

THE PREDICTION OF PPA AND KIP-KULIAH SCHOLARSHIP RECIPIENTS USING NAIVE BAYES ALGORITHM

Asri Mulyani*¹, Dede Kurniadi², Muhammad Rikza Nashrulloh³, Indri Tri Julianto⁴, Meta Regita⁵

^{1,2,5}Teknik Informatika, Jurusan Ilmu Komputer, Institut Teknologi Garut, Indonesia

^{3,4}Sistem Informasi, Jurusan Ilmu Komputer, Institut Teknologi Garut, Indonesia

Email: ¹asrimulyani@itg.ac.id, ²dede.kurniadi@itg.ac.id, ³rikza@itg.ac.id, ⁴indritrijulianto@itg.ac.id

(Naskah masuk: 26 April 2022, Revisi : 01 Mei 2022, diterbitkan: 20 Agustus 2022)

Abstract

The aim of the research is was to predict the scholar recipient for Peningkatan Prestasi Akademik (PPA) and the Kartu Indonesia Pintar Kuliah (KIP-K). The prediction results of scholarship recipients will provide information in the form of the possibility of acceptance and non-acceptance of scholarship applicants. To achieve this goal, this study uses the Naive Bayes algorithm, where this algorithm predicts future opportunities based on past data by going through the stages of reading training data, then calculating the number of probabilities and classifying the values in the mean and probability table. The data analysis includes data collection, data processing, model implementation, and evaluation. The data needed for analysis needs to use data from the applicants for Academic Achievement Improvement (PPA) scholarship and the Indonesia Smart Education Card (KIP-K) scholarship. The data used for training data were 145 student data. The results of the study using the Naive Bayes algorithm have an accuracy of 80% for PPA scholarships and 91% for KIP-K scholarships.

Keywords: KIP-Kuliah, Naive Bayes, PPA, Prediction, Scholarship.

PREDIKSI PENERIMA BEASISWA PPA DAN KIP-KULIAH MENGGUNAKAN ALGORITMA NAIVE BAYES

Abstrak

Tujuan dari penelitian adalah untuk memprediksi pendaftar Peningkatan Prestasi Akademik (PPA) dan Kartu Indonesia Pintar Kuliah (KIP-K). Hasil prediksi penerimaan beasiswa akan memberikan suatu informasi berupa kemungkinan diterima dan tidak diterimanya para pendaftar beasiswa. Untuk mencapai tujuan tersebut penelitian ini menggunakan algoritma *Naive Bayes* dimana algoritma ini memprediksi peluang masa depan berdasarkan data masa lalu dengan melalui tahapan membaca data latih, kemudian mengitung jumlah probabilitas serta mengklasifikasikan nilai dalam tabel *mean* dan probabilitas. Analisis data yang dilakukan meliputi pengumpulan data, pemrosesan data, implementasi model dan evaluasi. Data yang dibutuhkan untuk kebutuhan analisis menggunakan data pendaftar beasiswa Peningkatan Prestasi Akademik (PPA) dan beasiswa Kartu Indonesia Pintar Kuliah (KIP-K). Data yang digunakan untuk data *training* sebanyak 145 data Mahasiswa. Hasil penelitian menggunakan algoritma *Naive Bayes* memiliki akurasi 80% untuk beasiswa PPA dan 91% untuk beasiswa KIP-Kuliah.

Kata kunci: Beasiswa, KIP-kuliah, Naive Bayes, PPA, Prediksi.

1. PENDAHULUAN

Pendidikan merupakan salah satu potensi dalam mewujudkan cita-cita bangsa Indonesia yang tertuang dalam Pembukaan UUD 1945 alinea 4 “mencerdaskan kehidupan bangsa”. Di Indonesia, faktor ekonomi menjadi salah satu hambatan dalam menyelesaikan pendidikan tinggi. Maka dari itu, perlu adanya beasiswa yang disediakan oleh badan-badan seperti pemerintah, perusahaan, atau yayasan untuk membantu kelancaran dalam pelaksanaan Pendidikan Tinggi [1]. Undang-undang Nomor 20

Tahun 2003 tentang Sistem Pendidikan Nasional, Bab V pasal 12 (1.c), menyebutkan bahwa setiap peserta didik pada setiap satuan pendidikan berhak mendapatkan beasiswa bagi yang berprestasi yang orang tuanya tidak mampu membiayai pendidikannya [2]. Pemberian beasiswa ini diberikan sesuai dengan prasyarat yang diberikan oleh pihak pemberi beasiswa baik pemerintah, perusahaan atau yayasan [3]. Dengan adanya program beasiswa, diharapkan semangat mahasiswa untuk berprestasi dapat

meningkat, dengan begitu mahasiswa berhak mendapatkan penghargaan berupa beasiswa [4].

Sekolah Tinggi Teknologi Garut (STT-Garut) merupakan perguruan tinggi swasta yang ada di Kabupaten Garut, yang mana setiap tahunnya menyalurkan bantuan Beasiswa, baik Beasiswa yang diperoleh dari lembaga pemerintah, perusahaan, ataupun yayasan, dengan bentuk berupa bantuan biaya pendidikan bagi mahasiswa kurang mampu secara ekonomi tetapi memiliki potensi akademik. Namun jumlah alokasi Beasiswa yang disalurkan setiap tahunnya terbatas dan tidak sebanding dengan jumlah mahasiswa yang mengajukan Beasiswa, sehingga hal tersebut mengakibatkan harus adanya proses penyeleksian yang ketat oleh bagian pengelola beasiswa, agar penyaluran Beasiswa sesuai dan tepat sasaran. Berdasarkan hasil observasi data dari pengelola Beasiswa Peningkatan Prestasi Akademik (PPA) dan Kartu Indonesia Pintar (KIP-K) di STT-Garut, pendaftar beasiswa PPA dan KIP-K setiap tahunnya mengalami peningkatan, namun dengan proses penyeleksiannya masih manual dengan bantuan pengolahan aplikasi Microsoft Excel. Hal ini dapat menyebabkan proses penyeleksian manual kurang efektif dan tepat sasaran dalam segi mencapai tujuan maupun dalam segi waktu rekap penyeleksian.

Penelitian ini merujuk pada penelitian sebelumnya yang berkaitan dengan penerimaan beasiswa seperti yang dilakukan oleh Kurniadi, dkk dengan menggunakan Algoritma *K-Nearest Neighbor* dengan hasil akurasi 95,83%, penelitian ini memprediksi siswa yang memiliki peluang terbesar untuk mendapatkan beasiswa PPA [1]. Sedangkan penelitian berikutnya menerapkan Algoritma *Naive Bayes* untuk mengklasifikasikan penerima beasiswa PPA dan BBM (Beasiswa Bantuan Mahasiswa) dengan tingkat akurasi 83,05% untuk PPA dan 75% untuk BBM [5]. Disamping itu *Naive Bayes* juga digunakan dalam memprediksi hasil pembelajaran mata kuliah di masa depan dengan tingkat hasil akurasi 78.6% [6]. Sedangkan Penelitian lainnya menggunakan Algoritma C4.5 dengan tingkat akurasi 69,73%, penelitian ini membuat sistem pendukung keputusan penerimaan beasiswa [7]. Algoritma Fuzzy Sugeno juga pernah digunakan dalam penelitian untuk menghasilkan suatu sistem aplikasi berbasis web dengan memodelkan cara berfikirnya manusia dalam seleksi calon penerimaan beasiswa dengan tingkat akurasi 92,3% [8]. Kemudian penelitian berikutnya menggunakan algoritma neural network backpropagation dalam sistem pendukung keputusan untuk prediksi penerima beasiswa, penentuan nilai akhir mata kuliah dan prediksi kinerja mahasiswa dengan tingkat akurasi masing-masing 90%, 93,43% dan 80,95% [9]–[11].

Berdasarkan penelitian rujukan dan latar belakang yang telah diuraikan sebelumnya, maka tujuan penelitian ini adalah membahas khusus prediksi penerimaan beasiswa PPA dan beasiswa KIP-K di Sekolah Tinggi Teknologi Garut dengan

menggunakan Algoritma *Naive Bayes*. Pemilihan *Naive Bayes* karena menghasilkan tingkat akurasi cukup baik berdasarkan dari penelitian sebelumnya yang telah diuraikan, serta mampu mempelajari data kasus sebelumnya yang digunakan sebagai data uji [12]. Kontribusi dari penelitian ini diharapkan dapat membantu pengelola Beasiswa di STT-Garut dalam menentukan mahasiswa yang diterima dan ditolak khususnya untuk Beasiswa Jenis PPA dan KIP-K.

2. METODE PENELITIAN

Metode penelitian menggunakan pendekatan metode *Knowledge Discovery in Database (KDD)*. *KDD* merupakan metode untuk memperoleh pengetahuan dari *database* yang ada. Dalam *database* terdapat tabel-tabel yang saling berhubungan atau berelasi. Hasil pengetahuan yang diperoleh dalam proses tersebut dapat digunakan sebagai basis pengetahuan (*knowledge base*) untuk keperluan pengambilan keputusan [13]. Adapun untuk tahapan penelitian disajikan pada Gambar 1.



Gambar 1. Tahapan Penelitian

Pengumpulan dalam penelitian ini diambil langsung dari data Beasiswa Peningkatan Prestasi Akademik (PPA) dan Kartu Indonesia Pintar Kuliah (KIP-K) di STT-Garut. Data yang terkumpul untuk kebutuhan data *training* yaitu untuk program beasiswa PPA dengan jumlah 45 mahasiswa dan KIP-K dengan jumlah 100 mahasiswa.

Pemrosesan data merupakan proses memilih variabel apa saja yang akan digunakan. Variabel yang digunakan untuk program beasiswa PPA yaitu Indeks Prestasi Kumulatif (IPK), jumlah tanggungan dan penghasilan. Sedangkan untuk beasiswa Kartu Indonesia Pintar Kuliah (KIP-K) data yang digunakan terdiri dari tes tulis, tes wawancara, kondisi ekonomi, prestasi akademik, prestasi non akademik, hasil survei.

Tahap pemodelan menggunakan Algoritma *Naive Bayes* yaitu sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan [14], [15].

Tahap evaluasi, menggunakan *Confusion matrix*. *Confusion matrix* menurut Han dan Kamber yang dikutip dalam penelitian Fibrianda, dkk [16] dapat diartikan sebagai suatu alat yang memiliki fungsi untuk melakukan analisis apakah *classifier* tersebut baik dalam mengenali *tuple* dari kelas yang berbeda. Nilai dari *True-Positive* dan *True-Negative* memberikan informasi kepada *classifier* dalam melakukan klasifikasi data bernilai benar, sedangkan *False-Positive* dan *False-Negative* memberikan informasi ketika *classifier* salah dalam melakukan

klasifikasi data. Berikut gambar *Confusion Matrix* disajikan pada Gambar 2.

		Predicted class		Total
		yes	no	
Actual class	yes	TP	FN	P
	no	FP	TN	N
Total		P'	N'	P + N

Gambar 2. *Confusion Matrix*

3. HASIL DAN PEMBAHASAN

Bagian ini akan dibahas mengenai hasil penelitian yang telah dilakukan dengan penyajian bahasan sesuai tahapan yang telah diuraikan pada bagian 2 Metode penelitian dengan menerapkan Algoritma Naive Bayes.

3.1. Pengumpulan Data

Jumlah data keseluruhan yang digunakan sebagai data training dan *testing* adalah 145 mahasiswa yang terdiri dari data program beasiswa-Peningkatan Prestasi Akademik (PPA) dengan data sejumlah 45 Siswa dan data program beasiswa Kartu Indonesia Pintar Kuliah (KIP-K) dengan jumlah data 100 mahasiswa.

3.2. Pemrosesan Data

Data yang telah terkumpul, dilakukan pembersihan data kemudian data dipilih sesuai kebutuhan. Algoritma *Naive-Bayes* untuk beasiswa Peningkatan Prestasi Akademik (PPA) data yang digunakan terdiri dari IPK (Indeks Prestasi Kumulatif), jumlah tanggungan dan penghasilan. Sedangkan untuk beasiswa KIP-K data yang digunakan terdiri dari *test* tulis, *test* wawancara,

kondisi ekonomi, prestasi akademik, prestasi non akademik, dan hasil survei. Berikut data dan atribut yang digunakan untuk kebutuhan Beasiswa PPA disajikan pada Tabel 1.

Tabel 1. Data Hasil Pemrosesan Beasiswa PPA

Nama	IPK	Tang- gungan	Penghasi- lan	Status
Mhs 1	3.29	6	5.000.000	Ditolak
Mhs 2	3.50	5	2.500.000	Diterima
Mhs 3	3.32	4	2.061.200	Diterima
Mhs 4	3.48	4	5.576.000	Ditolak
Mhs 5	3.52	2	1.500.000	Diterima
Mhs 6	3.78	5	1.000.000	Diterima
Mhs 7	3.40	5	3.583.000	Diterima
Mhs 8	3.80	4	2.000.000	Diterima
Mhs 9	3.48	3	750.000	Diterima
Mhs 10	3.00	4	1.500.000	Diterima
.....
Mhs 45	3.57	5	5.168.407	Ditolak

Hasil pemrosesan data PPA, terkumpul sejumlah 45 data Mahasiswa dengan perbandingan 28 data Mahasiswa yang Diterima Beasiswa PPA dan 17 data Mahasiswa yang Ditolak Beasiswa PPA.

Sedangkan Data hasil pemrosesan beasiswa KIP-K terkumpul sejumlah 100 data Siswa dengan perbandingan 50 data Mahasiswa yang Diterima Beasiswa KIP-K dan 50 data Mahasiswa yang Ditolak Beasiswa KIP-K. Berikut Data Beasiswa KIP-K hasil pemrosesan beserta atribut yang akan digunakan seperti disajikan pada Tabel 2 berikut:

Tabel 2. Data Hasil Pemrosesan Beasiswa KIP-K

Nama Siswa	Tes Tulis	Wawancara	Kondisi Ekonomi	Prestasi Akademik	Prestasi Non Akademik	Hasil Survei	Status
Sis 1	73	84	Rp 200.000	60	60	79	Diterima
Sis 2	63	86	Rp 125.000	80	10	98	Diterima
Sis 3	53	84	Rp 160.000	20	10	80	Ditolak
Sis 4	36	84	Rp 300.000	10	10	85	Ditolak
Sis 5	45	85	Rp 230.000	10	10	75	Ditolak
Sis 6	63	86	Rp 333.333	10	10	75	Ditolak
Sis 7	55	87	Rp 80.000	100	10	97	Ditolak
Sis 8	92	95	Rp 166.666	100	0	91	Diterima
Sis 9	40	84	Rp 250.000	10	30	81	Ditolak
Sis 10	60	87	Rp 400.000	80	0	70	Ditolak
.....
Sis 100	90	78	Rp 250.000	60	30	87	Diterima

3.3. Pemodelan Algoritma

Pada tahapan ini, dilakukan implementasi perhitungan algoritma Naive Bayes yang merupakan salah satu metode data mining yang termasuk algoritma klasifikasi paling populer [12]. Teorema Bayes memiliki bentuk umum [17]:

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

Keterangan:

X = data dengan *class* yang belum diketahui

H = hipotesis data X merupakan suatu \neg -class spesifik

P(H|X) = Probabilistik-hipotesis H berdasarkan kondisi x

P(H) = Probabilitas-hipotesis H

P(X|H) = Probabilitas-X berdasarkan kondisi tersebut

P(X) = Probabilitas-dari X

Berikut implementasi perhitungan algoritma *Naive Bayes* pada beasiswa PPA dan KIP-K:

1. Perhitungan Algoritma Beasiswa PPA.

Berikut ini contoh langkah-langkah penerapan algoritma Naive Bayes menggunakan uji coba data pertama yang terdapat pada Tabel 3 Beasiswa PPA.

Langkah-1: Menghitung jumlah kelas dibagi label

- P (Y = Diterima) = 30/45
- P (Y = Ditolak) = 15/45

Langkah-2: Menghitung jumlah kasus yang sama dengan kelas yang sama

- P (IPK = 3,29 | Y = Diterima) = 1/30 [jumlah data IPK 3,29 dengan status-diterima dibagi dengan jumlah data diterima]
- P (IPK = 3,29 | Y = Ditolak) = 2/15 [jumlah data IPK 3,29 dengan status-ditolak dibagi dengan jumlah data ditolak]
- P (jumlah_tanggungan = 6 | Y = Diterima) = 2/30 [jumlah-data-jumlah_tanggungan 6 dengan status diterima dibagi dengan jumlah data diterima]
- P (jumlah_tanggungan = 2 | Y = Ditolak) = 4/15 [jumlah data-jumlah_tanggungan = 6 dengan status ditolak dibagi dengan jumlah data ditolak]
- P (Penghasilan = 5.000.000 | Y = Diterima) = 1/30 [jumlah data-Penghasilan 5.000.000 dengan status diterima dibagi dengan jumlah data diterima]
- P (Penghasilan = 5.000.000 | Y = Ditolak) = 2/15 [jumlah-data Penghasilan 5.000.000 dengan status ditolak dibagi dengan jumlah data ditolak]

Langkah-3: Menghitung seluruh jumlah variabel

- P (P (IPK = 3,29 | Y = Diterima) x P(jumlah_tanggungan= 6 | Y =-Diterima) x P(Penghasilan = 5.000.000 | Y = Diterima))

$$= \frac{1}{30} \times \frac{2}{30} \times \frac{1}{27} \times \frac{30}{45}$$

$$= \frac{60}{1215000}$$

$$= 0.000049$$

- P (P (IPK = 3,29 | Y = Ditolak) x P (jumlah_tanggungan = 6 | Y = Ditolak) x P(Penghasilan = 5.000.000 | Y = Ditolak) = $\frac{2}{15} \times \frac{4}{15} \times \frac{2}{45} = \frac{240}{151875} = 0.001580$

Langkah-4: Membandingkan-Variabel “Diterima dan Ditolak”

Dari perhitungan yang dilakukan

= (P | Ditolak) > (P | Diterima)

= 0.001580 > 0.000049

= Ditolak

Maka, hasil prediksi data pada Tabel 3 data pertama adalah Ditolak.

2. Perhitungan Algoritma Beasiswa KIP-K.

Berikut ini contoh langkah-langkah penerapan algoritma Naive Bayes menggunakan data-data yang terdapat pada Tabel 4 Beasiswa KIP-K dengan uji coba menggunakan data pertama.

Langkah-1: Menghitung Jumlah-Kelas

- P (Y = Diterima) = -50/100
- P (Y = Ditolak) = -50/100

Langkah-2: Menghitung jumlah Kelas dengan kelas yang Sama

- P (Tes tulis = 73 | y = Diterima) = 3/50 [jumlah data testulis 73 dengan status- \neg diterima dibagi dengan jumlah data yang diterima]
- P (Tes tulis = 73 | y = Ditolak) = 1/50 [jumlah data testulis 73 dengan status-ditolak dibagi dengan jumlah data yang ditolak]
- P (Tes Wawancara = 84 | y = Diterima) = 1/50 [jumlah data *test* wawancara 75 dengan status-diterima dibagi dengan jumlah data yang diterima]
- P (Tes Wawancara = 84 | y = Ditolak) = 3/50 [jumlah data *test* wawancara 75 dengan status-ditolak dibagi dengan jumlah data yang ditolak]
- P (Kondisi Ekonomi = 200.000 | y = Diterima) = 3/50 [jumlah data kondisi ekonomi 200.000 dengan status-diterima dibagi dengan jumlahdata yang diterima]
- P (Kondisi Ekonomi = 200.000 | y = Ditolak) = 1/50 [jumlah data kondisi ekonomi 200.000 dengan status-ditolak dibagi dengan jumlah-data yang ditolak]
- P (Prestasi Akademik = 60 | y = Diterima) = 7/50 [jumlah data Prestasi Akademik 80 dengan status-diterima dibagi dengan jumlah-data yang diterima]
- P (Prestasi Akademik =60 | y = Ditolak) = 5/50 [jumlah data Prestasi Akademik 80 dengan status-ditolak dibagi dengan jumlah-data yang ditolak]
- P (Prestasi Non Akademik = 60 | y= Diterima) = 3/50 [jumlah data Prestasi Non Akademik 60 dengan status-diterima dibagi dengan jumlah-data yang diterima]
- P (Prestasi Non Akademik =60 | y=Ditolak) = 1/50 [jumlah data Prestasi Non Akademik 60

dengan status-ditolak dibagi dengan-jumlah data yang ditolak]

- $P(\text{Hasil Survei} = 79 | y = \text{Diterima}) = 4/5$ [jumlah data Hasil Survei 79 dengan status-diterima dibagi dengan jumlah-data yang diterima]
- $P(\text{Hasil Survei} = 79 | y = \text{Ditolak}) = 7/5$ [jumlah data Hasil Survei 79 dengan status-ditolak dibagi dengan jumlah-data yang ditolak]

Langkah-3: Menghitung seluruh jumlah-variabel

- Diterima = $P(\text{P}(\text{Test tulis} = 73 | Y = \text{Diterima}) \times P(\text{test wawancara} = 75 | Y = \text{Diterima}) \times P(\text{Kondisi Ekonomi} = 200.000 | Y = \text{Diterima}) \times P(\text{Prestasi Akademik} = 60 | Y = \text{Diterima}) \times P(\text{Prestasi Non Akademik} = 60 | Y = \text{Diterima}) \times P(\text{Hasil Survei} = 79 | Y = \text{Diterima})$

$$\text{diterima} = \frac{3}{50} \times \frac{1}{50} \times \frac{3}{50} \times \frac{7}{50} \times \frac{3}{50} \times \frac{4}{50} \times \frac{50}{100} \\ = \frac{37800}{1562500000000} = 0,000000024192$$

- Ditolak = $P(\text{P}(\text{Tes-tulis} = 77 | Y = \text{Ditolak}) \times P(\text{tes-wawancara} = 75 | Y = \text{Ditolak}) \times P(\text{Kondisi Ekonomi} = 125.000 | Y = \text{Ditolak}) \times P(\text{Prestasi-Akademik} = 80 | Y = \text{Ditolak}) \times P(\text{Prestasi Non}$

Akademik = 0 | Y = Ditolak) x P(Hasil Survei = 82 | Y = Ditolak)

$$\text{ditolak} = \frac{1}{50} \times \frac{3}{50} \times \frac{1}{50} \times \frac{5}{50} \times \frac{1}{50} \times \frac{7}{50} \times \frac{50}{100} \\ = \frac{5250}{1562500000000} = 0.00000000336$$

Langkah-4: membandingkan Variabel

$$= (P | \text{Diterima}) > (P | \text{Ditolak}) \\ = 0,000000024192 > 0.00000000336 \\ = \text{Diterima}$$

Maka, hasil-prediksi data pada Tabel 4 data pertama adalah Diterima.

3.4. Testing dan Evaluasi

Untuk mengevaluasi performa model menggunakan algoritma Naïve Bayes maka dilakukan pengujian data sampel yang ada dengan data sampel itu sendiri untuk melakukan peninjauan keakuratan prediksi. Adapun perbandingan hasil pengujian untuk data prediksi PPA disajikan pada Tabel 3, sedangkan untuk data hasil pengujian beasiswa KIP-K disajikan pada Tabel 4.

Tabel 3. Perbandingan Hasil Pengujian Data Beasiswa PPA

Nama	IPK	Tanggungan	Penghasilan	Status	Prediksi
Mhs 1	3.29	6	5.000.000	Ditolak	Ditolak
Mhs 2	3.50	5	2.500.000	Diterima	Diterima
Mhs 3	3.32	4	2.061.200	Diterima	Diterima
Mhs 4	3.48	4	5.576.000	Ditolak	Ditolak
Mhs 5	3.52	2	1.500.000	Diterima	Diterima
Mhs 6	3.78	5	1.000.000	Diterima	Diterima
Mhs 7	3.40	5	3.583.000	Diterima	Diterima
Mhs 8	3.80	4	2.000.000	Diterima	Diterima
Mhs 9	3.48	3	750.000	Diterima	Diterima
Mhs 10	3.00	4	1.500.000	Diterima	Diterima
.....
Mhs 45	3.57	5	5.168.407	Ditolak	Ditolak

Tabel 4. Perbandingan Hasil Pengujian Beasiswa KIP-K

Nama Siswa	Tes Tulis	Wawancara	Kondisi Ekonomi	Prestasi Akademik	Prestasi Non Akademik	Hasil Survei	Status	Prediksi
Sis 1	73	84	Rp 200.000	60	60	79	Diterima	Diterima
Sis 2	63	86	Rp 125.000	80	10	98	Diterima	Diterima
Sis 3	53	84	Rp 160.000	20	10	80	Ditolak	Ditolak
Sis 4	36	84	Rp 300.000	10	10	85	Ditolak	Ditolak
Sis 5	45	85	Rp 230.000	10	10	75	Ditolak	Ditolak
Sis 10	60	87	Rp 400.000	80	0	70	Ditolak	Ditolak
....
Sis 100	90	78	Rp 250.000	60	30	87	Diterima	Diterima

Dalam pengujian dan evaluasi model algoritma *Naive Bayes* digunakan pengujian *Confusion matrix*. Berikut ini pengujian Beasiswa PPA dan KIP-K berdasarkan hasil *training* yang bertujuan untuk mengetahui hasil perbandingan *output* antara data aktual dan prediksi. Adapun hasil perhitungan *confusion matrix* untuk Beasiswa PPA disajikan pada Tabel 5 dan Beasiswa KIP-K pada Tabel 6.

Pada Tabel 5, menghasilkan *confusion matrix* dari penyeleksian beasiswa Peningkatan Prestasi Akademik (PPA). Jumlah *True-Positive* 27, jumlah

True-Negative 9, jumlah *False-Positive* 4 dan jumlah *False-Negative* 5.

Tabel 5. *Confusion Matrix* Beasiswa PPA

Actual	Prediction	
	Diterima	Ditolak
Diterima	27	5
Ditolak	4	9

Menghitung *Recall* atau *Sensitivity* dengan menggunakan rumus *Recall* [18]:

Menghitung *Recall* atau *Sensitivity* dengan menggunakan rumus *Recall* [18]:

$$\begin{aligned} \text{Sensitivity (Recall)} &= \frac{TP}{P} = \frac{TP}{TP+FN} \quad (2) \\ &= \frac{27}{27+5} = \frac{27}{32} = 0.84375 \end{aligned}$$

Kemudian, Menghitung *specificity* dengan menggunakan rumus *specificity* [18]:

$$\begin{aligned} \text{Specificity} &= \frac{TN}{N} = \frac{TN}{TN+FP} \quad (3) \\ &= \frac{9}{9+4} = \frac{9}{13} = 0,69231 \end{aligned}$$

Kemudian, Menghitung *precision* dengan menggunakan rumus *precision* [18]:

$$\begin{aligned} \text{Precision} &= \frac{TP}{TP+FP} \quad (4) \\ &= \frac{27}{27+4} = \frac{27}{31} = 0.87097 \end{aligned}$$

Setelah diketahui hasil perhitungan *Recall*, *specificity* dan *precision*, selanjutnya menghitung nilai akurasi menggunakan rumus akurasi [18]:

$$\begin{aligned} \text{Accuracy} &= \frac{TP+TN}{P+N} = \frac{TP+TN}{TP+TN+FP+FN} \quad (5) \\ &= \frac{27+9}{27+9+4+5} = \frac{36}{45} = 0.8 \end{aligned}$$

Jadi tingkat akurasi untuk program beasiswa Peningkatan Prestasi Akademik (PPA) adalah 80%.

Tabel 6. *Confussion Matrix* Beasiswa KIP-K

Actual	Prediction	
	Diterima	Ditolak
Diterima	45	5
Ditolak	4	46

Pada Tabel 6, menghasilkan *confussion matrix* dari penyeleksian beasiswa Kartu Indonesia Pintar Kuliah (KIP-K). Jumlah *True-Positive* 45, jumlah *True-Negative* 46, jumlah *False-Positive* 4 dan jumlah *False-Negative* 5.

Menghitung *Recall* atau *Sensitivity* dengan menggunakan rumus *recall* [18]:

$$\begin{aligned} \text{Sensitivity (Recall)} &= \frac{TP}{P} = \frac{TP}{TP+FN} = \frac{45}{45+5} = \frac{45}{50} \\ &= 0.90 \end{aligned}$$

Kemudian, Menghitung *specificity* dengan menggunakan rumus *specificity* [18]:

$$\text{Specificity} = \frac{TN}{N} = \frac{TN}{TN+FP} = \frac{46}{46+4} = \frac{46}{50} = 0.92$$

Kemudian, Menghitung *precision* dengan menggunakan rumus *precision* [18]:

$$\text{Precision} = \frac{TP}{TP+FP} = \frac{45}{45+4} = \frac{45}{49} = 0.92$$

Setelah mengetahui hasil-perhitungan *Recall*, *specificity* dan *precision*, selanjutnya menghitung akurasi untuk mengetahui-nilai yang telah didapatkan klasifikasi-penyeleksian-untuk setiap program-beasiswa yang telah dilakukan. untuk menghitung nilai-akurasi menggunakan rumus [16]:

$$\begin{aligned} \text{Accuracy} &= \frac{TP+TN}{P+N} = \frac{TP+TN}{TP+TN+FP+FN} \\ &= \frac{45+46}{45+46+4+5} = \frac{91}{100} = 0.91 \end{aligned}$$

Jadi, hasil akurasi beasiswa Kartu Indonesia Pintar Kuliah (KIP-K) adalah 91%.

Berikut rekapitulasi persentase Perhitungan *Confussion Matrix* untuk beasiswa PPA dan KIP-K meliputi Akurasi (*Acc*), *Precision* (*Pre*), *Recall* (*Rec*), dan *Specificity* (*Spe*) yang disajikan pada Tabel 7.

Tabel 7. Persentase Perhitungan *Confussion Matrix*

Beasiswa	Acc	Pre	Rec	Spe
PPA	80	87	84	69
KIP-K	91	92	90	92

4. KESIMPULAN

Algoritma *Naïve-Bayes* merupakan algoritma yang cukup baik digunakan untuk memprediksi penerima beasiswa Peningkatan Pretasi Akademik (PPA) dan Kartu Indonesia Pintar Kuliah (KIP-K). Hal tersebut dibuktikan berdasarkan hasil uji coba, untuk *Dataset* PPA menghasilkan tingkat akurasi 80%, sedangkan untuk *Dataset* KIP-K menghasilkan tingkat akurasi 91%.

DAFTAR PUSTAKA

- [1] D. Kurniadi, E. Abdurachman, H. L. H. S. Warnars, and W. Suparta, "The prediction of scholarship recipients in higher education using k-Nearest neighbor algorithm," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 434, no. 1, p. 012039, 2018, doi: 10.1088/1757-899X/434/1/012039.
- [2] M. N. Sutoyo, "Implementasi Metode MADM Model Yager untuk Seleksi Penerima Beasiswa PPA," *JUITA J. Inform.*, vol. 5, no. 2, p. 81, 2018, doi: 10.30595/juita.v5i2.1630.
- [3] M. B. Kurniawan, "Sistem Pendukung Pengambilan Keputusan Pemberian Beasiswa Dengan Menggunakan Metode AHP." Universitas Sanata Dharma, Yogyakarta, pp. 1–67, 2019.
- [4] H. Sulistiani, "Penerapan Algoritma Klasifikasi Sebagai Pendukung Keputusan Pemberian Beasiswa Mahasiswa," pp. 300–305, 2018, doi: 10.31227/osf.io/yaavj.

- [5] F. A. Harimurti and E. Riksakomara, "Klasifikasi Penerimaan Beasiswa Menggunakan Metode Naïve Bayes Classifier (Studi Kasus Universitas Trunojoyo Madura)," 2017.
- [6] A. Alwi and Munirah, "The Concept Of Naive Bayes And Its Simple Use For Prediction Final Score Of Student Examination Using R Language," *JUTIF*, vol. 3, no. 1, pp. 133–140, 2022.
- [7] A. Merdekawati, "Sistem Pendukung Keputusan Penerimaan Beasiswa Menggunakan Algoritma C4.5 (Studi Kasus : Baitul Maal)," *J. Pendidik. Teknol. dan Kejuru.*, vol. 15, no. 1, pp. 113–123, 2018, doi: 10.23887/jptk-undiksha.v15i1.13067.
- [8] H. S. A. K. Muhammad Fikri Khaid, Sujalwo, "Model Logika Fuzzy Sugeno Berbasis Web Untuk Seleksi Penerima Beasiswa," *J. Ilm. SINUS*, vol. 16, no. 1, p. 25, 2018, doi: 10.30646/sinus.v16i1.329.
- [9] A. Pujianto, K. Kusriani, and A. Sunyoto, "Perancangan Sistem Pendukung Keputusan Untuk Prediksi Penerima Beasiswa Menggunakan Metode Neural Network Backpropagation," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 2, p. 157, 2018, doi: 10.25126/jtiik.201852631.
- [10] D. Kurniadi, A. Mulyani, Y. Septiana, and I. M. Yusuf, "Prediction of courses score using Artificial Neural Network with Backpropagation algorithm," in *IOP Conference Series: Materials Science and Engineering*, Mar. 2021, vol. 1098, no. 3, p. 032110, doi: 10.1088/1757-899X/1098/3/032110.
- [11] D. Kurniadi, H. Leslie, H. Spits, and W. Suparta, "Predicting Student Performance with Multi-Level Representation in an Intelligent Academic Recommender System using Backpropagation Neural Network," *ICIC Express Lett. Part B Appl.*, vol. 12, no. 10, pp. 883–890, 2021, doi: 10.24507/iciicelb.12.10.883.
- [12] P. R. Sihombing and A. M. Arsani, "Comparison of Machine Learning Methods in Classifying Poverty in Indonesia in 2018," *J. Tek. Inform.*, vol. 2, no. 1, pp. 51–56, 2021, doi: 10.20884/1.jutif.2021.2.1.52.
- [13] Y. Mardi, "Data Mining : Klasifikasi Menggunakan Algoritma C4.5," *J. Edik Inform.*, vol. 2, no. 2, pp. 213–219, 2017.
- [14] E. Manalu, F. A. Sianturi, and M. R. Manalu, "Penerapan Algoritma Naive Bayes Untuk Memprediksi Jumlah Produksi Barang Berdasarkan Data Persediaan Dan Jumlah Pemesanan Pada Cv. Papadan Mama Pastries," *J. Mantik Penusa*, vol. 1, no. 2, 2017.
- [15] T. H. Apandi and C. A. Sugianto, "Algoritma Naive Bayes untuk Prediksi Kepuasan Pelayanan Perekaman e-KTP (Naive Bayes Algorithm for Satisfaction Prediction of e-ID)," *JUITA (Jurnal Inform. UMP)*, vol. 7, no. November, pp. 125–128, 2019.
- [16] M. F. Fibrianda and A. Bhawiyuga, "Analisis Perbandingan Akurasi Deteksi Serangan Pada Jaringan Komputer Dengan Metode Naïve Bayes Dan Support Vector Machine (SVM)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 9, pp. 3112–3123, 2018.
- [17] S. Rizal, P. Studi, T. Informatika, and U. Yudharta, "Penerapan Algoritma Naïve Bayes Untuk Prediksi Penerimaan Siswa Baru Di Smk Al-Amien Wonorejo," *Explor. IT J. Keilmuan dan Apl. Tek. Inform.*, vol. 10, no. 1, pp. 14–21, 2018, doi: 10.35891/explorit.v10i1.1671.
- [18] S. Widaningsih, "Perbandingan Metode Data Mining Untuk Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma C4,5, Naïve Bayes, KNN Dan SVM," *J. Tekno Insentif*, vol. 13, no. 1, pp. 16–25, 2019, doi: 10.36787/jti.v13i1.78.