

SISTEM PENDUKUNG KEPUTUSAN SELEKSI PENERIMA BEASISWA DENGAN METODE NAÏVE BAYES CLASSIFIER

Qori Alfina Pratiwi¹ Jati Sasongko Wibowo²

¹Program Strata 1 Program Studi Teknik Informatika pada Fakultas Teknologi Informasi Dan Industri Universitas Stikubank (UNISBANK) Semarang.

address, telp/fax of institution/affiliation, e-mail: goryalfina7@gmail.com¹ jatisw@edu.unisbank.ac.id²

ARTICLE INFO

Article history:

Received Desember 28 2022

Received in revised form Januari 8 2023

Accepted Januari 25 2023

Available online Juli 16 2023

ABSTRACT

Lot of problems arise in selecting scholarship recipients in a large number of submissions, the existence of several searches used, and the selection of files for scholarship applicants is still manual, so a system is needed that can speed up, help, and make it easier in the decision-making process to lighten work. student section. In supporting decisions this system will use the Naïve Bayes Classifier Method to determine what is acceptable and not acceptable. The NBC method can analyze and make improvements to old data, and the resulting data will provide simpler probability values that can later be used to make decisions. From the results of the research that has been carried out, it can be realized that the application of the data mining algorithm using the Naïve Bayes Classifier can be carried out to select scholarship recipients at Stikubank University Semarang. The result of the selection of Unisbank Semarang scholarship recipients is the accuracy value. 72% of the 135 data which is divided into 100 training data and 35 test data.

Keywords: Scholarships, Recipients, Naïve Bayes Classifier, Data Mining

1. Introduction

Definisi dari Sistem Pendukung Keputusan (SPK) yakni suatu sistem informasi yang interaktif dengan mempersiapkan manipulasi data serta informasi pemodelan. Sistem tersebut dapat dipakai sebagai dasar pengambilan keputusan dengan keadaan semistruktur dan terstruktur, dimana tidak ada satupun yang tahu bagaimana secara pasti keputusan semestinya dibuat (Budiman & Adi Prasetyo, 2019). Sistem Pendukung Keputusan atau yang sering disebut SPK dapat memberikan pemecahan permasalahan, pemilihan data yang sesuai serta melakukan penentuan terhadap pendekatan yang dapat dimanfaatkan untuk proses memudahkan serta mempercepat dalam mengambil keputusan, dimana berkaitan dengan penyeleksian penerima beasiswa. (Santoso, Azhar, Husain, & Muliadi, 2021)

Semua institusi pendidikan, khususnya universitas, menawarkan berbagai peluang, termasuk program beasiswa, kepada calon mahasiswa. Program ini ditujukan bagi calon mahasiswa yang memenuhi kriteria untuk dapat menyelesaikan pendidikan ke tingkat yang lebih tinggi.

Definisi beasiswa yakni sebuah bantuan biaya, dimana asalnya tidak dari diri sendiri atau orang tua sebagai proses untuk mencukupi keberlangsungan dalam proses pendidikan seseorang. Namun beasiswa diserahkan universitas kedutaan, pemerintah, non-pemerintah, dan lembaga-lembaga lain, atau juga dapat berkesempatan agar dapat menaikkan kapasitas sumber daya manusianya melewati jalur pendidikan (Perdana, Hasibuan, & Fadlina, 2022). Setiap beasiswa memiliki kriterianya masing-masing tergantung dari jenis beasiswanya, seperti di Universitas Stikubank Semarang yang telah memiliki beberapa kriteria untuk memastikan mahasiswa yang akan mendapatkan beasiswa. Kriteria yang dibutuhkan berdasarkan pendapatan orang tua, prestasi, tanggungan dari orang tua serta masih banyak lagi. Dari beberapa kriteria di atas akan memiliki index parameter yang lebih detail. Sehingga penerapannya tergantung pada sifat beasiswa.

Permasalahan yang banyak muncul dalam melakukan seleksi penerima beasiswa adalah banyaknya pengajuan, adanya beberapa kriteria yang digunakan, serta penyeleksian terhadap berkas mahasiswa pengaju beasiswa masih manual, sehingga dibutuhkan sebuah sistem yang dapat mempercepat, membantu, dan menjadikan lebih mudah dalam proses pengambilan keputusan untuk meringankan kerja bagian kemahasiswaan.

Pada penelitian ini, Sistem pendukung keputusan ini menggunakan metode klasifikasi Naive Bayesian. Metode NBC dapat menganalisis dan memperbaiki data yang terhadulu serta data hasil dari proses pemberian nilai suatu probabilitas sederhana, dimana akan dimanfaatkan sebagai proses pengambilan keputusan. Untuk memastikan bahwa NBC dianggap mampu menghasilkan data yang dinilai akurat serta dimanfaatkan sebagai dasar yang menentukan apakah seorang calon Mahasiswa akan diterima atau ditolak beasiswa.

2. Research Method

Metode Penelitian menggunakan model CRISP_DM (Cross Industry Standard Process for Data Mining) dengan 6 tahapan: Modeling atau Model Building, Data Understanding atau Pemahaman Data, Bussiness Understanding atau Pemahaman Proses Bisnis, Deployment atau Penyebaran. Data Preparation atau Pengolahan Data, Evaluation atau Pengujian dan Evaluasi. Dengan Instrumen penelitian meliputi: Kebutuhan Software atau perangkat lunak, Kebutuhan hardware atau Perangkat Keras serta Kebutuhan Data. Prosedur pengumpulan data menggunakan metode wawancara, observasi, serta pengumpulan dokumen Universitas Stikubank. Disusun mennggunakan uji Confusion Matrix.

Table 1. The Confusion Matrix

Predicted Condition	Positive	Negative
Positive	True Positive (TP)	False Positive (FP)
Negative	False Negative (FN)	True Negative (TN)

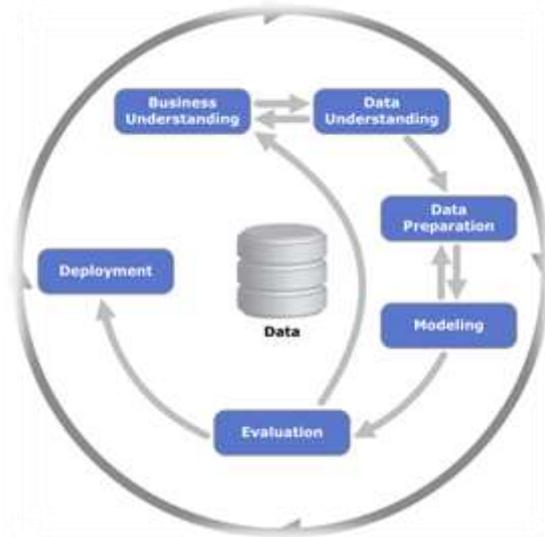


Figure 1. Proses CRISP-DM

3. Results and Analysis

Hasil penelitian ini bertujuan untuk memudahkan dalam melakukan seleksi penerima beasiswa di Universitas Stikubank Semarang Tahun 2020 hingga 2022 dengan perhitungan algoritma naïve bayes. Data latih yang digunakan adalah data pendaftar beasiswa di Universitas Stikubank Semarang dengan jumlah sebanyak 100 data.

Metodenya dalam menyeleksi penerima beasiswa yakni dengan Naïve Bayes Classifier, dimana metodenya bisa digunakan sebagai penanganan trend secara angung atau faktor musiman. Tahapan-tahapan yang terdapat dalam data mining dalam seleksi penerima beasiswa di Universitas Stikubank Semarang antara lain:

a. Pemahaman Data (Data Understanding)

Dimana datanya diperoleh dari mahasiswa yang melakukan pengajuan untuk mendaftar beasiswa, kemudian dalam dataset yang ada terdiri dari 8 atribut antara lain No. Pendaftaran, Nama Siswa, Penghasilan Orang Tua, Pekerjaan Orang Tua Kepemilikan Rumah, Jumlah Tanggungan, status dan juga Prestasi Siswa.

b. Pemahaman Bisnis (Business Understanding)

Berdasar pada data yang menerima beasiswa di 2020 – 2022 mahasiswa Universitas Stikubank Semarang yang dinyatakan diterima berjumlah 61 mahasiswa, dan melakukan pemanfaatan kepada teknik klasifikasi dalam data mining model naïve bayes, sehingga prosesnya dalam menyeleksi beasiswa bisa dijadikan lebih efektif serta efisien.

c. Pengolahan Data (Data Preparation)

Data yang didapatkan sejumlah 135 data, terdiri dari mahasiswa yang diterima dalam proses penyeleksian ataupun mahasiswa yang belum diterima, namun datanya masih terdapat anomaly serta duplikat ataupun inkonsisten data. Agar data yang dihasilkan memiliki kualitas baik, terdapat teknik preprocessing yakni:

- 1) Data Size Reduction and Dcretization, sebagai dasar untuk mendapatkan sataset beserta atributnya serta record dengan jumlah kecil namun sifatnya

bersifat informatif. Selanjutnya data training akan dilakukan proses untuk menghapus data duplikasi dengan tools Ms. Excel 2013.

- 2) Data Intergration and Transformation, sebagai proses peningkatan efisiensi serta akurasi algoritma yang ada. Kemudian datanya menggunakan data kategorikal.
- 3) Data Validation, inkonsistensi serta anomali data bisa terjadi karena terdapat missing values serta dikarenakan datanya yang digunakan terdapat nilai yang belum sesuai yang biasa dinamakan dengan outlier ataupun noise.

Sesudah melakukan preprocessing data dari Universitas Stikubank Semarang ada 135 menunjukkan bahwa tidak ada data yang terduplikasi, sebanyak 61 Mahasiswa diterima dan sebanyak 74 Mahasiswa tidak diterima. Berikut adalah data yang telah dilakukan preprocessing data.

d. Pemodelan (Modeling)

Dalam bagian ini biasanya dinamakan learning, dikarenakan datanya akan dikelompokkan dalam model serta akan dihasilkan beberapa aturan yang ada. selanjutnya, pembuatan model menggunakan algoritma naïve bayes. Contoh mengelompokkan naïve bayes dari 100 data latih yang diterima sebanyak 48 mahasiswa dan yang Tidak diterima 52 mahasiswa. Selanjutnya data latih yang sudah didapatkan tersebut akan diuji dengan menggunakan data uji seperti berikut:

- 1) Mahasiswa dengan Nama Yoseph Silvance memiliki atribut dengan Pekerjaan Orang Tua Petani Pendapatan Orang Tua Rendah, Jumlah Tanggungan Banyak, Kepemilikan Rumah Menumpang, Prestasi Ada, dan Status Tidak Diterima.

$$\begin{aligned}
 P(X = \text{Diterima} | \text{Diterima}) / X &= 48/100 \\
 P(X_1 = \text{Pekerjaan Orang Tua} | \text{Diterima}) &= 4/48 \\
 P(X_2 = \text{Penghasilan Orang Tua} | \text{Diterima}) &= 36/48 \\
 P(X_3 = \text{Jumlah Tanggungan} | \text{Diterima}) &= 10/48 \\
 P(X_4 = \text{Kepemilikan Rumah} | \text{Diterima}) &= 8/48 \\
 P(X_5 = \text{Prestasi} | \text{Diterima}) &= 22/48 \\
 P(X_6 = \text{Status} | \text{Diterima}) &= 48/100 \\
 P(X | \text{Diterima}) &= 48/100 * 4/48 * 36/48 * 10/48 * 8/48 * 22/48 = 0,000994647 \\
 P(X = \text{Tidak} | \text{Tidak}) / X &= 52/100 \\
 P(X_1 = \text{Pekerjaan Orang Tua} | \text{Tidak}) &= 9/52 \\
 P(X_2 = \text{Penghasilan Orang Tua} | \text{Tidak}) &= 37/52 \\
 P(X_3 = \text{Jumlah Tanggungan} | \text{Tidak}) &= 13/52 \\
 P(X_4 = \text{Kepemilikan Rumah} | \text{Tidak}) &= 7/52 \\
 P(X_5 = \text{Prestasi} | \text{Tidak}) &= 10/52 \\
 P(X_6 = \text{Status} | \text{Tidak}) &= 52/100 \\
 P(X | \text{Tidak}) &= 52/100 * 9/52 * 37/52 * 13/52 * 7/52 * 10/52 = 0,001189267
 \end{aligned}$$

Karena $P(X | \text{Tidak Diterima})$ lebih besar dibandingkan dengan $P(X | \text{Diterima})$, maka hasil seleksi penerima beasiswa untuk mahasiswa dengan Nama Yoseph Silvance dinyatakan Tidak Diterima.

- 2) Mahasiswa dengan Nama Sutanto memiliki atribut dengan Pekerjaan Orang Tua Petani, Penghasilan Orang Tua Rendah, Jumlah Tanggungan Banyak, Kepemilikan Rumah Sendiri, Prestasi Tidak Ada, dan Status Tidak Diterima.

$$\begin{aligned}
 P(X = \text{Diterima} | \text{Diterima}) / X &= 48/100 \\
 P(X_1 = \text{Pekerjaan Orang Tua} | \text{Diterima}) &= 4/48 \\
 P(X_2 = \text{Penghasilan Orang Tua} | \text{Diterima}) &= 36/48 \\
 P(X_3 = \text{Jumlah Tanggungan} | \text{Diterima}) &= 10/48
 \end{aligned}$$

$$\begin{aligned}
 P(X_4 = \text{Kepemilikan Rumah}|\text{Diterima}) &= 10/48 \\
 P(X_5 = \text{Prestasi}|\text{Diterima}) &= 26/48 \\
 P(X_6 = \text{Status}|\text{Diterima}) &= 48/100 \\
 P(X|\text{Diterima}) &= 48/100 * 4/48 * 36/48 * 10/48 * 10/48 * 26/48 = 0,001469365
 \end{aligned}$$

$$\begin{aligned}
 P(X = \text{Tidak}|\text{Tidak}) / X &= 52/100 \\
 P(X_1 = \text{Pekerjaan Orang Tua}|\text{Tidak}) &= 9/52 \\
 P(X_2 = \text{Penghasilan Orang Tua}|\text{Tidak}) &= 37/52 \\
 P(X_3 = \text{Jumlah Tanggungan}|\text{Tidak}) &= 13/52 \\
 P(X_4 = \text{Kepemilikan Rumah}|\text{Tidak}) &= 22/52 \\
 P(X_5 = \text{Prestasi}|\text{Tidak}) &= 42/52 \\
 P(X_6 = \text{Status}|\text{Tidak}) &= 52/100 \\
 P(X|\text{Tidak}) &= 52/100 * 9/52 * 37/52 * 13/52 * 22/52 * 42/52 = 0,015698327
 \end{aligned}$$

Karena $P(X|\text{Tidak})$ lebih besar dibandingkan dengan $P(X|\text{Diterima})$, maka hasil seleksi penerima beasiswa untuk mahasiswa dengan Nama Sutanto dinyatakan Tidak diterima.

- 3) Mahasiswa dengan Nama Lia Uswatun Nuha memiliki atribut dengan Pekerjaan Orang Tua Tidak Bekerja, Penghasilan Orang Tua Rendah, Jumlah Tanggungan Sedikit, Kepemilikan Rumah Tidak Memiliki, Prestasi Tidak Ada, dan Status Diterima.

$$\begin{aligned}
 P(X = \text{Diterima}|\text{Diterima}) / X &= 48/100 \\
 P(X_1 = \text{Pekerjaan Orang Tua}|\text{Diterima}) &= 5/48 \\
 P(X_2 = \text{Penghasilan Orang Tua}|\text{Diterima}) &= 36/48 \\
 P(X_3 = \text{Jumlah Tanggungan}|\text{Diterima}) &= 38/48 \\
 P(X_4 = \text{Kepemilikan Rumah}|\text{Diterima}) &= 23/48 \\
 P(X_5 = \text{Prestasi}|\text{Diterima}) &= 22/48 \\
 P(X_6 = \text{Status}|\text{Diterima}) &= 48/100 \\
 P(X|\text{Diterima}) &= 48/100 * 5/48 * 36/48 * 38/48 * 23/48 * 22/48 = 0,01605
 \end{aligned}$$

$$\begin{aligned}
 P(X = \text{Tidak}|\text{Tidak}) / X &= 52/100 \\
 P(X_1 = \text{Pekerjaan Orang Tua}|\text{Tidak}) &= 6/52 \\
 P(X_2 = \text{Penghasilan Orang Tua}|\text{Tidak}) &= 37/52 \\
 P(X_3 = \text{Jumlah Tanggungan}|\text{Tidak}) &= 39/52 \\
 P(X_4 = \text{Kepemilikan Rumah}|\text{Tidak}) &= 22/52 \\
 P(X_5 = \text{Prestasi}|\text{Tidak}) &= 42/52 \\
 P(X_6 = \text{Status}|\text{Tidak}) &= 52/100 \\
 P(X|\text{Tidak}) &= 52/100 * 6/52 * 37/52 * 39/52 * 22/52 * 42/52 = 0,03140
 \end{aligned}$$

Karena $P(X|\text{Tidak})$ lebih besar dibandingkan dengan $P(X|\text{Diterima})$, maka hasil seleksi penerima beasiswa untuk mahasiswa dengan Nama Lia Uswatun Nuha dinyatakan Tidak Diterima.

- 4) Mahasiswa dengan Nama Ana Alfiana memiliki atribut dengan Pekerjaan Orang Tua Lainnya, Penghasilan Orang Tua Rendah, Jumlah Tanggungan Sedikit, Kepemilikan Rumah Tidak Memiliki, Prestasi Tidak Ada, dan Status Diterima.

$$\begin{aligned}
 P(X = \text{Diterima}|\text{Diterima}) / X &= 48/100 \\
 P(X_1 = \text{Pekerjaan Orang Tua}|\text{Diterima}) &= 21/48 \\
 P(X_2 = \text{Penghasilan Orang Tua}|\text{Diterima}) &= 36/48 \\
 P(X_3 = \text{Jumlah Tanggungan}|\text{Diterima}) &= 38/48 \\
 P(X_4 = \text{Kepemilikan Rumah}|\text{Diterima}) &= 23/48
 \end{aligned}$$

$$\begin{aligned}
P(X5 = Prestasi|Diterima) &= 22/48 \\
P(X6 = Status|Diterima) &= 48/100 \\
P(X|Diterima) &= 48/100 * 21/48 * 36/48 * 38/48 * 23/48 * 22/48 = 0,06742 \\
P(X = Tidak|Tidak) / X &= 52/100 \\
P(X1 = Pekerjaan Orang Tua|Tidak) &= 21/52 \\
P(X2 = Penghasilan Orang Tua|Tidak) &= 37/52 \\
P(X3 = Jumlah Tanggungan|Tidak) &= 39/52 \\
P(X4 = Kepemilikan Rumah|Tidak) &= 22/52 \\
P(X5 = Prestasi|Tidak) &= 42/52 \\
P(X6 = Status|Tidak) &= 52/100 \\
P(X|Tidak) &= 52/100 * 21/52 * 37/52 * 39/52 * 22/52 * 42/52 = 0,10989
\end{aligned}$$

Karena $P(X|Tidak)$ lebih besar dibandingkan dengan $P(X|Diterima)$, maka hasil seleksi penerima beasiswa untuk mahasiswa dengan Nama Ana Alfiana dinyatakan Tidak Diterima.

Hasil perhitungan mahasiswa lain menyesuaikan dan didapatkan hasil klasifikasi seleksi penerima beasiswa dari data uji menggunakan metode Naïve Bayes Classifier

e. Evaluasi (Evaluation)

Dari hasil perhitungan, kesimpulan yang di peroleh meliputi:

- 1) Hasil perhitungan dari status Diterima dengan hasil Diterima sebanyak 9 data dari 30 data, yang selanjutnya disebut true positive (TP).
- 2) Hasil perhitungan dari status Diterima dengan hasil Tidak Diterima sebanyak 10 data dari 30 data, yang selanjutnya disebut false positive (FP).
- 3) Hasil perhitungan dari status Tidak Diterima dengan hasil Diterima sebanyak 5 data dari 30 data, yang selanjutnya disebut false negative (FN).
- 4) Hasil perhitungan dari status Tidak Diterima dengan hasil Tidak Diterima sebanyak 6 data dari 30 data, yang selanjutnya disebut true negative (TN).

Hasil akurasi sistem dengan Naïve Bayes Classifier dihitung confusion matrix, sebagai berikut:

Tabel 2 Hasil confusion matriks

Status \ Hasil	Diterima	Tidak Diterima
Diterima	4	9
Tidak Diterima	1	21

$$Akurasi = \frac{TP + TN}{TP + FP + FN + TN}$$

$$Akurasi = \frac{4 + 21}{4 + 9 + 1 + 21}$$

$$Akurasi = 72\%$$

Hasil dari perhitungan dengan Naïve Bayes Classifier dalam mengklasifikasikan seleksi dalam menerima beasiswa di Universitas Stikubank Semarang adalah akurasi sebesar 72%. Dengan begitu hasil sistem pendukung keputusan ini sudah cukup baik, sehingga sistem dapat dijadikan acuan dalam proses seleksi berikutnya.

f. Deployment

Selanjutnya melakukan penilaian dengan detail terhadap hasil dari modelnya, lalu melakukan implementasi berdasarkan keseluruhan model yang ada. kemudian akan

disesuaikan pada model dan memberikan hasil disesuaikan pada sasaran awal dalam tahapan CRISP-DM.

4. Conclusion

Dalam penelitian ini, telah dilakukan proses penyeleksian penerimaan beasiswa menggunakan teknik Naïve Bayes classifier pada Universitas Stikubank Semarang. Dataset terdiri dari 8 atribut yaitu No. Pendaftaran, Nama, Penghasilan Orang Tua, Pekerjaan Orang Tua, Jumlah Tanggungan, Prestasi, Kepemilikan Rumah, dan Status. berdasarkan 135 data yang dikelompokkan ke dalam 100 data latih serta 35 data uji, mendapatkan nilai akurasi sebesar 72% dengan parameter uji Confusion Matrix. Artinya dari perolehan akurasi tersebut didapatkan sebanyak 25 dari 35 penerima beasiswa sudah sesuai dengan sistem pendukung keputusan yang dibuat. Dengan demikian metode naïve bayes classifier yang telah diterapkan sudah cukup baik dan sesuai dengan sistem pendukung keputusan dalam seleksi penerima beasiswa.

References

- Budiman, A., & Adi Prasetyo, M. H. (2019). Sistem Pendukung Keputusan Penerima Dana Desa Dengan Menggunakan Metode Promethee Di Kecamatan Godean Kabupaten Sleman. *Jurnal Informatika Interaktif*, 131-137.
- Harimurti, F. A. (2017). Klasifikasi Penerimaan Beasiswa Menggunakan Metode Naive Bayes Classifier (Studi Kasus Universitas Trunojoyo Madura). Surabaya: Tugas Akhir Departemen Sistem Informasi Institut Teknologi Sepuluh Nopember.
- Idris, M. (2019). Implementasi Data Mining Dengan Algoritma Naive Bayes Untuk Memprediksi Angka Kelahiran. *Jurnal Pelita Informatika*, 421-428.
- Kurnia, A. U., Budi, A. S., & Susilo, P. H. (2020). Sistem Pendukung Keputusan Penerimaan Beasiswa Menggunakan Metode Naive Bayes. *JOUTICA*, 5(2), 397-402.
- Perdana, A., Hasibuan, N. A., & Fadlina. (2022). Sistem Pendukung Keputusan Seleksi Penerima Beasiswa Berprestasi Pada Yayasan Pendidikan Jaya Untuk Anak Karyawan Dengan Menerapkan Metode ORESTE dan ROC (Rank Order Centroid). *Jurnal of Machine Learning and Data Analytics (MALDA)*, 17-26.
- Safitri, T. N. (2020). Penerapan Algoritma Naive Bayes Untuk Penentuan Calon Penerima Beasiswa Pada SD Negeri 6 Ketapang. *JUISI*, 43-52.
- Santoso, H., Azhar, R., Husain, & Muliadi. (2021). Sistem Pendukung Keputusan Penerimaan Calon Binaan Inkubator Wirausaha Menggunakan Metode Perbandingan Eksponensial. *Jurnal Bumigora Information Technology (BITE)*, 151-160.
- Wahyuningsih, D., & Patima, E. (2018). Penerapan Naive Bayes Untuk Penerimaan Beasiswa. *Jurnal Telematika*, 11(1), 135-147.
- Wicaksono, T. G. (2017). Sistem Pemilihan Penerima Beasiswa Bidik Misi Dengan Metode Fuzzy Tsukamoto. *Duta.com: Jurnal Ilmiah Teknologi Informasi dan Komunikasi*, 17-50.