

# SCHOLARSHIP ELIGIBILITY DECISION SUPPORT USING NAIVE BAYES CASE STUDY : SANDJOJO FOUNDATION SCHOLARSHIP

<sup>1</sup>Tan, Yohanes Christianto Aryo Wicaksono, <sup>2</sup>YB. Dwi Setianto  
<sup>1,2</sup>Program Studi Teknik Informatika Fakultas Ilmu Komputer, Universitas  
Katolik Soegijapranata  
<sup>1</sup>christ.yohanes@gmail.com, <sup>2</sup>setianto@unika.ac.id

## ABSTRACT

*Until now there are no basic rules to determine the candidates that worthy to receive Sandjojo Foundation scholarship at Soegijapranata Catholic University. Therefore, the scholarship's recipients are not well-targeted. This project can help to solve problems by using Naive Bayes algorithm as a media for decision makers with criteria such as based on GPA of the students' registran, total parents' income, and must be an active student in organization. By using the registrant's previous data, Naive Bayes algorithm can study the data for each criteria that are inputted to determine whether the applicant is eligible or not to receive the scholarship from Sandjojo Foundation. This project made using web base application and has been tested four (4) times with various kinds of test data and training data, with the accuracy results are around 50% until 65% and the time is lower than three (3) seconds.*

**Keywords:** naive bayes, scholarship, decission support system.

## PENDAHULUAN

Penentuan Penerima beasiswa oleh UPT bidang 3 Unika Soegijapranata sampai saat ini masih dilakukan secara manual, hal ini mengakibatkan proses pengambilan keputusan tidak dapat dilakukan dengan cepat, akurat, dan efisien. Sering terjadi penentuan penerima beasiswa yang kurang tepat sasaran, untuk itu projek ini diyakini akan membantu otomatisasi pengambilan keputusan kelayakan penerima beasiswa Yayasan Sandjojo.

Dengan keluaran projek ini diyakini dapat menambah ketepatan penerima beasiswa Yayasan Sandjojo. Menggunakan metode Naive Bayes penentuan penerima beasiswa Yayasan Sandjojo dapat dipermudah karena metode mempelajari data penerima beasiswa yang sudah ada. Kemudian dikonversi dalam susunan suatu array 2 dimensi yang diolah dan menghasilkan nilai probabilitas sebagai penentuan class tiap data.

Project ini akan mengelompokan data menjadi 2 kelas yaitu mahasiswa layak dan tidak layak untuk menerima beasiswa "Yayasan Sandjojo" dengan faktor penentu IPK, penghasilan orang tua, dan keaktifan mahasiswa dalam organisasi.

## TINJAUAN PUSTAKA

Dalam menentukan kelayakan mahasiswa sebagai mahasiswa yang dianggap layak menerima beasiswa adalah suatu hal yang bisa dibbilang sepele hal itu merupakan hal yang penting.

Dalam jurnal yang berjudul "Algoritma Naive Bayes Untuk Mencari Perkiraan Waktu Studi Mahasiswa" oleh Jananto [1], menggunakan metode Naive Bayes untuk menentukan lama studi mahasiswa berdasarkan IPK, jenis kelamin, kota lahir, tipe sekolah, kota sekolah. Jananto [1]

melakukan ini dengan step, Mining data, Konversi data untuk tiap faktor, dan yang terakhir implementasi Naive Bayes. Hasil yang diperoleh dari 66 data tingkat kesalahan 20 record dengan presentase 66% benar. seperti yang dilakukan Arief Jananto projek ini memiliki kesamaan dalam hal parameter IPK sebagai salah satu faktor penentu dalam perhitungan Naive Bayes [1].

Jurnal “METODE NAIVE BAYES UNTUK PENENTUAN PENERIMA BEASISWA BIDIKMISI UNIVERSITAS MULAWARMAN” [2]. Penerima beasiswa bidikmisi di Universitas Mulawarman sering kali kurang tepat sasaran dan memakan waktu lambat. Penerima beasiswa menjadi tidak konsisten penyeleksipun tidak tranparan pada penerima, dibuatlah Sistem Pendukung Keputusan (SPK) untuk menyeleksi. SPK didalam jurnal ini ada beberapa komponen. Data Mangement sebagai Database, Model Management untuk perhitungan, Communication sebagai media interface, dan Knowledge Management. Hasil yang didapatkan akurasi prediksi benar dari SPK mencapai 85,56% dari 90 data uji dari jumlah data training 278 data. SPK jurnal ini menggunakan parameter penghasilan orang tua yang sama dengan project ini [2]. Project ini menggunakan data uji sebanyak 368 data training.

Berdasarkan dari jurnal Mantik Penusa dengan judul “PENERAPAN ALGORITMA NAIVE BAYES UNTUK MEMPREDIKSI JUMLAH PRODUKSI BARANG BERDASARKAN DATA PERSEDIAAN DAN JUMLAH PEMESANAN PADA CV. PAPADAN MAMA PASTRIES” dalam produksi roti dapat kelebihan dan kekurangan stok roti [3]. Manalu, et al. [3] membuat sistem yang mudah digunakan, akurat, efisien dan hemat waktu dengan menerapkan metode Naive Bayes dengan tujuan dapat mengantisipasi kelebihan atau kekurangan produksi. Manalu, et al. [3] menyimpulkan “Metode ini membutuhkan data traning yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian.”. Tahapan baca data training, hitung, menentukan value terbesar untuk dijadikan kesimpulan. Algoritma Naive Bayes membantu memprediksi produksi stok roti. Project ini memiliki kesamaan software penyimpanan data MySQL. Perbedaannya project ini hanya ada 1 macam data mahasiswa dalam jurnal terdapat 3 macam data roti tawar kupas, roti tawar gandum, dan roti tawar cokelat.

Pada jurnal “SISTEM PENDUKUNG KEPUTUSAN KLASIFIKASI TINGKAT KEGANASAN KANKER PAYUDARA DENGAN METODE NAIVE BAYES CLASSIFIER”, kanker Payudara salah satu jenis kanker yang sering ditemukan pada kebanyakan wanita [4]. Via, et al. [4] membuat keluaran sistem untuk mengklasifikasi tingkat keganasan kanker payudara dengan metode Naive Bayes Classifier. Karena terbukti memiliki akurasi dan kecepatan yang tinggi. Tahapan yang digunakan definisi masalah, pengumpulan data, pengolahan data, menentukan alternatif solusi. Via, et al. [4] bertujuan untuk mengklasifikasi keganasan kanker payudara dengan hasil akurasi 97,82% benar dan 2,18% salah dengan 137 data test. Persamaan project ini mengklasifikasikan 2 class. Perbedaan dengan project ini menggunakan 3 attribut, Via, et al. [4] menggunakan 9 attribut.

Berdasarkan Jurnal “Sistem Pendukung Keputusan Penentuan Kelayakan Calon Tenaga Kerja Indonesia Menggunakan Metode Naive Bayes”, perusahaan PT Karya Tama membutuhkan tenaga kerja yang layak [5]. Pada tahun 2012/2013 terdapat 542 pendaftar dan 385 diterima.

Wasiati dan Wijayanti [5] merancang suatu aplikasi berbasis dekstop penentu kelayakan tenaga kerja dengan metode Naive Bayes. Kriteria penentu usia, pendidikan, tinggi badan, berat badan, dan hasil test. Hasil perancangan aplikasi penentu kelayakan tenaga kerja perusahaan dapat menyeleksi pendaftar dengan mudah dan cepat dengan akurasi benar 73,89% benar dengan 180 data uji dan 362 data training. Persamaan project ini pada tampilan aplikasi menu login, menu akurasi, menu input pendaftar. Perbedaan project ini berbasis website [5].

Berdasarkan Aritikel Ilmiah “Implementasi Algoritma Naive Bayes dalam Proses Analisis Efisiensi Jasa Servis Bengkel Daihatsu Cabang Salatiga” untuk mencapai efisiensi jasa servis di bengkel Daihatsu cabang Salatiga artikel ini menggunakan algoritma Naive Bayes dan teknik Data Mining untuk mendapatkan informasi keluhan di bengkel tersebut [6]. Jatmikanto dan Pakereng [6] mengolah riwayat data dengan data mining, wawancara ke customer, merancang aplikasi, uji coba. Hasil dari artikel ini membagi menjadi 3 class dengan akurasi keberhasilan 80% dan kesalahan 20% dari 10 data uji dengan 684 data training. Artikel ini sama dengan project ini dengan terdapatnya keluaran berupa aplikasi berbasis Desktop, artikel ini menggunakan databse SQL Server Express 2008 sebagai database.

Berdasarkan artikel dalam jurnal “Sistem Persediaan Obat pada Puskesmas menggunakan Metode Naive Bayes (Studi Kasus Puskesmas Talang Tinggi Seluma)” puskesmas Talang Tinggi masih menggunakan sistem yang bersifat manual, ketika memesan obat harus melakukan rekap satu persatu, hal ini akan menyebabkan kekeliruan karena terdapat 200 jenis obat [7]. Furqan, et al. [7] merancang sistem prediksi persediaan sebelum obat itu habis yang diolah menggunakan metode Naive bayes dengan harapan mudah digunakan. Furqan, et al. [7] merancang sistem prediksi persediaan stok obat dengan survey, analisa, desain, persiapan hardware dan software, implementasi sistem, Pemeliharaan sistem. Furqan, et al. [7] menggunakan 3 parameter sebagai penentu yaitu satuan obat, golongan obat, dan jumlah obat. Hasil yang didapatkan dalam jurnal ini membantu proses Prediksi obat. Project ini sama menggunakan 3 attribute sebagai penentu. Perbedaan project ini pada bahasa pemrograman yang digunakan oleh jurnal yaitu Visual Basic 6.0.

Berdasarkan “Sistem Prediksi Kesiapan Siswam dalam Menghadapi Ujian Nasional menggunakan Metode Naive Bayes” sekolah tidak mengetahui siswa yang kurang siap untuk menghadapi ujian nasional yang menjadi kewajiban tiap siswa untuk dinyatakan lulus sekolah. Putri & Irsyadi [8] merancang Sistem Prediksi Kesiapan Siswa SMP dalam Menghadapi Ujian Nasional Menggunakan Metode Naive Bayes dengan tahapan pengumpulan data dengan data mining, pengolahan data perhitungan Theorema Bayes. Hasil jurnal ini membantu memprediksi kesiapan siswa menghadapi ujian nasional dengan akurasi 93,35% benar dengan percobaan sebanyak 6 kali dengan data uji yang berbeda, menggunakan kurang lebih 145 data training setiap test dan kurang lebih 130 data test. Project ini sama menggunakan tampilan website. Perbedaan dengan project ini tidak menggunakan metode data mining untuk pengumpulan data training.

Jadi projek ini menghasilkan keluaran aplikasi berbasis website dengan menggunakan metode Naive Bayes seperti yang dilakukan Jananto, 2013 [1]; Dahri et al. [2]; Manalu et al. [3];

Via et al. [4]; Wasiati & Wijayanti [5]; Jatmikanto & Pakereng [6]; Furqan et al. [7]; Putri & Irsyadi [8]; untuk menentukan kelayakan penerima beasiswa Yayasan Sandjojo dengan ketentuan IPK, penghasilan orang tua, dan keaktifan mahasiswa dalam organisasi seperti halnya atribut yang digunakan sebagai penentu [9]. Dengan 368 data training dan 114 data uji yang didapat dari UPT bidang 3.

## **METODOLOGI PENELITIAN**

Dalam pembuatan project ini langkah – langkah yang harus dilalui untuk mendapatkan hasil yang diinginkan:

- a. Pengumpulan Data  
Data didapatkan secara langsung bukan dari hasil wawancara atau menyebarkan pertanyaan. Meminta secara langsung kepada UPT bidang 3 di Universitas Katolik Soegijapranata yang mengurus beasiswa Yayasan Sandjojo. Project ini menggunakan 2 periode pendaftaran yaitu 2016 dan 2017.
- b. Pengumpulan dan Analisa Jurnal  
Dalam pembuatan project ini membutuhkan landasan teori untuk memperkuat pembuatan project ini. Jurnal – jurnal diambil dari website <https://scholar.google.co.id/> dan terlampir pada bagian daftar pustaka. Project ini menggunakan 8 jurnal sebagai referensi.
- c. Desain dan Implementasi  
Tahap yang dilalui untuk mencapai tujuan project ini adalah penerjemahan kedalam bahasa pemrograman. Design tampilan untuk mempermudah pengguna melihat data. Project ini menggunakan bahasa pemrograman PHP untuk membuat sistem pendukung pengambil keputusan dengan metode Naive Bayes yang memerlukan database. Project ini memerlukan Myql untuk menyimpan data training.
- d. Uji Coba  
Testing untuk menguji akurasi dari program dari data test. Project ini menggunakan 3 metode test antara lain test fungsi, test performa, dan test akurasi.
- e. Membuat Laporan  
Pembuatan laporan ini sebagai berkas untuk menjelaskan langkah – langkah dalam mengerjakan project, sekaligus untuk laporan dalam menyelesaikan project akhir.

## **PEMBAHASAN**

Project ini dibuat menggunakan bahasa pemrograman PHP karena aplikasi di universitas Katolik Soegijapranata kebanyakan menggunakan aplikasi berbasis website dengan pemrograman PHP. Project ini menggunakan database Mysql untuk menyimpan data training dan data uji, data dalam Mysql mudah untuk diolah karena akan dimasukan dalam struktur data Array 2 dimensi dan server di Universitas Katolik Soegijapranata menggunakan Mysql sebagai penyimpanan data di aplikasi Sintak. Project ini menggunakan algoritma Naive Bayes karena Thomas Bayes sebagai

ilmuan penemu Theorema Bayes mengemukakan memprediksi peluang dimasa depan berdasarkan pengalaman di masa sebelumnya yang dapat dilihat pada fungsi (1) [10].

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

Keterangan :

$X$  = Data dengan class yang belum diketahui

$H$  = Hipotesis data  $X$  merupakan suatu class spesifik

$P(H|X)$  = Probabilitas hipotesis  $H$  berdasarkan kondisi  $X$  (posteriori prob.)

$P(H)$  = Probabilitas hipotesis  $H$  (prior prob.)

$P(X|H)$  = Probabilitas  $X$  berdasarkan kondisi tersebut

$P(X)$  = Probabilitas dari  $X$

Naive Bayes dalam project ini menggunakan 3 parameter yaitu IPK, penghasilan orang tua, dan keaktifan mahasiswa dalam organisasi. Metode Naive Bayes akan membagi data menjadi 2 claster yaitu layak dan tidak layak dengan penambahan peringkat berdasarkan perhitungan hasil value layak. keputusan apabila hasil dari data input value layak semakin tinggi maka data tersebut akan muncul paling awal sebagai data prioritas. Project ini menggunakan range IPK sebagai berikut :

**Tabel 1.** Pengelompokan berdasarkan IPK

Kategori	Minimal	Maximal
I1	0	2.75
I2	2.76	3.25
I3	3.26	2.50
I4	3.51	4.00

Pembagian IPK dimaksudkan untuk mengelompokan data training dan data uji. Penghasilan orang tua di kelompokan berdasarkan total gaji, pengelompokan total penghasilan orang tua yaitu:

**Tabel 2.** Pengelompokan berdasarkan total penghasilan orang tua

Kategori	Minimal	Maximal
P1	0	999.999
P2	1.000.000	2.899.999
P3	2.900.000	7.899.999
P4	7.900.000	~

Gaji orang tua dihitung dengan penjumlahan gaji ayah dan gaji ibu hasil inputan dari pendaftar beasiswa dengan rumus sebagai berikut :

$$\text{Gaji Total} = \text{Gaji Ayah} + \text{Gaji Ibu}$$

**Tabel 3.** Pengelompokan berdasarkan jumlah sertifikat

Kategori	Keterangan	Jumlah
S1	Tidak ada	0
S2	Ada sedikit	1 – 6
S3	Ada banyak	> 7

Project ini akan menghitung setiap file sertifikat yang masuk untuk diperhitungkan jumlah sertifikat setiap data. Perhitungan setiap attribute akan memengaruhi hasil *Theorema Bayes*. Hasil dari setiap attribute dari setiap kelas akan dikalikan untuk mendapatkan hasil akhir yang nantinya menjadi penentu keputusan data masuk ke class apa layak atau tidak layak. Contoh perhitungan dengan kasus diberikan :

**Tabel 4.** Contoh data

IPK	Gaji Ayah	Gaji Ibu	Jumlah sertifikat
3.84	Rp. 2.000.000	Rp. 1.500.000	8

Jumlah database : 368 Data

Jumlah database dengan kelas Diterima : 144

Jumlah database dengan kelas Ditolak : 224

Total penghasilan = Gaji Ayah + Gaji Ibu = 2.000.000 + 1.500.000 = 3.500.000

Perhitungan setiap parameter di kelas Layak (Diterima)

$$P(I4 | Layak) = 39 / 144 = 0,2708333333$$

$$P(P3 | Layak) = 53 / 144 = 0,3680555556$$

$$P(S3 | Layak) = 11 / 144 = 0,0763888889$$

Perhitungan setiap parameter di kelas Tidak Layak (Ditolak)

$$P(I4 | Tidak Layak) = 50 / 224 = 0,2232142857$$

$$P(P3 | Tidak Layak) = 122 / 224 = 0,5$$

$$P(S3 | Tidak Layak) = 12 / 244 = 0,0491803279$$

Perhitungan peluang setiap kelas

$$P(\text{Layak}) = P(I4 | Layak) \times P(P3 | Layak) \times P(S3 | Layak)$$

$$P(\text{Layak}) = 0,2708333333 \times 0,3680555556 \times 0,0763888889$$

$$P(\text{Layak}) = 0,0076145753$$

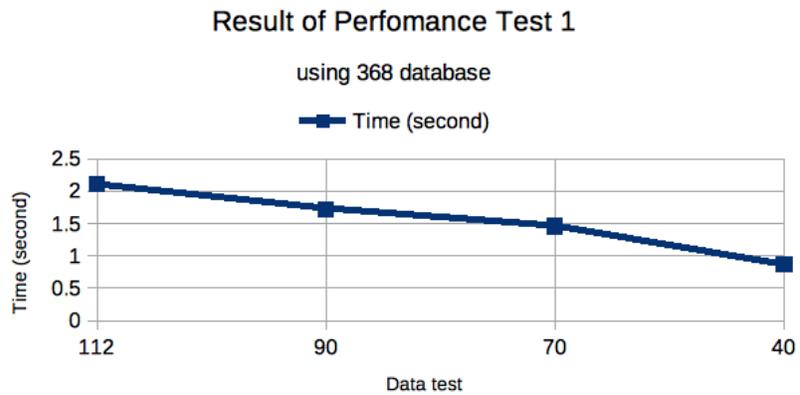
$$P(\text{Tidak Layak}) = P(I4 \mid \text{Tidak Layak}) \times P(P3 \mid \text{Tidak Layak}) \times P(S3 \mid \text{Tidak Layak})$$

$$P(\text{Tidak Layak}) = 0,2232142857 \times 0,5 \times 0,0491803279$$

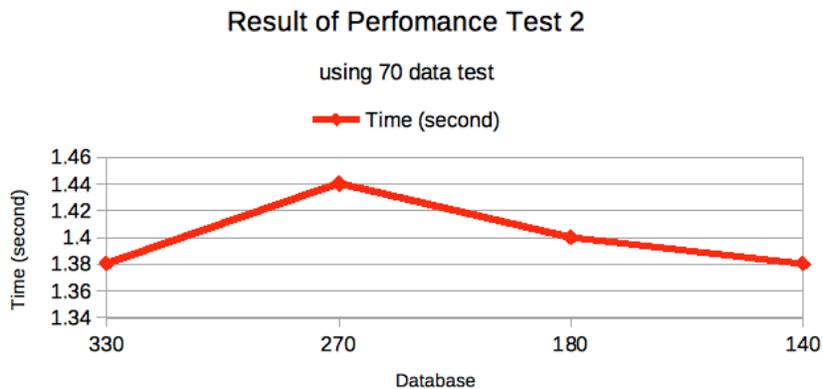
$$P(\text{Tidak Layak}) = 0,0054888759$$

Perhitungan probabilitas tiap kelas menunjukkan hasil probabilitas kelas “Layak” lebih besar dari probabilitas kelas “Tidak Layak” jadi keputusan data ini masuk ke kelas “Layak” Menerima Beasiswa Yayasan Sandjojo.

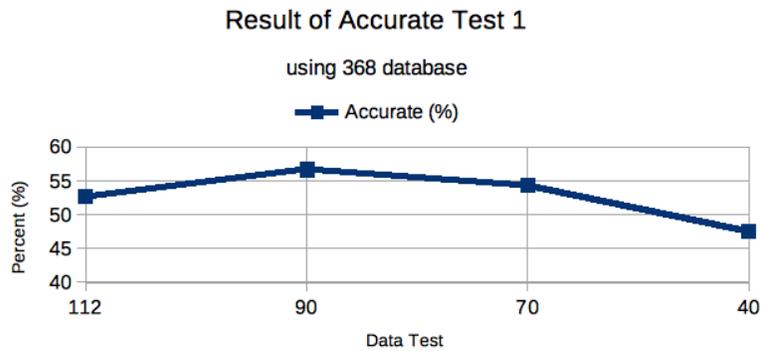
Hasil uji performa dan uji akurasi setiap uji yang dilakukan sebanyak 4 kali dengan kasus data yang berbeda dan hasil yang berbeda pada setiap uji performa dan uji akurasi sebagai berikut :



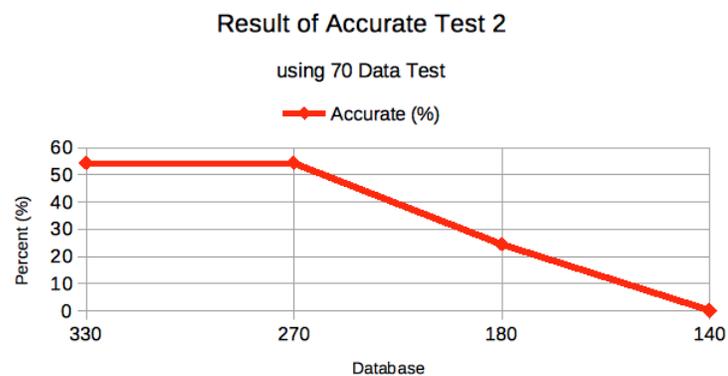
**Gambar 1.** Uji perfoma 1



**Gambar 2.** Uji perfoma 2



**Gambar 3.** Uji akurasi 1



**Gambar 4.** Uji akurasi 2

Dari gambar grafik diatas dapat dilihat kecepatan dalam menentukan kelayak penerima beasiswa dengan metode Naive Bayes kurang dari 3 detik dan akurasi yang dihasilkan mencapai 60% - 65%.

## KESIMPULAN

Dari hasil implementasi dan pengujian dapat disimpulkan hal – hal sebagai berikut :

1. Dalam projek ini penentu kelayakan penerima beasiswa Yayasan Sandjojo dalam tampilan sudah sesuai dengan rancangan design awal yang di buat.
2. Berdasarkan 4 pengujian performance dapat disimpulkan bahwa Metode Naive Bayes dapat menentukan kelayakan penerima beasiswa kurang dari 2.5 detik. Pengujian akurasi Metode Naive Bayes dalam menentukan kelayakan penerima beasiswa Yayasan Sandjojo lebih dari 50% dengan 4 pengujian.

Saran yang diyakini dapat meningkatkan akurasi dalam pengambil keputusan dengan metode Naive Bayes antara lain :

1. Penambahan kriteria

## 2. Penambahan database / data training Bab, Subbab, dan Subsubbab

### DAFTAR PUSTAKA

- [1] A. Jananto, "Algoritma Naive Bayes untuk Mencari Perkiraan Waktu Studi Mahasiswa," *Dinamik*, vol. 18, no. 1, Art. no. 1, 2013, Accessed: May 15, 2021. [Online]. Available: <https://www.unisbank.ac.id/ojs/index.php/fti1/article/view/1669>
- [2] D. Dahri, F. Agus, and D. M. Khairina, "Metode Naive Bayes Untuk Penentuan Penerima Beasiswa Bidikmisi Universitas Mulawarman," *Informatika Mulawarman : Jurnal Ilmiah Ilmu Komputer*, vol. 11, no. 2, Art. no. 2, Sep. 2016, doi: 10.30872/jim.v11i2.211.
- [3] E. Manalu, F. A. Sianturi, and M. R. Manalu, "Penerapan Algoritma Naive Bayes Untuk Memprediksi Jumlah Produksi Barang Berdasarkan Data Persediaan dan Jumlah Pemesanan pada Cv. Papadan Mama Pastries," *J.Mantik Penusa, Mantik Penusa*, vol. 1, no. 2, Art. no. 2, Dec. 2017, Accessed: May 15, 2021. [Online]. Available: <http://e-jurnal.pelitanusantara.ac.id/index.php/mantik/article/view/257>
- [4] Y. V. Via, B. Nugroho, and A. Syafrizal, "SISTEM PENDUKUNG KEPUTUSAN KLASIFIKASI TINGKAT KEGANASAN KANKER PAYUDARA DENGAN METODE NAÏVE BAYES CLASSIFIER," *SCAN - Jurnal Teknologi Informasi dan Komunikasi*, vol. 10, no. 2, Art. no. 2, Jan. 2015, doi: 10.33005/scan.v10i2.609.
- [5] H. Wasiati and D. Wijayanti, "Sistem Pendukung Keputusan Penentuan Kelayakan Calon Tenaga Kerja Indonesia Menggunakan Metode Naive Bayes (Studi Kasus: Di P.T. Karyatama Mitra Sehati Yogyakarta)," *IJNS - Indonesian Journal on Networking and Security*, vol. 3, no. 2, Art. no. 2, Jan. 2014, doi: 10.1123/ijns.v3i2.154.
- [6] T. Jatmikanto and M. A. I. Pakereng, "Implementasi Algoritma Naive Bayes dalam Proses Analisis Efisiensi Jasa Servis Bengkel Daihatsu Cabang Salatiga," p. 1.
- [7] M. Furqan, S. H. Wibowo, S. Kom, M. Kom, and D. Abdullah, "SISTEM PERSEDIAAN OBAT PADA PUSKESMAS MENGGUNAKAN METODE NAÏVE BAYES (STUDI KASUS PUSKESMAS TALANG TINGGI SELUMA)," *JURNAL TEKNIK INFORMATIKA UNIVERSITAS MUHAMADIYAH BENGKULU*, p. 10, 2016.
- [8] M. P. Putri and F. Y. A. Irsyadi, "SISTEM PREDIKSI KESIAPAN SISWA SMP DALAM MENGHADAPI UJIAN NASIONAL MENGGUNAKAN METODE NAÏVE BAYES," *e-Jurnal Mitra Pendidikan*, vol. 2, no. 5, Art. no. 5, May 2018.
- [9] "Beasiswa Sandjojo – Informasi Beasiswa." <https://www.unika.ac.id/infobeasiswa/category/info-beasiswa/beasiswa-sandjojo/> (accessed May 15, 2021).
- [10] C. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer-Verlag, 2006. Accessed: May 15, 2021. [Online]. Available: <https://www.springer.com/gp/book/9780387310732>