PENERAPAN *ALGORITMA NAIVE BAYES* UNTUK SELEKSI BEASISWA PEMPROV DAN BAWAKU DI UNIVERSITAS KOMPUTER INDONESIA

APPLICATION OF NAIVE BAYES ALGORITHM FOR SELECTION OF PROVIDING AND BAWAKU SCHOLARSHIP IN UNIVERSITY COMPUTER INDONESIA

Andi Carol Junior[1], Imelda Pangaribuan[2]
[1] Email : Andi.Carol13@gmail.com
[2] Email : Imelda@email.unikom.ac.id

Abstrak – UNIKOM **(**Universitas Komputer Indonesia) merupakan Universitas Program Sarjana yang menggelar program beasiswa pemerintahan daerah Jawa Barat (Pemprov Jabar) setiap tahunnya bagi mahasiswa berprestasi dan beasiswa Bantuan Wali Kota Bandung (BAWAKU) setiap tahunnya bagi mahasiswa kurang mampu. Banyaknya peserta yang mendaftar membuat staff kemahasiswaan atau Admin FO keteteran dan kebingungan dalam menyalurkan dan memberikan beasiswa kepada peserta. Tidak hanya total beasiswa yang banyak, namun jumlah peserta beasiswa yang mengikuti seleksi beasiswa pemprov dan bawaku juga banyak. Bahkan mahasiswa pernah mendapatkan dua beasiswa secara bersamaan, biasanya disebabkan karena kekeliruan Admin FO atau peserta beasiswa tidak melaporkan telah mendapat beasiswa. Berdasarkan permasalahan yang telah disepakati sebelumnya maka dibutuhkan pendukung keputusan yang dapat membantu Admin FO dalam menentukan penerima beasiswa. Penerapan Algoritma naïve bayes diharapkan berhasil memprediksi persentase probabilitas seleksi beasiswa. Metode tersebut mengklasifikasikan data mahasiswa berdasarkan beberapa kriteria yang disepakati. Dari 201 data mahasiswa yang mendaftarkan beasiswa Pemprov Jabar yang diuji dengan Algoritma naïve bayes, maka menghasilkan persentase 100% untuk prediksinya, dan dari 144 data mahasiswa yang mendaftarkan beasiswa Bawaku yang diuji dengan Algoritma naïve bayes, maka menghasilkan persentase 100%, dimana data beasiswa yang digunakan dari periode 2016, 2017 dan 2018 yang berhasil diklasifikasikan dengan benar.

Kata kunci *:* UNIKOM, Bawaku, sistem pendukung keputusan, Beasiswa, naïve bayes

Abstract – UNIKOM *(University Computer Indonesian) is a University's of the Undergraduate Program which holds a West Java regional government scholarship program (West Java Provincial Government) annually for outstanding students and the Mayor of Bandung Mayor (BAWAKU) scholarship every year for underprivileged students. The large number of participants who register makes student staff or Admin FO keteteran and confusion in channeling and providing scholarships to participants. Not only are there a large number of scholarships, but the number of scholarship participants participating in the provincial and my scholarship selection is also large. Even students have received two scholarships at the same time, usually because of the Admin FO mistake or scholarship participants did not report having received a scholarship. Based on the previously agreed issues, a decision support is needed that can assist the FO Admin in determining the scholarship recipients. The implementation of the Naïve Bayes algorithm is expected to be successful in predicting the percentage probability of scholarship selection. The method classifies student data based on several agreed criteria. From 201 data of students who register West Java Provincial Government scholarships tested using the Naïve Bayes Algorithm, it yields a percentage of 100% for his predictions, and from 144 data of students who register a Bawaku scholarship that is tested with the Naïve Bayes Algorithm, then the percentage yields 100%, wherein the scholarship data used from the period 2016, 2017 and 2018 which were successfully classified correctly.*.

Keyword : *UNIKOM, Bawaku, decision support system, scholarship, naïve bayes*

I. PENDAHULUAN

UNIKOM (Universitas Komputer Indonesia) merupakan Universitas Program Sarjana yang menggelar program beasiswa pemerintahan daerah Jawa Barat (Pemprov Jabar) setiap tahunnya bagi mahasiswa berprestasi dan beasiswa Bantuan Wali Kota Bandung (BAWAKU) setiap tahunnya bagi mahasiswa kurang mampu. Untuk memperoleh beasiswa ini harus disesuaikan dengan kriteria yang telah di buat oleh pihak (UNIKOM), Dengan beberapa cakupan yang dapat mengajukan beasiswa yaitu mahasiswa onGoing. Kriteria mahasiswa onGoing pemprov yang ditetapkan. Surat aktif mahasiswa, fomulir beasiswa, Tempat tinggal/ Regional, surat keterangan berprestasi, IPK >= 3.0, Kriteria mahasiswa onGoing BAWAKU yang ditetapkan. Surat aktif mahasiswa, fomulir beasiswa, Tempat tinggal/ Regional, surat keterangan SKTM (surat keterangan tidak mampu), IPK >= 3.00.

Objek penelitian ini mengambil data sample Beasiswa Pemprov Jabar dari priode 2016, 2017 dan 2018 sebagai acuan perhitungan Naïve Bayes dimana data beasiswa yang diperoleh dari keseluruhan periode 2016, 2017 dan 2018 sebanyak 451 peserta beasiswa pemprov. Untuk lebih mendatailnya data dari setiap periode telah diperoleh di bagian FO UNIKOM, dimana priode 2016 terdapat 142 peserta dan yang dinyatakan lulus sebanyak 17 peserta, priode 2017 terdapat 159 peserta dan yang dinyatakan lulus sebanyak 34 peserta, periode 2018 terdapat 150 peserta dan yang dinyatakan lulus sebanyak 25 orang. Dalam hal ini sudah terlihat bahwasannya setiap periode memiliki kuota penerimaan beasiswa yang berbeda karena dalam Pengajuan proposal dari pihak UNIKOM ke PEMPROV JABAR yang sudah di sahkan, pihak PEMPROV lah yang telah menentukan seberapa banyak kuota yang bisa diberikan sebagai Beasiswa.

Adapun untuk Beasiswa WABAKU dari priode 2016, 2017 dan 2018 sebagai acuan perhitungan Naïve Bayes dimana data beasiswa yang diperoleh dari keseluruhan periode 2016, 2017 dan 2018 sebanyak 320 peserta beasiswa pemprov. Untuk lebih mendatailnya data dari setiap periode telah diperoleh di bagian FO UNIKOM, dimana priode 2016 terdapat 108 peserta dan yang dinyatakan lulus sebanyak 20 peserta, priode 2017 terdapat 101 peserta dan yang dinyatakan lulus sebanyak 13 peserta, periode 2018 terdapat 111 peserta dan yang dinyatakan lulus sebanyak 23 orang. Dalam hal ini sudah terlihat bahwasannya setiap periode memiliki kuota penerimaan beasiswa yang berbeda karena dalam Pengajuan proposal dari pihak UNIKOM ke PEMPROV JABAR/WABAKU yang sudah di sahkan, pihak PEMPROV /WABAKU lah yang telah menentukan seberapa banyak kuota yang bisa diberikan sebagai Beasiswa.

Dari persoalan uraian dilatar belakang, dapat ditarik permasalahan yang dapat diidentifikasi sebagai berikut : (1) Banyaknya peserta yang mendaftar membuat staff kemahasiswaan atau Admin FO keteteran dan kebingungan dalam menyalurkan dan memberikan beasiswa kepada peserta. (2) Ada juga mahasiswa mendapatkan dua beasiswa sekaligus, bisa disebabkan karena kekeliruan Admin FO atau peserta beasiswa tidak melaporkan bahwa telah mendapat beasiswa.

Adapun penelitian terdahulu yang banyak menggunakan metode ini dibidang pendidikan dalam pengambilan keputusan pemberian beasiswa Sisilia Daeng Bakka Mau .2014 berjudul "Sistem Pendukung Keputusan Seleksi Beasiswa Menggunakan Teorema Bayes dan Dempster-Shaferc" [1]. Hasil dari penelitian ini adalah aplikasi pemberian beasiswa berhasil dikembangkan, memiliki persamaan dalam hasil perangkingan proses penyeleksian beasiswa menggunakan dua metode. Hasil metode terbaik adalah Teorema Bayes dan hasil perbandingan Probabilitas. Diasrina Dahri, Fahrul Agus, Dyna Maurisa Khairina. 2016 "Metode Naïve Bayes untuk Penentuan Penerimaan Beasiswa Bidikmis Universitas Mulawarman" [2]. Hasil dari penelitian ini berupa aplikasi Seleksi Penerima beasiswa Bidikmisi Universitas Mulawarman dengan mengunakan metode naïve bayes. Objek penelitian ini berhasil membantu menyeleksi penerimaan beasiswa bidikmis, dengan tinkat keakuratan sebesar 85,56% [2].

Tujuan dari Objek penelitian ini adalah: (1) Metode ini diharapkan sanggup menanganin masalah dalam menentukan seleksi beasiswa yang paling sesuai dilingkungan Unikom.(2) Objek penelitian ini untuk mengimplementasikan Algoritma aive bayes.

II. KAJIAN PUSTAKA

A. Data mining

Definisi dari berbagai sumber dari data mining adalah proses pengambilan informasi dari berbagai data-data yang sangat signifikan atau besar dengan pemanfaatan algotima dan teknik pengumpulan data dalam bidang statistic [2]. Data mining merupakan proses menganalisa data dari perspektif yang berbeda dan mengolahnya menjadi informasi – informasi penting yang dapat dipakai untuk meningkatkan keuntungan, memperkecil biaya pengeluaran, ataupun keduanya [3]. Definisi Data mining merupakan set kegiatan dalam pengumpulan, pemakaian data historical untuk menghasilkan pola atau hubungan dalam data berukuran besar atau banyak [3]. Dari beberapa definisi yang ada dapat disimpulkan Data mining merupakan proses set kegiatan untuk mengumpulkan data yang berukuran besar atau banyak yang nantinya di ekstrasi menjadi informasi –informasi penting yang dapat digunakan [3].

B. Tahapan–tahapan Data mining

Tahap – tahap data mining [6]: 1. data cleaning proses menghapus noise dan data yang tidak beraturan [6]. 2. data integration proses penggabungan data dari sebagian sumber atau database kedalam database baru [6]. 3. data selection data yang ada pada database tidak semuanya terpakai, oleh karena itu hanya beberapa data saja yang sesuai untuk dianalisis [6]. 4. data transformation proses penggabungan data kedalam format yang sesuai untuk diproses di data mining [6]. 5. proses mining suatu proses utama saat metode diterapkan untuk menentukan pengetahuan berharga dan tersembunyi dari data [6]. 6 pattern evaluation untuk mengidentifikasi beberapa pola – pola menarik kedalam knowladge base yang ditemukan [6]. 7. knowledge presentation visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan.

C. Metode Naive Bayes

Definisi naïve bayes merupakan metode probabilistic pengklasifikasian sederhana berdasarkan Teorema Bayes dimana pengklasifikasian dilakukan melalui data training sejumlah data secara efisien [4]. Adapun Naïve bayes mengasumsikan bahwa nilai dari sebuah input atribut pada kelas yang diberikan tidak tergantung dengan nilai atribut yang lain [5]. Teorema Bayes sendiri dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai Teorema Bayes [4].

D. Persamaan Metode Data Mining

Adapun rumus atau persamaan teorema naïve bayes [4]:

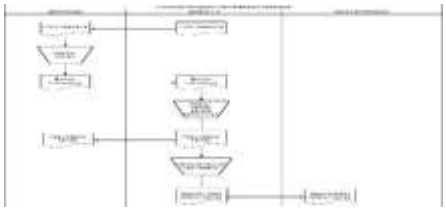
$$P(H|X) = \frac{P(X|H).P(H)}{P(X)} \quad (1)$$

Persamaan diatas merupakan rumus dasar metode naïve bayes [4].

III. METODE PENELITIAN

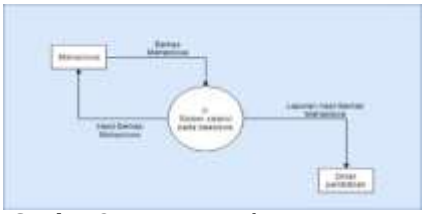
Objek penelitian bertepatan di UNIKOM (Universitas Komputer Indonesia) yang beralamat di Jln. Dipatiukur No. 112 – 116, kelurahan lebakgede, kecamatan coblong, kota bandung, jawa barat 40132.

Analisis procedure sistem yang sedang berjalan pada prosedure Seleksi beasiswa Pemprov Jabar dan Bawaku, memiliki tujuan untuk mengungkap penjelasan bagaimana cara kerja prosedure tersebut dan ketidak pastian yang dialami prosedure tersebut untuk dijadikan sistem yang baru agar terkomputerisasi, perancangan sistem yang berjalan di lakukan berdasarkan kejadian - kejadian yang ada dan dapat dibuat flowmap, prosedur sistem ini di deskripsikan sebagai berikut : (1) Mahasiswa memberikan berkas yang telah disiskan ke bagian Admin FO, (2) Admin Fo menyeleksi berkas dengan cara pembukuan (3) Kemudian hasil dari seleksi diberikan kembali ke Mahasiswa, dan Admin Fo membuat akn laporan hasil seleksi yang nantinya diberikan kepada Dinas pendidikan.



Gambar 1. Flow map Prosedur Seleksi Beasiswa

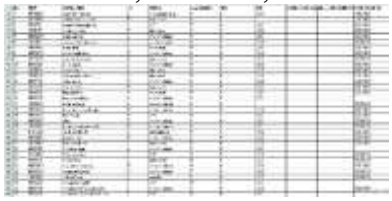
Pada tahap ini merupakan diagram konteks dari sistem seleksi beasiswa Pemprov Jabar dan BAWAKU yang sedang berjalan dimana entitas luarnya yaitu Mahasiswa dan Dinas Pendidikan. Sedangkan entitas dalamnya yaitu Admin Fo. Data yang menjadi masukan dalam sistem ini yaitu berkas mahasiswa. Sedangkan informasi yang dihasilkan oleh sistem ini adalah Hasil berkas mahasiswa dan Laporan hasil berkas mahasiswa.



Gambar 2. Diagram Konteks Beasiswa

IV. HASIL DAN PEMBAHASAN

Tahap ini akan menjelaskan bagaimana pemodelan penerapan algoritma naïve bayes pada data sample yang didapat, data yang menjadi sample adalah data beasiswa, tahap ini juga akan menjelaskan bagaimana proses penyeleksian dari data sample, pemodelan dan perhitungan pada data training yang menghasilkan peluang diterima atau tidaknya beasiswa pada setiap mahasiswa. Objek penelitian yang dilakukan pada bagian ekstrasi data beasiswa, data mahasiswa yang menerima beasiswa sesuai kebutuhan klasifikasi. Data yang dibutuh kan misalkan : NPM/NIM, NAMA_MHS, JK(jenis kelamin), PRODI, ID_JENJANG, SMT(semester), IPK, KODE_PEKERJAAN, JML_TANGGUNGAN, PENGHASILAN_ORTU, PRESTASI, KET_BEASISWA, ALAMAT, TELEPON.



Gambar 3. Data Beasiswa Pemprov

Pada Gambar 3. menunjukan hasil ekstrasi dimana ada 14 Atribute yang dipilih yang nantinya akan digunakan ke tahap selanjutnya. Data yang digunakan sebanyak 142 data sample dari periode 2016. Selanjutnya Pelebelan/kelas data merupakan tahap kedua, dimana data diberikan lebel, yaitu data mahasiswa tersebut menerima beasiswa alias "YA" atau tidak menerima beasiswa alias "TIDAK". Kemudian setelah melakukan Pelebelan, tahap selanjutnya adalah Preprocessing yang meliputi reduksi dan proses diskretisasi agar data yang telah di berikan Pelebelan layak untuk dimining, ini merupakan teknik klasifikasi dari data maining, variable – variable yang mengalami Preprocessing

diskretiasi adalah IPK, ALAMAT, PRESTASI, KETERANGAN, data – data ini mengalami diskretisasi yaitu dibuat dalam range pada skala tertentu.

Gambar 4. Data Beasiswa Bawaku

Pada Gambar 4. menunjukan hasil ekstrasi dimana ada 14 Atribute yang dipilih yang nantinya akan digunakan ke tahap selanjutnya. Data yang digunakan sebanyak 108 data sample dari periode 2016.Selanjutnya Pelebelan/kelas data merupakan tahap kedua, dimana data diberikan lebel, yaitu data mahasiswa tersebut menerima beasiswa alias "YA" atau tidak menerima beasiswa alias "TIDAK". Kemudian setelah melakukan Pelebelan, tahap selanjutnya adalah Preprocessing yang meliputi reduksi dan proses diskretisasi agar data yang telah di berikan Pelebelan layak untuk dimining, ini merupakan teknik klasifikasi dari data maining, variable – variable yang mengalami Preprocessing diskretiasi adalah IPK, ALAMAT, SKTM, KETERANGAN, data – data ini mengalamin diskretisasi yaitu dibuat dalam range pada skala tertentu.

Table 1. Hasil Pengujian Klasifikasi Beasiswa Pemprov Jabar

Data training/Sample	Akurasi Model %		Keakurasian
	YA	TIDAK	
25	0	25	100
70	0	70	100
142	17	125	100
201	76	125	100

Table 1. merupakan table akurasi untuk masing – masing pengujian dengan berbagai porsi data. Dari table tersebut dapat diketahui bahwa setiap pengujian memiliki akurasi paling tinggi dengan menggunakan data training sebanyak 201.

Table 2. Hasil Klasifikasi Beasiswa Bawaku

Data training/Sample	Akurasi Model %		Keakurasian
	YA	TIDAK	
25	0	25	100
75	0	75	100
108	20	88	100
144	56	88	100

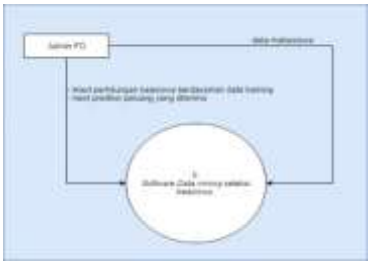
Table 2. merupakan table akurasi untuk masing – masing pengujian dengan berbagai porsi data. Dari table tersebut dapat diketahui bahwa setiap pengujian memiliki akurasi paling tinggi dengan menggunakan data training sebanyak 144. Perancangan kebutuhan perangkat keras tahap ini membahas tentang kebutuhan perangkat keras untuk dapat menjalankan software data mining ini, berikut adalah apa saja yang harus terpenuhi di perangkat keras untuk menjalankan software ini:

1. Prosesor Komputer : Intel CORE i5
2. Harddisk : 1TB
3. Memori RAM Komputer : 4GB
4. Layar Monitor : 14inch
5. Keyboard : Standard
6. Mouse : Optical

Perancangan kebutuhan perangkat lunak ini mendefinisikan tentang kebutuhan apa saja yang dapat menjalankan software data mining ini, berikut adalah perangkat lunak apa harus terpenuhi untuk menjalankan software ini:

1. Windows 10.
2. Java SE Development Kit 7.
3. Microsoft Accesses

Berikut adalah cara kerja software data mining secara garis besar dalam menerima data input , kemudian di proses , dan menjadi output sebagai hasil tampil dari data yang telah di proses.

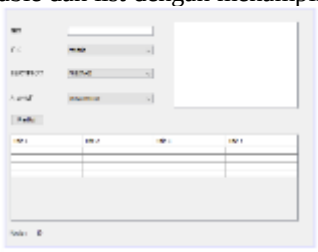


Gambar 5. Diagram Konteks

Pada diagram ini akan membahas tentang keterkaitan antara fungsi – fungsi software yang saling berkaitan satu sama lain dalam menerima data input, proses, dan menghasilkan output yang diinginkan.

Gambar 6. Data Flow Diagram Level 1

Di dalam alur input data mahasiswa di dalam software terdapat beberapa kondisi untuk penginputan data mahasiswa: atribut Nim hanya bisa diinput oleh angka, atribut Nim hanya bisa di input sebanyak 8 digit, atribut Nim tidak bisa di input denga huruf., Atribut IPK di input sesuai pilihan, Atribut SERTIFIKAT di input sesuai pilihan, Atribut ALAMAT di input sesuai pilihan, Untuk tombol prediksi, sesuai dengan hasil input dimana akan terhitung probabilitas dari setiap atribut, Hasil prediksi akan terlihat di table dan list dengan menampilkan probabilitas nya.



Gambar 7. Implementasi Antarmuka

V. KESIMPULAN DAN SARAN

Berdasarkan objek penelitian tentang seleksi beasiswa didapati sekumpulan kesimpulan sebagai berikut: 1. Naïve Bayes Classifier memiliki beberapa kelebihan yaitu, cepat dalam proses perhitungan, algoritma yang sederhana dan akurasi yang tinggi. Namun probabilitas pada Naïve Bayes Classifier tidak bisa mengukur seberapa besar tingkat keakuratan sebuah prediksi, selain itu metode naïve bayes juga memiliki kelemahan pada seleksi atribut yang banyak, yang mengakibatkan memakan banyak waktu. 2. Bahwa peserta beasiswa masih banyak mengalami kegagalan penyeleksian dikarenakan pengumpulan berkas yang kurang lengkap atau kriteria yang tidak terpenuhi seperti Regional, SKTM dan Sertifikat prestasi, mengakibatkan perhitungan probabilitasnya menjadi kecil dan hasil yang dikeluarkan adalah tidak diterima. 3. Adapun peluang diterima atau probabilitas yang dihasilkan tinggi pada seleksi ini dengan memenuhi kreteria yang telah di tentukan sebagai contoh IPK > 3.0, sesuai Regional yang di tentukan, serta kelengkapan dari SKTM (surat keterangan tidak mampu) dan Sertifikat berprestasi. 4. Dari data beasiswa Pemprov yang dijadikan data training, Algoritma naive bayes berhasil mengklasifikasikan 201 data yang telah diuji. Sehingga Algoritma naive bayes berhasil memprediksi besar keakuratan sebesar 100%. 5 Dari data beasiswa BAWAKU yang dijadikan data training, Algoritma naive bayes berhasil mengklasifikasikan 144 data yang diuji. Sehingga Algoritma naive bayes berhasil memprediksi besar persentase keakuratan sebesar 100%.

Dapat ditarik beberapa saran mengenai Objek penelitian tentang penerapan algoritma naive bayes untuk seleksi beasiswa pemprov dan bawaku di universitas komputer indonesia yaitu pengujian klasifikasi alangkah baiknya dilakukan dengan menggunakan metode klasifikasi yang lain seperti C4.5 atau pun Particle Swamp Optimization (PSO) untuk melihat hasil metode mana yang lebih akurat dalam memprediksi penyeleksian beasiswa. Diharapkan dengan penerapan Algoritma Naive Bayes dapat memperkecil kesalahan dalam penyeleksian sehingga dapat melakukan penyeleksian 2 beasiswa secara bersamaan.

DAFTAR PUSTAKA

Artikel jurnal:

[1] Mau, S. D. 2014. Sistem Pendukung Keputusan Seleksi Beasiswa Menggunakan Teorema Bayes dan Dempster-Shafer. Jurnal Pekommas, Vol. 17 No. 1, 23-32.

[2] Dahri. D., Agus. F., Khairina. D. M. 2016. Metode Naïve Bayes untuk Penentuan Penerimaan Beasiswa Bidikmis Universitas Mulawarman. Jurnal Informatika Mulawarman Vol. 11 No. 2, ISSN 1858-4853. September 2016.

[3] Karismariyanti. M. 2011. Simulasi Pendukung Keputusan Penerima Beasiswa Menggunakan Metode Composite Performance Index. Jurnal Teknologi Informasi Vol.1, No.2, November 2011.

[4] Hadiyani. Eka Pratiwi. 2013. "Rancang Bangun Sistem Pendukung Keputusan Untuk Pemilihan Anggota Terbaik AIESEC Surabaya Dengan Menggunakan Metode Naive Bayes". Fakultas Sains dan Teknologi. Universitas Airlangga.

[5] Tan, Pang-Ning., Steinbach, Michael., dan Kumar, Vipin. 2006. Introduction to Data Mining. Pearson Education. Boston

Buku:

[6] Fayyad. Usama. 1996. “Advances in Knowledge Discovery and Data Mining”, MIT Press.

Prosiding seminar:

[7] YH Putra, GGA Manu, Y Afrizal - Jurnal Tata Kelola dan Kerangka Kerja Teknolog Informasi, 2015, DOI: 10.13140/RG.2.2.10127.18088 · License: CC BY-ND 4.0.