JURNAL

PENGEMBANGAN SISTEM REKOMENDASI CALON PENERIMA BEASISWA DENGAN METODE NAIVE BAYES

Development of Recommendations of prospective scholarship recipients with Naive Bayes method

Oleh:

YOHAN ADITYA

12.1.03.03.0178

Dibimbing oleh :

1. Dr. SURYO WIDODO, M.Pd.
2. Rina Firliana, S.Kom. M.Kom.

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS TEKNIK

UNIVERSITAS NUSANTARA PGRI KEDIRI

TAHUN 2017

SURAT PERNYATAAN
ARTIKEL SKRIPSI TAHUN 2017

Yang bertanda tangan di bawah ini:

Nama Lengkap	: YOHAN ADITYA
NPM	: 12.1.03.03.0I78
Telepun / HP	: 0823 3317 8202
Alamat Surel (Email)	: yohan.aditya07@gmail.com
Judul Artikel	: Pengembangan Sistem Rekomendasi calon Penerima Beasiswa dengan Metode Naive Bayes
Fakultas – Program Studi	: Teknik – Sistem Informasi
Nama Perguruan Tinggi	: UNIVERSITAS NUSANTARA PGRI KEDIRI
Alamat Perguruan Tinggi	: Jl. KH. Achmad Dahlan No. 76 Kediri (64112)

Dengan ini menyatakan bahwa :

a. artikel yang saya tulis merupakan karya saya pribadi (bersama tim penulis) dan bebas plagiarisme;
b. artikel telah diteliti dan disetujui untuk diterbitkan oleh Dosen Pembimbing I dan II.

Demikian surat pernyataan ini saya buat dengan sesungguhnya. Apabila di kemudian hari ditemukan ketidaksesuaian data dengan pernyataan ini dan atau ada tuntutan dari pihak lain, saya bersedia bertanggungjawab dan diproses sesuai dengan ketentuan yang berlaku.

Mengetahui		Kediri, 1 Februari 2017
Pembimbing I	Pembimbing II	Penulis,
Dr. SURYO WIDODO, M.Pd. NIDN. 0002026403	RINA FIRLIANA, M.Kom. NIDN. 0731087703	YOHAN ADITYA NPM 12.1.03.03.0178

PENGEMBANGAN SISTEM REKOMENDASI CALON PENERIMA BEASISWA DENGAN METODE NAIVE BAYES

Yohan Aditya
12.1.03.03.0178
Teknik – Sistem Informasi
yohan.aditya07@gmail.com
Dr. Suryo Widodo, M.Pd Dan Rina Firliana, S.Kom. M.Kom.
UNIVERSITAS NUSANTARA PGRI KEDIRI

ABSTRAK

Yohan Aditya: Rekomendasi Calon Penerima Beasiswa Menggunakan Dengan Metode *Naive Bayes* di Progam Studi Sistem Informasi Fakultas Teknik Universitas Nusantara PGRI Kediri , Skripsi, Fakultas Teknik, Sistem Informasi UN PGRI Kediri, 2017.

Sistem Pemberian beasiswa merupakan progam kerja yang ada di setiap Perguruan Tinggi. Proses penentuan beasiswa memiliki beberapa tahapan diantaranya tahap pendaftaran dan penyeleksian. Karena proses penyeleksian masih dilakukan secara manual, pengambilan keputusan penentuan beasiswa secara strategis sulit direalisasikan akibat persepsi yang hiterogen sejalan dengan kepentingan masing-masing individu yang terlibat dalam penentuan beasiswa mengakibatkan muncul permasalahan beasiswa yang tidak tepat sasaran.

Data mining merupakan penambangan data atau upaya penggalian informasi dalam suatu *database* yang bejumlah besar. Salah satu metode dalam *data mining* yaitu adalah klasifikasi merupakan sebuah model dalam *data mining* yang dapat memprediksi catagorical label, didalam klasifikasi terdapat algoritma *Naive Bayes* dimana metode ini dapat memprediksi data selanjutnya.Dengan menggunakan algoritma *Naive Bayes* tersebut maka maka memudahankan dalam merekomendasikan nama-nama calon penerima beasiswa sehingga diharapkan akan lebih tepat sasaran.

Kata kunci : *Data mining*, *Naive Bayes*, Beasiswa

I. Latar Belakang

Pendidikan mempunyai peranan yang sangat penting dalam kehidupan masyarakat karena pendidikan merupakan upaya memperbaiki taraf hidup manusia. Hal ini juga telah dijelaskan dalam UU No.20/2003 tentang pendidikan yang menyatakan bahwa pendidikan merupakan usaha sadar dan terencanan untuk mewujudkan suasana belajar dan proses pembelajaran agar peserta didik secara aktif dalam mengembangkan potensi dirinya.

Pendidikan merupakan suatu kebutuhan primer yang sejak dini hingga dewasa hendaknya dirasakan oleh seluruh masyarakat. Hal ini sesuai dengan amanat Undang-Undang Dasar negara kita, anjuran agama, dab menjadi penentu kemajuan suatu bangsa. Pendidikan juga merupakan variabel vital untuk pembangunan suatu bangsa (Abdillah &

Emigawati, 2009). Suatu bangsa bisa maju dengan cepat dibandingkan dengan negara lain karena penyebaran pengetahuan (knowladge) yang merata keseluruh lapisan masyarakatnya. Institusi yang paling bertangung jawab untuk penyebaran pengetahuan adalah institusi pendidikan (Abdillah, Student learning center strategy based on e-learning and blog, 2013).

Pasal 31 ayat 1 Undang-Undang Dasar Negara Republik Indonesia 1945 telah menjelaskan bahwa setiap warga negara berhak mendapat dan berharap untuk selalu berkembang dalam pendidikan. Namun besarnya biaya pendidikan dan kemampuan ekonomi orang tua menjadi suatu hambatan bagi banyak pelajar untuk meneruskan jenjang pendidikan ke perguruan tinggi.

Untuk mengatasi permasalahan diatas, maka pemerintah maupun Perguruan Tinggi memberikan solusi untuk mengatasi hal tersebut salah satunya dengan memberikan berbagai beasiswa sesuai dengan kriteria yang ditetapkan oleh pemerintah maupun perrguruan tinggi.

Beasiswa adalah penghasilan bagi yang menerimanya. Hal ini sesuai dengan ketentuan pasal 4 ayat (1) UU PPh/2000. Disebutkan pengertian penghasilan adalah tambahan kemampuan ekonomis dengan nama dan dalam bentuk apa pun yang diterima atau diperoleh dari sumber yayasan pendidikan atau luar yayasan pendidikan yang dapat digunakan untuk konsumsi atau kekayaan wajib pajak (WP). Karena beasiswa diartikan menambah kemampuan ekonomis bagi penerimanya, berarti beasiswa merupakan penghasilan.

Pemberian beasiswa merupakan progam kerja yang ada di setiap Perguruan Tinggi. Proses penentuan beasiswa memiliki beberapa tahapan diantaranya tahap pendaftaran dan penyeleksian. Karena proses penyeleksian masih dilakukan secara manual, pengambilan keputusan penentuan beasiswa secara strategis sulit direalisasikan akibat persepsi yang hiterogen sejalan dengan kepentingan masing-masing individu yang terlibat dalam penentuan beasiswa mengakibatkan muncul permasalahan beasiswa yang tidak tepat sasaran.

Berdasarkan permasalahan diatas dapat diambil alternatif solusi yaitu membangun sisten penentuan beeasiswa dengan metode *Naive Bayes Clasification* dengan kriteria-kriteria yang sudah ditentukan.

II. Metode *Naive Bayes Classification*

Klasifikasi Bayes didasarkan pada teorema Bayes. Studi yang membandingkan algoritma-algoritma klasifikasi telah menemukan sebuah klasifikasi Bayes yang sederhana yang dikenal sebagai klasifikasi Naive Bayes

yang dapat dibandingkan perfomance-nya dengan klsifikasi keputusan dan jaringan saraf tiruan. Klasifikasi Bayes juga telah memperlihatkan keakurasian yang tinggi dan kecepatan yang baik ketika dijalankan pada database yang besar (Nagenda & Rajendra, 2012)

Metode *Naive Bayes* adalah metode yang dapat digunakan untuk melakukan klasifikasi dokumen secara statistik. Metode ini dapat memprediksi probabilitas keanggotaan suatu kelas dari suatu data. Teorema *Bayes* memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya. Pada teorema *Bayes*, *X* dijabarkan oleh kumpulan *n* atribut dengan adalah beberapa hipotesis, sehingga data *X* termasuk sebuah *class C* (Han, Kamber, & Pei, 2011). Dengan teorema *Bayes*, sebagai berikut :

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

Klasifikasi *Naïve Bayes* yang mengacu pada teorema *Bayes* mempunyai persamaan sebagai berikut

$$P(C_i|X) = \frac{P(X|Ci).P(Ci)}{P(X)}$$

Menurut (Han, Kamber, & Pei, 2011) proses dari pengklasifikasian *Naïve Bayes* adalah sebagai berikut :

1. Variabel *D* adalah kumpulan dari data dan label yang terkait dengan *class*. Setiap data diwakili oleh vektor atribut *n*-dimensi, X = ($x_1.x_2. \dots .x_n$) dengan *n* dibuat dari data atribut, berturut-turut, $A_1.A_2. \dots .A_n$.
2. Misalkan terdapat *class*, $C_1.C_2. \dots .C_n$. Diberikan sebuah data *X*, kemudian pengklasifikasian akan memprediksi *X* ke dalam kelompok yang memiliki probabilitas posterior tertinggi berdasarkan kondisi *X*. Artinya, pengklasifikasian *naïve bayes* memprediksi bahwa data *X* termasuk *class* C_1, jika dan hanya jika $P(Ci|X) > P(Cj|X)$ untuk $1 \leq m, j \neq i$
3. Ketika *P(X)* konstan untuk semua class maka hanya *P(X|Ci) P(Ci)* yang dihitung. Jika probabilitas *class prior* sebelumnya tidak diketahui, maka diasumsikan bahwa *class-nya* sama, yaitu $P(C_1) = P(C_2) = \dots = P(C_m)$, untuk menghitung dan . Perhatikan bahwa probabilitas *class prior* dapat diperkirakan oleh $P(Ci) = \frac{(|C(i,D|)}{|D|}$ dimana $|C_{i,D}|$ adalah jumlah data *training* dari *class* C_i, dan *D* adalah jumlah total data *training* yang digunakan.
4. Apabila diberikan kumpulan data yang mempunyai banyak atribut, maka mengurangi perhitungan *P(X|Ci)* , *naïve bayes* mengasumsikan pembuatan *class independen* yang bersyarat. Anggap bahwa nilai-nilai atribut tersebut bersifat independen

satu sama lain dan diantara atribut tidak terdapat relasi depedensi, maka :

$$P(\mathbf{X} \mid C_i) = \prod_{k=1}^{n} P(x_k \mid C_i) =$$
$$P(x_1 \mid C_i) \times P(x_2 \mid C_i) \times \ldots \times P(x_n \mid C_i)$$

Untuk klasifikasi dengan data kontinyu digunakan rumus *Densitas Gauss* :

Error! Reference source not found.

Keterangan :

P : Peluang

Xi : Atribut ke *i*

xi : Nilai atribut ke *i*

Y : Kelas yang dicari

yi : Sub kelas *Y* yang dicari

μ: *mean*, menyatakan rata-rata dari seluruh atribut

σ: Deviasi standar, menyatakan varian dari seluruh atribut.

III. HASIL DAN KESIMPULAN

Hasil

1. Preprocessing data training

Data Training digunakan oleh algoritma klassifikasi untuk membentuk sebuah model *classifier*. Model ini merupakan representasi pengetahuan yang akan digunakan untuk prediksi kelas data baru yang belum pernah ada.

Berikut ini adalah tabel 4.1 data mentah yang akan digunakan sebelum pengolahan NBC seperti berikut ini :

Tabel 4.1 Data Mentah

NO	NPM	NAMA	IPK	MHS BEKERJA / TDK	ANAK KE	JMLH TANGGUNGA	PEKERJAAN ORTU	PENGHASILAN ORTU	Status Beasiswa
1	14.1.03.03.01xx	Esti Agus Triana	3.60	Tidak	3	2	Swasta	3.000.000	TIDAK
2	14.1.03.03.00xx	Galuh Prahadi Gumelar	3.10	Tidak	1	2	Wiraswasta	1.000.000	DAPAT
3	14.1.03.03.00xx	Fergy Druta Pradana	3.10	Tidak	1	2	Wiraswasta	1.000.000	DAPAT
4	14.1.03.03.00xx	Limek Setyawan	3.40	Bekerja	2	2	Swasta	2.000.000	TIDAK
5	14.1.03.03.00xx	Fajar Wintoro	3.45	Tidak	2	2	Petani	1.000.000	DAPAT
6	14.1.03.03.00xx	Dwisd Permana	3.45	Tidak	2	1	Petani	1.000.000	DAPAT
7	14.1.03.03.01xx	M.Ibnu Karom	3.10	Tidak	1	2	Swasta	1.200.000	DAPAT
8	14.1.03.03.00xx	Tri Handoko	3.30	Tidak	3	2	Driver	1.200.000	DAPAT
9	14.1.03.03.00xx	Diana Sulistyani	3.75	Bekerja	1	2	Swasta	5.000.000	TIDAK
10	14.1.03.03.00xx	Erik Made Saputro	3.45	Tidak	2	2	Wiraswasta	1.300.000	DAPAT
11	14.1.03.03.00xx	Puji Nurvita	3.80	Bekerja	2	2	Petani	1.000.000	DAPAT
12	14.1.03.03.00xx	Siti Oktaviana R	3.50	Bekerja	1	2	Swasta	5.000.000	TIDAK
13	14.1.03.03.00xx	Dimas Wilan Agustiana	3.50	Bekerja	2	2	Wiraswasta	2.000.000	TIDAK
14	14.1.03.03.00xx	Mahendika Fernando	3.50	Bekerja	1	2	Wiraswasta	2.000.000	TIDAK
15	14.1.03.03.00xx	Moch Afandi	3.45	Bekerja	3	1	Wiraswasta	950.000	DAPAT
16	14.1.03.03.00xx	Fitra Cahaya Pramana	3.55	Bekerja	1	1	Petani	800.000	DAPAT
17	14.1.03.03.01xx	Dika Hari Prasnowo	2.20	Bekerja	1	2	Swasta	1.200.000	TIDAK
18	14.1.03.03.01xx	Christanto hendri setyawan	2.90	Bekerja	1	2	Polisi	2.500.000	TIDAK
19	14.1.03.03.00xx	Danny Ria Pratama	2.55	Bekerja	1	2	Wiraswasta	1.000.000	TIDAK
20	14.1.03.03.00xx	Ade Frasdiarto	2.70	Bekerja	2	1	Satpam	1.500.000	TIDAK
21	14.1.03.03.00xx	Ahmad Husein Arrifa'i	2.90	Bekerja	3	3	Petani	1.000.000	DAPAT
22	14.1.03.03.01xx	Bella Riang Pratama	3.60	Bekerja	2	2	Wiraswasta	1.000.000	DAPAT
23	14.1.03.03.01xx	Dika Hari Prasnowo	2.20	Bekerja	1	2	Swasta	1.200.000	TIDAK
24	14.1.03.03.01xx	Fajar Imaduddin	2.90	Bekerja	1	2	Swasta	900.000	DAPAT
25	14.1.03.03.00xx	One Rokhisi. S	2.55	Bekerja	1	3	Wiraswasta	1.000.000	TIDAK

Sebelum data mentah diproses untuk dijadikan sebagai data training terlebih dahulu diubah menjadi bentuk numerik atau ditransformasikan, untuk keterangan transformasi dapat dilihat pada tabel 4.2 berikut ini :

Tabel 4.2 Keterangan Transformasi Data

No	Data	Jenis Data	Transformasi
1	IPK	< =2,95	3
		2,96 - 3,51	2
		>= 3,52	1
2	Mhs berkerja/tidak	Berkerja	1
		Tidak Berkerja	2
3	Pekerjaan Orang Tua	Swasta	1
		Polisi	2
		Satpam	3
		Supir	4
		Wiraswasta	5
		Petani	6
4	Penghasilan Orang Tua	<= Rp. 2.299.000	2
		Rp. 2.300.000 - Rp 4.999.999	3
		>= Rp. 5.000.000	1
5	Status Beasiswa	Dapat	1
		Tidak	2

Dari tabe keterangan transfrormasi diatas selanjunya data mentah pada tabel 4.1 akan

di bentuk matriks transformasi pada tabel 4.3 data mentah ditransformasi kebentuk numerik yang dapat dilihat pada tabel berikut ini :

Tabel 4.3 Data mentah ditransformasi ke bentuk numerik

NO	NPM	NAMA	IPK	MHS BEKERJA / TDK	ANAK KE	JMLH TANGGUNGAN	PEKERJAAN ORTU	PENGHASILAN ORTU	Status Beasiswa
1	.1.03.03.01	Esti Agus Triana	1	2	3	2	1	3.000.000	2
2	.1.03.03.00	Galuh Prahadi Gumelar	2	2	1	2	5	1.000.000	1
3	.1.03.03.00	Fergy Druta Pradana	2	2	1	2	5	1.000.000	1
4	.1.03.03.00	Limek Setyawan	2	1	2	2	1	2.000.000	1
5	.1.03.03.00	Fajar Wintoro	2	2	2	2	6	1.000.000	1
6	.1.03.03.00	Devid Permana	2	2	2	1	6	1.000.000	1
7	.1.03.03.01	M.Ibnu Karom	2	2	1	2	1	1.200.000	1
8	.1.03.03.00	Tri Handoko	2	2	3	2	4	1.200.000	1
9	.1.03.03.00	Diana Sulistyani	1	1	1	2	1	5.000.000	2
10	.1.03.03.00	Erik Made Saputro	2	2	2	2	5	1.300.000	1
11	.1.03.03.00	Puji Nurvita	1	1	2	2	6	1.000.000	1
12	.1.03.03.00	Siti Oktaviana R.	2	1	1	2	1	5.000.000	2
13	.1.03.03.00	Dimas Wulan Agustiana	2	1	2	2	5	2.000.000	1
14	.1.03.03.00	Maharditka Fernando	2	1	1	2	5	2.000.000	1
15	.1.03.03.00	Moch Afandi	2	1	3	1	5	950.000	1
16	.1.03.03.00	Fitra Cahaya Pramana	1	1	1	1	6	800.000	1
17	.1.03.03.01	Dika Hari Prasnowo	3	1	1	2	1	1.200.000	2
18	.1.03.03.01	Christanto hendri setyawan	3	1	1	2	2	2.500.000	2
19	.1.03.03.00	Danny Ria Pratama	3	1	1	2	5	1.000.000	2
20	.1.03.03.00	Ade Frasdiarto	3	1	2	1	3	1.500.000	2
21	.1.03.03.00	Ahmad Husein Amfa'i	3	1	3	3	6	1.000.000	2
22	.1.03.03.01	Bellia Riang Pratama	1	1	2	2	5	1.000.000	1
23	.1.03.03.01	Dika Hari Prasnowo	3	2	1	2	1	1.200.000	2
24	.1.03.03.01	Fajar Imaduddin	3	2	1	2	1	900.000	1
25	.1.03.03.00	One Rokhsi. S	3	2	1	3	5	1.000.000	2

Kemudian dari data diatas selanjutnya dicari nilai korelasi terhadap variabel output seperti yang dapat dilihat pada tabel 4.4 korelasi .

Tabel 4.4 Korelasi

Correlations

Control Variables			IPK	Status_Kerja	Anak_Ke	JML_Tanggungan_Ortu	Pekerjaan_Ortu	Penghasilan_Ortu
Status_Beasiswa	IPK	Correlation	1.000	-.300	.334	-.097	.017	.760
		Significance (1-tailed)	.	.077	.055	.326	.469	.000
		df	0	22	22	22	22	22
	Status_Kerja	Correlation	-.300	1.000	-.041	.234	-.136	-.221
		Significance (1-tailed)	.077	.	.424	.136	.264	.150
		df	22	0	22	22	22	22
	Anak_Ke	Correlation	.334	-.041	1.000	-.036	.226	-.084
		Significance (1-tailed)	.055	.424	.	.434	.144	.349
		df	22	22	0	22	22	22
	JML_Tanggungan_Ortu	Correlation	-.097	.234	-.036	1.000	.076	-.070
		Significance (1-tailed)	.326	.136	.434	.	.362	.372
		df	22	22	22	0	22	22
	Pekerjaan_Ortu	Correlation	.017	-.136	.226	.076	1.000	-.433
		Significance (1-tailed)	.469	.264	.144	.362	.	.017
		df	22	22	22	22	0	22
	Penghasilan_Ortu	Correlation	.760	-.221	-.084	-.070	-.433	1.000
		Significance (1-tailed)	.000	.150	.349	.372	.017	.
		df	22	22	22	22	22	0

Berdasarkan hasil korelasi pada tabel 4.4 maka diperoleh hasil kriteria yaitu variabel indeks prestasi komulatif dan penghasilan orang tua. Dari hasil kriteria tersebut digunakan sebagai kriteria yang selanjutnya digunakan dalam perhitungan seleksi beasiswa dengan metode *naive bayes* karena memiliki nilai korelasi mulai dari 0,5 keatas (Sarwono, 2006).

Langkah langkah dari *naive bayes* sudah dipaparkan sebelumnya akan dijabarkan menggunakan *data testing*. *Data Testing* digunakan untuk mengukur sejauh mana *classifier* berhasil melakukan klasifikasi dengan benar dan untuk menentukan data yang nantinya akan dianalisis dengan metode *naive bayes*. Maka, tahapan awal adalah dengan membaca *data training* dari data mahasiswa. Adapun *data training* yang digunakan dapat dilihatpada tabel *data training.*

Tabel 4.5 *Data Training*

NO	NPM	NAMA	IPK	PENGHASILAN ORTU	Status Beasiswa
1	14.1.03.03.01	Esti Agus Triana	1	3	2
2	14.1.03.03.00	Galuh Prahadi Gumelar	2	2	1
3	14.1.03.03.00	Fergy Druta Pradana	2	2	1
4	14.1.03.03.00	Limek Setyawan	2	2	1
5	14.1.03.03.00	Fajar Wintoro	2	2	1
...	...	...	...	...	...
...	...	...	...	...	...
...	...	...	...	...	...
23	14.1.03.03.01	Dika Hari Prasnowo	3	2	2
24	14.1.03.03.01	Fajar Imaduddin	3	2	1
25	14.1.03.03.00	One Rokhsi. S	3	2	2

Dari tabel 4.5 *Data Training* maka akan dicari *nilai probabilitas*. Berikut ini perhitungan menggunakan *naive bayes* :

1. Mencari probabilitas untuk data kontinyu mengacu pada rumus (2.2)

Langkah pertama menghitung *nilai probabilitas* dari setiap dapat dilihat pada tabel 4.6 berikut ini :

Tabel 4.6 *prior probabilitas P(Ci)* dari setiap Beasiswa

Data Penerima Beasiswa	Banyak Data	P(X)
1 (DAPAT)	13/25	0,52
2 (TIDAK)	12/25	0,48

Dari perhitungan tersebut maka dapat diketahui probabilitas Ya sebesar 0,52 dan Tidak sebesar 0,48. Setelah itu mencari nilai probabilitas untuk kriteria IPK, untuk nilai dari kriteria IPK dapat dilihat pada tabel 4.7 berikut ini.

Tabel 4.7 Probabilitas P(IPK|Beasiswa)

IPK	Banyak Data	Ya	Tidak	Probabilitas	
				Dapat	Tidak
1 (< =2,95)	5	3	2	0,2308	0,1667
2 (2,96 - 3,51)	12	8	4	0,6154	0,3333
3 (> = 3,52)	8	2	6	0,1538	0,5
Jumlah	25	13	12		

Pada Tabel 4.7 menghasilkan probabilitas untuk IPK 1 (<= 2,95) nilai Dapat sebesar 0,2308 dan nilai tidak sebesar 0,1667. IPK 2 (2,96 - 3,51) nilai Dapat sebesar 0,6154dan nilai tidak sebesar 0,3333. IPK 3 (> = 3,52) nilai Dapat sebesar 0,1538dan nilai tidak sebesar 0,5.

Berikut ini perhitungan untuk mencari probabilitas kriteria pendapatan orang tua dapat dilihat pada tabel berikut ini.

Tabel 4.8 Probabilitas P(Pendapatan Orang Tua|Beasiswa)

Pendapatan Orang Tua	Banyak Data	Dapat	Tidak	Probabilitas	
				Ya	Tidak
2 (< = Rp. 2.299.000)	2	0	2	0	0,1667
3 (Rp. 2.300.000 - Rp 4.999.999)	18	13	5	1	0,4176
1 (> = Rp. 5.000.000)	5	0	5	0	0,4176
Jumlah	25	13	12		

Pada Tabel 4.8 menghasilkan probabilitas untuk Pendapatan Orang Tua 2 (< = Rp. 2.299.000) nilai Dapat sebesar 0 nilai tidak sebesar 0,1667. Pendapatan Orang Tua 3 (Rp. 2.300.000 - Rp 4.999.999) nilai Dapat sebesar 1 dan nilai tidak sebesar 0,4176. Pendapatan Orang Tua 1 (> = Rp. 5.000.000) nilai Dapat sebesar 0 dan nilai tidak sebesar 0,4176.

Berdasarkan perhitungan data training diatas apabila dimasukan sebuah data dengan kondisi seperti dibawah ini.

NPM	NAMA	IPK	Pendapatan Orang Tua
14.1.03.03.01xx	Desvanda Verdinan Hariyadi	3.25	1.100.000

Maka perhitungan menggunakan metode *naive bayes* sebagai berikut :

P(IPK="2"|Penerima Beasiswa=" Dapat") = 8/13 =0,6154

P(IPK="2"|Penerima Beasiswa ="Tidak")= 4/12 =0,3333

P(Pendapatan Orang Tua="2"|Penerima Beasiswa =" Dapat") = 13/13 = 1

P(Pendapatan Orang Tua ="2"|Penerima Beasiswa ="Tidak") = 5/12 = 0,4176

Menghitung posterior probabilities P(Ci) sampel data X terhadap class ke-i sesuai rumus (2.3).

Maka nilai P(X|Dapat) = 0,6154 x 1 = 0,6154

Nilai P(X|Tidak) = 0,3333 x 0,4176 = 0,1392

Menghitung posterior probabilities P(Ci) sampel data X terhadap class ke-i

P(X|Dapat) = 0,52 x 0,6154 = 0,3200

P(X|Tidak) = 0,48 x 0,1392 = 0,0668

Dari perhitungan *probabilitas* masing-masing calon penerima beasiswa terhadap data X, maka didapat bahwa *data testing* dengan data diatas memiliki *nilai probabilitas* Dapat sebesar 0,3200 dan Tidak sebesar 0,0668. Maka dapat disimpulakan bahwa mahasiswa tersebut termasuk kedalam golongan calon penerima Beasiswa.

Hasil pengujian dari ke 25 data training dan 1 data testing yang ada menggunakan Confusion Matrix menunjukan :

Tabel 4.9 Hasil Pengujian

No.	Penerima Beasiswa riil	Penerima Beasiswa versi NBC	Sesuai
1.	Tidak	Tidak	Sesuai
2.	Dapat	Dapat	Sesuai
3.	Dapat	Dapat	Sesuai
4.	Tidak	Tidak	Sesuai
5.	Dapat	Dapat	Sesuai
6.	Dapat	Dapat	Sesuai
7.	Dapat	Dapat	Sesuai
8.	Dapat	Dapat	Sesuai
9.	Tidak	Tidak	Sesuai
10.	Dapat	Dapat	Sesuai
....	....	....	....
....	....	....	....
....	....	....	....
23.	Tidak	Tidak	Ya
24.	Dapat	Tidak	Tidak Sesuai
25.	Tidak	Tidak	Sesuai

Kinerja sistem $= \frac{11 + 12}{11 + 12 + 0 + 2} \times 100\% = 92\ \%$

Dari perhitungan diatas antara data riil dengan perhitungan menggunakan *naive bayes* menghasilkan tinggkat akurasi 92%. Sehingga metode *naive bayes* dapat diterapkan dalam rekomendasi untuk menentukan calon Penerima Beasiswa.

Kesimpulan

Berikut simpulan yang dapat diambil dalam penulisan skripsi ini adalah :

Pengembangan Sistem rekomendasi calon penerima beasiswa menggunakan algoritma *Naive Bayes* dapat mempermudah dalam memilih calon penerima beasiswa sehingga Kepala Prodi bisa mengambil keputusan untuk merekomendasikan mahasiswa yang terpilih menjadi calon penerima beasiswa tersebut.

IV. DAFTAR PUSTAKA

Abdillah, L. (2013). Student learning center strategy based on e-learning and blog. *Seminar Nasional Sains dan Teknologi (SNST)*, 15-20.

Abdillah, L., & Emigawati. (2009). Analisis laporan tugas akhir mahasiswa diploma 1 dari sudut pandang kaidah karya ilmiah dan pengunaan teknologi informasi. *Jurnal Ilmiah MATRIX Vol 11*, 19-36.

Anandita, E. R. (2014). Klasifikasi Tebu Dengan menggunakan Algoritma

Naive Bayes Clasification pada Dinas Kehutanan dan Perkebunan Pati . *Skripsi*, 1-13.

Bala, M. :. (2012). Study of Application of Data Mining Techniques in Education. *International Journal of Research in Science and Technology (IJRST), Vol.No. 1*.

Bustami. (2013). Penerapan Algoritma Naive Bayes untuk Mengklasifikasi Data Nasabah Asuransi. *TECHSI: Jurnal Penelitian Teknik Informatika*, 127-146.

Goole, S., & Chanana, N. (2012). Data Mining Trens in Past, Current, and Future. *International Journal Computing & Business Research*.

Han, J., & Kamber, M. (2006). *Data Mining Concepts and Techniques.* San Francisco: Diane Cerra.

Han, J., Kamber, M., & Pei, J. (2011). *Data Mining: Concepts And Techniques 3RD Edition.* Morgan Kaufmann.

Indriyawan, E., & kk. (2011). *Mastering Delphi XE*. Yogyakarta: CV.Andi Offset.

Jiawei, H., & Micheline, K. (2012). *Data Mining Concept and Technique.*

Kendall Dalam Wicaksono, A. (2014). Aplikasi Sistem Pendukung Keputusan Pemilihan Mobil Bekas Menggunakan Metode Topsis.

Moh.Muthohir, & Vicent, S. (2013). Penentuan Rangking Calon Penerima Beasiswa menggunakan Fuzzy Query pada Database. *Jurnal Teknologi Informasi, Vol. 9, No. 2, ISSN 1414-9999*, 40-48.

Nagenda, K., & Rajendra, C. (2012). Customer behaviour analysis using CBA (Data Mining Approach). *IJECCE*, 65-68.

Neel, B. (2011). Predictive Data mining and discovering hidden values of Data warehouse. *ARPN Journal of System and Softwere, Vol 1 No. 1*, 1-5.

Nugroho, A. (2004). *Konsep Pengembangan Sistem Basis Data*. Bandung.

Nugroho, A. (2011). *Perancangan dan Implementasi Sistem Basis Data.* Yogyakarta: Andi.

Nurul, R., sofi, D., & Jajuli, M. (2015). Implementasi algoritma K-Means dalam Pengklasteran mahasiswa pelamar beasiswa. *Jurnal Ilmiah Teknologi Terapan ISSN : 2407-3911 Vol. 1,No.2*, 62-68.

Prasetyo, E. (2012). Algoritma K Means. *Clustering*, 44-45.

Priyanti, D., & Iriani, S. (2013). Sistem Informasi Data Penduduk Pada Desa Bogoharjo Kecamatan Ngadirojo

Kabupaten Pacitan. *IJSN Volume 2 No 4 -ISSN : 2302-5700*, 55-61.

Putra, A., & Yunika, D. H. (2011). Penentuan Penerima Beasiswa dengan menggunakan Fuzzy Multiple Atribute Decession Making. *Jurnal Sistem Informasi Vol 3 ISSN Online 2355-4614*, 289-293.

Ridwan, M., Suyono, H., & M.Sarosa. (2013). Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier. *Jurnal EECCIS vol.7, No. 1*, 59-64.

Ridwan, M., Suyono, H., & Sarosa, M. (2013). Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier. *Jurnal EECCIS vol.7, No. 1*, 59-64.

Rohmawati, N., Defiyanti, S., & Jajuli, M. (2015). Implementasi Alhgoritma K-Means dalam Pengklasteran Mahasiswa pelamar Beasiswa. *Jurnal Ilmiah Teknologi Informasi Terapan Vol I No. 2*, 62-68.

Rosa, A., & Shalahuddin, M. (2014). *Rekayasa Perangkat Lunak Terstruktur dan Berorientasi Objek.* Bandung: Informatika Bandung.

Sarwono, J. (2006). *Metode Penelitian Kantitatif dan Kualitatif.* Yogyakarta: Graha Ilmu.

Whitehorn, M., & Marklyn, B. (2003). *Seluk Beluk Database Relasional Edisi ke 2*. Jakarta: Erlangga.