

# Metode *Naive Bayes* Dalam Menentukan Program Studi Bagi Calon Mahasiswa Baru

**Wildani Eko Nugroho<sup>1\*</sup>, Ali Sofyan<sup>2</sup>, Oman Somantri<sup>3</sup>**

<sup>1,2</sup>Program Studi Teknik Komputer, Politeknik Harapan Bersama Tegal

<sup>3</sup>Program Studi Teknik Informatika, Politeknik Negeri Cilacap

<sup>1,2</sup>Jl. Mataram No. 9 Pesurungan Lor, Kota Tegal, 52147, Indonesia

<sup>3</sup>Jln. Dr. Soetomo No.1 Karangcengis Sidakaya, Kabupaten Cilacap, 53212, Indonesia

E-mail: wild4n1@gmail.com<sup>1</sup>, alisofyan.poltek@gmail.com<sup>2</sup>, oman\_mantri@yahoo.com<sup>3</sup>

---

## Info Naskah:

Naskah masuk: 1 Desember 2020

Direvisi: 15 Maret 2021

Diterima: 18 Maret 2021

---

## Abstrak

Pada sebuah perguruan tinggi dalam menentukan program studi bagi calon mahasiswa adalah hal yang sering dilakukan untuk memfokuskan calon mahasiswa agar sesuai dengan kompetensi yang dimiliki. Hal ini merupakan harapan yang sangat penting, karena calon mahasiswa dapat mengembangkan kompetensi diri sesuai dengan kemampuan akademis yang dimiliki. Pada metode penelitian ini menggunakan beberapa tahapan, antara lain dengan data cleaning, data collection, menentukan kriteria, menentukan probabilitas, dan pengujian akhir. Metode *Naive Bayes* dengan studi kasus pada Madrasah Aliyah Swasta PAB 6 Helvetia dan pengujian data 100 data siswa dengan tingkat akurasi 90% merupakan penelitian sebelumnya. Tujuan dari penelitian tersebut adalah untuk membuat klasifikasi jurusan berdasarkan kriteria sedangkan pada penelitian ini mempunyai tujuan membuat klasifikasi program studi untuk calon mahasiswa baru. Pada penelitian ini digunakan metode sama tapi jumlah data record berbeda, data pengujian sejumlah 1671 record data mahasiswa, data tersebut diperoleh dari jumlah data record sebanyak 2256. Dari jumlah data record sebanyak 2256, setelah dilakukan data cleaning dan data collection maka diperoleh data pengujian sebanyak 1671. Di dalam data pengujian terdapat beberapa nilai probabilitas yang berisi berbagai kriteria dan atribut yang digunakan untuk menentukan klasifikasi program studi bagi calon mahasiswa baru. Dari hasil pengujian dengan metode yang sama dengan berbeda jumlah record datanya, maka tingkat keakurasian dari 90% menjadi 96% dengan hasil nilai akurasi 96,68%. Dari nilai akurasi tersebut menunjukkan bahwa hasil klasifikasi yang diperoleh menunjukkan pada program studi DIII Farmasi.

---

## Keywords:

naïve bayes;  
classification;  
prospective students.

---

## Abstract

In a university, determining a study program for prospective students is something that is often done to focus on prospective students so that they are in accordance with their competencies. This is a very important hope, because prospective students can develop self-competence according to their academic abilities. This research method uses several stages, including data cleaning, data collection, determining criteria, determining probability, and final testing. The *Naive Bayes* method with a case study at the Private Madrasah Aliyah PAB 6 Helvetia and testing of 100 student data with an accuracy rate of 90% is a previous research. The purpose of this study was to make a classification of majors based on the criteria, while in this study the aim of making a classification of study programs for prospective new students. In this study, the same method was used but the number of data records was different, the test data was 1671 student data records, the data was obtained from 2256 data records. From the total data records were 2256, after data cleaning and data collection were carried out, 1671 test data were obtained. In the test data, there are several probability values that contain various criteria and attributes used to determine the classification of study programs for prospective new students. From the test results with the same method with different number of data records, the accuracy rate is from 90% to 96% with an accuracy value of 96.68%. From this accuracy value shows that the classification results obtained show the Pharmacy DIII study program.

---

\*Penulis korespondensi:

**Wildani Eko Nugroho**

E-mail: wild4n1@gmail.com

## 1. Pendahuluan

Perkembangan teknologi yang merambah dalam berbagai aspek adalah salah satu peran pendidikan. Hal tersebut merupakan salah satu peran yang sangat penting juga untuk sebuah berbagai perguruan tinggi. Teknologi juga mempengaruhi calon mahasiswa baru dalam menentukan program studi. Dalam menentukan program studi, yang dilakukan calon mahasiswa baru adalah proses untuk memfokuskan calon mahasiswa dalam bidang kompetensi tertentu. Sistem yang berjalan pada Politeknik Harapan Bersama Tegal dalam menentukan program studi tidak sepenuhnya efektif, karena calon mahasiswa baru diberikan pilihan pada saat melakukan registrasi pendaftaran[1].

Menentukan program studi bagi perguruan tinggi adalah hal yang sangat penting juga, karena berkaitan erat dengan pilihan dan minat masyarakat. Peminat adalah kemampuan yang cenderung meningkatkan fokus pada hal tertentu. Di Indonesia, banyak perguruan tinggi yang sudah menyediakan berbagai macam pilihan program studi yang diminati masyarakat. Politeknik Harapan Bersama Tegal merupakan salah satu perguruan tinggi swasta yang menyediakan berbagai macam program studi [2].

Klasifikasi merupakan topik yang diangkat dalam penelitian ini. Selain digunakan untuk melakukan klasifikasi pada pengujian data yang bersifat numerik, klasifikasi juga dapat digunakan untuk data yang bersifat kategorikal. Selain kategori dan numerical, klasifikasi juga dapat digunakan untuk mengidentifikasi gambar – gambar yang sesuai dengan kriteria yang diinginkan [3].

*Naïve bayes* adalah model yang digunakan untuk klasifikasi pada penelitian ini. *Naïve bayes* dapat bertugas untuk menetapkan label yang relevan. Tugas ini dilakukan dengan memperluas pengklasifikasian satu label seperti pengklasifikasi *naïve bayes* [4]. Secara khusus, klasifikasi *naïve bayes* dapat melakukan evaluasi pada data uji eksperimental dan pemodelan prediktif [5].

Metode *naïve bayes* adalah pengklasifikasi sederhana namun sangat praktis, dengan berbagai aplikasi termasuk *filter spam*, *diagnostic kanker*, dan pengenalan wajah, untuk menyebutkan beberapa contoh implementasi saja. Pertimbangan situasi di mana pengguna meminta layanan klasifikasi dari server pengklasifikasi *naïve bayes*, pengguna dan server tidak ingin mengungkapkan data pribadi mereka satu sama lain. Penelitian ini memfokuskan pada membangun klasifikasi *naïve bayes* yang menjadi privasi yang tahan terhadap serangan yang mudah dilakukan tetapi sulit dideteksi [6].

Pemantauan dan pemeliharaan modul *photovoltaic* sangat penting untuk operasional yang efisien dan handal. Hal ini merupakan salah satu implementasi metode *naïve bayes* untuk mengklasifikasi hotspot dalam modul *photovoltaic*. Adapun kategori hotspot yang dideteksi untuk melatih metode ini dilakukan uji coba pada sistem PV 42,24 kWp, yang menunjukkan bahwa tingkat pengenalan rata – rata sekitar 94,1% dicapai untuk kumpulan 375 sample data [7].

Sistem pendukung keputusan adalah penelitian yang berkaitan dengan minat dalam pemilihan program studi yaitu untuk menentukan jurusan pada SMA Islam di

wilayah kecamatan Bumiayu kabupaten Brebes. Dimana dalam menentukan jurusan siswa perlu dukungan dari berbagai pihak sekolah, maka dibangunlah sistem berbentuk aplikasi sistem pendukung keputusan untuk menentukan jurusan. Keterangan awal pemilihan atau penentuan jurusan oleh sistem adalah proses akhir penentuan jurusan [8].

Klasifikasi artikel berita dengan *naïve bayes* ialah penelitian lain yang menjadi rujukan. Peneliti ini menjelaskan tentang *naïve bayes classifier* yang sudah mengalami modifikasi dalam pembobotan kata berdasarkan posisi dalam berita. Eksperimen dilakukan dengan 900 dokumen dan dibagi menjadi 9 kategori dan tiap kategori terdapat 100 dokumen. Dokumen *training* dan dokumen testing pada penelitian ini merupakan hasil kombinasi dari jumlah data latih dari *naïve bayes classifier* yang sudah memiliki efektifitas. Kombinasi dokumen *training* dan dokumen testing secara berurutan adalah sebagai berikut: 5:95, 10:90, 15:85, 20:80, 25:75, dan 30:70. Hasil akurasi klasifikasi dengan menggunakan metode *naïve bayes classifier* secara berturut-turut adalah 54%, 65%, 65%, 69%, 71%, 76%. Menggunakan *naïve bayes classifier* maka didapat secara berturut-turut nilai akurasi klasifikasi adalah 57%, 68%, 69%, 70%, 72%, 78%. Dari hasil ini menunjukkan akurasi hasil klasifikasi berdasarkan pembobotan posisi kata mengalami peningkatan[9].

Penelitian sebelumnya pernah dilakukan dengan proses mining, dimana proses tersebut digunakan untuk menentukan jurusan dengan metode *naïve bayes*. Adapun proses mining adalah digunakan untuk mencari informasi menentukan jurusan siswa dengan metode *naïve bayes*. Data yang adalah 100 data siswa dimana data tersebut berdasarkan kriteria yang meliputi nilai rata-rata mata pelajaran IPA, IPS, hasil rekomendasi para guru, dan nilai angket dari calon siswa. Hasilnya didapatkan nilai akurasi penentuan jurusan sebanyak 90% dan *error* 10% [10].

Pada penelitian ini, akan mengembangkan penelitian sebelumnya, yaitu dengan pengujian data sebanyak 1671 record data. Dari jumlah *record* data ini kemudian dibagi menjadi 2 bagian, yang pertama adalah data training dengan jumlah *record* data sebanyak 1158 dengan prosentase 70%, sedangkan yang kedua adalah data testing dengan jumlah *record* data sebanyak 513 dengan prosentase 30%. Dari hasil pengujian dengan metode yang sama tetapi jumlah *record* data yang berbeda, maka akan menghasilkan tingkat keakurasian[10].

Pokok permasalahan yang didapatkan dari penelitian ini adalah bagaimana nilai akurasi pada metode *naïve bayes* dapat mengklasifikasikan dalam menentukan program studi bagi calon mahasiswa baru yang diharapkan di unit penerimaan mahasiswa baru pada Politeknik Harapan Bersama Tegal. Selain itu permasalahan lain adalah evaluasi proses penentuan program studi bagi calon mahasiswa baru sudah dilaksanakan tetapi belum menjamin ada mutu peningkatan prestasi pembelajaran bagi mahasiswa pada saat mengikuti proses belajar pada program studi yang diharapkan, sehingga dapat dijadikan pengambilan keputusan pada saat proses yudisium nilai akhir semester dan atau yudisium kelulusan.

Tujuan dari penelitian ini adalah membuat klasifikasi menentukan program studi bagi calon mahasiswa baru dengan metode *naïve bayes* pada unit penerimaan mahasiswa baru Politeknik Harapan Bersama, dimana hasil klasifikasi tersebut digunakan sebagai acuan peningkatan mutu dari setiap program studi.

## 2. Metode

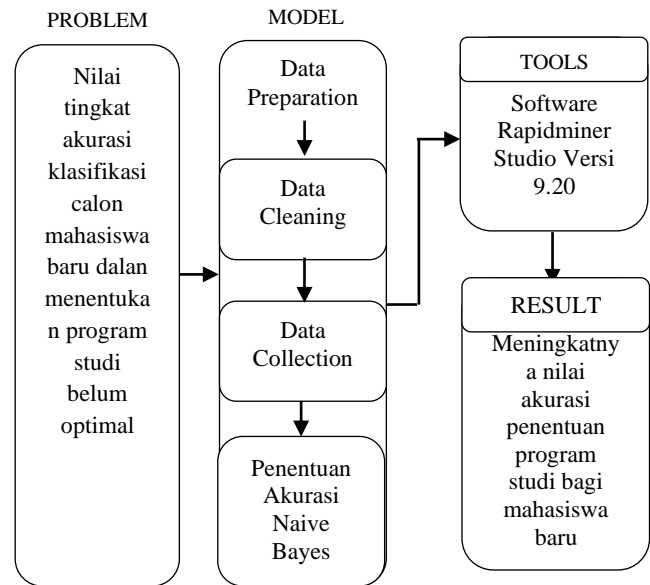
### 1.1. Metode Penelitian

Merupakan salah satu rangkaian yang dilakukan untuk merancang mengenai langkah-langkah tahapan penelitian. Adapun langkah tahapan penelitian yang dilakukan, antara lain pada bagian pertama melakukan pencarian data pada bagian unit penerimaan mahasiswa baru. Pencarian data dapat berupa melihat secara langsung proses penerimaan mahasiswa baru dan wawancara dengan kepala unit penerimaan mahasiswa baru. Setelah mencari data langkah selanjutnya mengumpulkan data dan diteruskan menganalisa variabel yang ada dan serta melakukan perancangan. Langkah terakhir dilakukan uji variabel dan menganalisa hasil pengujian.

Data calon mahasiswa tahun 2019 ada beberapa yang diubah ke dalam bentuk kriteria yang memiliki nilai data string dan data numerik, serta dikelompokkan datanya dengan *naïve bayes*. Adapun tahapan mengolah data yang dilakukan dalam metode penelitian ini adalah sebagai berikut:

- 1) *Data cleaning*, yaitu pada tahap ini dilakukan eliminasi pada beberapa kriteria yang tidak ada pengaruhnya pada nilai akurasi klasifikasi *naïve bayes*.
- 2) *Data collection*, yaitu pada tahapan proses merubah data menjadi data *training* dan *testing*. Data yang digunakan adalah data calon mahasiswa baru.
- 3) Menentukan kriteria, pada tahap hal yang dilakukan adalah menentukan kriteria yang digunakan untuk memilih kriteria data berdasarkan data yang sudah terkumpul.
- 4) Menentukan probabilitas setiap kriteria, pada tahap ini menentukan kriteria yang dijadikan acuan dalam mengklasifikasikan program studi bagi calon mahasiswa baru.
- 5) Pengujian akhir, merupakan pengujian tahap penerapan metode *naïve bayes* dengan beberapa data yang disiapkan.

Dari tahapan diatas dapat dituangkan dalam kerangka penelitian dengan metode yang diusulkan seperti pada Gambar 1.



Gambar 1. Kerangka Pikir Penelitian

### 1.2. Naïve Bayes

Metode *naïve bayes* adalah salah satu model klasifikasi probalistik sederhana guna menghitung kumpulan probabilitas dan menjumlahkan frekuensi dan kombinasi dari *dataset* [11].

*Naïve bayes* mempunyai beberapa keuntungan antara lain hanya membutuhkan jumlah training data yang kecil. *Training data* digunakan menentukan estimasi parameter dalam proses klasifikasi [12].

Adapun persamaan teorema metode *naïve bayes* adalah:

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

Di mana:

$X$  : *class data* yang belum diketahui.

$H$  : spesifik *class* yang merupakan hipotesis data

$P(H|X)$  : posteriori probabilitas (probabilitas hipotesis  $H$  berdasarkan kondisi  $X$ )

$P(H)$  : berisi nilai probabilitas hipotesis  $H$

$P(X|H)$  : berisi nilai probabilitas  $X$  berdasarkan kondisi pada hipotesis  $H$

$P(X)$  : berisi nilai probabilitas hipotesis  $X$

Seperti diketahui bahwa proses klasifikasi dengan metode *naïve bayes* perlu sejumlah petunjuk guna menentukan kelas yang cocok untuk contoh data yang di analisis[13]. Dengan demikian metode *naïve bayes* di atas disesuaikan dengan :

$$P(C|F1..Fn) = \frac{P(C)P(F1..Fn|C)}{P(F1..Fn)} \quad (2)$$

Di mana:

- Nilai variable  $C$  menggambarkan kelas
- Nilai variable  $F1-Fn$  digunakan untuk menggambarkan karakteristik yang dibutuhkan untuk klasifikasi

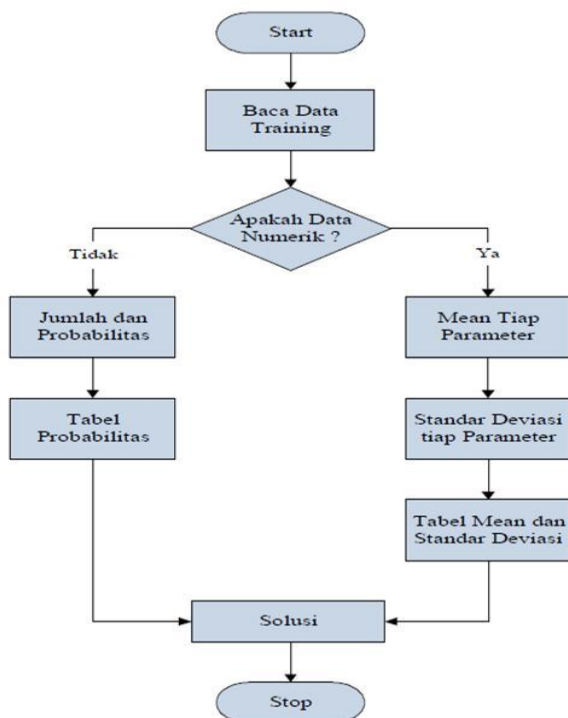
Dari pemaparan rumus ini maka dijelaskan peluang masuk sebuah contoh karakteristik tertentu ke dalam nilai kelas C adalah muncul kelas C di kali dengan peluang munculnya contoh karakteristik pada kelas C dibagi dengan peluang munculnya contoh karakteristik kelas C secara global. Kemudian rumus ini disederhanakan menjadi berikut [14]:

$$Posterior = \frac{prior \times likelihood}{evidence} \quad (3)$$

Dari hasil penyederhanaan rumus diatas maka untuk nilai *evidence* digunakan untuk setiap kelas pada setiap sampel dan selalu tetap, dan untuk nilai posterior buat perbandingan dengan nilai posterior lainnya. Sedangkan untuk menentukan suatu nilai ke kelas apa suatu sampel data dapat diklasifikasikan dapat dijabarkan dengan rumus *naïve bayes* sebagai berikut:  $(C|F1, \dots, Fn)$  [15], adapun aturan perkalian dari rumus ini adalah:

$$\begin{aligned} P(C|F1, \dots, Fn) &= P(C)P(F1, \dots, Fn|C) \\ &= P(C)P(F1|C)P(F2, \dots, Fn|C, F1) \\ &= P(C)P(F1|C)P(F2|C, F1)P(F3, \dots, Fn|C, F1, F2) \\ &= (C)P(F1|C)P(F2|C, F1)P(F3|C, F1, F2)P(F4, \dots, Fn|C, F1, F2, F3) \\ &= P(C)P(F1|C)P(F2|C, F1)P(F3|C, F1, F2) \dots P(Fn|C, F1, F2, F3, \dots, Fn-1) \end{aligned}$$

Dari penjelasan dan penjabaran diatas bahwa dari rumus tersebut dapat menyebabkan semakin banyak dan kompleksnya faktor-faktor syarat yang mempengaruhi nilai probabilitas yang mengakibatkan untuk perhitungan dan analisis hamper mustahil dilakukan satu persatu [16]. Dari rumus dan berbagai penjelasan diatas, tampilan alur penggunaan metode *naïve bayes* seperti Gambar 2.



Gambar 2. Alur Metode *Naïve Bayes*

Keterangan gambar diatas adalah sebagai berikut:

- 1) Membaca training data.
- 2) Perhitungan jumlah probabilitas dengan ketentuan:
  - a. Mencari hasil nilai means dan deviasi standar dari nilai parameter data numerik. Adapun untuk menghitung nilai mean dapat menggunakan persamaan sebagai berikut [17]:

$$\mu = \frac{\sum_{i=1}^n X_i}{n} \text{ atau } \mu = \frac{X_1 + X_2 + X_3 + \dots + X_n}{n} \quad (4)$$

Di mana:

$\mu$ : jumlah hitung nilai rata-rata

$x_i$ : nilai sampel ke -i

$n$ : jumlah sampel

sedangkan untuk perhitungan nilai deviasi standar dapat menggunakan persamaan berikut:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n-1}} \quad (5)$$

Di mana:

$\sigma$ : standar deviasi

$x_i$ : nilai x ke -i

$\mu$ : nilai mean

$n$ : jumlah sampel

- b. Mencari perhitungan nilai probabilitas dengan menghitung kategori jumlah data yang sama di bagi kategori jumlah data tersebut.
- 3) Memperoleh hasil perhitungan nilai dalam bentuk tabel means, deviasi standar dan nilai probabilitas.
- 4) Hasil solusi.

### 3. Hasil dan Pembahasan

Pada tahapan ini peneliti melakukan beberapa hal, yaitu; preposisi data, yakni pengolahan data mentah yang berupa data penerimaan calon mahasiswa baru tahun 2019/2020. Pada tahapan preposisi data ada beberapa atribut yang tidak dipakai. Jumlah *record dataset* 2256. Pada pengolahan dataset penelitian ini menggunakan metode *naïve bayes*, dimana dalam dataset yang digunakan terdiri dari data *string* atau kategorikal yang bersifat konstan.

Data *string* atau kategorikal pada metode *naïve bayes* dibedakan menjadi 2 jenis, yaitu data numerik yang mempunyai sifat kontinyu, sehingga hasil perbedaan pada saat menentukan nilai probabilitas pada setiap kriteria akan kelihatan, sedangkan berikutnya adalah data *string* atau kriteria dengan data numerik.

#### 3.1 PreProcessing Data

Adapun tahapan pertama yang dilakukan pada pengolahan dataset penelitian, adalah melakukan *data cleansing*. Pada tahapan ini terdapat sebuah kriteria yang dieliminasi atau tidak dipakai karena tidak berpengaruh pada hasil akurasi klasifikasi metode *naïve bayes*. Dari *dataset* yang berjumlah sebesar 2256, setelah dilakukan proses eliminasi terhadap *dataset*, maka hasil dari proses eliminasi dataset sebesar 1671, sedangkan atribut yang digunakan sejumlah 23. Pada Tabel 1 adalah atribut yang digunakan dalam data *cleaning*.

Tabel 1. Attribute Data Cleaning

No	Attribute	No	Attribute
1.	No Pendaftaran	13.	Kota Sekolah
2.	Nama	14.	Provinsi Sekolah
3.	Kelas	15.	Matematika
4.	Gelombang	16.	Bahasa Indonesia
5.	Alamat	17.	Bahasa Inggris
6.	Jenis Kelamin	18.	Kejuruan
7.	Tempat Lahir	19.	Nilai Kelulusan
8.	Tanggal Lahir	20.	Jurusan
9.	Usia	21.	Program Studi
10.	Agama	22.	Tahun Lulus
11.	Nama Sekolah	23.	Masa Tunggu
12.	Alamat Sekolah		

Tabel 2. Contoh Dataset Training dan Testing

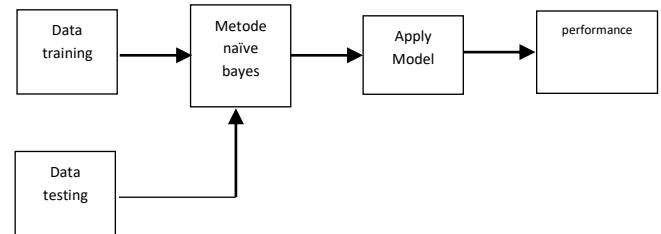
No Pendaftaran	Nama	Kelas	Gelombang
190117502	Puji Astuti	Reguler Malam	Gelombang 1
190117504	Safitri Dewi	Reguler Pagi	Gelombang 1
190127502	Ramang Darussalam	Reguler Pagi	Gelombang 1
190127503	Aldi Afriyanto	Reguler Malam	Gelombang 1
190127504	Nanda Binta Aunillah	Reguler Malam	Gelombang 1
190127505	Dwi Amalia Agustiani	Reguler Malam	Gelombang 1
190137501	Susmoro	Reguler Malam	Gelombang 1
190207501	Wahyu Wibowo	Reguler Pagi	Gelombang 1
190207502	Danu Hanafi Arif	Reguler Malam	Gelombang 1
190207503	Nur Mila Rizkiana	Reguler Malam	Gelombang 1

Tahapan selanjutnya adalah data collection, yaitu data yang digunakan untuk data *training* dan data *testing*, hal ini adalah data calon mahasiswa baru yang sudah mengalami proses data cleaning atau eliminasi atribut dan kriteria. Tahapan selanjutnya adalah menentukan kriteria, pada tahap menentukan kriteria ini digunakan kriteria data yang berdasarkan data yang telah terkumpul. Kriteria digunakan untuk menentukan data *training* dan *testing* dengan jumlah prosentase untuk data *training* 70% dan data *testing* 30%.

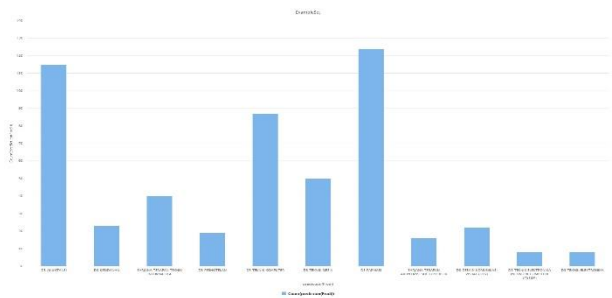
Tahapan berikutnya menentukan probabilitas setiap kriteria, pada tahap ini menentukan kriteria yang dijadikan acuan dalam mengklasifikasi program studi bagi calon mahasiswa baru, atribut tersebut ada pada tabel 1. Pengujian data, pada tahap ini merupakan tahapan penerapan metode *naive bayes* dengan beberapa data yang siap untuk diujikan. Pada Tabel 2 merupakan contoh data yang telah siap untuk dilakukan tahapan selanjutnya.

Dari *dataset* pada Tabel 2 kemudian siap untuk dijadikan *training* dan *testing* dengan metode *naive bayes*. Tetapi sebelum memasuki tahap klasifikasi menggunakan metode *naive bayes* dengan rapid miner, terlebih dahulu harus dilakukan pemilihan kriteria dan pembagian *dataset* menjadi dua bagian yang digunakan untuk data *training* dan

data *testing*. Adapun jumlah record data yang dilakukan pengujian sejumlah 1671 *record*. *Dataset* tersebut dibagi menjadi dua bagian dengan prosentase 70% untuk data *training* dan 30% untuk *data testing*. Adapun gambaran dari proses pengujian menggunakan metode *naive bayes* adalah sebagai berikut:



Gambar 3. Desain Pengujian Data Training dan Testing Pada Metode Naive Bayes



Gambar 4. Grafik Hasil Eksperimen Data Training dan Testing Pada Metode Naive Bayes

Tabel 3. Perbandingan Hasil Penelitian

No	Jumlah Record Data	Metode	Hasil Akurasi
1.	100 Record	Naive Bayes	90%
2.	1671 Record	Naive Bayes	96.68%

Desain eksperimen diatas menggambarkan bahwa dataset yang digunakan dipecah menjadi dua bagian dengan rasio prosentase 70% digunakan untuk data *training* dengan jumlah record data 1158 dan 30% data *testing* dengan jumlah record data 513. Hasil nilai akurasi yang diperoleh dari desain eksperimen diatas dengan menggunakan metode *naive* adalah 96.68%.

Selain hasil dari eksperimen menggunakan metode *naive bayes*, ada juga hasil dari grafik eksperimen menggunakan *naive bayes*. Pada Gambar 4 merupakan tampilan grafik hasil eksperimen data *training* dan data *testing* pada metode *naive bayes*. Dari Gambar 4 merupakan hasil eksperimen data *training* dan data *testing*. Grafik diatas dapat disimpulkan bahwa pemilihan atau menentukan program studi bagi calon mahasiswa baru menunjuk pada program studi DIII Farmasi. Perbandingan antara penelitian terdahulu dengan sekarang yang sedang dikerjakan maka hasilnya lebih baik dari hasil penelitian sekarang, perbandingan hasil penelitian dapat dilihat pada tabel 3.

Pada Tabel 3 menunjukkan bahwa antara penelitian yang terdahulu dengan yang sedang berjalan mengalami perubahan lebih baik, karena dilihat dari hasil akurasi yang diperoleh antara penelitian terdahulu memperoleh nilai akurasi 90%, sedangkan hasil penelitian yang sekarang

mengalami kenaikan nilai akurasi sejumlah 96.68%. Hal ini mengalami kenaikan sebesar 6.68%. Sehingga hasil klasifikasi dalam menentukan program studi bagi calon mahasiswa baru adalah program studi DIII Farmasi.

#### 4. Kesimpulan

Pada penelitian yang dilakukan maka disimpulkan bahwa adanya peningkatan mutu dalam penerimaan mahasiswa baru agar kriteria yang diingatkan setiap program studi terpenuhi. Banyaknya atribut yang masih belum sesuai, membuat proses training dan testing data harus dilakukan berkali-kali. Dari berbagai percobaan dengan berbagai model eksperimen yang terdapat pada metode naïve bayes mendapatkan hasil nilai akurasi 96.68%. Hasil klasifikasi dalam menentukan program studi bagi calon mahasiswa baru adalah program studi DIII Farmasi.

Adapun saran berdasarkan hasil penelitian ini yang menggunakan metode *naïve bayes* untuk menentukan program studi bagi calon mahasiswa baru, peneliti memberikan saran sebagai berikut: proses penerimaan mahasiswa baru yang belum jelas dan selektif mengakibatkan mahasiswa tersebut banyak yang tidak mengikuti mata kuliah tiap - tiap mata kuliah tiap-tiap program studi, tidak bias mengikuti perkuliahan yang ada di setiap program studi, tidak bisa mengikuti, perkuliahan yang ada di setiap program studi mengakibatkan menghambat mahasiswa dalam proses yudisium akhir semester dan yudisium proses kelulusan.

#### Ucapan Terimakasih

Ucapan terimakasih tak terhingga kepada semua pihak yang sudah membantu dan mendukung proses penelitian ini. Ucapan terimakasih khususnya kepada Direktorat Jendral Perguruan Tinggi yang telah memberikan dana penelitian melalui hibah penelitian skema Penelitian Dosen Pemula (PDP).

#### Daftar Pustaka

- [1] M. A. Al and M. Lirboyo, "Program studi sistem informasi fakultas teknik universitas nusantara PGRI Kediri 2017," 2017.
- [2] B. Amil *et al.*, Title," *J. Chem. Inf. Model.*, vol. 21, no. 1, pp. 1–9, 2020, [Online]. Available: <https://doi.org/10.1016/j.tmaid.2020.101607><https://doi.org/10.1016/j.ijsu.2020.02.034><https://onlinelibrary.wiley.com/doi/abs/10.1111/cjag.12228><https://doi.org/10.1016/j.ssci.2020.104773><https://doi.org/10.1016/j.jinf.2020.04.011>
- [3] C. Gonzalez-Arias, C. C. Viafara, J. J. Coronado, and F. Martinez, "Automatic classification of severe and mild wear in worn surface images using histograms of oriented gradients as descriptor," *Wear*, vol. 426–427, no. November 2018, pp. 1702–1711, 2019, doi: 10.1016/j.wear.2018.11.028.
- [4] H. C. Kim, J. H. Park, D. W. Kim, and J. Lee, "Multilabel naïve Bayes classification considering label dependence," *Pattern Recognit. Lett.*, vol. 136, pp. 279–285, 2020, doi: 10.1016/j.patrec.2020.06.021.
- [5] M. Andrejiova and A. Grincova, "Classification of impact damage on a rubber-textile conveyor belt using Naïve-Bayes methodology," *Wear*, vol. 414–415, pp. 59–67, 2018, doi: 10.1016/j.wear.2018.08.001.
- [6] C. zhi Gao, Q. Cheng, P. He, W. Susilo, and J. Li, "Privacy-preserving Naive Bayes classifiers secure against the substitution-then-comparison attack," *Inf. Sci. (Ny)*, vol. 444, pp. 72–88, 2018, doi: 10.1016/j.ins.2018.02.058.
- [7] K. A. K. Niazi, W. Akhtar, H. A. Khan, Y. Yang, and S. Athar, "Hotspot diagnosis for solar photovoltaic modules using a Naive Bayes classifier," *Sol. Energy*, vol. 190, no. July, pp. 34–43, 2019, doi: 10.1016/j.solener.2019.07.063.
- [8] T. Yudha and D. Kusuma, "Sistem Pendukung Keputusan Berbasis Web untuk Menentukan Penjurusan ( IPA / IPS / Bahasa ) pada SMA Islam Bumiayu ( Web Based Decision Support System to Determine Majoring ( Science , Social and Language ) at Islam High School of Bumiayu )," *Juita*, vol. 1, pp. 177–189, 2011.
- [9] J. Sains *et al.*, "Klasifikasi Berita Indonesia Menggunakan Metode Naive Bayesian Classification dan Support Vector Machine dengan Confix Stripping Stemmer," vol. 4, no. 2, 2015.
- [10] S. U. Khan, M. Niazi, and R. Ahmad, "Factors influencing clients in the selection of offshore software outsourcing vendors: An exploratory study using a systematic literature review," *J. Syst. Softw.*, vol. 84, no. 4, pp. 686–699, Apr. 2011, doi: 10.1016/j.jss.2010.12.010.
- [11] A. Saleh, "Klasifikasi Metode Naive Bayes Dalam Data Mining Untuk Menentukan Konsentrasi Siswa," *KeTIK*, pp. 200–208, 2015.
- [12] T. Wong, "A hybrid discretization method for naïve Bayesian classifiers," *Pattern Recognit.*, vol. 45, no. 6, pp. 2321–2325, 2012, doi: 10.1016/j.patcog.2011.12.014.
- [13] J. Wu, S. Pan, Z. Cai, X. Zhu, and C. Zhang, "Dual instance and attribute weighting for Naive Bayes classification," *Proc. Int. Jt. Conf. Neural Networks*, no. 1994, pp. 1675–1679, 2014, doi: 10.1109/IJCNN.2014.6889572.
- [14] M. J. Sánchez-Franco, A. Navarro-García, and F. J. Rondán-Cataluña, "A naïve Bayes strategy for classifying customer satisfaction: A study based on online reviews of hospitality services," *J. Bus. Res.*, vol. 101, no. December, pp. 499–506, 2019, doi: 10.1016/j.jbusres.2018.12.051.
- [15] D. Mondal, D. K. Kole, and K. Roy, "Gradation of yellow mosaic virus disease of okra and bitter melon based on entropy based binning and Naive Bayes classifier after identification of leaves," *Comput. Electron. Agric.*, vol. 142, no. October, pp. 485–493, 2017, doi: 10.1016/j.compag.2017.11.024.
- [16] A. H. Mirza, "Application of Naive Bayes Classifier Algorithm in Determining New Student Admission Promotion Strategies," *J. Inf. Syst. Informatics*, vol. 1, no. 1, pp. 14–28, 2019, doi: 10.33557/journalisi.v1i1.2.
- [17] V. R. Balaji, S. T. Suganthi, R. Rajadevi, V. Krishna Kumar, B. Saravana Balaji, and S. Pandiyan, "Skin disease detection and segmentation using dynamic graph cut algorithm and classification through Naive Bayes classifier," *Meas. J. Int. Meas. Confed.*, vol. 163, p. 107922, 2020, doi: 10.1016/j.measurement.2020.107922.