

# ANALISIS DATA MINAT CALON MAHASISWA UNIVERSITAS PAMULANG DENGAN MENGGUNAKAN ALGORITMA NAIVE BAYES CLASSIFIER

Eko Suharyanto<sup>1</sup> dan Afrizal Zein<sup>2</sup>

Sistem Informasi, Universitas Pamulang  
Jl. Puspitek Raya No. 10 Buaran, Viktor, Pamulang Selatan, Indonesia  
e-mail : [dosen01830@unpam.ac.id](mailto:dosen01830@unpam.ac.id), [dosen01495@unpam.ac.id](mailto:dosen01495@unpam.ac.id)

## Abstrak

Universitas Pamulang merupakan universitas yang bergerak di bidang pendidikan dengan berbagai jurusan yang ada, mulai dari Ekonomi, Hukum, Teknik, Sastra dan Pendidikan. Unpam mengalami kesulitan dalam mengelola data-data calon mahasiswa yang masuk dalam menentukan minat mahasiswa dalam memilih jurusan yang ada, sehingga jurusan mana yang paling diminati oleh calon mahasiswa tidak bisa diketahui oleh Kampus dan menyebabkan kesulitan dalam menentukan jurusan mana yang akan dipasarkan. Untuk mendukung dalam menentukan jurusan mana yang akan dipasarkan dan diprioritaskan digunakanlah *Metode Naive Bayes Classifier* membantu memberikan solusi dalam memprediksi minat calon mahasiswa terhadap jurusan yang ada pada data pada semester sebelumnya. Data penelitian berupa data sekunder yang diperoleh dari IT-Center Universitas Pamulang sebesar 5142. Data tersebut selanjutnya akan dianalisis untuk memetakan pola minat calon mahasiswa di Universitas Pamulang dengan menggunakan metode *Metode Naive Bayes Classifier*.

Hasil analisis *Metode Naive Bayes Classifier* menunjukkan bahwa minat calon mahasiswa yang akan masuk ke UNPAM didominasi oleh Fakultas Ekonomi jurusan Manajemen 45,19%, Teknik Informatika 22,09%, Akuntansi 13,71%, Hukum Perdata 13,57% dan Sistem Informasi 5,48% Sehingga, dapat membantu dalam menentukan jurusan apa yang paling banyak diminati oleh calon mahasiswa sesuai dengan kebutuhan dan keinginan serta menghasilkan alternative pilihan jurusan yang lain.

**Kata Kunci:** Analisis Data, Naive Bayes, dan Classifier

## Abstract

*Pamulang University is a university engaged in education with various majors, ranging from Economics, Law, Engineering, Literature and Education. Unpam has difficulty managing the data of prospective students who enter in determining student interest in choosing existing majors, so which majors are most in demand by prospective students cannot be known by the campus and cause difficulties in determining which majors will be marketed. To support in determining which majors will be marketed and prioritized, the Naive Bayes Classifier Method is used to help provide solutions in predicting the interest of prospective students in the majors in the data in the previous semester. The research data in the form of secondary data obtained from the IT-Center of Pamulang University is 5142. The data will then be analyzed to map the patterns of interest of prospective students at Pamulang University using the Naive Bayes Classifier Method.*

*The results of the analysis of the Naive Bayes Classifier Method show that the interest of prospective students who will enter UNPAM is dominated by the Faculty of Economics majoring in Management 45.19%, Informatics Engineering 22.09%, Accounting 13.71%, Civil Law 13.57% and Information Systems 5,48% Thus, it can assist in determining what majors are most in demand by prospective students according to their needs and desires and produce alternative choices of other majors.*

**Keywords:** Data Analysis, Naive Bayes, and Classifier

## 1. PENDAHULUAN

Universitas Pamulang merupakan salah satu universitas yang berada di Tangerang Selatan yang sudah cukup ternama. Apabila data calon mahasiswa yang masuk dalam periode tertentu dianalisa maka dapat diketahui pola yang sangat membantu dalam memajukan penerimaan mahasiswa guna memajukan kampus. Dengan adanya data calon mahasiswa yang masuk Setiap hari, datanya semakin besar. Data tidak hanya dapat berfungsi sebagai arsip, tetapi juga dapat dimanfaatkan dan diolah menjadi informasi yang bermanfaat untuk meningkatkan minat calon mahasiswa.

Permasalahan yang sering timbul yaitu sering sekali calon mahasiswa kesulitan dalam memilih jurusan yang ada, karena pemahaman mengenai jurusan serta target yang akan diambil dari calon mahasiswa untuk penyesuaian dari jurusan yang dipilih. Universitas Pamulang belum memanfaatkan data calon mahasiswa yang masuk dan biasanya hanya menjadi arsip yang tidak dimanfaatkan. Hal ini tentu menjadi kelemahan Unpam, disebabkan data ketertarikan calon mahasiswa dalam memilih jurusan tidak diketahui. Pada dasarnya pendataan calon mahasiswa memiliki Informasi yang berguna untuk mengambil keputusan dan mendapatkan pengetahuan baru tentang pola minat pelanggan terhadap produk yang di jual.

Pengolahan data dapat dilakukan dengan menggunakan teknologi tertentu. Salah satu teknik yang digunakan dalam pengolahan data adalah algoritma Naive Bayes Classifier (NBC). Sebagaimana penelitian Wijaya dan Dwiasnati yang berjudul Implementasi Data Mining Penjualan Obat Menggunakan Algoritma Naive Bayesian, pertanyaan yang ditujukan untuk menganalisis penentuan produk mana yang diklasifikasikan sebagai dijual dan tidak dijual, dengan sebuah kesimpulan mengetahui pola penjualan obat yang efektif dan tidak efektif berdasarkan jenis variasi obat yang baru (Syarli, 2022). Terdapat juga penelitian yang dilakukan oleh Ramadhan dkk (2018) dengan judul implementasi Data Mining untuk prestasi

akademik siswa dinilai menggunakan algoritma classifier Naive Bayes.

Tujuannya adalah Bentuklah tabel probabilitas yang akan dijadikan sebagai dasar untuk proses pengklasifikasian prestasi lulusan mahasiswa akademik yang akan diklasifikasikan, dan berdasarkan sejarah nilai-nilai yang telah diklasifikasikan, untuk memberikan rekomendasi nilai optimal untuk proses wisuda yang paling ceat dan tepat waktu. Difoto oleh siswa. Algoritma *Naive Bayes Classifier* (Jadhav et al (2016)) mengemukakan bahwa pengklasifikasi Naive Bayes merupakan model yang berdiri sendiri, ini membahas klasifikasi sederhana berdasarkan teorema Bayes. *Naive Bayes* adalah algoritma yang dapat mengklasifikasikan variabel menggunakan metode probabilistik dan statistik (Widianto, 2019).

Penelitian sebelumnya yang telah dilakukan di Unpam hanya mengambil judul perencanaan pendaftaran calon mahasiswa berbasis web, sedangkan untuk analisa data calon mahasiswa belum pernah dilakukan sebuah penelitian sehingga perlu adanya penelitian ini guna membantu menganalisis minat calon mahasiswa terhadap jurusan di Universitas Pamulang dan wadah koordinasi kegiatan penguatan fungsi keluarga secara terpadu (Fujianti,. 2021).

## 2. TINJAUAN PUSTAKA

Dalam penelitian ini, penulis mencari beberapa *referensi* dari penelitian-penelitian yang telah dilakukan sebelumnya yang berkaitan dengan latar belakang dan rumusan masalah yang ada pada penelitian ini sehingga saling berkaitan. Berikut beberapa penelitian terdahulu yang bisa dijadikan penulis referensi yaitu:

- a. Implementasi dari Algoritma Naive Bayes buat mengetahui minat beli pelanggan untuk sarang burung walet (Jurnal Sains Komputer dan Teknologi Informasi e – ISSN: 2655-7460. Volume 4 No.1, November 2021).

Dalam penelitian ini bertujuan untuk memahami minat beli konsumen

terhadap sarang burung walet berdasarkan jumlah permintaan yang semakin bertambah.

- b. Sistem Prediksi Bunga Penjualan Jaket Grosir Murah Kediri Menggunakan Metode Naive Bayes (2011)

Penelitian ini bertujuan untuk meramalkan minat pasar supaya terus tumbuh dan eksis atas dasar membanjirnya produk terbaru. Penelitian ini memakai algoritma pengklasifikasi Naive Bayes untuk memprediksi minat pasar. Hasil yang diperoleh dari penelitian ini adalah pendekatan naive Bayesian untuk memprediksi minat terhadap produk jaket menggunakan 50 record data memiliki akurasi 90%, sedangkan akurasi 100% dan recall 60%.

Berdasarkan penelitian yang ada, topik yang diambil berbeda dengan topik yang sudah ada sebelumnya. Judul topik yang digunakan adalah Analisis data ketertarikan customer terhadap produk yang ada di Universitas Pamulang dengan menggunakan Algoritma NBC. Perbedaan dari penelitian sebelumnya adalah objek dan pembahasan belum pernah dilakukan penelitian sebelumnya.

### 3. METODOLOGI PENELITIAN

Pada step ini akan dilakukan analisis kebutuhan dan spesifikasi masalah yang akan dipecahkan. Mulai dengan mengidentifikasi data set calon mahasiswa yang meliputi data nama calon mahasiswar, nomor telepon, jenis kelamin, tujuan belajar, Jurusan yang diminati, dan media yang digunakan untuk mencari informasi seputar Kampus Universitas Pamulang.

#### Metoda Pemilihan Sampel.

Metode pemilihan sampel adalah metode yang digunakan untuk mengambil sampel sebagai bahan penelitian. Dalam masalah penelitian ini penulis akan menggunakan sampel data 2 semester untuk memberikan data mahasiswa yang masuk di Universitas Palmulang.

#### Metoda Pengumpulan Data

Dalam pengumpulan data metode yang digunakan adalah metode untuk memperoleh data dan informasi yang

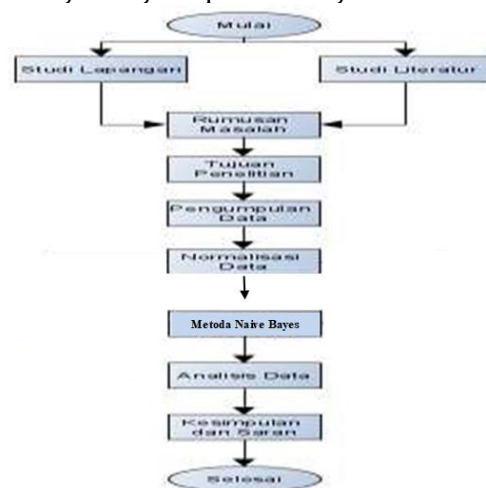
diperlukan untuk penelitian. Didalam penelitian saat ini digunakan dua jenis data untuk mendukung persiapan studi (Purnomo dan Syamsul 2018), yaitu:

1. Data Asli adalah sampel data yang diperoleh langsung dari objek penelitian, meliputi data bentuk bangunan dan data nama ruangan.
2. Data Sekunder adalah data kedua yang diperoleh secara tidak langsung dengan media melalui perantara, yaitu berupa dokumen-dokumen dan studi kepustakaan, yang sebelumnya direkam oleh pihak lain.

#### Perancangan Penelitian

Tujuan pengembangan dan desain Penelitian ini untuk menjelaskan arah dari penelitian dan tujuan yang ingin diraih.

Pemecahan masalah dari penelitian juga efektif jika tujuan penelitian jelas.



**Gambar 1.** Tahapan Perancangan Penelitian Analisis Data

Setelah semua data lengkap, data yang diperoleh dianalisis menggunakan metode teknik pengelolaan data. Analisis data dan model yang digunakan oleh penulis dalam melaksanakan penelitian ini dirancang untuk memperoleh jawaban dari pertanyaan-pertanyaan yang telah tercantum didalam *Question Identification*. Analisis data merupakan salah satu kegiatan penelitian berupa proses menyusun dan mengelola data untuk menginterpretasikan data yang telah diperoleh. Menurut Sugiyono (2016: 147), implikasi analisis data adalah sebagai berikut “Analisis data adalah

kegiatan yang mengikuti pengumpulan data dari semua responden atau sumber data lainnya. Kegiatan dalam analisis data adalah mengelompokkan data berdasarkan variabel dan responden Data ditabulasi, data untuk setiap penelitian disajikan, perhitungan dilakukan untuk menjawab rumus-rumus soal, dan dilakukan perhitungan untuk menguji hipotesis yang telah dikemukakan.” Analisis data digunakan untuk mengolah data menjadi informasi, data tersebut akan menjadi lebih mudah untuk dipahami dan diinterpretasikan. Data yang akan dianalisis adalah data penelitian dari studi lapangan dan studi pustaka.

**Pemodelan (Modelling)**

Pada penelitian ini, klasifikasi menggunakan algoritma Naive Bayes akan diuji untuk mengetahui hasil komputasi dari analisis dan untuk mengetahui apakah fungsi tersebut bekerja dengan baik. Bantuan dengan program PHP.

**Metode Analisis Data**

Bayes adalah teknik peramalan berbasis probabilitas sederhana berdasarkan penerapan teorema Bayes (atau aturan Bayes) dan dengan asumsi independensi yang kuat (naif). Artinya, di Naive Bayes, model yang digunakan adalah "model fitur independen".

**4. HASIL DAN PEMBAHASAN**

Salah satu konsep yang efektivitas dan memenuhi produktivitas adalah merupakan konsep yang harus dipatuhi oleh para pelaku usaha. Dasar dari konsep ini dapat mengarahkan bisnis untuk mencapai profitabilitas yang maksimal. Langkah untuk memenangkan persaingan adalah dengan menerapkan strategi pemasaran yang sesuai dengan target dari segi kualitas, harga dan daya saing produk yang dihasilkan. Di era sekarang ini yang banyak COVID-19, strategi untuk pemasaran yang tepatguna adalah pemasaran secara online. Pemasaran via internet melalui media online merupakan langkah bagus yang dapat dilaksanakan oleh Oleh sebab itu, Unpam mengaplikasikan strategi pemasaran yang memungkinkan lulusan yang baik diterima

oleh lingkungan sekitar masyarakat, melalui pemasaran via internet, bisa melalui media sosial yang juga dikenal dengan pemasaran berbasis teknologi digital.

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

Dimana:

X = Data class yang masih belum diketahui

H = Hipotesis data X

P(H|X) = Probabilitas hipotesis

P(H) = Probabilitas hipotesis H (Prior probability)

P(X|H) = Probabilitas X / H

P(X) = Probabilitas dari X.

Penjabaran selanjutnya untuk rumus Naïve Bayes tersebut dapat dilakukan melalui penjabaran (C|X1...,Xn) menggunakan aturan perkalian sebagai berikut.

$$P(C|X1, ..., Xn) = P(C)P(X1, ..., Xn|C) \\
P(C)P(X1|C)P(X2, ..., Xn|C, X1) \\
P(C)P(X1|C)P(X2|C, F1)P(X3, ..., Xn|C, X1, X2) \\
P(C)P(X1|C)P(X2|C, X1)P(X3, ..., Xn|C, X1, X2, X3) \\
P(C)P(X1|C) ... P(Xn|C, X1, X3, ..., Xn - 1) \quad (2)$$

Dapat diperhatikan bahwa jika semakin kompleks faktor yang bisa mempengaruhi nilai probabilitas, semakin kecil kemungkinan untuk menghitung nilai-nilai ini secara individual. Akibatnya, komputasi menjadi semakin sulit, sehingga di sinilah asumsi ketergantungan yang sangat tinggi dapat digunakan, bahwa setiap atribut dapat saling independen. (Xi|Xj) === (Xi) untuk I j, jadi: (Xi | , Xj) = (Xi|C) ... (3) Didaalam sebuah metode Naive Bayesian, data latih dan data uji perlu diklasifikasikan, dan semakin banyak data latih yang terlibat, semakin baik. Hasil yang baik memberikan prediksi. Komputasi P(Ci), yang merupakan probabilitas prior untuk setiap subkelas C, akan dibangkitkan menggunakan persamaan berikut: Dimana: Si: total data latih dari kategori Ci S: Jumlah total data pelatihan.

**Tabel 1. Perhitungan Probabilitas Prior (P(Ci))**

X2	Jumlah kejadian "Dipilih"		Probabilitas	
	minat	Tidak minat	Minat (C0)	Tidak minat (C1)
informasi	12	9	0.190476	0.243243
ig	28	10	0.444444	0.27027
web	4	1	0.063492	0.027027
brostur	3	3	0.047619	0.081081
baliho	2	1	0.031746	0.027027
twitter	2	1	0.031746	0.027027
fb	2	1	0.031746	0.027027
teman	8	10	0.126984	0.27027
seminar	3	1	0.047619	0.027027
majalah	1	1	0.015873	0.027027
Jumlah	63	37	1	1

Dataset akan menggunakan metode Naive Bayes untuk multistage processing, sehingga perlu ditentukan nilai datanya yaitu dari 100 data latih yang digunakan diketahui kelas C0 (Puas) sebanyak 63 data, dan kelas C1 (tidak memuaskan) adalah 37 data. Perhitungan prior probability dapat diselesaikan dengan rumus (3), yaitu:

**Tabel 2. Probabilitas Prior**

Perhitungan Prior	Hasil	
P (C0)	63/100	0,63
P (C1)	37/100	0,37

$$P(X2|C)P(X3|C) \dots$$

$$P(C|X1, \dots, Xn) = P(X1|C)$$

$$= \prod_{i=1}^n P(Xi|C) \quad (3)$$

Perhitungan probabilitas posterior dilakukan pada 100 data latih, dengan X sebagai vektor untuk menentukan status keluarga, yaitu kategori X, informasi X, harga X, target X, kategori X, minat X. Dengan demikian P(X|Ci) dapat diubah menjadi:

**Tabel 3. Probabilitas Jenis Fakultas**

X1	Jumlah kejadian "Dipilih"		Probabilitas	
	minat	Tidak minat	Minat (C0)	Tidak minat (C1)
Jenis kelas	8	3	0.126984	0.08108108
Teknik	34	17	0.539683	0.45945946
Ekonomi	4	3	0.063492	0.08108108
Pasca	5	4	0.079365	0.10810811
Pendidikan	4	3	0.063492	0.08108108
Sastra	6	5	0.095238	0.13513514
Hukum	2	2	0.031746	0.05405405
Mipa	63	37	1	1

Tabel 4. Probabilitas informasi

**Tabel 4. Probabilitas Informasi**

X2	Jumlah kejadian "Dipilih"		Probabilitas	
	minat	Tidak minat	Minat (C0)	Tidak minat (C1)
informasi	12	9	0.190476	0.243243
ig	28	10	0.444444	0.27027
web	4	1	0.063492	0.027027
brostur	3	3	0.047619	0.081081
baliho	2	1	0.031746	0.027027
twitter	2	1	0.031746	0.027027
fb	2	1	0.031746	0.027027
teman	8	10	0.126984	0.27027
seminar	3	1	0.047619	0.027027
majalah	1	1	0.015873	0.027027
Jumlah	63	37	1	1

X3	Jumlah kejadian "Dipilih"		Probabilitas	
	minat	Tidak minat	Minat (C0)	Tidak minat (C1)
harga	1	3	0.015873	0.081081
Sangat tinggi	9	1	0.142857	0.027027
tinggi	51	31	0.809524	0.837838
sedang	1	1	0.015873	0.027027
rendah	1	1	0.015873	0.027027
Sangat rendah	63	37	1	1

X4	Jumlah kejadian "Dipilih"		Probabilitas	
	minat	Tidak minat	Minat (C0)	Tidak minat (C1)
target	3	1	0.047619	0.027027
kuliah	4	3	0.063492	0.081081
bekerja	45	23	0.714286	0.621622
skill	8	8	0.126984	0.216216
hobby	3	2	0.047619	0.054054
liburan	63	37	1	1

X5	Jumlah kejadian "Dipilih"		Probabilitas	
	minat	Tidak minat	Minat (C0)	Tidak minat (C1)
Jurusan	35	3	0.620635	0.864865
Manajemen	15	5	0.079365	0.135135
Teknik Informatika	6	3	1	1

**Pendefinisian Variabel**

Berdasarkan Tabel 4. berapa dapat didefinisikan data uji X adalah sebagai berikut:

X = {Xjeniskelas=private,  
 Xinformasi=web, Xharga=tinggi,  
 Xtarget=bekerja, Xkelas=online}

Pendefinisian Probabilitas Prior P(Ci) Hasil pendefinisian Probabilitas prior berdasarkan persamaan (5) menghasilkan nilai untuk kelas puas (C0) sebesar 0.63 dan untuk Tidak Puas (C1) sebesar 0.37.

Perhitungan Probabilitas Data Uji Berdasarkan data uji pada tabel 4.3 dapat dilakukan klasifikasi kedalam kelas puas (C0) dengan ketentuan nilai masing-masing kriteria yaitu: 0.127, 0.190, 0.016, 0.048, 0.921.

Kemudian nilai dari masing-masing kriteria tersebut dikalikan P(X|C0) = 0.127\*0.190\*0.016\*0.048\*0.921 = 1.70678.

Untuk menghitung klasifikasi kedalam kelas tidak puas (C1) dengan ketentuan nilai masing-masing kriteria yaitu: 0.081, 0.243, 0.081, 0.027, 0.865. Kemudian nilai dari masing-masing kriteria tersebut dikalikan.

$$P(X|C1) = 0.081 * 0.243 * 0.081 * 0.027 * 0.865 = 3.72354$$

Rangking	Fakultas	Jurusan	Total Penerimaan	Presentasi
1	Ekonomi	Manajemen	2252	45,19 %
2	Teknik	Teknik Informatika	1099	22,06 %
3	Ekonomi	Akuntansi	683	13,71 %
4	Hukum	Perdata	676	13,57 %
5	Teknik	Sistem Informasi	273	5,48 %

Gambar 2. Simulasi Pengelompokan Minat

No	Nama Mahasiswa	Prodi	Fakultas	Peker.Ayah	Peker.Ibu	Peng
1	ADE IRMA MANURUNG	SISTEM INFORMASI	TEKNIK	PETANI/NELAYAN/SUPIR	IBU RUMAH TANGGA	2 - 2,9
2	LISKA BATU AGARA	PENDIDIKAN EKONOMI	FAKULTAS PENDIDIKAN	PETANI/NELAYAN/SUPIR	IBU RUMAH TANGGA	500 rit ribu
3	FADILLAH ANI HASIBUAN	SASTRA INGRIS	SASTRA	TUKANG/TIDAK TETAP/LAIN-LAIN	IBU RUMAH TANGGA	500 rit ribu

Gambar 3. Data Latih

Hasil Analisa Naïve Bayes didapat minat calon mahasiswa yang akan masuk ke UNPAM didominasi oleh Fakultas Ekonomi jurusan Manajemen

Simulasi Pengelompokan Minat Calon Mahasiswa UNPAM

Nama Calon Mahasiswa: Andin

Fakultas Yang Diminati: Ekonomi

Pekerjaan Ayah: PNS/TNI/POUR/BUMN

Pekerjaan Ibu:

Gambar 4. Hasil Analisa

## 5. SIMPULAN

Berdasarkan hasil penelitian dan pembahasan maka kesimpulan dalam penelitian ini adalah terdapat 5 program studi yang menjadi pilihan oleh siswa/I SMA/K dan dengan program studi antara lain antara lain Fakultas Ekonomi prodi Manajemen, Fakultas Ilmu Komputer Prodi Teknik Informatika, Fakultas Ekonomi Prodi Akuntansi, Fakultas Hukum Prodi Perdata dan Fakultas Ilmu Komputer Prodi Sistem Informasi

## DAFTAR PUSTAKA

- Baskoro, Ambarwati, L., dan Wicaksana, I. S. (2013). Belajar Data Mining Dengan RapidMiner. *Jurnal*, 125.
- Karthika, S., & Sairam, N. (2015). A Naïve Bayesian Classifier for Educational Qualification. *Indian Journal of Science and Technology*, 8(16), 1–5. <http://doi.org/10.17485/ijst/2015/v8i16/62055>.
- Medkhar, D. S., Bote, M. P., & Deshmukh, S. D. (2013). Heart Disease Prediction System Using Naive Bayes. *International Journal Of Enhanced Research In Science Technology & Engineering*, 2(3), 1–5. <http://doi.org/10.1093/infdis/139.2.232>
- JADHAV, A., PANDITA, A., PAWAR, A., & SINGH, V. (2016). Classification of Unstructured Data using Naïve Bayes Classifier and Predictive Analysis for RTI Application. *ABHIYANTRIKI: An International Journal of Engineering & Technology*, 3(6), 1–6.
- Widianto, Haldi Mochammad. 2019. Algoritma Naïve Bayes. (Online), (<https://binus.ac.id/bandung/2019/12/alg-oritma-naive-bayes/>), diakses 08 Desember 2020).
- Fujianti, L., Susilowati, Soemarsono, Sri Irvianti, Kenigi Harisandi. (2021). Meningkatkan Keahlian Pembukuan Berbasis Handphone Bagi Umkm Posdaya Cempaka. *SULUH : Jurnal Abdimas*. Vol. 3 (1): 81 –88.

**H.Purnomo, LN Syamsul (2018)**, Teknologi Informasi: Teori, Konsep, Dan Implementasi, STIMATA, Jl. LA. Sucipto 249 A Malang, 2018.

**Sugiyono, Metode Penelitian Pendidikan, (Bandung: Alabeta, 2016)**, hal. 14 63  
Usman Rainse dan Abidin, Metodologi Penelitian Sosial dan Ekonomi Teori dan Aplikasi, (Bandung: Alfabeta, 2012), Hal. 19-20.

**Meilani, B. D., & Susanti, N. (2015)**. Aplikasi Data Mining Untuk Menghasilkan Pola Kelulusan Siswa Ddengan Metode Naive Bayes. Jurnal Ilmiah NERO, 1(3), 182–189.

**Saleh, A. (2015)**. Implementasi Metode Klasifikasi Naive Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga. CITEC Journal, 2(3), 207–217.

**Suryadi, A., & Harahap, E. (2018)**. Sistem Rekomendasi Penerimaan Mahasiswa Baru Menggunakan Naive Bayes Classifier Di Institut Pendidikan Indonesia. JOUTICA, 3(2), 171–182.