



PENERAPAN MODEL NAIVE BAYES UNTUK MEMPREDIKSI POTENSI PENDAFTARAN SISWA DI SMK TAMAN SISWA TELUK BETUNG BERBASIS WEB

Sherlyn Eka Yuliana Putri¹, Ade Surahman

S1Informatika¹(Universitas Teknokrat Indonesia)

S1Teknik Komputer²(Universitas Teknokrat Indonesia)

sherlyneka06@gmail.com¹, adesurahman@teknokrat.ac.id²

Received: (Juni 2020) Accepted: (Juni 2020) Published: (Juni 2020)

Abstract

Taman Siswa Vocational School at the beginning of each new school year will hold the admission of prospective new students. Competition in attracting prospective new students to enter SMK Taman Siswa can occur. With the competition between SMKs, the number of new student registrations every year will increase and it is possible that there will be a decrease in prospective new students, so the number of new students at the SMK is not ideal. The problem is the inability to predict the number of prospective applicants for new students in SMK Taman Siswa, causing unprepared effective policies for the school. Based on these problems, the authors designed a system of applying the Naive Bayes Model to Predict the Potential Student Registration at SMK Taman Students Teluk Betung Based on WEB. A system that can predict the potential for new student registrations to enter the Taman Siswa Vocational School. The final result obtained is a prediction information system for student registration potential using the Naive Bayes model which has the ability to predict the potential number of students registering and not registering at SMK Taman Siswa. This application was tested and showed an accuracy of 87% for predictions of potential student registration.

Keywords: *Implementation, Naive Bayes, Registration, Students, Website.*

Abstrak

SMK Taman Siswa setiap awal tahun ajaran baru akan menyelenggarakan penerimaan calon peserta didik baru. Persaingan dalam menarik minat calon siswa baru untuk masuk SMK Taman Siswa dapat terjadi. Dengan adanya persaingan antar SMK, jumlah pendaftaran siswa baru setiap tahunnya akan mengalami peningkatan dan tidak menutup kemungkinan akan mengalami penurunan calon peserta didik baru, sehingga jumlah siswa baru di SMK tersebut tidak ideal. Permasalahannya adalah ketidakmampuan dalam memprediksi jumlah calon pendaftar peserta didik baru di SMK Taman Siswa, menyebabkan tidak dapat di siapkannya kebijakan yang efektif bagi pihak sekolah. Berdasarkan masalah tersebut, maka penulis merancang sebuah sistem Penerapan Model Naive Bayes Untuk Memprediksi Potensi Pendaftaran Siswa Di SMK Taman Siswa Teluk Betung Berbasis WEB. Sistem yang dapat melakukan prediksi potensi pendaftaran siswa baru yang akan masuk ke SMK Taman Siswa. Hasil akhir yang diperoleh adalah sebuah sistem informasi prediksi potensi pendaftaran siswa dengan menggunakan model Naive Bayes yang memiliki kemampuan untuk memprediksi potensi banyaknya siswa yang mendaftar dan tidak mendaftar di SMK Taman Siswa. Aplikasi ini diuji dan menunjukkan akurasi sebesar 87% untuk prediksi potensi pendaftaran siswa.

Kata Kunci: *Penerapan, Naive Bayes, Pendaftaran, Siswa, Website.*

To cite this article:

Sherlyn Eka Yuliana Putri, Ade Surahman. (2020). PENERAPAN MODEL NAIVE BAYES UNTUK MEMPREDIKSI POTENSI PENDAFTARAN SISWA DI SMK TAMAN SISWA TELUK BETUNG BERBASIS WEB. *Jurnal Ilmiah Mahasiswa Rekayasa Perangkat Lunak*, Vol(1), 93-99.

PENDAHULUAN

Persaingan antar Sekolah Menengah Atas (SMA) atau Sekolah Menengah Kejuruan (SMK) Swasta semakin meningkat belakangan ini. Perkembangan teknologi juga ikut andil dalam mempengaruhi perkembangan pendidikan di Indonesia, khususnya SMK Swasta. Secara tidak langsung menimbulkan persaingan dalam menarik minat calon siswa baru untuk masuk SMK Swasta tersebut. Dengan adanya

persaingan antar SMK jumlah pendaftaran siswa baru setiap tahunnya akan mengalami peningkatan dan tidak menutup kemungkinan akan mengalami penurunan calon peserta didik baru, sehingga jumlah siswa baru di SMK tersebut tidak ideal. Pada setiap awal tahun ajaran baru pihak SMK Taman Siswa akan menyelenggarakan penerimaan calon peserta didik baru. Ketidakstabilan dalam pendaftaran calon peserta didik baru di SMK Taman Siswa juga bisa terjadi. Untuk hal itu akan lebih baik jika pihak penyelenggara mampu melakukan antisipasi untuk meminimalisir kelemahan yang ada dengan melakukan prediksi potensi pendaftaran siswa baru yang akan masuk ke SMK Taman Siswa sehingga dapat disiapkan dan diberlakukan kebijakan yang efektif bagi sekolah.

Berdasarkan uraian masalah diatas, peneliti mengusulkan pembuatan sistem yang dapat memberikan informasi prediksi potensi pendaftaran siswa baru dengan menggunakan metode Naive Bayes dan parameter yaitu usia, jenis kelamin, alamat, asal pendidikan, alamat pendidikan, jurusan, agama, prediksi potensi pendaftaran. Peneliti berharap dengan dilakukan penelitian ini dapat memberikan kebutuhan data yang sesuai dalam menginformasikan prediksi potensi pendaftaran mahasiswa baru.

Tujuan penelitian ini adalah untuk membangun sistem prediksi potensi pendaftaran siswa baru di SMK Taman Siswa dan hasil prediksi potensi pendaftaran mahasiswa baru di SMK Taman Siswa memiliki akurasi yang baik

TELAAH PUSTAKA

1. Data Mining

Data Mining merupakan proses pengekstraksian informasi dari sekumpulan data yang sangat besar Saleh (2015). Pengelolaan Data Mining melalui penggunaan algoritma dan teknik penarikan dalam bidang statistik, pembelajaran mesin dan sistem manajemen basis data (Taruna, 2008). Data mining adalah proses menganalisa data dari perspektif yang berbeda dan menyimpulkan data menjadi informasi-informasi penting yang dapat meningkatkan kualitas data yang akan dipakai. Dapat ditarik kesimpulan bahwa Data Mining merupakan proses mengumpulkan data berukuran besar yang akan diekstraksi kemudian menjadikan data tersebut menjadi informasi – informasi data yang nantinya dapat digunakan.

2. Naive Bayes

Naive Bayes Classifier merupakan sebuah metode klasifikasi probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan (Saleh, 2015). Teorema Bayes dikombinasikan dengan “Naive” yang berarti setiap atribut/variabel bersifat bebas (*independent*). Naive Bayes Classifier dapat dilatih dengan efisien dalam pembelajaran terawasi (*supervised learning*) (Gunawan, 2018). Untuk menjelaskan metode Naive Bayes, proses klasifikasi memerlukan sejumlah petunjuk untuk menentukan kelas apa yang sesuai bagi data yang dianalisis.

Persamaan metode Naive Bayes dapat ditunjukkan seperti pada gambar 1 berikut:

$$P(x_1, x_2, \dots, x_n | c_j) = P(c_j) \prod_i P(x_i | c_j)$$

Dimana :

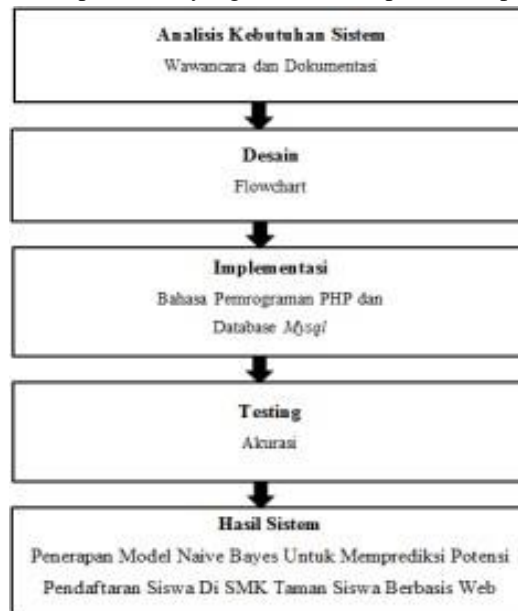
- $P(c_j)$ = Peluang hipotesis j (*Prior Probability*)
- $P(x_1, x_2, \dots, x_n | c_j)$ = Peluang atribut-atribut (inputan) jika diketahui hipotesis c_j (*Likelihood*)
- $P(x_i | c_j)$ = *Posterior Probability*

Gambar 1. Persamaan Naive Bayes

METODE PENELITIAN

Tahapan Pemikiran

Tahapan pemikiran yang melandasi penelitian yang dilakukan dapat dilihat pada Gambar 1 berikut:



Gambar 1. Tahapan Pemikiran

Analisis Kebutuhan Sistem

Dalam penelitian ini, peneliti melakukan analisis kebutuhan sistem dengan cara pengumpulan data dengan cara :

1. Wawancara
Wawancara dilakukan kepada pihak SMK Taman Siswa Teluk Betung untuk menggali informasi tentang masalah yang terjadi terkait dengan pendaftaran siswa baru.
2. Dokumentasi
Dokumentasi dilakukan untuk mengumpulkan data yang dibutuhkan untuk penelitian. Dalam penelitian ini, data-data penelitian didapatkan dari data siswa dengan Tahun ajaran 2017/2018 sebanyak 66 dataset Tahun ajaran 2018/2019 sebanyak 81 dataset, Jumlah data pendaftaran keseluruhan 147

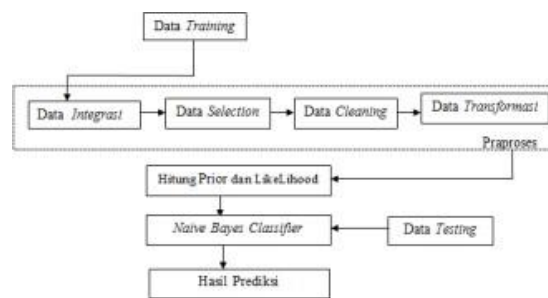
Berdasarkan wawancara dan dokumentasi yang dilakukan, peneliti meringkas bahwa masalah yang terjadi yaitu tidak stabilnya jumlah calon siswa (pendaftar) di SMK Taman Siswa Teluk Betung sehingga jumlah siswa menjadi tidak ideal. Hal ini terjadi karena belum tersedianya media atau sistem yang membantu memprediksi jumlah calon pendaftar sehingga kebijakan yang efektif tidak dapat disiapkan lebih awal oleh pihak sekolah.

Desain

Berdasarkan masalah yang ada, peneliti kemudian mendesain sebuah sistem untuk melakukan prediksi potensi pendaftaran siswa SMK Taman Siswa. Berikut analisis kebutuhan fungsional untuk sistem yang menggunakan pendekatan atau algoritma Naïve Bayes Classifier:

1. Admin dapat masuk kedalam sistem
2. Admin dapat memasukkan data siswa asli dan data siswa praproses
3. Sistem dapat menghitung nilai Prior
4. Sistem dapat menghitung nilai Likelihood
5. Admin dapat menyimpan, mengubah, menghapus data yang ada pada sistem
6. Sistem dapat menampilkan informasi data siswa
7. Sistem dapat menampilkan data siswa berdasarkan atribut yang diinginkan

Desain arsitektur dalam bentuk flowchart yang dirancang berdasarkan analisis kebutuhan fungsional adalah seperti pada gambar 2 berikut:



Gambar 2. Desain Flowchart

Dataset

Sumber data yang digunakan pada penelitian ini adalah data pendaftaran siswa SMK Taman Siswa tahun ajaran 2017/2018 dan 2018/2019. Jumlah data pendaftaran keseluruhan 147 dataset dari:

- a. Tahun ajaran 2017/2018 sebanyak 66 dataset
- b. Tahun ajaran 2018/2019 sebanyak 81 dataset

Dataset yang didapatkan terdiri dari 19 fitur atribut. Dari 147 dataset tersebut kemudian dikelola untuk dijadikan data *training* dan data *testing*. Data training dan data testing yang digunakan pada penelitian ini memiliki prosentase yang sama yaitu 100% untuk data *training* dan 100% untuk data *testing*.

HASIL DAN PEMBAHASAN

Data Integration

Pada tahap ini, dilakukan penggabungan data dari berbagai sumber. Pada penelitian ini, integrasi dilakukan dengan menggabungkan data tahun ajaran 2017/2018 sebanyak 66 dataset dan tahun ajaran 2018/2019 sebanyak 81 dataset.

Data Selection

Pada tahap ini, dilakukan seleksi atribut atau fitur. Dimana dari 19 fitur yang ada, diseleksi menjadi 8 fitur meliputi jenis kelamin, tempat tanggal lahir, agama, alamat, asal pendidikan, jurusan, pekerjaan orang tua, status pendaftaran seperti pada contoh gambar 3.

Jenis Kelamin	Tempat, Tanggal Lahir	Jurusan	Agama	Alamat Peserta Didik	Sekolah Asal	Pekerjaan Orang Tua	Status Pendaftar
					Nama Sekolah		
Laki-laki	Bandar Lampung, 07 Januari 2002	AK	Islam	Jl. Yos Sudarso Gg. Bakau I LK III RT. 41 Bumi Waras	SMP Negeri 11 Bandar Lampung	Buruh	Daftar
Perempuan	Bandar Lampung, 22 Desember 2001	AK	Islam	Jl. Slamet Riyadi II LK III RT. 44 Bumi Waras	MTs Hasanuddin Telukbetung	Wirausaha	Tidak daftar
Perempuan	Bandar Lampung, 15 Juni 2001	AK	Islam	Jl. Ikan Sepat No. 23 LK III RT. 47 Pesawahan Teluk Betung Selatan	SMP Tamansiswa Telukbetung	Karyawan Swaswta	Daftar

Gambar 3. Hasil data selection

Data Cleaning

Pada tahap ini dilakukan pembersihan dataset jika ditemukan data kosong atau salah isi. Namun dari 147 dataset yang digunakan, tidak ada yang kosong atau salah isi. Sehingga luaran dari tahap ini, dataset masih berjumlah 147.

Data Transformation

Pada tahap ini dilakukan transformasi data sehingga sesuai dengan kebutuhan algoritma. Berdasarkan data yang diberikan oleh SMK Taman Siswa peneliti mengubah jenis kelamin dari Laki-Laki dan Perempuan menjadi L/P, alamat peserta didik hanya keterangan kecamatan, usia peneliti dapatkan dari data tahun lahir peserta didik, jurusan Akuntansi Teknik Komputer dan Jaringan Administrasi Perkantoran menjadi AK TKJ AP, seperti pada gambar 4 berikut.

Jenis Kelamin	Usia	Jurusan	Agama	Alamat Peserta Didik	Sekolah Asal	Pekerjaan Orang Tua	Status Pendaftar
					Nama Sekolah		
L	14	AK	Islam	Bumi Waras	SMP Negeri 11 Bandar Lampung	Buruh	Daftar
P	15	AK	Islam	Bumi Waras	MTs Hasanuddin Telukbetung	Wirausaha	Tidak daftar
P	15	AK	Islam	Teluk Betung Selatan	SMP Tamansiswa Telukbetung	Karyawan Swaswta	Daftar

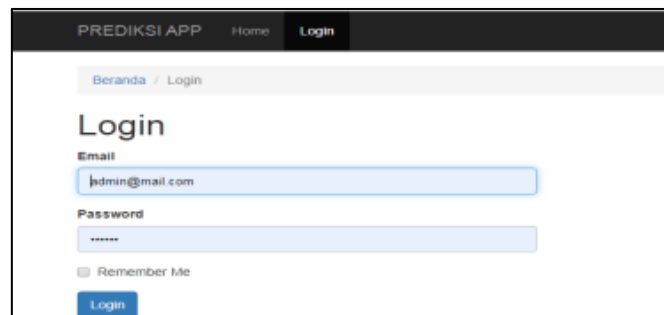
Gambar 4. Hasil data transformation

Implementasi Antar Muka Sistem

Berikut beberapa contoh tampilan antar muka sistem yang telah diimplementasikan.

1. Halaman Utama Sistem

Tampilan saat pertama kali aplikasi digunakan, pada form ini terdapat list fitur-fitur aplikasi yang dapat digunakan setelah user melakukan login pada area yang telah disediakan (gambar 5).



Gambar 5. Halaman Login

2. Halaman Master Data Siswa

Halaman ini menampilkan master data siswa yang didapatkan dari SMK Taman Siswa. Halaman master data menampilkan seluruh data siswa yang telah diinputkan melalui sistem seperti pada gambar 6 berikut.

Tahun	Nama Siswa	Nomor Induk	Jenis Kelamin	Tanggal Lahir	Agama	Alamat Siswa	Alamat Kecamatan	Nama Sekolah Asal	
2017	A. AL-FAN LUBIS	4783 / 0020170005	Laki-laki	Bandar Lampung	7 Jan 2002	Islam	J. Yos Sudarso Gg. Bekau LK II RT. 41	Kecamatan Bumi Waras Bandar Lampung	SMP Negeri 11 Bandar Lampung
2017	ALDILA SAHALA	4784 / 0011700004	Perempuan	Bandar Lampung	22 Des 2001	Islam	J. Sarnet Ploas LK II RT. 41	Kecamatan Bumi Waras	MTs Hasanuddin Tebukbetung

Gambar 6. Halaman Master Data Siswa

3. Halaman Data Siswa Proses

Halaman data siswa praproses berikut (gambar 7), menampilkan data yang telah diproses dari master data sebelumnya sesuai dengan 4 tahapan yaitu *data integration*, *data selection*, *data cleaning* dan *data transformation*.

#	Jenis Kelamin	Tempat Lahir	Agama	Alamat Siswa	Nama Sekolah Asal	Jurusan	Pekerjaan Orang Tua	Status Pendaftaran
1	Laki-laki	Bandar Lampung	Islam	Kecamatan Bumi Waras	SMP Negeri 11 Bandar Lampung	AK	Wirasaha	Daftar
2	Perempuan	Bandar Lampung	Islam	Kecamatan Bumi Waras	MtS Hasanudin Tebuketung	AK	Wirasaha	Tidak Daftar
3	Perempuan	Indarayu	Islam	Kecamatan Tebuketung Utara	SMP Negeri 27 Bandar Lampung	AK	Buruh	Daftar

Gambar 7. Halaman Data Siswa Praproses

4. Halaman Data Testing

Tampilan pada halaman data *testing* berikut (gambar 8) ditunjukkan untuk menampilkan data *testing* siswa yang akan dihitung prediksi potensi pendaftarannya menggunakan algoritma Naive Bayes.

#	Jenis Kelamin	Usia	Alamat Siswa	Nama Sekolah Asal	Jurusan	Agama	Pekerjaan Orang Tua	Status Pendaftaran
1	Laki-laki	17	Kecamatan Bumi Waras	SMP Negeri 11 Bandar Lampung	AK	Islam	Wirasaha	Daftar
2	Perempuan	17	Kecamatan Bumi Waras	MtS Hasanudin Tebuketung	AK	Islam	Wirasaha	Tidak Daftar
3	Perempuan	17	Kecamatan Tebuketung Utara	SMP Negeri 27 Bandar Lampung	AK	Islam	Buruh	Daftar
4	Perempuan	19	Kecamatan Bumi Waras	SMP Tamansiswa Tebuketung	AK	Islam	Buruh	Daftar
5	Perempuan	19	Kecamatan Tebuketung Selatan	SMP Tamansiswa Tebuketung	AK	Islam	Buruh	Daftar
6	Perempuan	19	Kecamatan Kintabon	SMP Tamansiswa Tebuketung	AK	Budha	Karyawan Swasta	Daftar

Gambar 8. Halaman Data Testing

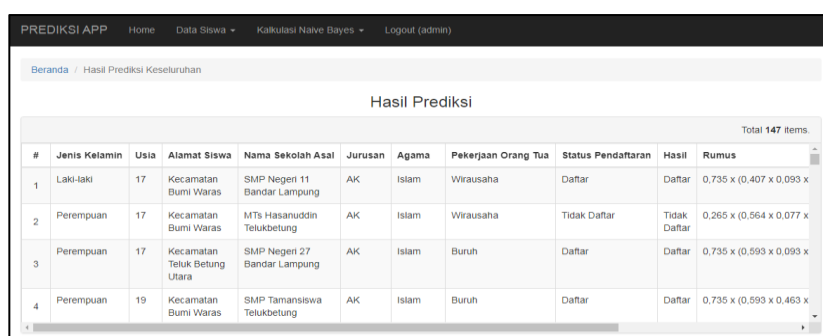
5. Halaman Hitung Prediksi

Sistem menyediakan halaman untuk hitung prediksi yang dibagi menjadi 2 yaitu prediksi persiswa dan prediksi keseluruhan. Untuk halaman prediksi persiswa, ditunjukkan jika ingin melakukan testing klasifikasi dengan dataset satu persatu. Pada halaman prediksi persiswa, sistem akan menunjukkan hasil perhitungan sesuai persamaan (gambar 1) dan hasilnya seperti pada gambar 9 berikut.

KEPUTUSAN: Tidak Daftar

Gambar 9. Tampilan Halaman Prediksi Persiswa

Tampilan pada hasil hitung prediksi secara keseluruhan secara umum sama dengan prediksi persiswa, hanya saja perhitungan dapat dilakukan untuk jumlah dataset yang banyak sekaligus. Berikut tampilan hasil prediksi keseluruhan pada gambar 10.



#	Jenis Kelamin	Usia	Alamat Siswa	Nama Sekolah Asal	Jurusan	Agama	Pekerjaan Orang Tua	Status Pendaftaran	Hasil	Rumus
1	Laki-laki	17	Kecamatan Bumi Waras	SMP Negeri 11 Bandar Lampung	AK	Islam	Wirausaha	Daftar	Daftar	$0,735 \times (0,407 \times 0,093 \times$
2	Perempuan	17	Kecamatan Bumi Waras	MTs Hasanuddin Telukbetung	AK	Islam	Wirausaha	Tidak Daftar	Tidak Daftar	$0,265 \times (0,564 \times 0,077 \times$
3	Perempuan	17	Kecamatan Teluk Betung Utara	SMP Negeri 27 Bandar Lampung	AK	Islam	Buruh	Daftar	Daftar	$0,735 \times (0,593 \times 0,093 \times$
4	Perempuan	19	Kecamatan Bumi Waras	SMP Tamansiswa Telukbetung	AK	Islam	Buruh	Daftar	Daftar	$0,735 \times (0,593 \times 0,463 \times$

Gambar 10. Tampilan Halaman Prediksi Keseluruhan

Pengujian

Untuk mengukur keberhasilan prediksi potensi pendaftaran siswa, maka dilakukan pengukuran akurasi dari 147 data yang diujikan. Hasil menunjukkan bahwa dari 147 dataset yang diuji, terdapat 128 dataset yang diprediksi dengan benar sehingga tingkat akurasi mencapai 87%.

SIMPULAN

Berdasarkan pembahasan dan evaluasi dari, maka dapat diperoleh kesimpulan bahwa prediksi potensi pendaftaran siswa di SMK Taman Siswa dapat dilakukan dengan tingkat akurasi 87%. Sedangkan saran bagi penelitian ini dan selanjutnya yaitu, jumlah dataset yang digunakan agar dapat ditingkatkan lagi agar akurasi dapat meningkat dan ragam pola dataset yang tersedia juga mampu mewakili banyak kondisi yang sesuai.

UCAPAN TERIMA KASIH

Pada bagian ini, penulis menyampaikan ucapan terima kasihnya kepada pihak-pihak yang membantu dan mendukung hingga terselesaikannya artikel ilmiah ini, semoga dapat menjadi referensi atau sumber yang mendukung penelitian berikutnya.

REFERENSI/DAFTAR PUSTAKA

- Andini T.I., Witanti W., dan Renaldi F. 2016. Prediksi Potensi Pemasaran Produk Baru dengan Metode Naive Bayes Classifier dan Regresi Linear, Seminar Nasional Aplikasi Teknologi Informasi (SNATI).
- Anhar. 2010. PHP dan MySql Secara Otodidak, Jurnal Ilmiah Ilmu Komputer. PT Trans Media Jakarta
- Dhanashree S.M., Mayur P.B., dan Shruti D.D. 2012. Prediction System For Heart Disease Using Naive Bayes, International journal of Advanced Computer and Mathematical Sciences
- Firman A., Wowor H., dan Najoan X. 2014. Sistem Informasi Perpustakaan Berbasis Web Application, Jurnal Sistem Informasi Bisnis
- Fithri D. L., dan Darmanto E. 2014. Sistem Pendukung Keputusan Untuk Memprediksi Kelulusan Mahasiswa Menggunakan Metode Naive Bayes, Vol.319-324.
- Gunawan I. 2018. Sistem Pakar Diagnosa Penyakit Kulit Pada Kucing Menggunakan Naive Bayes Berbasis Web, Lampung: Universitas Teknokrat
- Herkutanto C. 2013. Pengembangan Dan Analisis Kualitas Aplikasi Pencarian Gambar Berdasar Histogram Warna Berbasis Web, Yogyakarta: Fakultas Teknik Universitas Negeri Yogyakarta.
- Jogiyanto. 2009. Analisis dan Desain Sistem Informasi, Yogyakarta: Andi Offset
- Kusrini dan Luthfi, E. T. 2009. Algoritma Data Mining, Yogyakarta.