



# Sistem Pendukung Keputusan Untuk Mengevaluasi Internal Program Studi

Indhitya R. Padiku<sup>a\*</sup>, Achmad Widodo<sup>c</sup>

<sup>a</sup>Jurusan Sistem Informasi, Universitas Negeri Gorontalo

<sup>b</sup>Jurusan Teknik Mesin, Fakultas Teknik, Universitas Diponegoro

Naskah Diterima : 21 Februari 2015; Diterima Publikasi : 23 Maret 2015

---

## Abstract

The development of major and study program cannot be separated by some internal factors weather directly influenced number of new registration students or indirectly. It needs a method to both know and to analyze internal evaluation variables in major or study program. *Naive Bayes Classifier* (NBC) method is the simple form of *Bayesian network* that assume all features are independent each other. NBC shows us a great performance entirely in accuracy and error level classification. NBC is able to differentiate irrelevance attribute and also classified some attributes in prediction needs. This research hopefully can be useful for major internal evaluating and study program in order to increase the number of new registration students. The classification by influenced of variables to evaluate the condition of both major and study program for the new registration students.

**Keywords:** Decision Support System; NBC; Study Program Evaluation

## Abstrak

Perkembangan jurusan dan program studi tidak lepas dari faktor-faktor internal yang mempengaruhi jumlah calon mahasiswa baru baik langsung maupun tak langsung. Sehingga diperlukan metode untuk mengetahui kondisi dari setiap variabel-variabel yang menjadi bahan evaluasi internal di jurusan program studi.

Metode *Naive Bayes Classifier* (NBC) merupakan bentuk sederhana dari *Bayesian network* dengan asumsi bahwa semua fitur adalah independen satu sama lain. NBC menunjukkan kinerja yang unggul secara keseluruhan dalam hal akurasi dan tingkat kesalahan klasifikasi. NBC dapat mengklasifikasikan banyak atribut untuk keperluan prediksi.

Penelitian ini diharapkan dapat digunakan untuk evaluasi internal jurusan dan program studi dengan mengklasifikasikan variabel-variabel yang berpengaruh untuk mengetahui kondisi jurusan dan program studi sebagai salah satu upaya untuk meningkatkan jumlah calon mahasiswa baru dan juga dapat menjadi salah satu media untuk memperkenalkan program studi pada calon mahasiswa baru.

**Kata Kunci :** Sistem Pendukung Keputusan; NBC; Evaluasi Program Studi

---

## 1. Pendahuluan

Jurusan terdiri dari beberapa program studi (Prodi) yang melaksanakan fungsi pendidikan, penelitian serta pengabdian kepada masyarakat dan mengelola ilmu pengetahuan yang sesuai dengan bidang studi. Program studi harus mampu meningkatkan kualitas secara berkelanjutan yang nantinya akan diimplementasikan kepada mahasiswa. Beberapa faktor yang mempengaruhi perkembangan tiap program studi baik secara langsung maupun tak langsung dapat mempengaruhi jumlah mahasiswa baru, faktor-faktor utama tersebut diantaranya jumlah dosen, jumlah mahasiswa serta ketersediaan sarana prasarana (Vallejos, 2012).

Pada program studi mempunyai permasalahan yang dapat dianalisis berdasarkan data yang

dikumpulkan, teknik pengumpulan data memerlukan metode tertentu, penggunaan metode yang tepat dapat menjadi informasi dan pengetahuan sebagai bahan pengamatan dan pertimbangan dalam memprediksi perkembangan program studi. Pimpinan fakultas dapat melakukan evaluasi awal pada program studi, sehingga pimpinan dapat secara proporsional menentukan kuota jumlah mahasiswa pada penerimaan mahasiswa baru. Metode *naive bayes classifier* (NBC) merupakan bentuk sederhana dari sebuah *Bayesian network* yang merupakan salah satu metode data mining, kesederhanaan metode ini bergantung pada asumsi bahwa semua fitur adalah independen satu sama lain. NBC mempunyai keunggulan akurasi dan minimnya tingkat kesalahan klasifikasi (Koc, 2012).

Metode NBC adalah memanipulasi atribut saling bergantung, dan mewakili distribusi probabilitas gabungan lebih dari satu set diskrit, yaitu variabel stokastik (Ouali *et al.*, 2006). NBC dapat digunakan untuk memprediksikan objek kelas yang labelnya tidak diketahui atau dapat memprediksikan data yang akan muncul di masa depan (Chu dan Zao, 2011)

Keputusan perlu dilakukan oleh pengambil keputusan dengan preferensi tertentu berdasarkan informasi dan pengetahuan. Metode NBC dikembangkan dengan menggunakan rasio kemungkinan sebagai validasi metric model penilaian, dengan memprediksikan segala kemungkinan maka resiko dapat diminimalisir (Jiang *et al.*, 2007). Kinerja NBC kompetitif dalam proses klasifikasi walaupun menggunakan asumsi atribut yang independen (tidak ada kaitan antar atribut).

Pada penelitian ini digunakan metode NBC untuk mengevaluasi program studi sehingga dapat membantu untuk pengambilan keputusan sehingga dapat meningkatkan kinerja fakultas pada umumnya dan program studi. Hasil evaluasi tersebut dapat menjadi media untuk mempromosikan program studi, sehingga keminatan setiap calon mahasiswa baru meningkat, dan menjadi dasar pimpinan dalam penentuan secara proporsional jumlah mahasiswa yang akan di terima tiap tahunnya.

## 2. Kerangka Teori

### 2.1. Naïve Bayes Classifier (NBC)

NBC adalah metode classifier yang berdasarkan probabilitas dari Teorema Bayesian dengan asumsi bahwa setiap variabel  $X$  bersifat bebas (*independent*). Dengan kata lain, *Naïve Bayesian Classifier* mengansumsikan bahwa keberadaan sebuah atribut (*variabel*) tidak ada kaitannya dengan beradaan atribut yang lain (Abidin., 2012).

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

Keterangan :

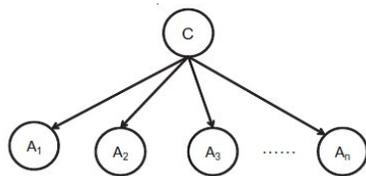
$X$  : data sampel dengan klas (label) yang tidak diketahui

$H$  : hipotesa bahwa  $X$  adalah data dengan klas (label)

$P(H)$  : peluang dari hipotesa  $H$  (*likelihood*)

$P(X)$  : peluang data sampel yang diamati (*evidence*)

$P(X|H)$  : peluang data sampel  $X$ , bila diasumsikan bahwa hipotesa  $H$  benar (*prior*)



Gambar 1. Struktur Naïve Bayes Classifier

Formulasi NBC untuk klasifikasi adalah :

$$P(Y|X) = \frac{P(Y) \prod_{i=1}^n P(X_i | Y)}{P(X)} \tag{2}$$

Keterangan :

$P(Y|X)$  : Peluang data dengan vector  $X$  pada kelas  $Y$

$P(Y)$  : Peluang awal kelas  $Y$

$\prod_{i=1}^n P(X_i | Y)$  : Peluang independen kelas  $Y$  dari semua fitur dalam vector  $X$

$P(X)$  : Nilai konstan atau nilai tetap

Berdasarkan nilai  $P(X)$  yang memang telah bernilai tetap maka nantinya hanya menghitung pada bagian  $P(Y) \prod_{i=1}^n P(X_i | Y)$ . Untuk peluang independen  $\prod_{i=1}^n P(X_i | Y)$  adalah pengaruh semua atribut dari data terhadap kelas  $Y$  yang dinitasikan seperti dibawah ini :

$$P(X|Y = y) = \prod_{i=1}^n P(X_i | Y = y) \tag{3}$$

Atribut yang masuk dalam atribut kategorikal dapat langsung di hitung, tetapi untuk atribut yang kontinyu sebelum di implementasikan dengan formula *naïve bayes classifier* harus melalui proses seperti berikut :

1. Untuk setiap atribut kontinyu diganti ke nilai interval diskret yaitu dengan mentransformasi atribut kontinyu ke dalam ordinal.
2. Mengasumsikan bentuk tertentu dari distribusi peluang untuk atribut kontinyu dan memperkirakan parameter distribusi dengan data pelatihan. Distribusi Gaussian biasanya dipilih untuk merepresentasikan peluang bersyarat dari atribut kontinyu pada sebuah kelas  $P(X_i|Y)$ , sedangkan distribusi Gaussian dikarakteristikan dengan dua parameter yaitu : mean ( $\mu$ ) dan varian ( $\sigma^2$ ).

(1) Untuk setiap kelas  $y_j$ , peluang bersyarat kelas  $y_j$ , untuk atribut  $X_i$  adalah :

$$P(X_i = x | Y = y_j) = \frac{1}{\sqrt{2\pi\sigma_{ij}^2}} \exp \left\{ -\frac{(x_i - \mu_{ij})^2}{2\sigma_{ij}^2} \right\} \tag{4}$$

### 2.2. Data Mining

Data mining atau *Knowledge Discovery in Database* (KDD) adalah proses menentukan informasi yang berguna serta pola-pola yang ada dalam data. Informasi ini terkandung dalam basis data yang berukuran besar yang sebelumnya tidak diketahui dan potensial bermanfaat (Han & Kamber, 2006). Data Mining merupakan salah satu langkah dari serangkaian proses *iterative* KDD.

Pengetahuan penemuan sebagai suatu proses terdiri dari urutan berulang dari langkah-langkah berikut:

1. Data pembersihan
2. Integrasi data
3. Pemilihan data

4. Transformasi data
5. Data mining
6. Evaluasi Pola
7. Presentasi Pengetahuan

Fungsi-fungsi dalam data mining membaginya ke dalam enam fungsi yaitu (Susanto & Suryadi, 2010) :

1. Fungsi Deskripsi (*description*)
2. Fungsi Estimasi (*estimation*)
3. Fungsi Prediksi (*prediction*)
4. Fungsi Klasifikasi (*classification*)
5. Fungsi Klasterisasi (*clustering*)
6. Fungsi Asosiasi (*association*)

Algoritma yang dapat digunakan diantaranya :

- a) *Generalizes Association Rules*
- b) *Quantitative Association Rules*
- c) *Assynchronous Parallel Mining (APM)*

Kumpulan objek data terdapat pada sebuah data set atau sering disebut sebagai *record*, *point*, *case*, data dan lain-lain, kumpulan objek tersebut mempunyai karakter dasar objek yang digambarkan dengan sejumlah atribut. Atribut sering juga disebut sebagai karakteristik, fitur atau variabel. Atribut mempunyai jenis yang beragam untuk setiap elemen data. Secara umum tipe atribut ada dua yaitu kategoris (kualitatif) dan numerik (kuantitatif), berikut ini Tabel 2.2. menjelaskan tipe-tipe atribut (Prasetyo, 2012) :

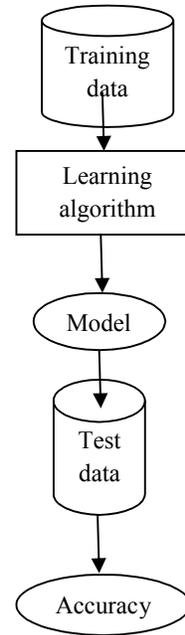
Tabel 1. Tipe atribut

Tipe Atribut	Keterangan
Kategoris (Kualitatif)	Nominal Nominal adalah nilai atribut yang memberikan nilai berupa nama, dengan nama tersebut sebuah atribut membedakan dirinya pada data yang satu dan data lainnya
	Ordinal Ordinal adalah atribut yang memberikan nilai berupa nama yang mempunyai arti informasi terurut.
Numerik (Kuantitatif)	Interval Atribut yang mempunyai perbedaan diantara dua nilai yang mempunyai makna yang berarti.
	Rasio Atribut yang mempunyai perbedaan diantara dua nilai dan rasio dua nilai mempunyai makna yang berarti.

### 2.3. Klasifikasi

Ada dua proses penting yang dilakukan saat melakukan klasifikasi. Proses yang pertama adalah *learning* yaitu proses pembelajaran menggunakan *training set*. Untuk kasus dengan menggunakan NBC, perhitungan probabilitas dari data berdasarkan data pembelajaran. Proses yang kedua adalah proses *testing* yaitu menguji model menggunakan data testing. Untuk menguji ketepatan klasifikasi dari

model yang telah terbentuk perlu menggunakan *data testing*. Kedua model untuk *data training* maupun *data testing* ketepatan klasifikasinya akan dibandingkan. *Data testing* adalah data yang digunakan dalam melakukan tes terhadap data. Beberapa data digunakan untuk memverifikasi bahwa data inputan yang diberikan dapat menghasilkan output yang diharapkan, data lain dapat digunakan untuk menguji dengan secara terpusat dan sistematis. *Data testing* dapat disimpan dan digunakan kembali atau hanya digunakan sekali.



Gambar 1. Tahapan proses klasifikasi (Abidin, 2012)

*Keterangan :*

Learning (training): Pembelajaran menggunakan data training

Testing : Menguji metode atau model menggunakan data testing

### 2.4. Akurasi klasifikasi

Akurasi menunjukkan kedekatan nilai hasil pengukuran dengan nilai sebenarnya. Untuk menentukan tingkat akurasi perlu diketahui nilai sebenarnya dari parameter yang diukur dan kemudian dapat diketahui seberapa besar tingkat akurasinya. Akurasi pada sebuah klasifikasi berpengaruh pada performa dari suatu klasifikasi. Untuk menganalisa performa tersebut dapat menggunakan sebuah matrik dengan membandingkan data kelas yang asli dengan prediksi dari data yang di input atau disebut dengan *confusion matrix*. Untuk setiap kolom dari matriks berhubungan dengan *output* klasifikasi dan yang berhubungan dengan input adalah pada setiap baris (Neila, 2012).

*Confusion matrix* terdapat pada weka *classifier* yang memberikan informasi dalam bentuk angka sehingga mempermudah untuk menghitung rasio

keberhasilan dari suatu klasifikasi dalam menganalisis performa algoritma. *Confusion matrix* adalah matrik yang berbentuk 2x2 untuk mendapatkan hasil klasifikasi dengan tepat, tiap kelas yang diprediksi mempunyai kemungkinan keluaran yang berbeda-beda seperti diantaranya *true positives* (TP) yaitu menunjukkan ketepatan klasifikasi dan *false positive* (FP) adalah jika nilai aslinya dan nilai yang diprediksi berbeda. Tabel *confusion matrix* seperti Tabel 2 dibawah ini.

Tabel 2. *confusion matrix*

		Predicted Class	
		C1	C2
Actual Class	C1	True Positive	False negative
	C2	False positive	True negatives

Dalam *confusion matrix* beberapa hal yang dilakukan dengan menggunakan data hasil klasifikasi diantaranya adalah :

- menghitung persentase kelas positif (*true positive*) yang diperoleh dalam klasifikasi.
- Precision* berfungsi menghitung persentase *false positive*
- Menghitung nilai rata-rata keberhasilan klasifikasi dengan cara membagi jumlah data yang terklasifikasi dengan benar dengan seluruh data yang diklasifikasi

Dengan algoritma klasifikasi dapat dilakukan dengan menghitung jumlah dari *record* yang di prediksi secara benar (akurasi) untuk mengevaluasi performa sebuah model yang dibangun. Akurasi dapat didefinisikan sebagai berikut :

$$\text{Akurasi} = \frac{\text{jumlah prediksi benar}}{\text{jumlah total prediksi}} \quad (5)$$

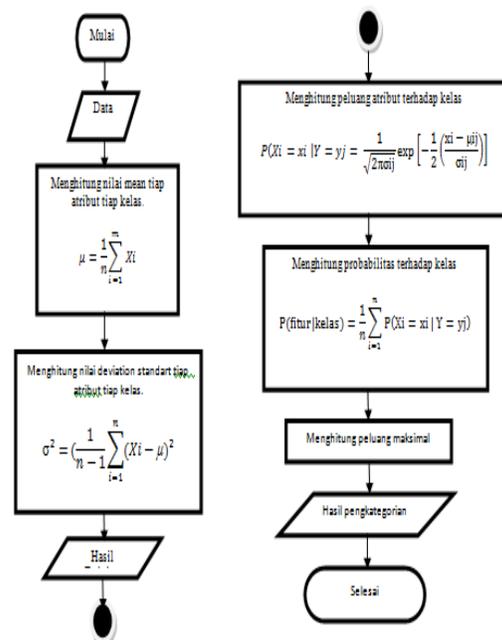
Keterangan :

Jumlah prediksi benar = hanya data yang diklasifikasikan benar atau seluruh jumlah data *true* (*true positive* dan *true negative*)

Jumlah total prediksi = seluruh jumlah data, baik yang diklasifikasikan *true* ataupun *false* (*true positive*, *true negative*, *false positive*, dan *false negative*)

### 3. Metodologi

Pada penelitian ini di implementasikan dengan menggunakan perhitungan *Naïve Bayes Classifier*. Pada *naïve bayes classifier* terdapat dua proses penting yaitu *training* dan *testing*. *Training* digunakan untuk melakukan proses pembelajaran berdasarkan data yang telah memiliki keputusan atau class label. Sedangkan tahap atau fase *testing* adalah proses penentuan keputusan itu sendiri berdasarkan hasil training yang telah dilakukan. Berikut ini alur proses NBC yang tertera pada Gambar 1.



Gambar 2. Alur proses NBC

Penjelasan Alur Proses NBC :

#### 1. Data

Data berasal dari hasil mining data seperti yang didasarkan dari langkah-langkahnya berikut ini:

- Data pembersihan untuk menghilangkan noise dan data yang tidak konsisten yaitu data yang tidak dibutuhkan dalam proses analisis seperti atau menyaring hanya fitur yang berpengaruh.
- Integrasi data yaitu data yang telah melalui pembersihan data dengan menghilangkan noise dan data yang tidak konsisten kemudian di kombinasikan dalam satu tabel utama
- Pemilihan data adalah data yang relevan dipilih yang sesuai dengan proses analisis yang diambil dari basis data
- Transformasi data yaitu data diubah ke dalam bentuk yang sesuai untuk proses analisis dengan melakukan ringkasan atau di agregasi
- Data mining yaitu proses menggunakan metode dalam data mining dalam penelitian ini menggunakan algoritma NBC. Pada penelitian ini menghasilkan klasifikasi seperti berikut :
- Evaluasi Pola untuk mengidentifikasi pola yang benar-benar menarik yang mewakili pengetahuan didasarkan pada langkah-langkah sebelumnya
- Presentasi Pengetahuan adalah visualisasi dan teknik representasi pengetahuan digunakan untuk menyajikan pengetahuan kepada pengguna

2. Menghitung nilai mean tiap atribut tiap kelas  
Untuk data bersifat numerik harus diketahui terlebih dahulu nilai mean untuk tiap atribut dari tiap kelasnya.
3. Menghitung nilai standar deviasi tiap atribut tiap kelas  
Setelah di dapatkan nilai mean kemudian dihitung nilai standar deviasinya untuk tiap atribut dari tiap kelasnya.
4. Hasil *training* adalah data yang telah di dapatkan dengan memperhatikan pola sebelumnya digunakan untuk di implemetasikan untuk formulasi NBC berikutnya.
5. Menghitung peluang atribut tiap kelas adalah setelah didapatkan nilai mean dan nilai standart deviasi kemudian di hitung peluang tiap kelas dengan dimasukkan data *test*.
6. Menghitung probabilitas terhadap kelas atau mencari masing-masing nilai likelihoodnya terhadap masing-masing kelas yang muncul.
7. Menghitung peluang maksimal adalah kelas yang nilai probabilitasnya paling maksimal, karena nilai probabilitas atau peluang hanya berkisar 0-1 maka dilihat nilai diantara keduanya dengan paling mendekati angka 1 yang menjadi hasil peluang maksimalnya.
8. Hasil kategori atau disebut hasil klasifikasi adalah kesimpulan berupa klasifikasi terakhir dari *record-record* yang dimasukkan.

#### 4. Hasil dan Pembahasan

##### 4.1. Hasil Penelitian

Dalam penelitian ini dihasilkan sebuah aplikasi yang didalamnya terdiri dari *front end system* dan *back end system*, isi dari *front end* adalah info yang berupa grafik perkembangan program studi dari tahun 2008-2013, serta login dan isi dari *back end system* adalah input data, data *training*, data *testing*, hasil pengolahan data serta menampilkan laporan pada program studi yang hanya bisa di akses oleh admin program studi dan dekan program studi. Dengan aplikasi tersebut dekan selaku pimpinan fakultas bisa mengevaluasi kekurangan dan kebutuhan program studi.

Sistem evaluasi ini menghasilkan *output* yaitu grafik hasil klasifikasi status program studi baik secara keseluruhan dilihat berdasarkan atribut yang berpengaruh dan juga grafik dilihat berdasarkan program studi. Dengan adanya grafik tersebut maka pimpinan dapat mengevaluasi kebutuhan dan dapat melihat kekurangan untuk program studi, sehingga dapat menjadi bahan pengambilan keputusan.

Dengan menggunakan formula NBC maka setiap nilai dari data yang di input diproses sehingga mendapatkan probabilitas kelas yang paling tinggi nilainya atau hasil klasifikasinya mendekati angka satu (1), yang seperti diketahui bahwa rentang dari probabilitas adalah hanya nol sampai satu (0-1).

Penggunaan metode NBC ini untuk menghitung hingga mendapatkan probabilitas atau peluang yang paling besar nilainya sehingga bisa di dapatkan *class* dari data yang menjadi data *test* tersebut. Untuk nilai *class* yang paling mendekati angka 1 merupakan hasil dari status prodi pada tahun 2014, yang tahun 2014 tersebut digunakan sebagai data *testing*. Untuk hasil inputan data *test* dapat di lihat pada Gambar 3 dibawah ini.

The screenshot shows a web interface for a Naive Bayes Classifier. It displays input data for the year 2014: Publikasi=11, Dosen=S3, Pegawai=PNS, Mahasiswa=55, Alumni=45, Kelas=2, Laboratorium=Terawat, and Seminar=8. The resulting classification is 'Baik'.

Tahun	publikasi	Dosen	Pegawai	Mahasiswa	Alumni	Kelas	Laboratorium	seminar	Status Prodi
2014	11	S3	PNS	55	45	2	Terawat	8	Baik

Gambar 3. Hasil *input* data *testing*

Dari Gambar 4 di atas terlihat nilai untuk *class* baik adalah 1 dan nilai cukup adalah 0 maka hasil perhitungan menggunakan NBC untuk data pada masing-masing atribut yaitu publikasi=11, dosen=S3, pegawai=PNS, mahasiswa=55, alumni=45, kelas=2, laboratorium=terawat dan seminar=8 maka status prodi yang di dihasilkan adalah termasuk pada kelas **Baik**.

Hasil dari keseluruhan data yang telah di analisis di program studi beserta hasil status prodi yang didapatkan di tampilkan dalam bentuk tabel dan grafik. Gambar 4 berikut menampilkan dalam bentuk tabel :

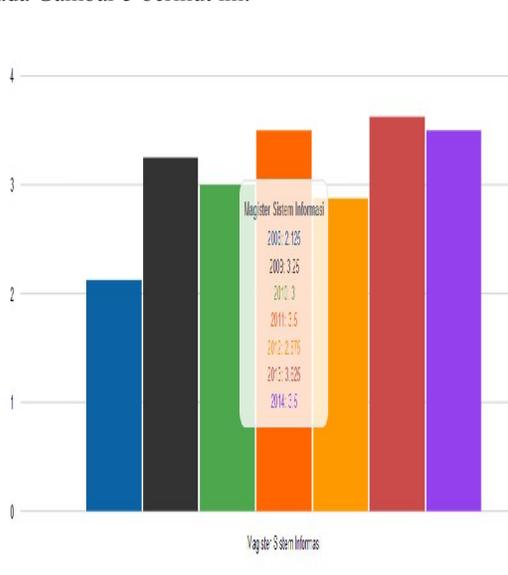
### Hasil Analisis

Menampilkan tabel hasil analisis Program Studi.

Prodi	Tahun	Publikasi	Dosen	Pegawai	Mahasiswa	Alumni	Kelas	Laboratorium	Seminar	Status Prodi
Magister Sistem Informasi	2008	0	S3	PNS	12	8	1	Terawat	2	Cukup
Magister Sistem Informasi	2009	8	S3	PNS	35	21	1	Terawat	18	Baik
Magister Sistem Informasi	2010	7	S3	PNS	45	16	2	Terawat	6	Baik
Magister Sistem Informasi	2011	16	S3	PNS	61	24	2	Terawat	15	Baik
Magister Sistem Informasi	2012	3	S3	PNS	65	70	2	Terawat	0	Cukup
Magister Sistem Informasi	2013	18	S3	PNS	45	45	2	Terawat	11	Baik
Magister Sistem Informasi	2014	11	S3	PNS	55	45	2	Terawat	8	Baik

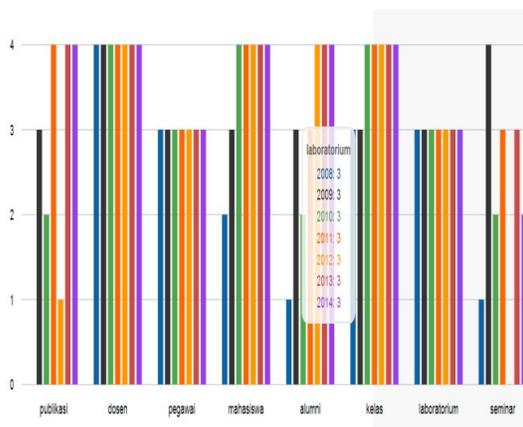
Gambar 4. Hasil *input* data *testing*

Untuk hasil analisis dalam bentuk grafik terlihat pada Gambar 5 berikut ini.



Gambar 5. Tampilan grafik hasil analisis

Grafik pada gambar diatas adalah grafik yang menggambarkan perkembangan pada program studi Magister Sistem Informasi Universitas Diponegoro dari tahun 2008 sampai 2014. Sehingga dapat menjadi bahan pertimbangan pimpinan untuk pengambilan keputusan. Untuk melihat lebih detail lagi perkembangan untuk masing-masing atribut pada program studi.



Gambar 6. Tampilan Laporan dalam bentuk grafik

Pada Gambar 6 berupa tampilan laporan dalam bentuk grafik dapat dilihat perkembangan untuk masing-masing atribut yang berpengaruh pada program studi, sehingga pimpinan dapat mengevaluasi dan mengambil keputusan untuk atribut-atribut yang berpengaruh yang harus ditingkatkan untuk membangun program studi yang lebih baik kedepannya.

#### 4.2. Pembahasan

Seperti yang telah di gambarkan pada bab sebelumnya alur untuk penggunaan formula *naïve bayes classifier* adalah pertama menentukan mean dan varian untuk setiap atribut yang bersifat kontinyu, kemudian menghitung peluang tiap atribut, hingga mendapatkan hasil likelihood masing-masing *class* untuk mendapatkan status prodi.

Dari penggunaan formula NBC pada sistem evaluasi internal program studi menghasilkan nilai-nilai yang di analisis sehingga mengasilkan klasifikasi untuk status prodi. Berikut ini hasil-hasil dari setiap nilai yang di *input* pada masing-masing atribut. Sesuai dengan langkah-langkah untuk NBC :

1. Menentukan mean dan varian pada masing-masing atribut yang bernilai kontinyu terlihat pada Tabel 3 berikut ini.

Tabel 3. Hasil mean dan varian

Data	Mean	Varian
Publikasi   Baik	12,25	5,56
Maahasiswa   Baik	47,75	10,37
Alumni   Baik	26,5	12,767
Kelas   Baik	1,75	0,5
Seminar   Baik	12,5	5,196
Publikasi   Cukup	1,5	2,121
Mahasiswa   Cukup	40,5	40,30
Alumni   Cukup	39	43,84
Kelas   Cukup	1,5	0,707
Seminar   Cukup	1	1,41

2. Probabilitas setiap kelas dengan diberikan masukan yang menjadi data test.

- a. Untuk kelas Baik terlihat pada Gambar 7 berikut :

```
[publikasi] => 0.0623595583185
[dosen] => 1
[pegawai] => 1
[mahasiswa] => 0.00305301763381
[alumni] => 4.71890227121E-8
[kelas] => 0.75
[laboratorium] => 1
[seminar] => 0.0109418479246
```

Gambar 7. Probabilitas kelas Baik

- b. Untuk kelas Cukup terlihat pada Gambar 8 berikut :

```
[publikasi] => 1.08653989691E-10
[dosen] => 1
[pegawai] => 1
[mahasiswa] => 0.000729322702129
[alumni] => 0.00603714587471
[kelas] => 0.5
[laboratorium] => 1
[seminar] => 8.44767383233E-9
```

Gambar 8. Probabilitas kelas Cukup

3. Setelah di dapatkan nilai masing-masing kelas, maka kemudian mencari masing-masing likelihoodnya. Hasilnya terlihat pada Gambar 9:

[Baik] => 7.37267795957E-14

[Cukup] => 2.02071107453E-24

Gambar 9. Likelihood masing-masing kelas

4. Menghitung probabilitas dengan hasilnya yang mendapatkan nilai paling mendekati 1 adalah kelas cukup. Hasilnya terlihat pada Gambar 10 berikut :

[Baik] => 1

[Cukup] => 0

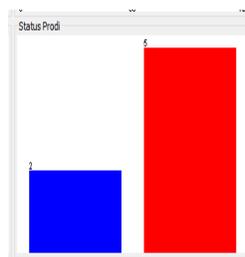
Gambar 10. Hasil kelas yang paling mendekati angka 1

Pada penelitian ini menghitung akurasi menggunakan *confusion matrix* pada weka *classifier*. Dari hasil *confusion matrix* 25 record di prediksi tepat pada class a atau cukup sedangkan 1 record diprediksikan tidak tepat untuk kelas cukup, karena terdapat pada kelas b atau kelas kurang. Pada class b diprediksikan tepat yaitu 16 record sedangkan 1 record di prediksikan tidak tepat karena diprediksikan ada pada kelas cukup. Dari hasil tersebut, dapat di hitung nilai akurasi :

$$\begin{aligned} \text{Presentase akurasi} &= \frac{\text{Danyaknya prediksi yang benar}}{\text{Total banyaknya prediksi}} \\ &= \frac{2+5}{2+1+1+5} \times 100\% \\ &= 100\% \end{aligned}$$

Dengan didapaknya presentase mencapai 100% untuk tingkat akurasi dari perhitungan *confusion matrix*, maka data program studi yang menggunakan metode *naïve bayes classifier* dinyatakan akurat.

Untuk lebih memahami dalam menganalisa hasil dari klasifikasi tiap atribut, diberikan hasilnya dalam bentuk visualisasi grafik seperti pada Gambar 11.



Gambar 11. Hasil klasifikasi weka

Keterangan :

Warna biru : grafik *class* cukup

Warna merah : grafik *class* baik

## 5. Kesimpulan

Berdasarkan penelitian yang dilakukan, Penggunaan metode NBC dalam sistem pendukung keputusan untuk mengevaluasi setiap atribut yang berpengaruh pada program studi, sehingga dapat diketahui keadaan atau status masing-masing program studi. Sistem evaluasi ini juga dapat menjadi bahan pertimbangan untuk pengambilan keputusan pimpinan dalam merekomendasikan kebutuhan untuk program studi sehingga dapat menciptakan fakultas menjadi lebih baik lagi. Penggunaan NBC sangat efektif digunakan pada data program studi karena mempunyai nilai akurasi mencapai 100%.

## Daftar Pustaka

- Abidin, T.F., Subianto, M., 2012. Praktikum Data Mining Naive Bayesian Classifier.
- Chu, P.S. and Zhao, X, 2011. Bayesian analysis for extreme climatic events: A review, Hawaii. *Journal of Atmospheric Research* 102. 243–262
- Han J. and Kamber, M, 2006. *Data Mining Concepts and Techniques*.
- Jiang, X. and Mahadevan, S, 2007. Bayesian risk-based decision method for model validation under uncertainty. *Journal of Reliability Engineering and System Safety* 92. 707–718.
- Koc, L., Mazzuchi, T.A., Sarkani, S., 2012. A network intrusion detection system based on a Hidden Naïve Bayes multiclass classifier, USA. *Journal of Expert Systems with Applications* 39. 13492–13500.
- Neila, R., Djamaludin, A., Yuliyardi, S., Peno, S., 2012. Teacher Quality Improvement Program: Empowering teachers to increasing a quality of Indonesian's education, Indonesia. *Journal of Procedia-Social and Behavioral Sciences* 69 (2012 ) 1836 – 1841.
- Ouali, A, Cherif, A.R., Krebs, M.O., 2006. Data mining based Bayesian networks for best classification. *Journal of Computational Statistics & Data Analysis* 51. 1278 – 1292.
- Vallejos, M, Alvarado, J.M., Puente, A., 2012, College performance prediction test, Spain. *Journal of Procedia - Social and Behavioral Sciences* 31. 846 – 851