

Rancang Bangun Sistem Pendukung Keputusan Penerimaan Beasiswa Menggunakan Metode *Naive Bayes Classifier*

Riani Dewi H¹, Yunita², Novi Indrawati³

Mahasiswa Program Studi Teknik Informatika Fakultas Teknik Universitas Trunojoyo Madura
Kampus UTM, Jl. Raya Telang Po Box Kamal 69162

E-mail : rdewihariyati@gmail.com¹, neta.ayun@gmail.com², novindriawati28@gmail.com³

Abstrak

Beasiswa merupakan salah satu program pemerintah yang bertujuan meringankan beban biaya pendidikan bagi siswa maupun mahasiswa. Di perguruan tinggi program beasiswa ini diberikan kepada mahasiswa yang masih aktif kuliah dan tidak melebihi masa studi normal. Jadi mahasiswa yang mengajukan beasiswa minimal semester 2 karena sudah mempunyai nilai IPK dan maksimal semester 8. Mahasiswa yang mendaftar beasiswa harus memenuhi syarat yang sudah ditentukan oleh pihak universitas. Dalam penelitian ini ada beberapa kriteria-kriteria yang digunakan untuk penyeleksian beasiswa, yaitu Penghasilan Orangtua, IPK, PLN, dan Tanggungan Keluarga. Metode yang digunakan untuk seleksi beasiswa adalah metode *Naive Bayes Classifier*, metode ini akan mengklasifikasi pendaftar beasiswa menjadi dua kelas yaitu kelas layak dan tidak layak. Diharapkan sistem ini dapat berfungsi optimal dan baik dalam melakukan seleksi beasiswa.

Kata Kunci : kriteria, seleksi, klasifikasi

Abstract

Scholarship is one of the government's program that aims to ease the burden of education expenses for student or university student. In college the scholarship awarded to student who are still active and doesn't exceed the normal study period. So, students who apply for scholarship at minimum of 2 semester because they have a value of GPA and the maximum of 8 semester. The students who enrolled of scholarship must be eligible that has been specified by the university. In this research, there are several criteria used for the selection of scholarship, that is Income of Parents, GPA, PLN, and Dependant of Family. The method used for scholarship selection was method of *Naive Bayes Classifier*; this method would classify the registrants into two classes, that is worthy and unworthy class. Should us that this system could function optimally and better for selection of scholarship.

Keyword : criteria, selection, classification

1. Pendahuluan

1.1 Latar Belakang

Hak setiap warga Negara adalah mendapatkan pengajaran, dan hak tersebut dicantumkan dalam Pasal 31 (1) Undang-Undang Dasar 1945. Mengacu pada pasal tersebut, pemerintah pusat dan pemerintah daerah wajib memberikan layanan dan kemudahan juga wajib menjamin terlaksananya pendidikan yang bermutu untuk setiap warga negara Indonesia tanpa adanya diskriminasi. Hal tersebut diutamakan juga untuk warga negara yang kurang mampu tetapi berprestasi dan berpotensi akademik yang baik. Oleh karena itu setiap lapisan lembaga pendidikan tidak tekecuali perguruan tinggi negeri akan memberikan beasiswa setiap semesternya untuk meringankan beban pendidikan mahasiswa.

Dalam penyeleksian beasiswa ada beberapa kriteria yang telah ditentukan oleh masing-masing lembaga. Pada penelitian ini kriteria yang digunakan adalah Penghasilan Orangtua, IPK, PLN, dan Tanggungan Keluarga.

Bidang kemahasiswaan penyeleksi beasiswa memerlukan suatu sistem yang dapat membantu memberikan solusi dalam penyeleksian mahasiswa penerima beasiswa. Oleh karena itu dibangun suatu sistem pendukung keputusan penyeleksian beasiswa

yang dapat membantu menyelesaikan permasalahan tersebut.

Dalam penelitian ini menggunakan metode *Naive Bayes Classifier* yang dianggap mampu memberikan hasil yang akurat dalam pengklasifikasian pendaftar beasiswa ke dalam dua kelas yaitu kelas layak dan kelas tidak layak.

1.2 Tujuan

Adapun tujuan dari penelitian ini adalah :

1. Membangun sebuah sistem pendukung keputusan untuk penyeleksian penerimaan beasiswa dengan kriteria Penghasilan Orangtua, IPK, PLN dan Tanggungan.
2. Membuat suatu sistem pendukung keputusan menggunakan metode *Naive Bayes Classifier*.

1.3 Penelitian Terdahulu

Sholihah[1] meneliti tentang penentuan Uang Kuliah Tunggal (UKT) yang ada di Universitas Trunojoyo Madura. Penelitian tersebut dilandaskan oleh beberapa faktor diantaranya adanya pemberian Uang Kuliah Tunggal (UKT) pada mahasiswa Universitas Trunojoyo Madura yang berbeda-beda berdasarkan beberapa kriteria yang

telah ditentukan. Ada empat golongan besaran nilai UKT yang ada di Universitas Trunojoyo Madura, golongan 1 bernilai 1.000.000, golongan 2 bernilai 1.800.000, golongan 3 bernilai 2.250.000, dan golongan 4 bernilai 2.500.000. Untuk menentukan besaran pada masing-masing mahasiswa peneliti membuat suatu sistem pendukung keputusan menggunakan metode *Naive Bayes Classifier*. Pada penelitian ini mampu memberikan hasil dengan tingkat akurasi yang cukup tinggi. Dari beberapa skenario, skenario satu memberikan akurasi paling tinggi dengan hasil pada data kategorik 90 % sedangkan pada skenario hasil pada dua data kategorik 78.67%.

Kusumadewi[2] pada penelitiannya mengenai pengklasifikasian status gizi seseorang menggunakan alat ukur antropometri sebagai variabel input. Penelitian tersebut didasarkan adanya suatu perbedaan status gizi pada masing-masing individu. Alat ukur menggunakan antropometri digunakan untuk menentukan keadaan gizi seseorang, untuk hasil yang optimal ada beberapa pedoman antropometri diantaranya penilaian terhadap usia dan berat badan, panjang badan, tinggi badan dan lingkaran lengan atas. Pada penelitian ini terdapat 5 golongan status gizi yaitu berat kurang, berat normal, obesitas kurang, obesitas sedang, dan obesitas berat. Untuk menentukan status gizi tersebut menggunakan metode *Naive Bayes Classifier*, dengan hasil yang cukup berakurasi tinggi yaitu bernilai 93,2%.

Hasibuan dkk, [3] pada penelitiannya mengenai seleksi penerimaan beasiswa, penelitian dilakukan di SMAN 2 Metro. Penelitian ini berdasarkan pada adanya kendala dan kesulitan dalam proses penyeleksian beasiswa di SMAN 2 Metro. Pada penelitian ini menggunakan metode *Analytical Hierarchy Process* (AHP) sebagai perhitungan manualnya dan menggunakan *expert choice*, untuk mendapatkan hasil keputusan yang konsisten. Kesalahan yang sering terjadi dalam perhitungan metode ini adalah terjadinya kesalahan pada penentuan bobot dan proses perbandingan secara berpasangan.

Fredrik [4] meneliti tentang masa studi mahasiswa berdasarkan jalur penerimaan menggunakan klasifikasi data mining. Penelitian ini dilakukan di STMIK Palangkaraya terhadap mahasiswa jurusan D3 Manajemen Informatika pada tahun kelulusan 2006-2008. Penelitian ini dapat digunakan untuk membantu peningkatan perbaikan kualitas perguruan tinggi tersebut. Peneliti menggunakan metode *Naive Bayes Classifier* untuk penelitian ini. Dengan metode *Naive Bayes Classifier* untuk pengklasifikasian dan prediksi dapat memberikan suatu hasil berupa informasi atau pengetahuan yang dapat dijadikan dasar pengambilan keputusan atau strategi pihak institusi dalam rangka meningkatkan kualitas perguruan tinggi.

Jananto[5] meneliti mengenai pencarian perkiraan masa studi mahasiswa. Penelitian dilakukan di Universitas Stikubank (UNISBANK) Semarang. Penelitian dilakukan untuk bisa mengetahui masa studi mahasiswa di suatu universitas, karena masa studi mahasiswa sangatlah penting untuk suatu perguruan tinggi demi meningkatkan perbaikan kualitas. Pada penelitian ini peneliti menggunakan metode *Naive Bayes Classifier*. Algoritma *Naive Bayes Classifier* menghitung perbandingan peluang antara jumlah dari masing-masing kriteria. Dari hasil uji coba diperoleh tingkat kesalahan prediksi berkisar 20% sampai dengan 50% dengan data training dan testing yang diambil secara random. Namun rata-rata tingkat kesalahan berkisar 20 % hingga 34%. Tinggi rendahnya tingkat kesalahan dapat disebabkan oleh jumlah record data dan tingkat konsistensi dari data training yang digunakan. Sedangkan hasil prediksi dari ketepatan lama studi dari mahasiswa angkatan 2008 adalah sebesar 254 mahasiswa diprediksi "Tepat Waktu" dan sisanya yaitu 4 orang diprediksi "Tidak Tepat Waktu".

2. Metode

Metode yang digunakan dalam penelitian pengambilan keputusan ini adalah metode *Naive Bayes Classifier*. Metode ini menggunakan konsep probabilitas yang bertujuan untuk melakukan klasifikasi data pada *class* tertentu.[6]

Metode *Naive Bayes Classifier* merupakan penyederhanaan dari teorema *Bayes*, penemu metode ini adalah seorang ilmuwan Inggris yang bernama *Thomas Bayes*. Probabilitas bersyarat adalah dasar dari teorema *Bayes* yang dinyatakan dalam persamaan 1.

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

Probabilitas X di dalam Y adalah probabilitas interseksi X dan Y dari probabilitas Y atau dengan bahasa lain $P(X|Y)$ adalah prosentase banyaknya X didalam Y.

Algoritma dalam metode *Naive Bayes Classifier* adalah algoritma teknik klasifikasi. Pada pengklasifikasian *Naive Bayes Classifier* diasumsikan bahwa ciri pada kelas tertentu tidak ada hubungannya dengan ciri dari kelas lainnya. Penjelasan teorema *Naive Bayes Classifier* seperti pada persamaan 2.

$$P(C|F_1..F_n) = \frac{P(C)P(F_1|C)P(F_2|C) \dots P(F_n|C)}{P(F_1..F_n)} \dots\dots\dots(2)$$

dimana C merupakan kelas, sementara variabel $F_1..F_n$ karakteristik petunjuk yang dibutuhkan untuk melakukan klasifikasi.

Selanjutnya adalah mencari nilai *posterior* yang digunakan sebagai perbandingan dengan nilai *posterior* kelas-kelas lainnya untuk menentukan ke kelas apa suatu sampel akan diklasifikasikan. Hal tersebut ditunjukkan pada persamaan 3 dan 4.

$$P(Y1, Y2, Y3 \dots Yn | X) = P(Y1 | X) * P(Y2 | X) * \dots * P(Yn | X) \dots \dots \dots (3)$$

$$P(X | Y1, Y2, Y3 \dots Yn) = P(X) * P(Y1 | X) * P(Y2 | X) * \dots * P(Yn | X) \dots \dots \dots (4)$$

Dalam metode *Naive Bayes Classifier* terdapat dua fase yaitu fase training dan testing.

1. Data Pelatihan

Jika jenis data bersifat kategorial untuk sebagian data yang telah diketahui kelasnya diproses untuk membentuk model perkiraan. Namun jika jenis data bersifat *numeric*, maka proses training harus melewati beberapa tahapan yaitu perhitungan nilai mean, *variance*, dan deviasi standart pada tiap kriteria untuk masing-masing golongan.

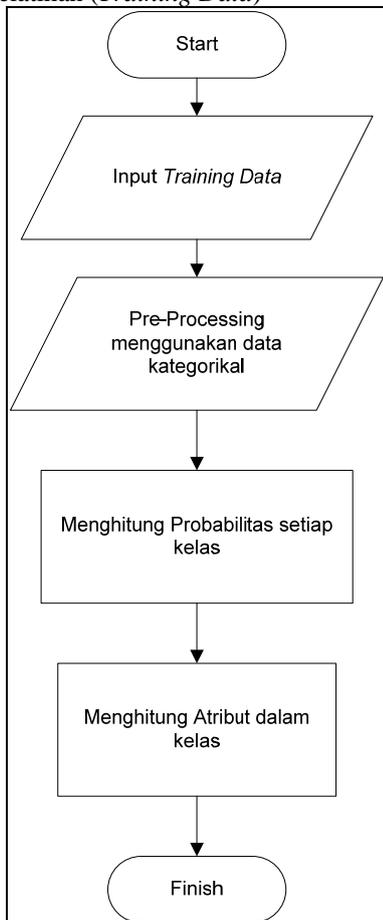
2. Data Percobaan

Proses ini merupakan proses perhitungan data yang mengacu pada data training. Pada proses ini ada beberapa tahapan, yaitu menghitung peluang kriteria terhadap golongan, menghitung peluang setiap golongan, dan menentukan nilai *maximal* pada masing-masing *posterior*.

A. Rancangan Sistem

a. Flowchart

Data Pelatihan (*Training Data*)

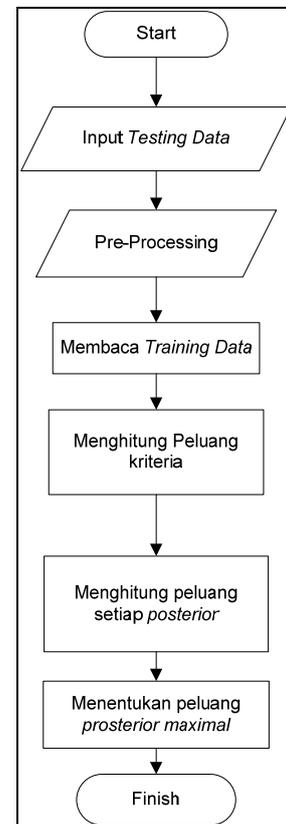


Gambar 1. Perancangan Training Data

Proses *Training Data* dalam metode *Naive Bayes Classifier* ditunjukkan pada Gambar 1 dengan penjelasan sebagai berikut :

1. Mulai.
2. Menginputkan data pelatihan atau *Training Data*.
3. *Pre-processing* data dengan parameter yang digunakan bersifat kategorikal.
4. Proses perhitungan dengan metode *Naive Bayes Classifier* untuk data kategorikal langkah pertama yang dihitung adalah probabilitas setiap kelasnya.
5. Selanjutnya menghitung probabilitas atribut dalam kelasnya.
6. Hasil probabilitas yang telah dihitung.
7. Selesai.

Data Percobaan (*Testing Data*)



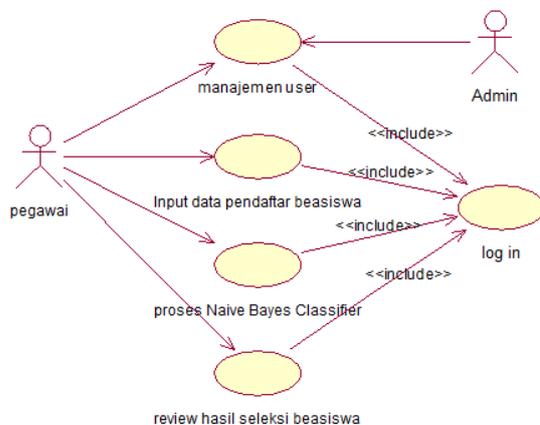
Gambar 2. Perancangan Testing Data

Proses *Testing Data* merupakan proses inti dari metode *Naive Bayes Classifier*. Pada proses ini akan dihasilkan keputusan akhir dari perhitungan metode *Naive Bayes Classifier*. Proses tersebut ditunjukkan pada Gambar 2 dengan penjelasan sebagai berikut :

1. Mulai.
2. Menginputkan *Testing Data* dimana data ini merupakan data pendaftar beasiswa.
3. *Pre-processing* data dengan parameter yang digunakan bersifat kategorikal.

4. Membaca *Training Data*, karena proses pada *Testing Data* mengacu pada *Training Data*.
5. Menghitung peluang kriteria terhadap masing-masing kelas.
6. Menghitung peluang pada setiap *posterior*.
7. Setelah menghitung peluang pada setiap *posterior*, dilanjutkan dengan menentukan peluang *maximal* dari *posterior*.
8. Hasil akhir berupa keputusan dari peluang *maximal* tersebut.
9. Selesai.

b. Use Case



Gambar 3. Use Case Sistem

Secara garis besar, use case rancangan sistem pendukung keputusan yang akan dibuat adalah seperti pada Gambar 3. Berikut ini penjelasan dari masing-masing tahapan pada alur sistem secara umum.:

1. *Actor* : admin dan pegawai
2. *Use Case*:
 - Manajemen user :
Merupakan *User Case* untuk olah data user sesuai dengan tingkatan level.
 - Menginputkan data pendaftar :
Merupakan *User Case* untuk menginputkan data pendaftar beasiswa yang dilakukan oleh pegawai.
 - Proses *Naive Bayes Classifier* :
Merupakan *User Case* untuk melakukan perhitungan penyeleksian terhadap data mentah pendaftar beasiswa menggunakan *Naive Bayes Classifier* untuk mendapatkan hasil yang layak diterima dan tidak layak.
 - Review hasil seleksi beasiswa :

Merupakan *Use Case* hasil seleksi dari perhitungan menggunakan *Naive Bayes Classifier*.

2.1 Metode Pengumpulan Data

Untuk menyelesaikan penelitian ini, berikut adalah metode pengumpulan data yang dilakukan oleh peneliti :

1. Penelitian Lapangan.

Penelitian lapangan dilakukan di Fakultas Teknik Universitas Trunojoyo Madura untuk memperoleh data dan informasi yang dibutuhkan untuk penelitian mengenai pendaftar beasiswa tahun 2014. Bentuk penelitian yang dilakukan antara lain:

a. Wawancara

Wawancara atau mengajukan tanya jawab dengan staf penyeleksi beasiswa Fakultas Teknik Universitas Trunojoyo Madura untuk menggali informasi lebih *detail* mengenai proses beasiswa.

2. Penelitian Kepustakaan

Pada penelitian kepustakaan ini dilakukan dengan studi literatur dengan pembelajaran melalui buku, *e-book*, jurnal, internet, dan sumber referensi lainnya.

2.2 Metode Analisa Data

Pada penelitian ini jenis data yang digunakan bersifat kategorik dengan beberapa kriteria yaitu Penghasilan Orangtua, PLN, PBB dan Tanggungan Keluarga. Setelah melakukan pengumpulan data dan menentukan jenis data kemudian data akan diolah menggunakan metode *Naive Bayes Classifier*.

Pada metode *Naive Bayes Classifier* akan diklasifikasi ke dalam dua *class* yaitu *class* layak dan *class* tidak layak untuk mendapatkan beasiswa.

3. Hasil dan Pembahasan

Perancangan sistem penentuan beasiswa di jelaskan bahwa proses yang di lakukan dalam metode *Naive Bayes Classifier* yaitu menentukan data training. Dalam metode ini terdapat data berbentuk numeric dan kategorikal, jika data berbentuk *numeric*, maka penghitungannya menggunakan nilai mean, deviasi standart dan *variance* nya. Jika data berbentuk kategorikal, maka akan dihitung probabilitas setiap kelas dan probabilitas setiap atribut dalam kelasnya. Dalam studi kasus ini menggunakan data kategorik. Sebagai contoh data training ditunjukkan pada Tabel 1:

Tabel 1. Data Training

No	Penghasilan Ortu	IPK	PLN	Jumlah Tanggungan	Kelayakan	Menu
1	mampu	tinggi	450	banyak	layak	Edit Hapus
2	mampu	tinggi	450	banyak	layak	Edit Hapus
3	tidak mampu	rendah	900	sedikit	tidak layak	Edit Hapus
4	menengah	sedang	900	banyak	layak	Edit Hapus
5	tidak mampu	sedang	1300	sedikit	layak	Edit Hapus
6	tidak mampu	tinggi	900	sedikit	layak	Edit Hapus
7	tidak mampu	tinggi	900	banyak	tidak layak	Edit Hapus
8	menengah	sedang	450	sedang	layak	Edit Hapus
9	mampu	tinggi	1300	sedang	layak	Edit Hapus
10	tidak mampu	rendah	900	banyak	layak	Edit Hapus
11	mampu	sedang	900	sedikit	tidak layak	Edit Hapus
12	tidak mampu	sedang	450	banyak	layak	Edit Hapus
13	tidak mampu	tinggi	450	sedang	layak	Edit Hapus
14	mampu	rendah	1300	banyak	tidak layak	Edit Hapus

Data training atau data pelatihan yang digunakan dalam penelitian ini sebanyak 25 data training. Pada penggunaan data training ini bisa menggunakan beberapa skenario untuk uji coba sistem.

```

::: PROBABILITAS :::
Kelayakan = Layak : 16/25
Kelayakan = Tidak Layak : 9/25

::: PENGHASILAN ORTU
Tidak Mampu = Layak : 9/16
Tidak Mampu = Tidak Layak : 2/9

Menengah = Layak : 4/16
Menengah = Tidak Layak : 2/9

Mampu = Layak : 3/16
Mampu = Tidak Layak : 5/9

::: IPK
Rendah = Layak : 2/16
Rendah = Tidak Layak : 4/9

Sedang = Layak : 6/16
Sedang = Tidak Layak : 3/9

Tinggi = Layak : 8/16
Tinggi = Tidak Layak : 2/9

```

Gambar 4. Pehitungan Probabilitas

Gambar 4 diatas merupakan perhitungan probabilitas yang dilakukan untuk menentukan probabilitas dari masing-masing kriteria pada setiap class. Perhitungan probabilitas ini mengacu pada jumlah data taraining atau data pelatihan yang telah ditetapkan sebelumnya.

DATA TESTING

Penghasilan Ortu :

IPK :

PLN :

Jumlah Tanggungan :

Gambar 5. Data Testing

Gambar 5 merupakan data testing atau data pengujian, data testing yang *diinputkan* ini nantinya akan ditentukan oleh sistem menggunakan metode *Naïve Bayes Classifier*.

```

::: SKORING
Layak = (0.5625 * 0.125 * 0.625 * 0.25)*(16/25)
Tidak Layak = (0.222222222222 * 0.444444444444 * 0.111111111111 *
0.333333333333)*(9/25)

Layak : 0.00703125
Tidak Layak : 0.0013168724279835

::: REKOMENDASI
Jika Penghasilan Ortu = tidak mampu | IPK = rendah | PLN = 450 | Tanggungan =
sedikit, maka Mendapat Beasiswa = Layak :)

```

Gambar 6. Hasil Data Testing

Gambar 6 menunjukkan perhitungan data testing, hasil perhitungan diatas merupakan perhitungan *posterior* pada masing-masing class yaitu class layak dan class tidak layak. Setelah menghitung masing-masing nilai *posterior* lalu menentukan nilai *posterior maximum* dari kedua class tersebut untuk menentukan keputusan apakah mahasiswa tersebut layak atau tidak untuk menerima beasiswa. Hal tersebut merupakan proses akhir dari metode *Naïve Bayes Classifier*.

4. Kesimpulan

Kesimpulan yang dapat diambil dari penelitian tersebut adalah :

1. Metode *Naïve Bayes Classifier* memberikan proses penyeleksian yang cepat dan algoritmanya mudah dimengerti.
2. Bahwa jumlah data training atau data pelatihan dapat mempengaruhi tingkat akurasi.
3. Dibutuhkan beberapa uji coba (skenario) data training atau data pelatihan unuk mendapatkan hasil dengan tingkat akurasi yang tinggi.

Saran yang dapat peneliti sampaikan antara lain :

1. Dalam penelitian ini hanya menggunakan 5 kriteria untuk penyeleksian mahasiswa penerima beasiswa. Untuk penelitian selanjutnya bisa ditambah beberapa kriteria lagi.
2. Jenis pengelompokan data dalam penelitian ini bersifat *kategorik*, untuk hasil yang lebih optimal mungkin bisa ditambah dengan perhitungan data *numeric*.
3. Untuk mengetahui hasil perbandingan perhitungan seleksi beasiswa penelitian selanjutnya bisa menggunakan metode yang lain.

Teknologi Informasi Sekolah Teknik Elektro dan Informatika.

Ucapan Terima Kasih

Ucapan terima kasih kami sampaikan kepada kedua orang tua kami, yang tak hentinya memberikan do'a serta semangat untuk kami dan kepada dosen pembimbing kami tercinta ibu Andharini Dwi Cahyani, S.Kom, M.Kom yang senantiasa membimbing dan memberikan semangat untuk menyelesaikan penelitian ini.

Daftar Pustaka

- [1] Sholihah, Ani.2013. *Sistem Penentuan Uang Kuliah Tunggal (UKT) menggunakan metode Naive Bayes Classifier Classifier*. Program Studi Teknik Informatika Universitas Trunojoyo.
- [2] Kusumadewi, Sri.2009.*Klasifikasi Status Gizi Menggunakan Naive Bayesian Classification*.Jurusan Teknik Informatika Universitas Islam Indonesia.
- [3] Said Hasibuan, Aulia Vitari dan Muhammad.2010.*Sistem Penunjang Keputusan Penerimaan Beasiswa Menggunakan Metode Analytical Hierarchy Process*. Studi Kasus Penerimaan Beasiswa Di Sman2 Metro.Magister Teknologi Informasi IBI Darmajaya
- [4] Ulysses, John Fredrik .125301917.*Data Mining Classification Untuk prediksi Lama Masa Studi Mahasiswa Berdasarkan Jalur Penerimaan Dengan Metode Naive Bayes*. Magister Teknik Informatika Universitas Atma Jaya Yogyakarta.
- [5] Jananto, Arief.2013. *Algoritma Naive Bayes untuk Mencari Perkiraan Waktu Studi Mahasiswa*. Program studi sistem informasi universitas Stikibank.
- [6] Natalius, Samuel.2010/2011. *Metoda Naive Bayes Classifier dan Penggunaannya pada Klasifikasi Dokumen*. Program Studi Sistem dan