PENERAPAN KLASIFIKASI METODE NAÏVE BAYES PADA PENERIMAAN ANGGOTA UPU-CC DI UNIVERSITAS POTENSI UTAMA

APPLICATION OF NAVE BAYES CLASSIFICATION METHOD ON ADMISSION OF UPU-CC MEMBERS AT POTENSI UTAMA UNIVERSITY

Charles Bronson Harahap[1], Hardianto[2], Linda Wahyuni[3], Cindy Paramitha[4]
Fakultas Teknik dan Ilmu Komputer
Universitas Potensi Utama
Jl. K.L Yos Sudarso Km. 6,5 No.3A-Medan, 20241, Indonesia
Email : [1]charlesharahap07@gmail.com, [2]hardiantoandre08@gmail.com,
[3]lindawahyuni391@gmail.com, [4]cindyparamitha96@gmail.com

ABSTRAK

Organisasi kemahasiswaan di kampus bisa disebut Unit Kegiatan Mahasiswa (UKM). Mahasiswa aktif dapat bergabung dengan salah satu organisasi kemahasiswaan tersebut, di dalam Universitas Potensi Utama terdapat beberapa jenis UKM salah satunya adalah Klub Komputer Universitas Potensi Utama (UPU-CC) dalam organisasi ini memiliki tujuan untuk menggabungkan ilmu komputer seperti pemrograman, multimedia, keamanan jaringan, teknologi pengetahuan. Masalah penerimaan anggota harus diuji terlebih dahulu dengan ilmu yang sesuai dengan bidang minatnya agar nantinya tidak salah kaprah. Peneliti memandang perlu adanya klasifikasi dalam pemilihan agar pemilihan anggota dapat berjalan dengan cepat dan akurat. Tujuan dari penelitian ini adalah untuk membuat klasifikasi dengan menerapkan metode Naïve Bayes dan menguji akurasi metode Naïve Bayes untuk penerimaan anggota UPU-CC. Ada 47 data yang akan diuji pada tahun 2019. Implementasi diuji menggunakan tools Weka 3.8. Akurasi metode nave bayes sebesar 97,8723% dengan nilai Correctly Classified Instances sebesar 46 data sebesar 97,8723% dan Incorrectly Classified Instances 1 data sebesar 2,1277%.

Kata kunci: *Naïve Bayes, Akurasi, dan UPU-CC*

ABSTRACT

Student organizations on campus can be called Student Activity Units (UKM). Active students can join one of these student organizations, within the University of Potential Main there are several types of SMEs, one of which is the University Potensi Utama Computer Club (UPU-CC) in this organization has the goal to combine computer science such as programming, multimedia, network security, knowledge technology. The problem with accepting members must be tested in advance with the knowledge in accordance with the area of interest so that later they are not mistaken. Researchers consider the need for classification in the election so that the selection of members can run quickly and accurately. The purpose of this study is to make a classification by applying the Naïve Bayes method and testing the accuracy of the Naïve Bayes method for acceptance of UPU-CC members. There are 47 data to be tested in 2019. Implementation is tested using Weka 3.8 tools. The accuracy of the naïve bayes method is 97.8723% with the Correctly Classified Instances value of 46 data of 97.8723% and the Incorrectly Classified Instances 1 data of 2.1277%.

Keywords*: Naïve Bayes, Accuracy, and UPU-CC*

1. PENDAHULUAN

Pada zaman sekarang mahasiswa dituntut harus terus aktif tidak hanya dalam dunia pendidikan saja akan tetapi harus dapat berperan aktif dalam lingkungan organisasi. Pada dunia pekerjaan sekarang beberapa perusahaan meminta untuk calon pelamarnya memiliki pengalaman dalam berorganisasi dengan tujuan agar calon pelamar tersebut dapat terlihat aktif dalam berkelompok dan dapat memungkinkan untuk bekerja secara team work.

Organisasi kemahasiswaan disebut Unit Kegiatan Mahasiswa (UKM). UKM adalah organisasi yang melibatkan mahasiswa di lingkungan kampus. Universitas Potensi Utama memiliki beberapa jenis UKM seperti Universitas Potensi Utama Computer Club (UPU-CC), Universitas Potensi Utama Robotic Club (UPU-RC), Universitas Potensi Utama English Club (UPU-EC), UKMI Ulul Albab UPU, Ikatan Mahasiswa Kristem Universitas Potensi Utama (IMK-UPU), Sanggar Seni Universitas Potensii Utama (Sanggar seni-UPU), Persatuan Mahasiswa Cinta Alam, Lingkungan Universitas Potensi Utama (Pascal-UPU)

Pada setiap UKM memiliki standar penerimaan anggota masing – masing agar tidak salah mengambil jalur terhadap minat dan keahlian masing – masing mahasiswa. Mahasiswa yang dapat mendaftar anggota bisa dari semester 1 sampai dengan semester 6. Pada semester 7 dan 8 mahasiswa tidak dapat mendaftar lagi menjadi anggota dikarenakan mahasiswa disemester akhir sudah harus fokus kepada Skripsi atau Tugas Akhir mahasiswa.

Klasifikasi adalah pengelompokkan berdasarkan atribut yang sama berdasarkan data atau informasi yang telah dikumpulkan(Berliana, Gito, and Agus 2018). Klasifikasi yang diurutkan berdasarkan kemiripan akan dapat diterapkan dengan berbagai kasus permasalahan dan dapat membentuk sebuah *rule* aturan terhadap data yang telah dikumpulkan(Sumpeno 2009). Klasifikasi juga sering digunakan dalam metode Data Mining untuk menentukan kemiripan data antar satu kelompok data(Suryadi and Harahap 2018)

Data mining merupakan proses menggali data dalam skala besar untuk menentukan data yang dapat diklasifikasikan antara satu kelas dengan kelas yang lain(Nugroho and Wibowo 2017). Data mining dapat menghubungkan data dengan sejumlah informasi untuk menemukan model yang dapat menjadi dasar dari pengambilan keputusan guna menyajikan data yang secara relavan(Mas'udia 2015).

Naïve Bayes merupakan metode yang diterapkan pada teorema bayes dalam menentukan klasifikasi. Naïve Bayes dapat di ketahui nilai dari setiap kelas dengan melalukan perhitungan probabilitas pada tiap kelas dan perhitungan statistik sederhana(Patmi Kasih 2017).

Berdasarkan Patmi kasih dan Inta Nur Farida dengan judul penelitian "Sistem Bantu Pemilihan dosen pembimbing Tugas Akhir berdasarkan kategori pilihan dan keahlian dosen menggunakan Naïve Bayes" bahwa Naïve Bayes dapat menghasilkan informasi nama – nama dosen pembimbing tugas akhir sesuai dengan kategori keahlian dengan kategori skripsi mahasiswa (Patmi Kasih 2017).

Menurut penelitian yang dilakukan Derick Iskandar dan Yoyon K.Suprapto dengan judul penelitian "Perbandingan akurasi klasifikasi tingkat kemiskinan antara algoritma C4.5 dan Naïve Bayes Clasifier" menyatakan bahwa metode classifier dan C4.5 merupakan metode yang dilakukan dengan penerepan data mining dengan nilai akurasi C4.5 sebesar 64% dan Naïve Bayes sebesar 61%(Iskandar and Suprapto 2013).

Menurut hasil penelitian Alfa Saleh dan Fina Nasari dengan judul penelitian "Penggunaan Teknik *Unsupervised Discretization* Pada Metode Naïve Bayes dalam Menentukan Jurusan Siswa Madrasah Aliyah" bahwa Naïve Bayes bertujuan untuk menemukan klasifikasi dari beberapa kriteria yang diuji dengan 120 data siswa dengan menerapkan Teknik dari unsupervised discretization nai yang awalnya hanya 90% dapat naik menjadi 92.8% (Saleh and Nasari 2018b).

Implementasi dari penerapan metode Naïve Bayes terhadap penentuan seleksi penerimaan anggota baru pada UKM UPU-CC yang akan dilakukan dengan perhitungan manual yang masing masing atribut kelas dihitung nilai probabilitas dan statistiknya. Kemudian dilakukan pengujian dengan menggunakan aplikasi weka 3.8. Pengujian tersebut dilakukan dengan tujuan memastikan nilai ke akuratan perhitungan dengan metode Naïve Bayes(Septiani 2017).

UKM UPU-CC merupakan kegiatan organisasi kemahasiswaan dengan standar komputerisasi yang bertujuan untuk menunjang pengetahuan di bidang ilmu komputer dari tingkat dasar hingga tingkat *expert* agar membentuk mahasiswa menjadi ahli dibidang komputer. Keahlian dibidang komputer dapat didukung dengan pengetahuan yang dibina sejak dini oleh karena itu anggota harus mengetahui tingkat dasar dan diuji sebelum masuk menjadi anggota.

UKM UPU-CC memiliki standar penilaian tersendiri, seperti yang sudah peneliti dapat informasi dari Ketua Umum UKM UPU-CC yang menilai dari segi tes tertulis, tes praktek sesuai bidang, dan interview secara personal dengan ketua masing – masing devisi yang telah dipilih oleh calon anggota baru. Hasil penilaian yang telah didapat akan dirundingkan lagi sehingga akan dapat kesimpulan apakah diterima atau ditolak. Oleh sebab itu, peneliti menganggap perlu untuk melakukan klasifikasi dengan menggunan metode Naïve Bayes agar tidak terjadi kecurangan, individualis dan dapat dengan mudah menentukan keputusan tersebut.

2. METODE PENELITIAN

Teknik Pengumpulan Data

Adapun metode penelitian yang dilakukan pada peniltian adalah sebagai berikut :

1. Analisis Masalah
   Permasalahan dalam penelitian ini adalah penentuan dalam pemilihan anggota UKM UPU-CC yang memiliki tahapan dalam seleksi anggota dari mulai pendaftaran hingga pengumuman diterima atau ditolak. Dari beberapa kriteria yang ada disimpulkan dan dipelajari alurnya sehingga dapat memiliki konsep untuk diterapkan dalam pengklasfikasian tersebut dengan memisahkan antar devisi yang ada sehingga tidak terjadi kesalahan pemberian nilai.
2. Pengumpulan Data
   Peneliti memperoleh data mahasiswa yang diterima dan ditolak berdasarkan data yang telah di input oleh sekertaris UPU-CC. Data tersebut merupakan data mentah dalam bentuk Excel yang dibuat untuk mendata anggota UPU-CC yang mendaftar. Dan sudah dapat dilihat hasil dari nilai tes yang mereka dapatkan. Data yang akan di training sebanyak 47 data dari 150 data yang ada.
3. Pengujian Data
   Pengujian data mahasiswa tersebut di lihat dengan menggunakan perhitungan Naïve Bayes dan di implementasikan dengan *tools* Weka 3.8 untuk melihat kesesuai dengan perhitungan manual yang di lakukan sehingga mengetahui hasil dari akurasi yang didapatkan dengan menerapkan metode Naïve Bayes tersebut.

Metode Naïve Bayes

Naive Bayes adalah sekelompok data yang diklasifikasi dengan menerapkan perhitungan statistik sederhana dan mencari nilai probabilitas dari masing – masing atribut(Muhamad et al. 2017). Independent yang saling yang saling berhubungan antara nilai variabel dengan kelasnya yang biasa kita sebut dengan teori dari teorema bayes (Rozi, Hamdana, and Iqbal Alfahmi 2018).

Naïve Bayes biasanya menangani data yang tidak relevan untuk meganalisa data set dengan skala yang cukup luas karena dapat berbentuk karakter maupun angka(Devita, Herwanto, and Wibawa 2018). Jumlah atribut dan banyaknya data dapat mempengaruhi hasil dari nilai akurasi yang didapatkan karena setiap kelas dari masing – masing atribut cukup berbeda(Ridwan, Suyono, and Sarosa 2013). Naive Bayes membutuhkan sejumlah data training yang akan diuji untuk

perhitungannya dan memiliki data testing untuk melihat ke akuratan data dari perhitungan yang didapat(Syarli and Muin 2016).

Penerapan Metode Naive Bayes

Penerapan dari teorema Bayes adalah (Muhamad et al. 2017):

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

Keterangan :
X : Nilai dari kelas data yang akan di input.
H : Hipotesis nilai spesifik dari suatu *class*
P(H|X) : Probabilitas hipotesis nilai H posteriori dari probabilitas nilai X (posteriori probabilitas)
P(H) : Prior Probabilitas hipotesis H (prior probabilitas)
P(X|H) : Nailai X berdasarkan probabilitas hipotesis nilai H
P(X) : Probabilitas nilai X
Pada penjelasan metode Naive Bayes, diingatkan bahwa proses dari sebuah klasifikasi perlu adanya sejumlah petunjuk agar dapat menentukan *class* yang tepat untuk sampel yang akan dianalisis tersebut. Oleh sebab itu, metode Naive Bayes di atas dimasukkan kedalam rumus sebagai berikut(Devita, Herwanto, and Wibawa 2018):

$$P(C|F1 \ldots . Fn) = \frac{P(C)P(F1 \ldots . Fn|C)}{P(F1 \ldots . Fn)} \quad (2)$$

A. Pada Variabel C menunjukkan nilai kelas, sedangkan variabel F1 ... Fn menunjukkan setiap karakteristik arahan yang dibutuhkan untuk melakukan klasifikasi. Rumus tersebut memberikan penjelasan bahwa peluang tiap variabel sampel karakteristik pada sebuah kelas C (Posterior) merupakan peluang yang timbul dari nilai kelas C (sebelum nilai diinputkan, bisa diartikan dengan prior), kemudian dikalikan dengan peluang pada sample yang berada dikelas C (dapat disebut dengan likelihood), setelah itu dibagi dengan peluang pada karakteristik sample secara global (dapat disebut dengan evidence). Rumus tersebut dapat didefinisikan sebagai berikut(Saleh and Nasari 2018a):

$$Posterior = \frac{prior\ x\ likelihood}{evidence} \quad (3)$$

Nilai *Evidence* akan selalu tetap pada setiap kelas dengan satu sample. Psetrior memiliki nilai yang nantinya dapat dibedakan dengan nilai dari setiap posterior kelas lainnya, untuk memastikan kelas mana yang dijadikan satu sample untuk diklasifikasikan. Pengertian lebih lanjut lagi dari rumus Bayes tersebut dijelaskan dengan pemaparan $(C|F1, \ldots, Fn)$ Penjelasannya dapat dilihat dari rumus berikut ini (Saleh 2015):

$$\begin{aligned}
&P(C|F_1, \ldots, F_n) = P(C)P(F_1, \ldots F_n|C) \\
&= P(C)\,P(F_1|C)P(F_2, \ldots Fn|C, F_1) \\
&= P(C)P(F_1|C)P(F_2|C, F_1)P(F_3, \ldots, F_n|C, F_1, F_2) \\
&= (C)P(F_1|C)P(F_2|C, F_1)P(F_3|C, F_1, F_2)P(F_4, \ldots F_n|C, F_1, F_2, F_3) \\
&= P(C)P(F_1|C)p(F_2|C, F_1)P(F_3|C, F_1, F_2) \ldots P(F_n|C, F_1, F_2, F_3, \ldots, F_{n-1})
\end{aligned} \quad (4)$$

Berdasarkan rumus diatas tersebut dapat ditarik kesimpulan bahwa semakin besarnya data dan semakin sulit atribut dapat berpengaruh terhadap hasil dari probabilitas, yang bisa saja permasalahan tersebut susah untuk dianalisis nilainya. Oleh sebab itu, perhitungan tersebut akan sulit dilakukan. Rumus ini memiliki sifat independensi naif, karena dapat terlihat dari F1,F2...Fn yang tidak saling terhubung atau independen. Dari pemaparan tersebut, maka munjul rumus sebagai berikut:

$$P(F_i|F_j) = \frac{P(F_i \cap F_j)}{P(F_j)} = \frac{P(F_i)P(F_j)}{P(F_j)} = P(F_i) \qquad (5)$$

Untuk i≠j , sehingga

$$P(F_i|C, F_j) = P(F_i|C) \qquad (6)$$

Persamaan tersebut merupakan penerapan metode Naive Bayes yang kemudian digunakan dalam melakukan klasifikasi. Klasifikasi tersebut memiliki data yang bersifat kontinyu digunakan rumus Densitas Gauss :

$$P(X_i = x_i|Y = y_j) = \frac{1}{\sqrt{2\pi\sigma_{ij}}} e^{-\frac{(x_i-\mu_{ij})^2}{2\sigma^2 ij}} \qquad (7)$$

Keterangan :

P : P yang berarti Peluang
Xi : Xi yang berarti Atribut dari nilai i
xi : Atribut dari nilai i
Y : Kelas data yang akan dihitung
yi : Sub kelas dari Y yang diinput
μ : mean, nilai rata – rata dari keseluruhan atribut
σ : Nilai deviasi standar, berdasarkan varian dari keseluruhan atribut

3. HASIL DAN PEMBAHASAN

Penerapan Metode Naive Bayes

Penerapan metode Naïve Bayes akan dilakukan dengan menilai dari setiap atribut dengan himpunan tersebut diganti dengan pembobotan yang berkarakter sehingga dapat dihitung dari nilai klasifikasi yang telah ditentukan untuk mendapatkan nilai range nya(Saleh and Nasari 2018a). Pada masing-masing atribut akan dinilai probabilitasnya masing – masing dengan data yang bersifat kontinyu untuk melakukan subjek dari hasil yang didapatkan (Saleh 2015). Pengklasifikasian data dengan metode Naïve Bayes membutuhkan data training untuk menghitung estimasi yang diperlukan(Saleh and Nasari 2018a).

Atribut Data

Pada masing – masing atribut data terlebih dahulu membuat ketentuan pembobotan dari astribut tersebut agar mudah dikenali. Adapun bobot dari masing – masing nilai tes tertulis, nilai tes praktek dan nilai wawancara dapat dilihat sebagai berikut :

Tabel 1. Pembobotan Atribut

Atribut *Class*	Sub Atribut *Class*	*Range* Nilai
Nilai Tes Tertulis	Sangat Bagus	> 80
	Bagus	=>70 s/d <=80
	Kurang Bagus	<70
Nilai Tes Praktek Sesuai Devisi	Sangat Bagus	> 80
	Bagus	=>70 s/d <=80
	Kurang Bagus	<70
Wawancara	Sangat Baik	> 80
	Baik	=>70 s/d <=80
	Kurang Baik	<70

Data Training

Pada proses pemilihan data training dapat dengan bentuk Excel. Dari data tersebut dapat diambil beberapa sample untuk dijadikan data training untuk dijadikan acuan dalam perhitungan dengan menggunakan metode Naïve Bayes. Data training yang digunakan berjumlah 47 sample data dan dapat dilihat pada tabel 2 berikut:

Tabel 2. Data Training

No	Nim	Nama Lengkap	Program Studi	Semester	Devisi	Status Pekerjaan	Nilai Tes Tertulis	Nilai Tes Praktek Sesuai Devisi	Wawancara	Hasil
1	1812140539	Richard Nainggolan	Teknik Informatika	5	Programming	Tidak Bekerja	Bagus	Bagus	Baik	Diterima
2	1813010255	Alisa Dwi Octavieni	Sistem Informasi	3	Knowledge of Technology	Tidak Bekerja	Sangat Bagus	Bagus	Baik	Diterima
3	1812010072	T.burha. fadillah	Teknik Informatika	3	Mutimedia	Tidak Bekerja	Sangat Bagus	Sangat Bagus	Baik	Diterima
4	1613010379	Siti Fatimah	Sistem Informasi	7	Programming	Bekerja	Bagus	Bagus	Kurang Baik	Tidak Diterima
5	1812010046	GIVAR ELVA RIZKY H	Teknik Informatika	3	Programming	Tidak Bekerja	Bagus	Sangat Bagus	Baik	Diterima
6	1912010014	Ismal Zikri	Teknik Informatika	1	Network Security	Tidak Bekerja	Sangat Bagus	Sangat Bagus	Baik	Diterima
7	1613010348	Rika Sukini	Sistem Informasi	7	Programming	Bekerja	Kurang Bagus	Bagus	Baik	Tidak Diterima
8	1712010149	Zulmi Fathur Rahman	Teknik Informatika	5	Programming	Tidak Bekerja	Kurang Bagus	Kurang Bagus	Kurang Baik	Tidak Diterima
9	1812010314	Indah sari lubis	Teknik Informatika	3	Network Security	Bekerja	Sangat Bagus	Sangat Bagus	Baik	Diterima
10	1812010325	Nurslama br siregar	Teknik Informatika	3	Network Security	Tidak Bekerja	Sangat Bagus	Sangat Bagus	Baik	Diterima
11	1812010299	Aliwi Aulia	Teknik Informatika	3	Programming	Tidak Bekerja	Bagus	Bagus	Sangat Baik	Diterima
12	1812010329	Restu Nofendi Gulo	Teknik Informatika	3	Programming	Tidak Bekerja	Bagus	Bagus	Baik	Diterima
13	1712010131	Joshua Saputra	Teknik Informatika	5	Network Security	Tidak Bekerja	Kurang Bagus	Kurang Bagus	Kurang Baik	Tidak Diterima
14	1713010060	M Arrien Ardian	Sistem Informasi	5	Knowledge of Technology	Tidak Bekerja	Bagus	Sangat Bagus	Baik	Diterima
15	1921010043	Wiyuda Bambang Ferdianto	Desain Komunikasi Visual	1	Mutimedia	Tidak Bekerja	Sangat Bagus	Sangat Bagus	Baik	Diterima
16	1913010263	Ardansyah	Sistem Informasi	1	Knowledge of Technology	Tidak Bekerja	Kurang Bagus	Kurang Bagus	Kurang Baik	Tidak Diterima
17	1712010025	Muhammad Habib	Teknik Informatika	5	Programming	Tidak Bekerja	Sangat Bagus	Sangat Bagus	Sangat Baik	Diterima
18	1921160085	Rika Aulia Hasibuan	Desain Komunikasi Visual	1	Mutimedia	Bekerja	Kurang Bagus	Kurang Bagus	Kurang Baik	Tidak Diterima
19	1812010509	Yedutun Bill Situmorang	Teknik Informatika	3	Network Security	Bekerja	Sangat Bagus	Sangat Bagus	Baik	Diterima
20	1812010038	Yudha Andhika Rahman	Teknik Informatika	3	Network Security	Bekerja	Sangat Bagus	Sangat Bagus	Baik	Diterima
21	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
47	1612010412	Immanuel Silaban	Teknik Informatika	7	Knowledge of Technology	Tidak Bekerja	Bagus	Kurang Bagus	Kurang Baik	Tidak Diterima

Atribut dan Probabilitas

Kriteria dan Probabilitas yang diterapkan harus dihitung dari masing – masing atribut dan kriteria. Data latih pada tabel 2 akan dilihat dari penjelasan pada setiap masing – masing atribut dan pada setiap atribut dapat dicari nilai setiap atribut berapa yang diterima dan berapa yang ditolak. Pada pencarian probabilitas tiap atribut digunakan perhitungan dengan Excel sehingga Hasil probabilitas tersebut dapat dilihat pada tabel 3 sebagai berikut :

Tabel 3. Nilai Probabilitas dari Tiap Atribut

Atribut	Hasil		Probabilitas	
Program Studi	Diterima	Ditolak	Diterima	Ditolak
Teknik Informatika	21	7	0,677	0,438
Sistem Informasi	9	5	0,290	0,313
Desain Komunikasi Visual	1	4	0,032	0,250
Jumlah	31	16	1	1
Semester	Diterima	Ditolak	Diterima	Ditolak
Semester 1	11	3	0,355	0,188
Semester 3	17	3	0,548	0,188
Semester 5	3	3	0,097	0,188
Semester 7	0	7	0,000	0,438
Jumlah	31	16	1	1
Devisi	Diterima	Ditolak	Diterima	Ditolak
Programming	13	4	0,419	0,250
Multimedia	8	5	0,258	0,313
Network Security	8	4	0,258	0,250
Knowledge Of Technology	2	3	0,065	0,188
Jumlah	31	16	1	1
Status Pekerjaan	Diterima	Ditolak	Diterima	Ditolak
Bekerja	5	6	0,161	0,375
Tidak Bekerja	26	10	0,839	0,625
Jumlah	31	16	1	1
Nilai Tes Tertulis	Diterima	Ditolak	Diterima	Ditolak
Sangat Bagus	26	0	0,839	0,000
Bagus	5	6	0,161	0,375
Kurang Bagus	0	10	0,000	0,625
Jumlah	31	16	1	1
Nilai Tes Praktek Sesuai Devisi	Diterima	Ditolak	Diterima	Ditolak
Sangat Bagus	26	1	0,839	0,063
Bagus	5	4	0,161	0,250
Kurang Bagus	0	11	0,000	0,688
Jumlah	31	16	1	1
Wawancara	Diterima	Ditolak	Diterima	Ditolak
Sangat Bagus	8	0	0,258	0,000
Bagus	23	4	0,742	0,250
Kurang Bagus	0	12	0,000	0,750
Jumlah	31	16	1	1

Probabilitas Atribut Program Studi

Pada tabel 2 dapat dilihat jumlah dari data training sebanyak 47 data, dari 47 data tersebut mimiliki atribut program studi dan memiliki sub atribut teknik informatika, sistem informasi, dan desain komunikasi visual dengan jumlah prodi Teknik informatika yang mendapatkan hasil diterima sebesar 21 orang dan 7 orang dan dihitung probabilitasnya dengan cara jumlah progam studi Teknik informatika yang diterima dibagi dengan jumlah keseluruhan data testing yang diterima sehingga memperoleh hasil nilai probabilitas yang diterima 0,677 dan yang ditolak 0,438. Pada sub atribut lainnya dapat dilihat pada tabel 3.

Probabilitas Atribut Semester

Berdasarkan tabel 3 dapat dilihat data training yang telah diuji nilai probabilitasnya dari 47 data yang ada dapat dilihat nilai diterima dari semester sebanyak 11 orang dan yang ditolak sebesar 3 orang, nilai dari semester yang diterima sebanyak 17 orang dan 3 orang ditolak, nilai dari semester 5 yang diterima sebesar 3 orang dan 3 orang ditolak, sedangkan untuk semester 7 tidak ada yang diterima karena semester 7 tidak dibenarkan untuk mengikuti UKM dikarenakan harus fokus Tugas Akhir maupun skripsi dan yang ditolak sebesar 7 orang.

Probabilitas Atribut Devisi

Pada atribut devisi dapat dilihat dari tabel 2 ada beberapa mahasiswa memilih devisi yang diinginkan akan tetapi mungkin saja belum sesuai dengan bidang keahliannya sehingga tidak berhasil diterima. Devisi programming memiliki 13 orang yang diterima dan 4 yang ditolak, devisi multimedia memiliki 8 orang yang diterima dan 5 ditolak, devisi network security memiliki 8 orang yang diterima dan 4 orang yang ditolak dan devisi knowledge of technology mimiliki 2 orang yang diterima dan 3 orang ditolak. Pada setiap atribut dapat dapat dihitung nilai probabilitasnya dengan nilai yang diterima pada tiap devisi dibagi nilai yang diterima dari jumlah keseluruhan data training yang dapat dilihat pada tabel 3.

Probabilitas Atribut Kerja

Pada tabel 2 data training menunjukkan atribut pekerjaan dapat berpengaruh terhadap nilai klasifikasi dari atribut pekerjaan dapat mempengaruhi nilai seperti nilai probabilitas mahasiswa yang bekerja ada 5 orang yang diterima dan 6 orang ditolak dengan nilai probabilitas yang diterima 0,161 dan ditolak 0,375, untuk yang tidak bekerja yang diterima 26 orang dan yang tidak bekerja 10 orang dengan nilai probabilitas yang diterima sebesar 0,839 dan 0,625 dapat dengan jelas pada tabel 3.

Probabilitas Atribut Nilai Tes Tertulis

Pada setiap mahasiswa yang akan mendaftar menjadi anggota sebelumnya harus ada tes terlebih dahulu yang salah satunya tes tertulis yang berisi pilihan berganda untuk mengetahui tingkat pengetahuan secara teoritis terlebih dahulu. Pada tabel 3 dapat dilihat jumlah nilai sangat bagus yang diterima sebanyak 26 orang dan nilai sangat bagus tidak mungkin ditolak sehingga memiliki nilai, untuk nilai bagus yang diterima sebanyak 5 orang dan 4 orang yang ditolak, dan nilai yang kurang bagus otomatis langsung ditolak sehingga tidak memiliki nilai yang diterima dan nilai yang ditolak sebanyak 11 orang. Nilai dari masing masing probabilitas dapat dilihat di tabel 3.

Probabilitas Atribut Nilai Tes Praktek Sesuai Devisi

Pada masing – masing devisi memiliki standar dan level untuk menguji pengetahuan calon anggota dengan menyiapkan tes secara praktek sesuai tingkat dari devisi tersebut yang hasilnya kemudian di diskusikan dan diambil kesimpulannya dengan nilai. Pada tabel 3 dapat dilihat nilai yang sangat bagus yang diterima sebanyak 26 orang dan nilai yang sangat bagus 1 orang, nilai yang bagus memiliki 5 orang yang diterima dan 4 orang yang ditolak, untuk nilai yang kurang bagus tidak akan diterima dengan jumlah mahasiswa yang diterima 0 dan yang ditolak 11 orang.

Probabilitas Atribut Wawancara

Pada tabel 3 dapat dilihat nilai dari masing – masing probabilitas, dengan adanya atribut wawancara bertujuan untuk mengetahui calon anggota yang akan masuk memiliki wawasan pengetahuan yang cukup luas atau tidak sehingga dapat menentukan nilai dari tes wawancara tersebut. Dari nilai yang sangat bagus diperoleh 8 orang yang diterima dan tidak ada ada yang ditolak, nilai bagus diperoleh 23 orang yang diterima dan 4 orang yang ditolak, untuk nilai yang kurang bagus berarti tidak memiliki pengetahuan yang cukup luas sehingga tidak ada yang diterima dan jumlah yang ditolak sebesar 12 orang.

Perhitungan Manual Metode Naïve Bayes dan C.45

Dari perhitungan nilai probabilitas diatas maka aka nada perhitungan untuk tiap masing masing kriteria (C) dengan tujuan untuk mengetahui masing – masing nilai dari kriteria tersebut. Perhitungan manual ini dilakukan dengan dengan menggunakan *Microsoft Excel* sehingga perhitungan tersebut akan dilakukan dengan cepat. Adapun data yang akan diuji sebanyak 47 data yang diambil dari data training tersebut.

Perhitungan Manual Metode Naïve Bayes

Perhitungan Manual Metode Naïve Bayes dapat dijelaskan pada tabel 3.

Tabel 3. Penerapan Perhitungan Metode Naïve Bayes

	Klasifikasi		P(X|Hasil = "Diterima"	P(X|Hasil = "Ditolak"
Xc1=	Program Studi	= Teknik Informatik	0,677	0,438
	Semester	= 5	0,097	0,188
	Devisi	= Programming	0,419	0,250
	Status Pekerjaan	= Tidak Bekerja	0,839	0,625
	Nilai Tes Tertulis	= Bagus	0,839	0,375
	Devisi	= Bagus	0,161	0,250
	Wawancara	= Baik	0,742	0,250
	#P(X|Ci) : P(X|panggil)		0,002	0,000
	#P(X|Ci)*P(Ci) = P(X|panggil) * P(panggil)		0,002	0,000

	Klasifikasi		P(X|Hasil = "Diterima"	P(X|Hasil = "Ditolak"
Xc3=	Program Studi	Informatika	0,677	0,438
	Semester	= 3	0,548	0,188
	Devisi	= Multimedia	0,258	0,313
	Status Pekerjaan	= Tidak Bekerja	0,839	0,625
	Nilai Tes Tertulis	= Sangat Bagus	0,839	0,000
	Sesuai Devisi	= Sangat Bagus	0,839	0,063
	Wawancara	= Baik	0,742	0,250
	#P(X|Ci) : P(X|panggil)		0,042	0,000
	#P(X|Ci)*P(Ci) = P(X|panggil) *		0,028	0,000

	Klasifikasi		P(X|Hasil = "Diterima"	P(X|Hasil = "Ditolak"
Xc2=	Program Studi	= Sistem Informasi	0,290	0,313
	Semester	= 3	0,548	0,188
	Devisi	Technology	0,065	0,188
	Status Pekerjaan	= Tidak Bekerja	0,839	0,625
	Nilai Tes Tertulis	= Sangat Bagus	0,839	0,625
	Devisi	= Bagus	0,839	0,250
	Wawancara	= Baik	0,742	0,250
	#P(X|Ci) : P(X|panggil)		0,004	0,000
	#P(X|Ci)*P(Ci) = P(X|panggil) * P(panggil)		0,003	0,000

	Klasifikasi		P(X|Hasil = "Diterima"	P(X|Hasil = "Ditolak"
Xc4=	Program Studi	Informasi	0,290	0,313
	Semester	= 7	0,000	0,438
	Devisi	= Programming	0,419	0,250
	Status Pekerjaan	= Bekerja	0,161	0,375
	Nilai Tes Tertulis	= Bagus	0,161	0,375
	Sesuai Devisi	= Bagus	0,161	0,250
	Wawancara	= Kurang Baik	0,000	0,750
	#P(X|Ci) : P(X|panggil)		0,000	0,001
	#P(X|Ci)*P(Ci) = P(X|panggil) *		0,000	0,0003

	Klasifikasi		P(X|Hasil = "Diterima"	P(X|Hasil = "Ditolak"
Xc5=	Program Studi	= Teknik Informatik	0,677	0,438
	Semester	= 3	0,548	0,188
	Devisi	= Programming	0,419	0,250
	Status Pekerjaan	= Tidak Bekerja	0,839	0,625
	Nilai Tes Tertulis	= Bagus	0,161	0,375
	Devisi	=Sangat Bagus	0,839	0,063
	Wawancara	= Baik	0,742	0,250
	#P(X|Ci) : P(X|panggil)		0,013	0,000
	#P(X|Ci)*P(Ci) = P(X|panggil) * P(panggil)		0,009	0,000

	Klasifikasi		P(X|Hasil = "Diterima"	P(X|Hasil = "Ditolak"
Xc7=	Program Studi	Informasi	0,290	0,313
	Semester	= 7	0,000	0,438
	Devisi	= Programming	0,419	0,250
	Status Pekerjaan	= Bekerja	0,161	0,375
	Nilai Tes Tertulis	= Kurang Bagus	0,000	0,625
	Sesuai Devisi	= Bagus	0,161	0,250
	Wawancara	= Baik	0,742	0,250
	#P(X|Ci) : P(X|panggil)		0,000	0,001
	#P(X|Ci)*P(Ci) = P(X|panggil) *		0,000	0,0002

Xc6=	Klasifikasi		P(X|Hasil = "Diterima"	P(X|Hasil = "Ditolak"
	Program Studi	= Teknik Informatik	0,677	0,438
	Semester	= 1	0,355	0,188
	Devisi	= Network Security	0,258	0,250
	Status Pekerjaan	= Tidak Bekerja	0,839	0,625
	Nilai Tes Tertulis	=Sangat Bagus	0,839	0,000
	Devisi	=Sangat Bagus	0,839	0,063
	Wawancara	= Baik	0,742	0,250
	#P(X|Ci) : P(X|panggil)		0,027	0,000
	#P(X|Ci)*P(Ci) = P(X|panggil) * P(panggil)		0,018	0,000

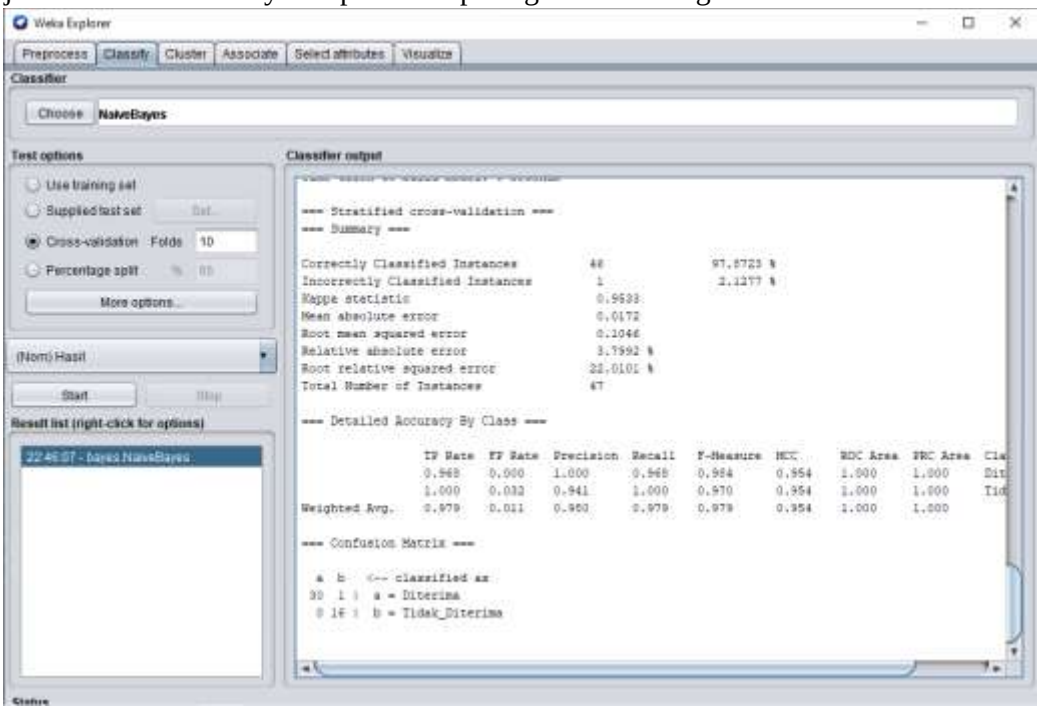
Xc8=	Klasifikasi		P(X|Hasil = "Diterima"	P(X|Hasil = "Ditolak"
	Program Studi	Informatika	0,677	0,438
	Semester	= 5	0,097	0,188
	Devisi	= Programming	0,419	0,250
	Status Pekerjaan	= Tidak Bekerja	0,839	0,625
	Nilai Tes Tertulis	= Kurang Bagus	0,000	0,625
	Sesuai Devisi	= Kurang Bagus	0,000	0,688
	Wawancara	= Kurang Baik	0,000	0,750
	#P(X|Ci) : P(X|panggil)		0,000	0,004
	#P(X|Ci)*P(Ci) = P(X|panggil) *		0,000	0,0014

Xc9=	Klasifikasi		P(X|Hasil = "Diterima"	P(X|Hasil = "Ditolak"
	Program Studi	Informatika	0,677	0,438
	Semester	= 3	0,548	0,188
	Devisi	= Network Security	0,258	0,250
	Status Pekerjaan	= Bekerja	0,161	0,375
	Nilai Tes Tertulis	= Sangat Bagus	0,839	0,000
	Devisi	= Sangat Bagus	0,839	0,063
	Wawancara	= Baik	0,742	0,250
	#P(X|Ci) : P(X|panggil)		0,008	0,000
	#P(X|Ci)*P(Ci) = P(X|panggil) * P(panggil)		0,005	0,0000

Xc10=	Klasifikasi		P(X|Hasil = "Diterima"	P(X|Hasil = "Ditolak"
	Program Studi	Informatika	0,677	0,438
	Semester	= 3	0,548	0,188
	Devisi	Security	0,258	0,250
	Status Pekerjaan	= Tidak Bekerja	0,839	0,625
	Nilai Tes Tertulis	= Sangat Bagus	0,839	0,000
	Sesuai Devisi	= Sangat Bagus	0,839	0,063
	Wawancara	= Baik	0,742	0,250
	#P(X|Ci) : P(X|panggil)		0,042	0,000
	#P(X|Ci)*P(Ci) = P(X|panggil) *		0,028	0,0000

Pengujian Metode Naïve Bayes

Pada perhitungan Metode Naïve Bayes secara manual diatas. Pada tahap proses pembuatan akurasi dan ketepatan sasaran akan diuji dengan menggunakan *tools* Weka 3.8. Hasil implementasi pengujian metode Naïve Bayes dapat dilihat pada gambar 1 sebagai berikut:



Gambar 1. Hasil Klasifikasi Metode Naive Bayes

Gambar 1 di atas dapat diketahui hasil Correctly Classified Instances adalah sebesar 97,8723% dan hasil untuk Incorrectly Classified Instances adalah sebesar 2,1277%. Data tersebut diuji dengan data sebanyak 47 data calon anggota UKM UPU-CC, ada sebanyak 46 data yang diterima dan berhasil diklasifikasikan dengan benar dan sebanyak 1 data ditolak dan tidak berhasil diklasifikasikan dengan baik. Dari kesalahan tersebut dapat dilihat pada karakter yang dberada pada nomor urut 17 yang tidak berhasil di identifikasi dengan benar.

4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan pada penerapan metode naïve bayes untuk penerimaan anggota UKM UPU-CC di Universitas Potensi Utama, dapat ditarik kesimpulan sebagai berikut:

1. Metode Naïve Bayes dan C4.5 berhasil diterapkan dengan hasil yang cukup maksimal untuk pennerpan pada permasalahan penerimaan anggota dengan nilai dari atribut program studi, semester, devisi yang dipilih, status pekerjaan, nilai tes tertulis, nilai tes praktek sesuai devisi yang dipilih dan nilai wawancara.
2. Berdasarkan data training sebanyak 47 data dan dilakukan perhitungan dengan aplikasi Weka dapat dilihat nilai akurasi sebesar 97,8723%

5. SARAN

Peneliti menyarankan untuk dapat mengembangkan lagi dengan metode lain atau dapat diuji dengan membandingkan Naïve Bayes dengan metode lain. Pada nilai atribut dan karakternya dapat ditambah dan dengan jumlah yang data yang lebih besar lagi agar terlihat nilai akurasi yang lebih baik dari yang dilakukan peneliti sebelumnya.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Universitas Potensi Utama yang telah memberi dukungan financial terhadap penelitian ini.

DAFTAR PUSTAKA

Berliana, Maulida, Mochzen Gito, and M Agus. 2018. "Analisis Pengaruh Penerimaan Grup Whatsapp Sebagai Sarana Komunikasi Dan Pembelajaran Antara Mahasiswa Dan Dosen Di Stt Wastukancana Purwakarta." 5(4): 455–64.

Devita, Riri Nada, Heru Wahyu Herwanto, and Aji Prasetya Wibawa. 2018. "Perbandingan Kinerja Metode Naive Bayes Dan K-Nearest Neighbor Untuk Klasifikasi Artikel Berbahasa Indonesia." *Jurnal Teknologi Informasi dan Ilmu Komputer* 5(4): 427.

Iskandar, Derick, and Yoyon K Suprapto. 2013. "Perbandingan Akurasi Klasifikasi Tingkat Kemiskinan Antara Algoritma C4 . 5 Dan Naïve Bayes Clasifier." *JAVA Journal of Electrical and Electronics Engineering* 11(1): 14–17.

Mas'udia, Putri Elfa. 2015. "Klasifikasi Tugas Akhir Untuk Menentukan Dosen Pembimbing Menggunakan Naïve Bayes Classifier (Nbc)." *Prosding SENTIA Politeknik Negeri Malang* 7: 135–40.

Muhamad, Husin et al. 2017. "Optimasi Naïve Bayes Classifier Dengan Menggunakan Particle Swarm Optimization Pada Data Iris." *Jurnal Teknologi Informasi dan Ilmu Komputer* 4(3): 180.

Nugroho, Mohamad Fajarianditya, and Setyoningsih Wibowo. 2017. "Fitur Seleksi Forward Selection Untuk Menetukan Atribut Yang Berpengaruh Pada Klasifikasi Kelulusan Mahasiswa Fakultas Ilmu Komputer UNAKI Semarang Menggunakan Algoritma Naive Bayes." *Jurnal Informatika Upgris* 3(1): 63–70.

Patmi Kasih, Intan Nur Farida. 2017. "Sistem Bantu Pemilihan Dosen Pembimbing Tugas Akhir Berdasarkan Kategori Pilihan Dan Keahlian Dosen Menggunakan Naïve Bayes." *Seminar Nasional Teknologi Informasi, Komunikasi dan Aplikasinya* 04(SNATIKA): 62–68.

Ridwan, Mujib, Hadi Suyono, and M Sarosa. 2013. "Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier." *Eeccis* 7(1): 59–64.

Rozi, Imam Fahrur, Elok Nur Hamdana, and Muhammad Balya Iqbal Alfahmi. 2018. "PENGEMBANGAN APLIKASI ANALISIS SENTIMEN TWITTER MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER (Studi Kasus SAMSAT Kota Malang)." *Jurnal Informatika Polinema* 4(2): 149.

Saleh, Alfa. 2015. "Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga." *Creative Information Technology Journal* 2(3): 207–17.

Saleh, Alfa, and Fina Nasari. 2018a. "Penerapan Equal-Width Interval Discretization Dalam Metode Naive Bayes Untuk Meningkatkan Akurasi Prediksi Pemilihan Jurusan Siswa." *Masyarakat Telematika Dan Informasi : Jurnal Penelitian Teknologi Informasi dan Komunikasi* 9(1): 1.

———. 2018b. "Penggunaan Teknik Unsupervised Discretization Pada Metode Naive Bayes Dalam Menentukan Jurusan Siswa Madrasah Aliyah." *Jurnal* Teknologi *Informasi dan Ilmu Komputer* 5(3): 353.

Septiani, Wisti Dwi. 2017. "Komparasi Metode Klasifikasi Data Mining Algoritma C4.5 Dan Naive Bayes Untuk Prediksi Penyakit Hepatitis." *None* 13(1): 76–84.

Sumpeno, Surya. 2009. "Klasifikasi Emosi Untuk Teks Bahasa Indonesia." *Seminar* (c).

Suryadi, Andri, and Erwin Harahap. 2018. "Sistem Rekomendasi Penerimaan Mahasiswa Baru Menggunakan Naive Bayes Classifier Di Institut Pendidikan Indonesia." *Joutica* 3(2): 171.

Syarli, S., and A. Muin. 2016. "Metode Naive Bayes Untuk Prediksi Kelulusan (Studi Kasus: Data Mahasiswa Baru Perguruan Tinggi)." *Jurnal Ilmiah Ilmu Komputer* 2(1): 22–26.