



## JURNAL INFORMATIKA DAN TEKNOLOGI KOMPUTER

Halaman Jurnal: <https://journal.amikveteran.ac.id/index.php/jitek>  
Halaman UTAMA Jurnal : <https://journal.amikveteran.ac.id/index.php>



### IMPLEMENTASI ALGORITME NAÏVE BAYES UNTUK MENENTUKAN KELAYAKAN CALON PENERIMA BEASISWA

Hendri Noviyanto<sup>a</sup>, Bayu Mukti<sup>b</sup>

<sup>a</sup> Fakultas Teknik Elektro dan Informatika / Sistem Komputer, [hendrinoviyantoo@gmail.com](mailto:hendrinoviyantoo@gmail.com), Universitas Surakarta

<sup>b</sup> Fakultas Teknik Elektro dan Informatika / Teknik Informatika, [bayu.unsa@gmail.com](mailto:bayu.unsa@gmail.com), Universitas Surakarta

#### ABSTRACT

Scholarships are a form of assistance given to students who have good academic abilities or to their parents who cannot afford to pay for their education. In its application, the selection process for prospective scholarship recipients is carried out manually by comparing the applicant files. This form of selection is very time-consuming and therefore ineffective and inefficient. Another obstacle is that the decision maker has difficulty in deciding who will get the scholarship, and there are many other obstacles faced. The purpose of this study is to assist decision-makers in the selection process for prospective scholarship recipients easily and quickly. The dataset in this study uses eight attributes with 100 instances. The classification method used is the Naïve Bayes Algorithm. The validation process uses the split test technique. This method is compared with other methods such as C4.5 and KNN. The results of this research process, the proposed method can predict prospective scholarship recipients correctly as evidenced by an accuracy value of 90%. However, other algorithms also get the same accuracy value in the 80:20 split test, the accuracy obtained is different when the 70:30 split test score with nave Baye is ranked first. Seeing the results obtained, the Naïve Bayes algorithm can be applied to a decision support system to determine prospective scholarship recipients properly.

**Keywords:** scholarships, data mining, eligibility, naïve bayes, predictions.

#### Abstrak

Beasiswa merupakan sebuah bentuk bantuan yang diberikan kepada mahasiswa yang memiliki kemampuan akademik baik atau bagi orang tuanya yang tidak mampu dalam membiayai pendidikan. Dalam pengaplikasiannya proses seleksi calon penerima beasiswa dilakukan secara manual dengan membandingkan antar berkas pendaftar. Bentuk seleksi yang demikian sangat menguras pikiran dan waktu sehingga tidak efektif dan efisien. Kendala lain adalah pihak pemberi keputusan kesulitan dalam memutuskan siapa saja yang akan mendapatkan beasiswa, dan masih banyak lagi kendala yang dihadapi. Tujuan penelitian ini adalah untuk membantuk pihak pengambil keputusan dalam proses seleksi calon penerima beasiswa dengan mudah dan cepat. Dataset pada penelitian ini menggunakan delapan buah atribut dengan 100 instance. Metode klasifikasi yang digunakan adalah Algoritme Naïve Bayes. Proses validasi menggunakan teknik splittest. Metode ini dibandingkan dengan metode lain seperti C4.5 dan KNN. Hasil dari proses penelitian ini, metode yang diajukan dapat memprediksi calon penerima beasiswa dengan tepat yang dibuktikan dengan nilai akurasi sebesar 90%. Namun, algoritme lain juga mendapatkan nilai akurasi yang sama pada splittest 80:20, akurasi yang didapatkan berbeda ketika nilai splittest 70:30 dengan naïve baye menduduki peringkat pertama. Melihat hasil yang diperoleh, algoritme Naïve Bayes dapat diaplikasikan ke dalam sistem pendukung keputusan untuk menentukan calon penerima beasiswa dengan baik.

**Kata Kunci:** beasiswa, data mining, kelayakan, naïve bayes, prediksi.

#### 1. PENDAHULUAN

Pendidikan merupakan sebuah sarana untuk meningkatkan taraf hidup seseorang, melalui sebuah pendidikan seseorang dapat berkembang dengan baik dan terarah. Sebuah pendidikan memerlukan ruang dan fasilitas untuk menunjang pengetahuan seperti sebuah perguruan tinggi. Perguruan tinggi memberikan fasilitas kepada para mahasiswanya untuk terus berkembang dalam meningkatkan kemampuan soft skill

maupun hard skill. Namun, belum semua mahasiswa mampu menikmati pendidikan yang sesuai dengan baik dan nyaman. Hal ini disebabkan dari beberapa keadaan latar belakang mahasiswa, salah satunya adalah faktor ekonomi. Faktor ekonomi dapat menyebabkan seorang mahasiswa atau penuntut ilmu harus putus dengan dunia pendidikan dikarenakan tidak adanya biaya.

Pada saat ini, pemerintah telah meluncurkan Kartu Indonesia Pintar Mahasiswa (KIP) yang sangat bermanfaat untuk para kalangan pelajar guna meringankan biaya pendidikan [1]. Kendati demikian, karena banyaknya jumlah warga negara Indonesia, pemerintahan dalam melakukan seleksi terhadap penerima KIP masih banyak kendala, termasuk masih terdapat golongan kurang mampu tidak mendapatkan KIP. Hal ini, mengharuskan Perguruan Tinggi menyiapkan dana untuk membantu mahasiswa yang masuk ke dalam kategori kurang mampu namun memiliki kualitas baik guna mendapatkan bantuan biaya pendidikan atau beasiswa.

Beasiswa merupakan sebuah solusi untuk mengatasi masalah biaya bagi mereka yang kurang mampu [2][3]. Sehingga diharapkan dengan adanya beasiswa dapat menciptakan pengaruh yang positif bagi mereka yaitu tidak lagi perlu khawatir dengan tagihan biaya pendidikan [4]. Beasiswa yang dilakukan secara selektif akan mendapatkan hasil yang maksimal, seperti penerima beasiswa akan bersungguh-sungguh dalam melaksanakan pembelajaran.

Salah satu permasalahan dalam penyaluran dana beasiswa adalah kurang tepatnya sasaran pemberian dana beasiswa, sehingga menyebabkan tidak maksimal dalam pendayagunaan dana beasiswa. Hal lain yang perlu diperhatikan adalah proses seleksi penerimaan beasiswa secara manual juga tidak efektif dan efisien. Proses data mining akan sangat membantu dalam proses penentuan calon penerima beasiswa. Data Mining (DM) merupakan sebuah teknik pengolahan data dengan memanfaatkan beberapa data yang cukup banyak. DM akan menggali informasi pada data dan menghasilkan *Knowledge Discovery in Database* (KDD) sebagai bahan pertimbangan keputusan [5].

Penelitian terkait dengan proses seleksi calon penerima beasiswa sudah sangat banyak dilakukan oleh beberapa peneliti. Seperti penelitian [6] membuat sebuah sistem pendukung keputusan untuk menentukan penerima beasiswa Program Indonesia Pintar pada taraf SD, SMP dan SMA. Penelitian ini memanfaatkan teknik clustering dengan memanfaatkan algoritme K-Means. Hasilnya adalah 2 kluster, yaitu kluster penerima beasiswa tetap sasaran dan kluster penerima beasiswa tidak tepat sasaran. Penelitian [4] membuat sebuah sistem informasi untuk seleksi penerimaan beasiswa di fakultas teknik menggunakan metode *Simple Additive Weighting* (SAW) dengan teknik testing menggunakan blacbox. Tingkat akurasi yang didapatkan menggunakan teknik ini sebesar 62%. Penelitian [1] menggunakan algoritme Decision Tree (C4.5 untuk melakukan prediksi seleksi beasiswa menggunakan jalur KIP. Nilai akurasi yang didapatkan menggunakan algoritme C4.5 sebesar 100%, dan Naïve Bayes sebesar 90,16%. Penelitian [2] menggunakan algoritme Naïve Bayes dengan perolehan nilai akurasi sebesar 90,48% untuk memprediksi calon siswa untuk mendapatkan beasiswa. Penelitian [3] dan [7] menyeleksi calon penerima beasiswa menggunakan algoritme C4.5 dengan mencari nilai entropy dan gain untuk membentuk pohon keputusan. Penelitian [8] menggunakan metode *fuzzy multiple attribute decision making* (FMADM) dengan *weighted product* untuk memberikan bobot setiap atribut yang akan dilanjutkan dengan pemeringkatan yang akan menyeleksi alternatif dan sejumlah alternatif. Penelitian [9] melakukan seleksi beasiswa menggunakan *Simple Additive Weighting* (SAW) dengan perolehan nilai akurasi sebesar 67%. Dari beberapa metode yang digunakan oleh para peneliti terdahulu, penelitian ini menggunakan metode *Naïve Bayes* yang akan dibandingkan dengan metode *Decision Trees*, dan KNN.

## 2. TINJAUAN PUSTAKA

### 2.1. Prediksi

Prediksi merupakan sebuah proses memperkirakan sesuatu secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi masa lalu dan sekarang yang dimiliki supaya selisih kesalahan bisa diperkecil [10]. Menurut kamus besar Bahasa Indonesia, prediksi adalah kegiatan memprediksi atau meramal atau memperkirakan nilai pada masa yang akan datang menggunakan data masa lalu.

## 2.2. Naïve Bayes

Naïve Bayes merupakan sebuah metode probabilitas untuk proses pengklasifikasian secara sederhana yang dilakukan dari data training ke sejumlah data secara efisien. Teorema Bayes dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal dengan teorema Bayes [11].

Berikut ini adalah rumus persamaan teorema naïve Bayes.

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)} \quad (1)$$

Persamaan diatas merupakan rumus dasar metode naïve Bayes.

## 3. METODOLOGI PENELITIAN

### 3.1. Dataset

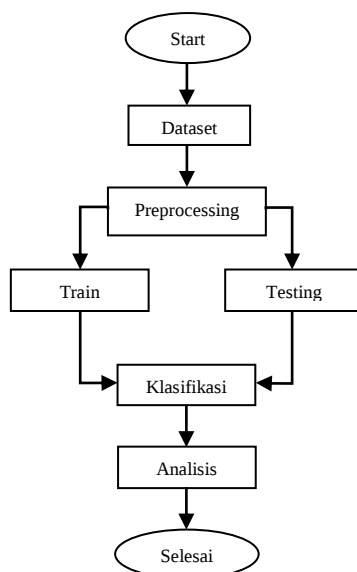
Pada penelitian ini dataset didapatkan dari pangkalan data Universitas Surakarta. Data yang digunakan adalah data mahasiswa semester 2 pada jenjang S1. Data yang digunakan sebesar 100 mahasiswa dengan 8 buah atribut yang digunakan. Atribut meliputi Jenis Kelamin, IPK, Jenkel Wali, Usia Wali, Pekerjaan Wali, Kepemilikan Rumah Wali, Penghasilan Wali, Jumlah Tanggungan Wali.

Tabel 1. Atribut dan Value

| No | Atribut                | Value                                       | Keterangan                                                                                          |
|----|------------------------|---------------------------------------------|-----------------------------------------------------------------------------------------------------|
| 1  | Jenis Kelamin          | Laki   Perempuan                            |                                                                                                     |
| 2  | IPK                    | Baik   Cukup   Kurang                       | Baik 3.5 – 4.0   Cukup 2.8 – 3.4   Kurang 1.0 – 2.7                                                 |
| 3  | Jenkel Wali            | Laki   Perempuan                            |                                                                                                     |
| 4  | Usia Wali              | Tua   Muda                                  | Tua $\geq 50^{th}$   Muda $\leq 50^{th}$                                                            |
| 5  | Pekerjaan Wali         | PNS   Pegawai Swasta   Pedagang   Serabutan |                                                                                                     |
| 6  | Kepemilikan Rumah Wali | Rumah Sendiri   Sewa   Menumpang            |                                                                                                     |
| 7  | Penghasilan Wali       | Rendah   Cukup   Sedang   Besar             | Rendah $\leq 1.500.000$   Cukup $\leq 2.500.000$   Sedang $\leq 4.000.000$   Besar $\geq 8.000.000$ |
| 8  | Jumlah Tanggungan Wali | Sedikit   Sedang   Banyak                   | Sedikit 1   Sedang $\leq 3$   $\geq 5$ orang                                                        |

### 3.2. Alur Penelitian

Pada penelitian ini, terdiri dari beberapa proses yang dilakukan seperti data selection, *preprocessing*, proses klasifikasi menggunakan algoritme *Naïve Bayes*. Langkah terakhir adalah proses analisis kinerja hasil dari klasifikasi. Gambaran umum alur penelitian ini dapat dilihat pada Gambar 1 di bawah ini.



Gambar 1. Alur Penelitian

Menurut Gambar 1 Alur penelitian, langkah pertama adalah pengumpulan dataset. Pada penelitian ini dataset diperoleh dari Universitas Surakarta. Langkah kedua adalah *preprocessing*, pada fase ini data diolah supaya siap untuk dilakukan pemrosesan atau pengklasifikasian. Data yang telah melewati preprocessing akan dibagi menjadi data training dan testing sebesar 80:20. Proses klasifikasi pada penelitian ini menggunakan algoritme Naïve Bayes yang dibantu software Rapidminer. Proses analisis dilakukan untuk mengetahui tingkat kesuksesan algoritme dalam mengklasifikasi data dengan baik. Analisis dilihat dari tingkat akurasi yang didapatkan.

#### 4. HASIL DAN PEMBAHASAN

Pada penelitian ini menggunakan software Rapidminer untuk membantu proses prediksi calon penerima beasiswa. Proses validasi data menggunakan teknik split data 70:30. 70% digunakan untuk data training dan 30% digunakan sebagai data testing. Dataset menggunakan data dari Universitas Surakarta dengan jumlah data yang digunakan sebesar 100 instances dan 8 attribute. Proses pengujian dari metode yang diusulkan akan dikomparasikan dengan beberapa metode seperti Decision Tree (C4.5), KNN dan Support Vector Machine.

Tabel 2. Hasil Proses Klasifikasi

|         | Algoritme   |        |     |
|---------|-------------|--------|-----|
|         | Naïve Bayes | C4.5   | KNN |
| Akurasi | 90%         | 86,67% | 80% |

Dari tabel 2 diatas, dapat kita ketahui bahwa nilai akurasi yang didapatkan algoritme Naïve Bayes lebih baik yaitu 90%, disusul oleh C4.5 86.67% dan K-NN 80%. Untuk data training 70% Naïve Bayes lebih unggul dibandingkan dengan algoritme pembandingnya.

Pada tabel 3 merupakan hasil uji coba menggunakan data yang displit menjadi 80:20, yaitu 80% untuk data training dan 20% untuk data testing. Hasilnya cukup bagus karena terdapat beberapa perubahan pada nilai akurasi yang didapatkan. Nilai akurasi yang didapatkan adalah 90% untuk ketiga algoritme yang digunakan. Hal ini dapat disimpulkan bahwa semakin banyak data training maka kemampuan algoritme dalam mengklasifikasikan juga semakin baik. Namun algoritme masih menjadi algoritme yang paling unggul karena nilai akurasinya sudah cukup bagus meskipun dengan data training yang tidak begitu banyak.

| Tabel 3. Hasil Proses Klasifikasi |             |      |     |
|-----------------------------------|-------------|------|-----|
|                                   | Algoritme   |      |     |
|                                   | Naïve Bayes | C4.5 | KNN |
| Akurasi                           | 90%         | 90%  | 90% |

Hasil yang ditampilkan pada tabel 3 memiliki kesamaan nilai akurasi, namun ketika dilihat pada tabel confusion matrik terdapat perbedaan dalam pengelompokannya meskipun nilai yang didapatkan tetap sama. Hasil confusion matrik dapat dilihat pada gambar 2, 3 dan 4.

#### PerformanceVector

```
PerformanceVector:
accuracy: 90.00%
ConfusionMatrix:
True:  Bukan Penerima  Penerima
Bukan Penerima: 6      0
Penerima:      2      12
```

Gambar 2. NB

#### PerformanceVector

```
PerformanceVector:
accuracy: 90.00%
ConfusionMatrix:
True:  Bukan Penerima  Penerima
Bukan Penerima: 8      2
Penerima:      0      10
```

Gambar 3. C4.5

#### PerformanceVector

```
PerformanceVector:
accuracy: 90.00%
ConfusionMatrix:
True:  Bukan Penerima  Penerima
Bukan Penerima: 7      1
Penerima:      1      11
```

Gambar 4. K-NN

Pada Gambar 2 terdapat 2 kesalahan dalam mengklasifikasikan bukan penerima beasiswa, sedangkan pada Gambar 3 terdapat 2 kesalahan dalam mengklasifikasikan penerima beasiswa dan pada Gambar 4 terdapat masing-masing 1 kesalahan pada proses pengklasifikasian.

| Row No. | Kelas          | prediction(K... | confidence(Bukan ... | confidence(Pene... | IPK    | Jenkel Wali | Usia Wali | Peko  |
|---------|----------------|-----------------|----------------------|--------------------|--------|-------------|-----------|-------|
| 1       | Bukan Penerima | Bukan Penerima  | 1                    | 0                  | Baik   | Laki        | Muda      | PNS   |
| 2       | Penerima       | Penerima        | 0.024                | 0.976              | Baik   | Laki        | Muda      | Peg.  |
| 3       | Penerima       | Bukan Penerima  | 1                    | 0                  | Baik   | Perempuan   | Tua       | Serd. |
| 4       | Penerima       | Penerima        | 0.024                | 0.976              | Baik   | Laki        | Muda      | Peg.  |
| 5       | Penerima       | Penerima        | 0.024                | 0.976              | Baik   | Laki        | Muda      | Serd. |
| 6       | Penerima       | Penerima        | 0.024                | 0.976              | Cukup  | Laki        | Muda      | Peg.  |
| 7       | Bukan Penerima | Bukan Penerima  | 1                    | 0                  | Baik   | Perempuan   | Tua       | PNS   |
| 8       | Penerima       | Penerima        | 0.024                | 0.976              | Kurang | Laki        | Tua       | Peg.  |
| 9       | Penerima       | Penerima        | 0.024                | 0.976              | Cukup  | Laki        | Muda      | Peg.  |
| 10      | Bukan Penerima | Bukan Penerima  | 1                    | 0                  | Baik   | Perempuan   | Tua       | PNS   |
| 11      | Penerima       | Penerima        | 0.024                | 0.976              | Baik   | Laki        | Tua       | Serd. |
| 12      | Penerima       | Penerima        | 0.024                | 0.976              | Baik   | Laki        | Muda      | Peg.  |

Gambar 5. Hasil Prediksi

Pada Gambar 5 terdapat kelas dan hasil prediksi yang dibandingkan, dari gambar tersebut kita dapat mengetahui prediksi mana saja yang mengalami kesalahan. Hal ini membuat peneliti dapat lebih detail dalam mengevaluasi hasil prediksi.

## 5. KESIMPULAN DAN SARAN

Pada penelitian ini, kita menggunakan Algoritme Naïve Bayes sebagai metode yang diusulkan. Metode ini kemudian dibandingkan dengan metode lain seperti C4.5 dan k-NN untuk melihat performa algoritme tersebut (lihat tabel 2). Naïve Bayes memiliki kelebihan dalam mengklasifikasikan data lebih baik dari algoritme pesaingnya meskipun dengan sedikit data. Namun ketika data yang digunakan untuk proses training data cukup banyak, nilai akurasi dari ketiga algoritme yang digunakan meningkat kecuali Naïve Bayes. Melihat hasil penelitian ini, algoritme Naïve Bayes dapat diaplikasikan ke dalam sistem pendukung keputusan untuk mendapatkan calon penerima beasiswa lebih baik. Pada penelitian yang akan datang, peneliti akan mengevaluasi atribut yang digunakan untuk mendapatkan nilai akurasi yang lebih baik.

## DAFTAR PUSTAKA

- [1] K. Khotimah, "Teknik Data Mining menggunakan Algoritma Decision Tree (C4.5) untuk Prediksi Seleksi Beasiswa Jalur KIP pada Universitas Muhammadiyah Kotabumi," J. SIMADA (Sistem Inf.

- dan Manaj. Basis Data), vol. 4, no. 2, pp. 145–152, 2021, doi: 10.30873/simada.v4i2.3064.
- [2] Wahyuningsih, Budiman, and I. Umami, “Implementasi Algoritma Naïve Bayes Untuk Menentukan Calon Penerima Beasiswa di SMK YPM 14 Sumobito Jombang,” vol. 4, no. 1, pp. 1–6, 2021.
  - [3] N. Islamuddin, “Data Mining Untuk Seleksi Kelayakan Calon Penerima Beasiswa di STMIK Bina Bangsa Kendari,” *Simkom*, vol. 4, no. 2, pp. 9–20, 2019, doi: 10.51717/simkom.v4i2.30.
  - [4] K. Irianti, A. Y. Husodo, B. Irmawati, P. Studi, T. Informatika, and U. Mataram, “SISTEM INFORMASI SELEKSI PENERIMAAN BEASISWA DI FAKULTAS TEKNIK - UNIVERSITAS MATARAM ( Information System for Scholarship Selection In Faculty of Engineering - University of Mataram ),” vol. 1, no. 2, pp. 149–157, 2019.
  - [5] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining*, Third Edit. United States: Elsevier, 2011.
  - [6] Darlinda and N. Utamajaya, “Sistem Pendukung Keputusan Penerima Beasiswa Program Indonesia Pintar Menggunakan Metode Algoritma K-Means Clustering,” *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 2, pp. 167–175, 2022, doi: 10.30865/jurikom.v9i2.3971.
  - [7] H. Yutanto and N. Setiawan, “Penerapan Data Mining Sebagai Model Seleksi Penerima Beasiswa Penuh ( Studi Kasus : Stie Perbanas Surabaya ),” *J. link*, vol. 27, no. 1, pp. 8–12, 2018.
  - [8] P. Oktavia, “Sistem Pendukung Keputusan Seleksi Penerima Beasiswa dengan Metode Weighted Product pada SMP Negeri 1 Parung Berbasis Web,” *J. Inform. Univ. Pamulang*, vol. 3, no. 2, p. 80, 2018, doi: 10.32493/informatika.v3i2.1432.
  - [9] F. N. Khasanah and S. Rofiah, “Sistem Seleksi Penerimaan Beasiswa Menggunakan Metode Pendukung Keputusan Simple Additive Weighting,” *Semin. Nas. APTIKOM*, pp. 118–125, 2019.
  - [10] B. Santosa, “Tutorial Support Vector Machine 1 Ide Dasar Support Vector Machine,” pp. 1–23, 2011.
  - [11] I. Pangaribuan, “Penerapan Algoritma Naive Bayes Untuk Seleksi Beasiswa Pemprov Dan Bawaku Di Universitas Komputer Indonesia Application of Naive Bayes Algorithm for Selection of Providing and Bawaku Scholarship in University Computer Indonesia,” 2019.