

PENERAPAN NAIVE BAYES UNTUK PENERIMAAN BEASISWA

Delpiah Wahyuningsih¹, Eli Patima²

^{1,2}Program Teknik Informatika
STMIK Atma Luhur Pangkalpinang
Email: delphibabel@atamaluhur.ac.id¹, elycst@gmail.com²

ABSTRAK

SMA Negeri 4 Pangkalpinang adalah salah satu sekolah negeri yang memiliki program beasiswa. Beasiswa yang ada di SMA Negeri 4 Pangkalpinang terbagi menjadi 2 yaitu Beasiswa Prestasi dan Beasiswa Kurang Mampu. Dalam memberikan beasiswa harus dilakukan proses secara selektif agar sesuai dengan jenis beasiswa yang didapat. Setiap beasiswa memiliki syarat-syarat atau ketentuan yang harus dipenuhi. Proses penerimaan beasiswa yang ada di SMA Negeri 4 Pangkalpinang masih bersifat inefisien dikarenakan dalam proses penerimaan beasiswa hasil seleksi kurang akurat karena hanya mengandalkan pengamatan dengan indera penglihatan (visual) dan apabila kondisi lelah maka rentan terjadi kesalahan (human error). Tujuan dari penelitian ini adalah menganalisis penerapan algoritma naive bayes untuk perancangan sistem penerimaan beasiswa pada SMA Negeri 4 Pangkalpinang, mengetahui kelemahan yang ada dalam proses penerapan algoritma naive bayes untuk penerimaan beasiswa pada SMA Negeri 4 Pangkalpinang. Naive Bayes merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari data set yang diberikan. Dalam penelitian ini akan menghasilkan sebuah Penerapan Algoritma Naive Bayes Untuk Penerimaan Beasiswa Pada SMA Negeri 4 Pangkalpinang. Hasil dari pengimplementasian sistem ini dapat memudahkan pihak sekolah dalam proses penerimaan beasiswa.

Kata kunci: Algoritma *Naive Bayes*, Penerimaan Beasiswa, Sistem Penunjang Keputusan

ABSTRACT

SMA Negeri 4 Pangkalpinang is one of the public schools that have a scholarship program. The scholarship in SMA Negeri 4 Pangkalpinang is divided into 2, namely Achievement Scholarship and Undergraduate Scholarship. In giving scholarships a selective process must be done to fit the type of scholarship earned. Each scholarship has terms or conditions that must be met. The process of acceptance of scholarship in SMA Negeri 4 Pangkalpinang still inefisien because in the process of acceptance of scholarship result of selection less accurate because only relying on observation with senses of sight (visual) and if condition tired then susceptible error (human error). The purpose of this research is to analyze the application of naive bayes algorithm for the design of scholarship receiving system at SMA Negeri 4 Pangkalpinang, to know the

weaknesses that exist in the process of applying naive bayes algorithm for scholarship acceptance at SMA Negeri 4 Pangkalpinang. Naive Bayes is a simple probabilistic classifier that computes a set of probabilities by summing the frequency and value combinations of the given data sets. In this research will produce an Application of Naive Bayes Algorithm for Scholarship Acceptance at SMA Negeri 4 Pangkalpinang. The results of the implementation of this system can facilitate the school in the process of receiving scholarships.

Keyword: Naive Bayes Algorithm, Scholarship Reception, Decision Support System

PENDAHULUAN

Beasiswa adalah pemberian berupa bantuan keuangan yang diberikan kepada perorangan yang bertujuan untuk digunakan demi keberlangsungan pendidikan yang ditempuh. Pemberian beasiswa dilakukan untuk membantu siswa dalam meraih cita-cita dan dengan harapan siswa tersebut memperoleh prestasi yang baik, selain itu juga untuk membantu siswa yang kurang mampu. Dalam memberikan beasiswa harus dilakukan proses secara selektif agar sesuai dengan jenis beasiswa yang didapat. Setiap beasiswa memiliki syarat – syarat atau ketentuan yang harus dipenuhi. Setiap sekolah ada beasiswa yang diberikan kepada siswa-siswa, salah satu sekolah yang memberikan beasiswa yaitu SMA N 4 Pangkalpinang.

Beasiswa yang ada di SMA N 4 Pangkalpinang terbagi menjadi 2 yaitu Beasiswa Prestasi dan Beasiswa Kurang Mampu. Beasiswa Prestasi dilakukan melalui program KIP yaitu Kartu Indonesia Pintar dan untuk siswa yang kurang mampu melalui SKTM (Surat Keterangan Tidak Mampu). Proses penerimaan beasiswa yang ada di SMA N 4 Pangkalpinang masih bersifat *inefisien* dikarenakan dalam proses penerimaan beasiswa hasil seleksi kurang akurat karena hanya mengandalkan pengamatan dengan indera penglihatan (*visual*) dan apabila kondisi lelah maka rentan terjadi kesalahan (*human error*).

Dengan perkembangan teknologi komputer yang pesat dapat memudahkan dalam hal melakukan pekerjaan, selain itu manfaatnya telah dirasakan oleh semua kalangan dari anak kecil sampai lanjut usia. Perkembangan teknologi komputer khususnya teknologi informasi dalam *android* dapat membentuk suatu aplikasi

untuk pendidikan khususnya beasiswa yang menjadi suatu revolusi dibidang informasi berbasis teknologi internet dan dapat dijadikan *alternatif* bagi pengembangan sistem informasi dengan biaya rendah.

Naive Bayes merupakan salah satu algoritma yang melakukan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris *Thomas Bayes* dalam memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai *Teorema Bayes*. Dengan menggunakan algoritma *naive bayes* dapat mencegah terjadinya kesalahan dalam melakukan perhitungan penerimaan beasiswa dan dapat menghasilkan perhitungan yang akurat (Rahman, Antony Anwari dan Suryanto Agus).

METODE PENELITIAN

Naive Bayes

Naive Bayes merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari data set yang diberikan. *Naive Bayes* merupakan pengklarifikasi dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris *Thomas Bayes*, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya. Metode *Naive Bayes* juga dinilai berpotensi baik dalam mengklasifikasikan dokumen dibandingkan dengan metode pengklasifikasian lain dalam hal akurasi dan efisiensi komputasi. (Bustami)

persamaan dari *teorema Bayes* adalah :

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

Dimana :

X : Data dengan class yang belum diketahui

H : Hipotesis data merupakan suatu class yang spesifik

P(H|X) : Probabilistik hipotesis H berdasar kondisi X (posteriori probabilistik)

P(H) : Probabilistik hipotesis H (prior probabilitas)

P(X|H) : Probabilistik hipotesis X berdasar kondisi pada hipotesis H

P(X) : Probabilitas X

Untuk menjelaskan teorima *Naive Bayes*, perlu diketahui bahwa proses klarifikasi memerlukan sejumlah petunjuk untuk menentukan kelas apa yang cocok bagi sampel yang dianalisis tersebut. Karena itu, teorema *Bayes* di atas disesuaikan sebagai berikut :

$$P(C|F_1 \dots F_n) = \frac{P(C)P(F_1 \dots F_n|C)}{P(F_1 \dots F_n)} \quad (2)$$

Dimana variabel *C* mempresentasikan kelas, sementara variabel *F1 ... Fn* mempresentasikan karakteristik petunjuk yang dibutuhkan untuk melakukan klasifikasi. Maka rumus tersebut menjelaskan bahwa peluang masuknya sampel karakteristik tertentu dalam kelas *C* (*Posterior*) adalah peluang munculnya kelas *C* (sebelum masuknya sampel tersebut, seringkali disebut *prior*), dikali dengan peluang kemunculan karakteristik – karakteristik sampel pada kelas *C* (disebut juga *likelihood*), dibagi dengan peluang kemunculan karakteristik – karakteristik sampel secara global (disebut juga *vidence*). Karena itu, rumus diatas dapat pula ditulis secara sederhana sebagai berikut :

$$Posterior = \frac{Prior \times likelihood}{vidence} \quad (3)$$

Nilai *Evidence* selalu tetap untuk setiap kelas pada satu sampel. Nilai dari *posterior* tersebut nantinya akan dibandingkan dengan nilai - nilai *posterior* kelas lainnya untuk menentukan ke kelas apa suatu sampel akan diklasifikasikan. Penjabaran lebih lanjut rumus *Bayes* tersebut dilakukan dengan menjabarkan (*C/F1, ... ,Fn*) menggunakan aturan perkalian sebagai berikut:

$$\begin{aligned} P(C|F_1, \dots, F_n) &= P(C) P(F_1, \dots, F_n|C) \\ &= P(C)P(F_1|C)P(F_2, \dots, F_n|C, F_1) \\ &= P(C)P(F_1|C)P(F_2|C, F_1)P(F_3, \dots, F_n|C, F_1, F_2) \\ &= P(C)P(F_1|C)P(F_2|C, F_1)P(F_3|C, F_1, F_2), P(F_4, \dots, F_n|C, F_1, F_2, F_3) \\ &= P(C)P(F_1|C)P(F_2|C, F_1)P(F_3|C, F_1, F_2) \dots P(F_n|C, F_1, F_2, F_3, \dots, F_{n-1}) \end{aligned} \quad (4)$$

Dapat dilihat bahwa silpenjabaran tersebut menyebabkan semakin banyak dan semakin kompleksnya factor – factor syarat yang mempengaruhi nilai probabilitas, yang hamper mustahil untuk dianalisa satu persatu. Akibatnya, perhitungan tersebut menjadi sulit untuk dilakukan. Disinilah digunakan asumsi independensi yang sangat tinggi (*naif*), bahwa masing- masing petunjuk (F1, F2, ..., Fn) saling bebas (*independen*) satu sama lain. Dengan asumsi tersebut, maka berlaku suatu kesamaan sebagai berikut:

$$P(F_i|F_j) = \frac{P(F_i \cap F_j)}{P(F_j)} = \frac{P(F_i)P(F_j)}{P(F_j)} = P(F_i)$$

Untuk $i \neq j$, sehingga

$$P(F_i|C, F_j) = P(F_i|C) \tag{5}$$

Dari persamaan di atas dapat disimpulkan bahwa asumsi independen sinaif tersebut membuat syarat peluang menjadi sederhana, sehingga perhitungan menjadi mungkin untuk dilakukan. Selanjutnya, penjabaran P (C/F1, ..., Fn) dapat disederhanakan menjadi :

$$\begin{aligned} P(C|F_1, \dots, F_n) &= P(C)P(F_1|C)P(F_2|C)P(F_3|C) \dots \\ &= P(C) \prod_{i=1}^n P(F_i|C) \end{aligned} \tag{6}$$

Persamaan di atas merupakan model dari teorema *Naive Bayes* yang selanjutnya akan digunakan dalam proses klasifikasi. Untuk klasifikasi dengan data kontinyu digunakan rumus *Densitas Gauss* :

$$P(X_i = x_i|Y = y_j) = \frac{1}{\sqrt{2\pi\sigma_{ij}}} e^{-\frac{(x_i - \mu_{ij})^2}{2\sigma_{ij}^2}} \tag{7}$$

Keterangan :

- P : Peluang.
- Xi : Atribut ke i.
- Xi : Nilai atribut ke i.
- Y : Kelas yang dicari.
- yi : Sub kelas I yang dicari.
- u : Mean, menyatakan rata – rata dari semua atribut.
- o : Deviasi standar, menyatakan varian dari seluruh atribut. (Sitohang, F)

Penelitian Terdahulu

Pertama dari Hesham Mohamed MA Al Laroussi, Implementasi Algoritma *Naive Bayes* Sebagai Proses Seleksi Penerima Beasiswa Libyan Embassy Berbasis Web. Penelitian ini sanggup menerapkan algoritma *Naive Bayes* dalam proses penyeleksian penerima beasiswa Libyan embassy dengan akurasi 88.3% dari 60 sampel mahasiswa yaitu sebanyak 7 hasil yang tidak sesuai dan 53 hasil yang sesuai. Kedua dari Rizal Amegia Saputra dan Shinta Ayuningtias, Penerapan Algoritma *Naive Bayes* Untuk Penelitian Calon Penerima Beasiswa Pada SMK Pasim Plus Sukabumi. Dalam penelitian ini digunakan penelitian dengan metode eksperimen dengan menggunakan data siswa yang mengajukan beasiswa di “SMK Pasim Plus Sukabumi” sejumlah 68 orang, yang terdiri dari 25 siswa yang dinyatakan lulus seleksi dan 43 diantaranya yang dinyatakan tidak lulus seleksi. Ketiga dari Wisnu Uriawan dan Gina Zaida Yusfira, Prediksi Penerima Beasiswa Pegawai Menggunakan Metode *Naive Bayes Classifier*. Dengan adanya aplikasi prediksi penerima beasiswa pegawai, proses penentuan yang dilakukan terlaksana dengan cara yang lebih objektif. Keempat dari Riani Dewi H, Yunita, dan Novi Indrawati, Rancang Bangun Sistem Pendukung Keputusan Penerimaan Beasiswa Menggunakan Metode *Naive Bayes Classifier*. Metode *Naive Bayes Classifier* memberikan proses penyeleksian yang cepat dan algoritmanya mudah dimengerti. Kelima dari Yusuf Zakariya, Implementasi Metode *Naive Bayes* Dalam Penentuan Kelayakan Penerima Bantuan Siswa Miskin. Sistem penentuan kelayakan penerima bantuan siswa miskin dengan metode *naive bayes* dapat dirancang dengan bahasa pemrograman *C Sharp* dan *database mysql*.

Data Training

Data Training merupakan data yang digunakan oleh algoritma klasifikasi untuk membentuk sebuah *model classifier*. Data training yang digunakan untuk proses perhitungan layak atau tidak layak nya pada SMA Negeri 4 Pangkalpinang adalah sebagai berikut :

Tabel 1. Data Training

No	Data	Keterangan	Range
1	Nilai Raport	Rendah	0 – 65
		Sedang	66 – 80
		Tinggi	81 – 100
2	Status Orangtua	Layak	Bercerai
		Tidak Layak	Tidak Bercerai
3	Jumlah Tanggungan	Rendah	0 – 2
		Sedang	3 – 4
		Tinggi	>= 5
4	Penghasilan Orangtua	Rendah	0 – 750.000
		Sedang	750.000 – 2.500.000
		Tinggi	>= 2.500.000

Data Testing

Data testing merupakan data yang akan menampilkan skor atau nilai dari data training. Dimana akan menghasilkan layak atau tidak layak nya siswa untuk mendapatkan beasiswa serta tampil data rekomendasi hasil layak atau tidak layak. Data testing yang diterapkan pada sistem beasiswa SMA Negeri 4 Pangkalpinang sebagai berikut:

```

:: SKORING

Layak = (Nilai Raport * Status Orangtua * Jumlah Tanggungan *
Penghasilan Orantua) * (Probabilitas Layak)

Tidak Layak = (Nilai Raport * Status Orangtua * Jumlah
Tanggungan * Penghasilan Orantua) * (Probabilitas
Tidak Layak)

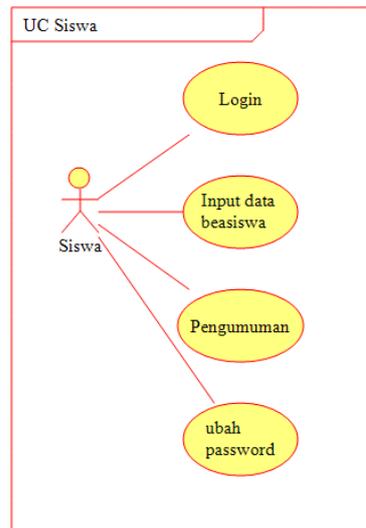
Layak = Hasil.....
Tidak Layak = Hasil.....

:: REKOMENDASI

Jika.....maka.....
    
```

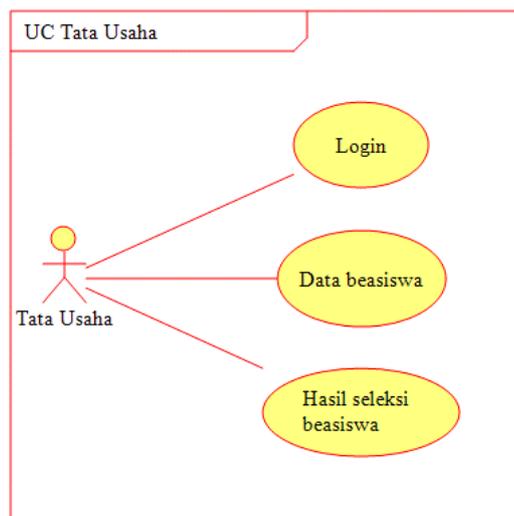
Gambar 1. Data Testing

Sistem Usulan dari sisi *Client*



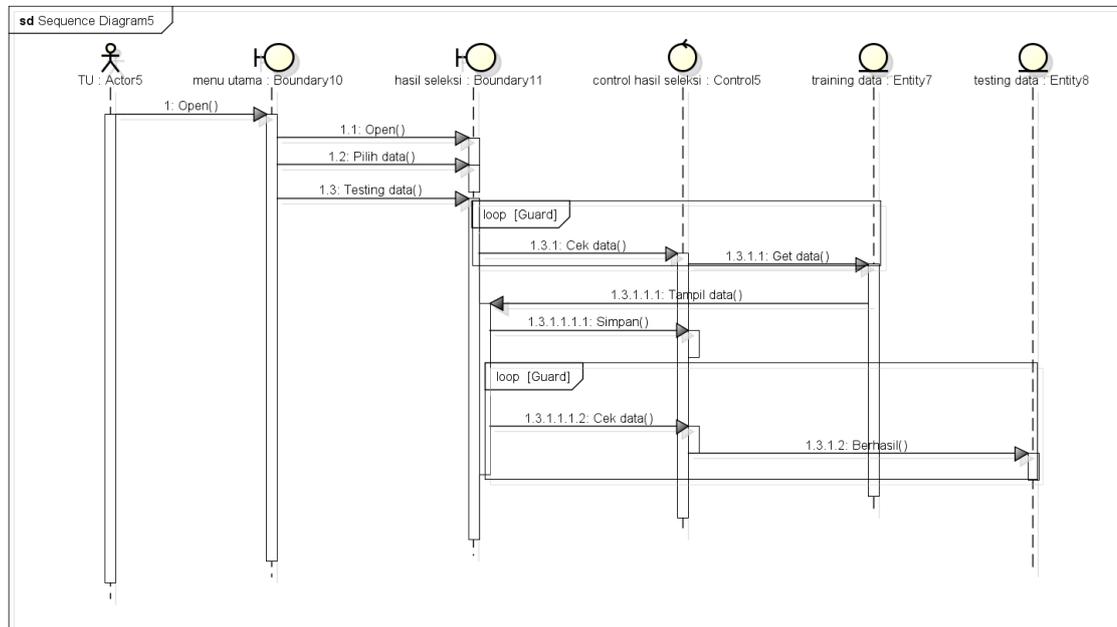
Gambar 2. Sistem usulan sisi *client*

Sistem Usulan dari sisi *Server*



Gambar 3. Sistem Usulan *Server*

Sequence Diagram



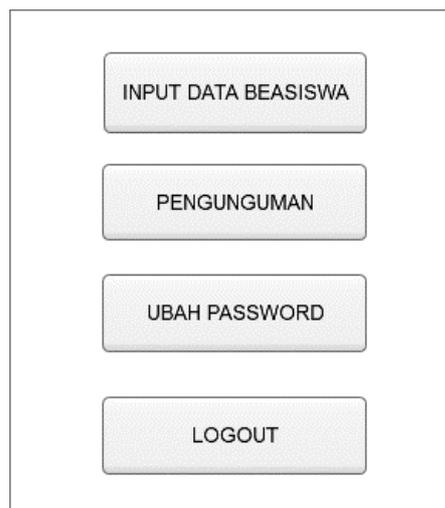
powered by Astah

Gambar 4. Sequence Diagram

HASIL DAN PEMBAHASAN

Tampilan dari sisi Client

Client membuka aplikasi yang tersedia pada *smartphone* berbasis android yang telah terinstal pada masing-masing siswa, masukkan *username* dan *password* kemudian tampil menu seperti gambar 5.



Gambar 5 Menu Client

Pertama, *Client* akan menginput data beasiswa yang akan di isi berupa data pribadi dan data pendukung beasiswa seperti gambar 6

The image shows a web form titled "INPUT DATA BEASISWA". It is divided into two main sections: "DATA DIRI" and "DATA BEASISWA".

DATA DIRI section includes:

- NIS: text input field
- Nama: text input field
- Jenis Kelamin: radio buttons for "Pria" and "Wanita"
- Alamat: text area with "enter text..." placeholder
- No. Telepon: text input field
- Jurusan: dropdown menu
- Kelas: dropdown menu
- TA: dropdown menu

DATA BEASISWA section includes:

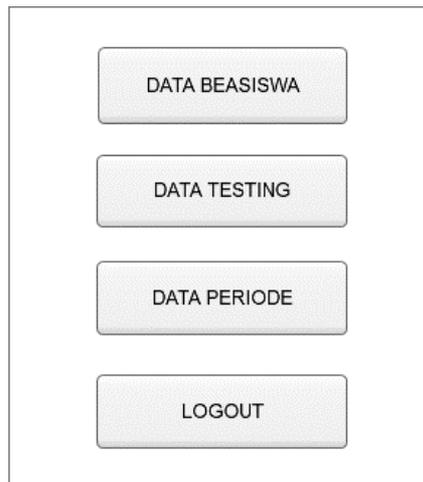
- Nilai Raport: text input field
- Status Orangtua: dropdown menu
- Jumlah Tanggungan: text input field
- Penghasilan Orangtua: text input field
- File pendukung: dropdown menu

At the bottom, there is a "File pendukung 1" section with a text input field and a "browse....." button. A "KIRIM" button is located at the very bottom of the form.

Gambar 6 Data Beasiswa

Tampilan dari sisi server

Sisi *server* yang mengelola data beasiswa yaitu data beasiswa masuk, periode serta mengelola data dengan naive bayes yaitu tahapan training (data beasiswa) dan testing (mengelola hasil dari tahapan training)



Gambar 7. Menu Server

Data beasiswa Data periode Data testing	Home logout							
No	NIS	Nama	Nilai Raport	Status Orang tua	Jumlah Tanggungan	Penghasilan Orangtua	Kelayakan	Aksi
1	xxxx	xxxx	999	xxxx	999	999	Layak/Tidak Layak	Ubah Hapus

Gambar 8 Data Beasiswa

Data beasiswa merupakan tahapan training ketika siswa mengisi data dan mengklik data simpan maka pada kodingan di proses sesuai dengan tabel 1.

Data beasiswa Data periode Data testing	Home logout
Tahun Ajaran	<input type="text"/>
Nilai Raport	<input type="text"/>
Status Orangtua	<input type="text"/>
Jumlah Tanggungan	<input type="text"/>
Penghasilan Orangtua	<input type="text"/>
<input type="button" value="Testing"/>	

Gambar 9 Data Testing

Pada gambar 9 menjelaskan dimana admin memilih data dan memproses ketahapan testing, ketika di proses sesuai dengan gambar 1 maka akan menampilkan hasil layak atau tidak layak

NIS	Nama	Kelayakan
12345	Safitri	Layak
23456	Agus Firza	Tidak Layak

Gambar 10 Hasil Data Testing

SIMPULAN

1. Sistem beasiswa ditentukan dengan beberapa kriteria yaitu nilai raport, status orangtua, jumlah tanggungan, dan penghasilan orangtua untuk membantu proses perhitungan.
2. Terapan *Naive Bayes* pada sistem ini dengan dua tahapan yaitu pertama menggunakan data training yang telah di transformasi ke dalam bentuk numerik, selanjutnya dicari nilai korelasi dan data tersebut ditesting kemudian akan menampilkan hasil dari perhitungan *Naive Bayes*.
3. *Naive Bayes* memberikan kemudahan kepada sekolah terutama bagian tata usaha dalam pengambilan keputusan untuk siswa yang berhak menerima beasiswa, hasil akhir dari metode ini mengeluarkan siapa yang layak atau tidak layak untuk mendapatkan beasiswa.

DAFTAR PUSTAKA

- Bustami. Penerapan Algoritma Naive Bayes untuk mengklasifikasikan data Nasabah asuransi. Teknik Informatika Universitas Malikussaleh.
- Hesham Mohamed MA Al Laroussi. 2015. Implementasi Algoritma Naive Bayes Sebagai Proses Seleksi Penerima Beasiswa Libyan Embassy Berbasis Web. Universitas Islam Negeri Maulana Malik Ibrahim Malang.
- Rahman, Antony Anwari dan Suryanto Agus. 2017. Implementasi Sistem Informasi Seleksi Penerimaan Beasiswa Dengan Metode Naive Bayes Classifier.

Jurnal Penelitian Pendidikan Indonesia (JPPI) Vol 2 Nomor 3. Universitas Negeri Semarang.

RianiDewi H, Yunita, dan Novi Indrawati. 2008. Rancang Bangun Sistem Pendukung Keputusan Penerimaan Beasiswa Menggunakan Metode Naive Bayes Classifier. Teknik Informatika Fakultas Teknik Universitas Trunojoyo Madura.

Rizal Amegia Saputra dan Shinta Ayuningtias. 2016. Penerapan Algoritma Naive Bayes Untuk Penelitian Calon Penerima Beasiswa Pada Smk Pasim Plus Sukabumi. STMIK Nusa Mandiri.

Sitohang, F.2013. Sistem pendukung keputusan Penerimaan beasiswa dengan metode TOPSIS. Pelita Informatika Budi Darma.

Wahyudi, Rizki dan Utami, Ema. 2016. Sistem Pakar E-Tourism Pada Dinas Pariwisata D.I.Y Menggunakan Metode Forward Chaining. Jurnal Ilmiah Dasi Vol. 17 No. 2. ISSN: 1411-3201.

Wisnu Uriawan dan Gina Zaida Yusfira. 2016. Prediksi Penerima Beasiswa Pegawai Menggunakan Metode Naive Bayes Classifier. Program studi Sistem Informasi STMIK LPKIA.

Yusuf Zakariya. 2017. Implementasi Metode Naive Bayes Dalam Penentuan Kelayakan Penerima Bantuan Siswa Miskin. Fakultas Teknik Universitas Nusantara PGRI Kediri