure 3 above, these are the results of the classification of KSE scholarship applicants at each of the partner universities of the Karya Salemba Empat Foundation in North Sumatra, namely UINSU 291, Unimed 217 and USU 391.

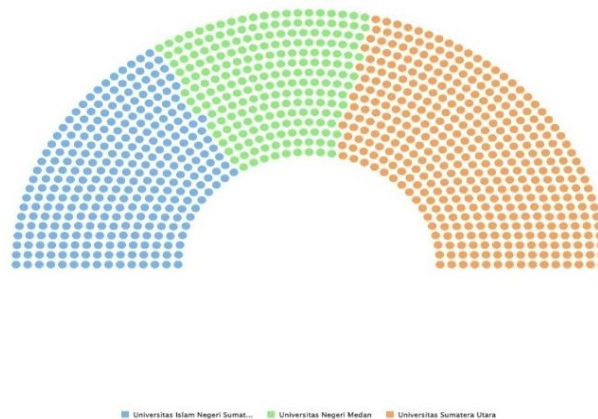


Figure 6. Parliament graphic

In picture 4, the graphic display above is a display of features from the results of processing KSE registrant data, but this graph is displayed in character form to see whether or not there is a correlation in the data analysis which is the study used in the table.



Figure 7. World Cloud Data Processing Results

The display of the 5 data processing graphs above is a feature displayed from the processing of KSE registrant data, but this graph is displayed in character form to see whether there is a high correlation in the data analysis which is the study used in the table.

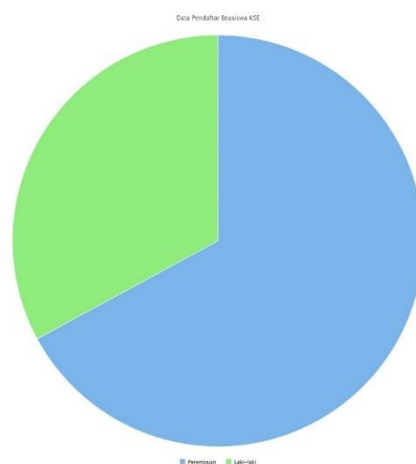


Figure 8. World Cloud Data Processing Results

As seen in Figure 6 in this Pie Chart is the result of data processing, namely the visualization of the gender data of KSE scholarship applicants. With details of 648 females and 251 males.

4. CONCLUSION

Data analysis allows the Karya Salemba Empat Foundation to understand the profiles and preferences of prospective scholarship applicants better, so they can develop more effective strategies in attracting companies as donors. By understanding company sentiment and perceptions of the KSE scholarship program, Karya Salemba Empat Foundation can identify areas that need to be developed to improve their reputation and remain competitive in Indonesia. Thus, the application of data science in interpreting data on the number of KSE scholarship registrations is expected to make a positive contribution in developing KSE scholarship branding strategies, optimizing resource allocation and increasing the attractiveness of companies to manage their CSR funds to the Karya Salemba Empat Foundation and can help the Karya Salemba Foundation Four to become more efficient, effective, and competitive strengthening their image as a respected education partner.

REFERENCES

- [1] H. Hairani and A. Z. Amrullah, "Pelatihan Pengenalan Data Science untuk Meningkatkan Kemampuan dalam Pengolahan Data," *Jurnal Abdidas*, vol. 1, no. 3, pp. 95–99, 2020, doi: 10.31004/abdidas.v1i3.31.
- [2] S. Suryaningsih and F. Ainun Nisa, "Kontribusi STEAM Project Based Learning dalam Mengukur Keterampilan Proses Sains dan Berpikir Kreatif Siswa," *Jurnal Pendidikan Indonesia*, vol. 2, no. 6, pp. 1097–1111, 2021, doi: 10.36418/japendi.v2i6.198.



- [3] D. Anggarwati, O. Nurdiawan, I. Ali, and D. A. Kurnia, "Penerapan Algoritma K-Means Dalam Prediksi Penjualan Karoseri," *Jurnal Data Science & Informatika (Jdsi)*, vol. 1, no. 2, pp. 58–62, 2021.
- [4] A. Lestari Tungadi, S. Michael Liem, and R. Vincentsius, "Pemanfaatan Data Science dan Teknologi LBS Dalam Optimalisasi Pemasaran Online Pada IKM," *Seminar Nasional Informatika (SEMNASIF)*, vol. 1, no. 1, pp. 28–32, 2020.
- [5] G. Kennard Taruna and S. Budi, "Penerapan Data Science pada Dataset Olympics," *Strategi*, vol. 4, no. November, pp. 2443–2229, 2022.
- [6] N. Ningsih, A. Aprianto, and A. Angeline, "Pendekatan Data Science untuk Deteksi Dini Diabetes Menggunakan Naive Bayes Classifier," *Journal of Information System, Graphics, Hospitality and Technology*, vol. 5, no. 1, pp. 26–31, 2023, doi: 10.37823/insight.v5i1.300.
- [7] E. Prabowo and Abd. C. Fauzan, "Implementasi Algoritma Naive Bayes Classifier Untuk Penentuan Kelayakan Penerima Beasiswa Peningkatan Prestasi Akademik (PPA) Bagi Mahasiswa," *ILKOMNIKA: Journal of Computer Science and Applied Informatics*, vol. 3, no. 2, pp. 165–177, 2021, doi: 10.28926/ilkomnika.v3i2.296.
- [8] I. Ismail, "Penerapan Data Science Menggunakan Artificial Neural Network (ANN) Metode Self Organizing Mapping (SOM) untuk Klasifikasi Industri," *Warta Akab*, vol. 45, no. 2, pp. 66–70, 2021, doi: 10.55075/wa.v45i2.62.
- [9] Angga Saputra Dinata and Nisar, "Penerapan Algoritma Naive Bayes Dalam Pengadaan Buku Referensi Pada Perpustakaan SMA Negeri 1 Trimurjo Berbasis Web," *Indonesian Journal of Science, Technology and Humanities*, vol. 1, no. 2, pp. 80–90, 2023, doi: 10.60076/ijstech.v1i2.131.
- [10] A. U. Kurnia, A. S. Budi, and P. H. Susilo, "Sistem Pendukung Keputusan Penerimaan Beasiswa Menggunakan Metode Naive Bayes," *Joutica*, vol. 5, no. 2, p. 397, 2020, doi: 10.30736/jti.v5i2.484.
- [11] N. Riyanah and F. Fatmawati, "Penerapan Algoritma Naive Bayes Untuk Klasifikasi Penerima Bantuan Surat Keterangan Tidak Mampu," *JTIM : Jurnal Teknologi Informasi dan Multimedia*, vol. 2, no. 4, pp. 206–213, 2021, doi: 10.35746/jtim.v2i4.117.
- [12] N. Kamaliah and A. F. Juliana Rahman, "Menjaga Kualitas Pembelajaran Praktikum Pengolahan Data Secara Daring pada Mata Pelatihan Analisis dan Interpretasi Data," *Andragogi: Jurnal Diklat Teknis Pendidikan dan Keagamaan*, vol. 9, no. 1, pp. 24–32, 2021, doi: 10.36052/andragogi.v9i1.208.
- [13] T. Y. C. A. Wicaksono and Y. D. Setianto, "Sistem Pendukung Pengambilan Keputusan Kesesuaian Beasiswa Menggunakan Naive Bayes," *Praxis*, vol. 2, no. 2, p. 187, 2020, doi: 10.24167/praxis.v2i2.2624.
- [14] F. D. Pratama, I. Zufria, and T. Triase, "Implementasi Data Mining Menggunakan Algoritma Naive Bayes Untuk Klasifikasi Penerima Program Indonesia Pintar," *Rabit : Jurnal Teknologi dan Sistem Informasi Univrab*, vol. 7, no. 1, pp. 77–84, 2022, doi: 10.36341/rabit.v7i1.2217.
- [15] I. Ahmad, S. Samsugi, and Y. Irawan, "Implementasi Data Mining Sebagai Pengolahan Data," *Jurnal Teknoinfo*, vol. 16, no. 1, p. 46, 2022, [Online]. Available: <http://portaldata.org/index.php/portaldata/article/view/107>
- [16] S. Iskandar, N. R. Refisis, and B. A. Ginting, "Metode Naive Bayes Classifier Dalam Penentuan Penerima Beasiswa Bidikmisi Di Universitas Negeri Medan," *Karimatika*, vol. 7, no. 1, pp. 10–23, 2021.
- [17] AHMAD ADRI, "Implementasi Data Mining Menggunakan Algoritma Apriori," *Paper Knowledge . Toward a Media History of Documents*, vol. 6, no. 2, pp. 1–77, 2021.
- [18] N. I. Nurhidayati, Y. Yahya, F. Fathurrahman, L. M. Samsu, and W. Amnia, "Implementasi Algoritma Naive Bayes Untuk Klasifikasi Penerima Beasiswa (Studi Kasus Universitas Hamzanwadi)," *Infotek : Jurnal Informatika dan Teknologi*, vol. 6, no. 1, pp. 177–188, 2023, doi: 10.29408/jit.v6i1.7529.
- [19] C. A. P. Dita, P. Chairunisyah, and ..., "Penerapan Naive Bayesian Classifier Dalam Penyeleksian Beasiswa PPA," *Journal of Computer ...*, vol. 2, no. 2, pp. 194–198, 2021, [Online]. Available: <https://ejournal.seminar-id.com/index.php/josyc/article/view/649>
- [20] B. S. Pranata and D. P. Utomo, "Penerapan Data Mining Algoritma FP-Growth Untuk Persediaan Sparepart Pada Bengkel Motor (Study Kasus Bengkel Sinar Service)," *Bulletin of Information Technology (BIT)*, vol. 1, no. 2, pp. 83–91, 2020.
- [21] R. A. Pangestu and S. Noris, "Analisa Data Mining Prediksi Lelang Suku Cadang Dengan Metode K-NearestNeighbor (Studi Kasus PT. Parmud Jaya Perkasa)," *Jurnal Informatika Multi*, vol. 1, no. 4, pp. 285–295, 2023.
- [22] Rayuwati, Husna Gemasih, and Irma Nizar, "IMPLEMENTASI AIGORITMA NAIVE BAYES UNTUK MEMPREDIKSI TINGKAT PENYEBARAN COVID," *Jural Riset Rumpun Ilmu Teknik*, vol. 1, no. 1, pp. 38–46, 2022, doi: 10.55606/jurritek.v1i1.127.
- [23] D. N. Yoliadi, "Data Mining Dalam Analisis Tingkat Penjualan Barang Elektronik Menggunakan Algoritma K-Means," *Insearch: Information System Research Journal*, vol. 3, no. 01, 2023, doi: 10.15548/isrj.v3i01.5829.
- [24] Ardiansyah, Risnita, and M. S. Jailani, "Teknik Pengumpulan Data Dan Instrumen Penelitian Ilmiah Pendidikan Pada Pendekatan Kualitatif dan Kuantitatif," *Jurnal IHSAN : Jurnal Pendidikan Islam*, vol. 1, no. 2, pp. 1–9, 2023, doi: 10.61104/ihsan.v1i2.57.
- [25] A. G. Prawiyogi, T. L. Sadiyah, A. Purwanugraha, and P. N. Elisa, "Penggunaan Media Big Book untuk Menumbuhkan Minat Membaca di Sekolah Dasar," *Jurnal Basicedu*, vol. 5, no. 1, pp. 446–452, 2021, doi: 10.31004/basicedu.v5i1.787.
- [26] E. Sutoyo and A. Almaarif, "Educational Data Mining for Predicting Student Graduation Using the Naive Bayes Classifier Algorithm," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 4, no. 1, pp. 95–101, 2020, doi: 10.29207/resti.v4i1.1502.
- [27] F. Aziz, A. Wahab, and P. Muna, "Journal Pharmacy and Application of Computer Sciences Penerapan Metode Machine Learning Untuk Mengklasifikasi Penanganan Perawatan Pasien Application of Machine Learning Method To Classification of Patient Care Treatment," *Journal Pharmacy and Application of Computer Sciences*, vol. 1, no. 2, pp. 57–62, 2023, [Online]. Available: <https://journalanr.arlisakamadani.com/index.php/jopacs/article/view/37/32>