



## Sistem Klasifikasi *Monitoring* dan Evaluasi Kelayakan Penerima Beasiswa UAD Menggunakan Algoritma *Naïve Bayes*

\*Dyllan Bagus Siswanto<sup>1</sup>, Dwi Normawati<sup>2</sup>

<sup>1,2</sup>Informatika, Fakultas Teknologi Industri, Universitas Ahmad Dahlan Yogyakarta

Jl. Ringroad Selatan, Kragilan, Tamanan, Kec. Banguntapan, Kabupaten Bantul,  
Daerah Istimewa Yogyakarta 55191

Email: <sup>1</sup>dyllanbgus8056@gmail.com, <sup>2</sup>dwi.normawati @tif.uad.ac.id,

### ABSTRACT

*Kartu Indonesia Pintar (KIP) scholarship program at Ahmad Dahlan University includes a monitoring and evaluation internal process conducted at the end of each semester to monitor scholarship recipients. This allows for an assessment of their eligibility to receive the scholarship in the upcoming semester. The current manual MONEVIN process results in a time-consuming and less objective eligibility analysis. This eligibility determination system is needed, utilizing data mining techniques based on previous KIP scholarship recipient data to make predictions. Naïve Bayes algorithm, a data mining technique employing mathematical probability calculations, is utilized. The process begins with preprocessing, followed by data mining, evaluation, system implementation, and testing using System Usability Scale. The research uses a dataset of 270 student records, employing a 9-fold cross-validation process to split the data. Implemented model is integrated into a website-based system accessible to Biro Kemahasiswaan dan Alumni (BIMAWA). Model testing employs the Confusion Matrix technique, resulting in an accuracy score of 0.985, precision of 0.987, recall of 0.985, and an F-score of 0.985, indicating a favorable classification outcome. The system's eligibility assessment is further tested using the SUS, yielding a score of 90. Therefore, it can be concluded that the developed system is suitable for use.*

*Keywords : KIP Scholarship in UAD; Data Mining; classification; Naïve Bayes; Confusion Matrix*

### ABSTRAK

Program beasiswa Kartu Indonesia Pintar (KIP) di Universitas Ahmad Dahlan memiliki kegiatan laporan *Monitoring* dan Evaluasi *Internal* yang dilakukan setiap akhir semester sebagai *monitoring* penerima beasiswa sehingga dapat mengevaluasi dengan dinyatakan layak atau tidak layak untuk menerima beasiswa tersebut di semester berikutnya. Proses MONEVIN masih dikerjakan manual yang mengakibatkan proses analisis penentuan kelayakan memakan waktu yang lama dan kurang objektif sehingga diperlukan sebuah sistem penentuan kelayakan menggunakan teknik *data mining* berdasarkan data penerima beasiswa KIP yang diketahui sebelumnya dan selanjutnya dapat melakukan prediksi. Algoritma *Naïve bayes* merupakan teknik *data mining* dengan perhitungan probabilitas matematika. Proses diawali dengan *preprocessing*, *data mining* hingga evaluasi dilanjutkan implementasi sistem dan pengujian menggunakan *System Usability Scale*. Data yang digunakan pada penelitian ini berjumlah 270 data mahasiswa dengan dilakukan proses *split data* menggunakan *9-fold cross validation*. Hasil model diimplementasikan ke dalam sistem berbasis website yang dapat diakses oleh pihak Biro Kemahasiswaan dan Alumni (BIMAWA) dalam pengklasifikasian kelayakan, pengujian model menggunakan teknik *Confusion Matrix* menghasilkan nilai *accuracy* 0,985, *precision* 0,987, *recall* 0,985 dan *F-score* 0,985 dengan hasil klasifikasi dinyatakan bagus dan pengujian kelayakan sistem menghasilkan *score SUS* sebesar 90 sehingga dapat disimpulkan sistem yang dibangun layak/baik untuk digunakan.

Kata kunci : Beasiswa KIP UAD; *Data Mining*; Klasifikasi; *Naïve Bayes*

## 1. PENDAHULUAN

Beasiswa merupakan pendanaan yang sumbernya bukan dari pihak terkait maupun orang tua dengan tujuan untuk meringankan biaya pendidikan dan memotivasi penerimanya dalam belajar (Alita et al., 2021). Program KIP merupakan salah satu beasiswa yang disediakan oleh UAD, tujuan dari program ini yaitu membantu masyarakat yang kurang mampu dalam meningkatkan akses ke perguruan tinggi juga melaksanakan visi dari presiden Indonesia terkait Sumber Daya Manusia (SDM) yang unggul (Kemdikbud, n.d.). Beasiswa ini mencakup biaya kuliah hingga lulus, sehingga proses seleksinya dilakukan setiap akhir semester agar dapat memantau mahasiswa yang benar-benar layak menerima beasiswa tersebut di semester berikutnya dan melakukan evaluasi kelayakan penerima Beasiswa apakah dapat dinyatakan layak menerima Beasiswa tersebut disemester depan, proses ini disebut Proses *Monitoring* dan evaluasi *internal* (MONEVIN).

Terdapat tujuh kriteria atau variabel yang dijadikan sebagai penentuan kelayakan yaitu IPK, seminar, pelatihan, prestasi, publikasi, perilaku cendekiawan dan organisasi (ormawa,

ortom, UKM, LSO) dengan hasil rekomendasi keputusan layak dan tidak layak.

Dewasa ini, proses MONEVIN oleh BIMAWA UAD masih dikerjakan manual dengan melakukan *monitoring* satu persatu dan menentukan hasil dari *monitoring* tersebut berupa evaluasi sebuah keputusan kelayakan yang mengakibatkan proses analisis penentuan kelayakan memakan waktu yang lama dan kurang objektif, oleh karena itu diperlukan sistem klasifikasi penentuan kelayakan pada penerima beasiswa.

*Data mining* suatu proses analisis dari tumpukan data yang dikumpulkan lalu menghasilkan pola tertentu sehingga dapat dibangun sebuah model yang dapat memprediksi menggunakan *machine learning* yang belum diketahui kelasnya (Abdullah et al., 2021; Damuri et al., 2021).

Klasifikasi merupakan teknik *data mining* yang melakukan pembagian suatu kelompok ke dalam kelas yang sudah ditentukan sebelumnya (Abdullah et al., 2021). Algoritma *Naïve Bayes* yaitu suatu metode klasifikasi untuk menghasilkan kelas keputusan dengan perhitungan peluang matematika dari

syarat kebenaran berdasarkan variabel (Damuri et al., 2021).

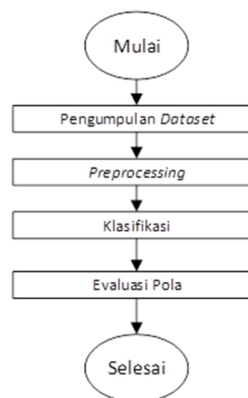
Penelitian pada tahun 2021 oleh Noviyanto et al. (2021), yaitu penentuan kelayakan calon penerima beasiswa menggunakan metode *Naïve Bayes* dan melakukan perbandingan dengan metode C4.5 dan KKN untuk mencari metode terbaik. Jumlah data yang digunakan sebanyak seratus data mahasiswa serta terdapat delapan variabel yang digunakan. Proses dilakukan sebanyak dua kali dengan perbedaan pada perbandingan antara data latih dan data uji, pada proses pertama dengan perbandingan 70:30 dengan menghasilkan metode dengan *accuracy score* terbaik adalah *Naïve Bayes* sedangkan pada proses kedua dengan perbandingan 80:20 semua metode menghasilkan *accuracy score* yang sama. Dapat disimpulkan bahwa metode *Naïve Bayes* dapat melakukan klasifikasi dengan jumlah data latih yang sedikit (Noviyanto et al., 2021).

Hasil penelitian sebelumnya hanya berupa analisa sehingga penelitian ini menghasilkan sebuah sistem berbasis *website* yang dapat melakukan prediksi dengan menggunakan data yang baru atau belum diketahui hasil prediksinya. Oleh karena itu berdasarkan penjelasan

sebelumnya, maka tujuan dari penelitian ini adalah membangun sebuah “Sistem Klasifikasi *Monitoring* dan Evaluasi Kelayakan Penerima Beasiswa UAD Menggunakan Algoritma *Naïve Bayes*”, dimana algoritma ini melakukan proses klasifikasi sederhana melalui *training set* sebagian data secara efisien (Sumiah et al., 2020). Data yang digunakan merupakan data penerima beasiswa KIP dari tahun 2020 hingga 2022 dengan bahasa pemrograman *PHP* digunakan untuk membangun sistem dan bahasa pemrograman *Python* digunakan untuk mengolah data.

## 2. METODE

Tahapan dari penelitian ini dapat digambarkan seperti pada Gambar 1 *flowchart* tahapan penelitian.



Gambar 1. *Flowchart* tahapan penelitian

Tahapan penelitian berisi penjelasan setiap tahapan pada proses penelitian yang dimulai dari

pengumpulan *dataset* hingga evaluasi pola.

## 2.1. Pengumpulan *Dataset*

Pengumpulan *dataset* yang dilakukan dengan menghubungi pihak BIMAWA, dengan mendapatkan data yang digunakan untuk penelitian ini. Data merupakan penerima beasiswa KIP mulai tahun 2020 hingga 2022 sebanyak 270 *record data* dengan rekomendasi layak dan tidak layak serta terdapat 16 atribut yang tersedia namun dilakukan seleksi atribut berdasarkan kebutuhan atau syarat yang digunakan untuk proses penentuan kelayakan berdasarkan wawancara dengan pihak BIMAWA sehingga hanya menggunakan tujuh atribut fitur dan satu atribut target. Tabel 1 atribut *dataset* yang digunakan pada penelitian ini didapat dari BIMAWA UAD selaku pengelola beasiswa.

Tabel 1. Atribut *dataset*

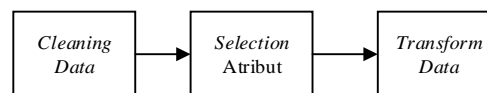
Nama	Deskripsi	Value
1. IPK	IPK mahasiswa	0 – 4.00
2. Seminar	Mengikuti Seminar	1:Ada, 0:Tidak
3. Pelatihan	Mengikuti Pelatihan Memiliki	1:Ada, 0:Tidak
4. Prestasi	Prestasi (Non/akademik) Memiliki	1:Ada, 0:Tidak
5. Publikasi	jurnal publikasi	1:Ada, 0:Tidak
6. Perilaku Cendekiawan	Mengikuti kegiatan	1:Ada, 0:Tidak

7. Organisasi	bersifat pengabdian Mengikuti ormawa/ortonom/UKM/LSO	1:Ada, 0:Tidak
8. Status ( <i>Class</i> )	Hasil Klasifikasi	Layak, Tidak Layak

## 2.2. *Preprocessing*

Tahapan awal *Data Mining* merupakan *preprocessing*, dimana proses ini melakukan persiapan data mentah sebelum dilakukan proses pengalihan informasi (Lonang et al., 2022). Data mentah terkadang mengandung nilai kosong atau inkonsisten data, sehingga teknik *preprocessing* diperlukan untuk menyelesaikan masalah tersebut.

Berdasarkan kondisi data yang ada, *preprocessing* dilakukan untuk memperbaiki data mentah. Menurut peneliti dan pihak BIMAWA terdapat beberapa *preprocessing* yang perlu dilakukan seperti pada Gambar 2 (Lonang et al., 2022).



Gambar 2. Tahapan *Preprocessing*

### a. *Cleaning Data*

Proses pertama pada *preprocessing*, tujuan dari proses ini untuk membersihkan data kotor atau bernilai kosong.

### b. *Selection* Atribut

Pemilihan atribut digunakan untuk menyeleksi atribut yang akan dipakai pada proses *data mining*.

### c. *Transform Data*

Perubahan nilai suatu data menjadi sebuah kategori, namun pada penelitian ini dilakukan proses normalisasi data untuk memastikan data tetap konsisten.

## 2.3. Klasifikasi

Klasifikasi memiliki dua hal penting yang harus dilakukan, pertama membangun model atau aturan sebagai *prototype* yang akan disimpan kedalam memori dan kedua berdasarkan model atau aturan tersebut dilakukan pengenalan/prediksi/klasifikasi pada suatu data lain yang belum diketahui kelasnya. Dalam menyelesaikan kasus klasifikasi terdapat metode atau model yang sudah dikembangkan oleh para ahli salah satunya adalah algoritma *Naïve Bayes* (Lonang et al., 2022).

Algoritma *Naïve Bayes* merupakan suatu teknik klasifikasi, metode ini menghasilkan kelas keputusan berdasar-kan perhitungan peluang matematika dengan syarat nilai keputusan benar berdasarkan variabel yang digunakan (Suyanto, 2019).

Metode klasifikasi ini yang paling populer karena dalam proses pengklasifikasian tidak perlu menggunakan data yang besar namun dapat menghasilkan nilai akurasi yang baik. *Naïve bayes* dipilih karena ketelitian pada Algoritma ini cukup tinggi dan proses pengklasifikasian cukup cepat. Dalam penentuan probabilitas atau peluang dapat ditemukan menggunakan Persamaan 1 sebagai berikut :

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

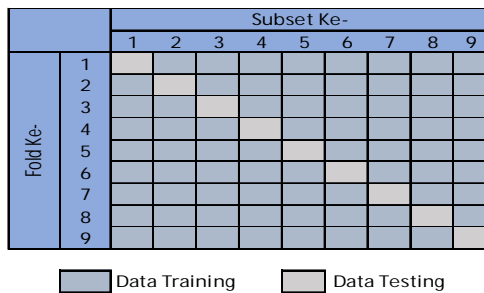
X adalah data kelas yang belum diketahui, H adalah Hipotesis data dari suatu kelas yang spesifik,  $P(H|X)$  adalah Probabilitas hipotesis H dengan syarat X (Posterior Probabilitas),  $P(X|H)$  adalah Probabilitas hipotesis X dengan syarat H,  $P(H)$  adalah Probabilitas hipotesis H (Prior Probabilitas),  $P(X)$  adalah Probabilitas hipotesis X.

## 2.4. *K-Fold Cross Validation*

*K-fold Cross Validation* yaitu pembagi secara acak data asli ke dalam sub sampel k, kemudian satu sub sampel sebagian data dijadikan data uji dan sisa data lainnya dijadikan data latih. Proses tersebut diulang sebanyak k dan menghasilkan perkiraan berdasarkan

rata-rata untuk setiap kelipatannya (Abidin et al., 2021).

Tujuan dari teknik ini yaitu mengevaluasi hasil perhitungan analisis (Nasution et al., 2019). Gambar 3 ilustrasi *9-fold cross validation* mengilustrasikan dilakukan pengujian sebanyak sembilan kali dimana nilai sembilan diambil berdasarkan banyak data sehingga data setiap fold mengalami balancing data dan nilai ke-k sama dengan pengujian ke-k. Pada setiap pengujian memiliki data uji sebanyak satu dan data latih sebanyak delapan.



Gambar 3. Ilustrasi 9-fold Cross Validation

## 2.5. Evaluasi Performa

Tahapan ini menggunakan teknik *Confusion matrix* yang merupakan metode penghitungan akurasi pada proses *data mining* yang menghasilkan dua kelas atau lebih (Hozairi et al., 2021; Normawati et al., 2021). *Confusion matrix* dapat digambarkan sebagai matriks 2x2 yang menampilkan hasil klasifikasi *biner* pada suatu *dataset* dengan empat istilah didalamnya yaitu

*True Positive*(TP), *True Negative*(TN), *False Positive* (FP) dan *False Negative* (FN). Seperti pada Tabel 2 *confusion matrix* (Andika et al., 2019; Mutawalli et al., 2019).

Tabel 2. *Confusion Matrix*

		TRUE VALUES	
		TRUE	FALSE
PREDICTION	TRUE	TP <i>Correct Result</i>	FP <i>Unexpected Result</i>
	FALSE	FN <i>Missing Result</i>	TN <i>Correct absence of result</i>

TP adalah Data positif yang dideteksi benar, FP adalah Data negatif yang dideteksi benar, FN adalah Data negatif yang dideteksi negatif, TN adalah Data negatif yang dideteksi benar.

Hasil dari Tabel 2 *confusion matrix* dapat digunakan untuk menghitung performa klasifikasi dengan mencari nilai dari *accuracy*, *precision*, *recall* dan *F Score* sebagai berikut (Andika et al., 2019; Shianto et al., 2019):

### a. Accuracy

*Accuracy* adalah jumlah rasio prediksi yang benar seperti pada Persamaan 2.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (2)$$

### b. Precision

*Precision* adalah rasio jumlah antara data bernilai benar dengan data prediksi benar seperti pada Persamaan 3.

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

### c. Recall

*Recall* adalah perbandingan rasio jumlah prediksi benar dengan seluruh rasio jumlah data yang benar seperti pada Persamaan 4.

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

### d. F Score

*F Score* adalah perhitungan dengan membandingkan antara *precision* dan *recall* seperti Persamaan 5.

$$F\ Score = 2 * \frac{Precision * recall}{Precision + recall} \quad (5)$$

## 2.6 Pengujian Sistem

*System Usability Scale* (SUS) adalah metode pengujian kelayakan sistem yang dilakukan kepada pengguna perangkat lunak dengan alat ukur, dengan sifatnya “*quick and dirty*” sehingga dapat diandalkan dalam pengujian (Ramadhan et al., 2019).

## 3. HASIL DAN PEMBAHASAN

Proses pertama mempersiapkan data yang akan diolah dengan menghilangkan *noise* atau data yang tidak diketahui nilainya dengan memberi nilai yang kosong dengan rata-rata nilai yang lain. Selanjutnya adalah pemilihan atribut yang akan digunakan, berdasarkan data mentah terdapat 16 atribut yang tersedia namun dilakukan seleksi atribut berdasarkan kebutuhan

atau syarat yang digunakan untuk proses penentuan kelayakan berdasarkan wawancara dengan pihak BIMAWA sehingga menggunakan tujuh atribut fitur dan satu atribut *class* atau target. Berikut merupakan *dataset* Beasiswa KIP yang akan digunakan, pada Tabel 3 *dataset* beasiswa KIP.

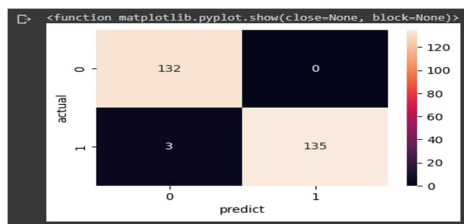
Tabel 3. *Dataset* Beasiswa KIP

Data ke-	IPK	Seminar Pelatihan	Prestasi	Publikasi	Perilaku Cendekiawan	Organisasi
1	3,87	1	1	1	1	0
2	3,89	1	1	0	1	1
3	3,77	1	1	1	1	1
...						
269	3,93	1	1	0	1	0
270	3,72	1	1	1	1	1

Tahap akhir dari *preprocessing* adalah transformasi data, dimana proses ini akan menormalisasikan data berdasarkan atribut fitur yang digunakan karena normalisasi digunakan untuk memastikan *dataset* tetap konsisten. Proses pengalihan informasi atau mencari pola berdasarkan *dataset* yang ada, setelah informasi telah didapat maka dibuatkan sebuah model yang dapat digunakan untuk melakukan prediksi di masa datang dengan data yang belum diketahui nilai dari target atau rekomendasinya. Implementasi pengolahan data pada bahasa

pemrograman *python*, pembuatan model ini menggunakan *library pickle* sehingga model tersebut dapat di ekstraksi dan dapat digunakan saat memprediksi data yang baru. Sistem yang dibangun menggunakan bahasa *PHP*, untuk menggabungkan antara pengolahan data dan sistem yang bangun menggunakan *library flask* pada *python*.

Evaluasi performa dari model yang dibangun menggunakan teknik *cross validation* dimana untuk menentukan beberapa nilai seperti *accuracy*, *precision*, *recall* dan *f score*. Pengujian tersebut dilakukan menggunakan proses *k-fold cross validation*. Pada penelitian ini memiliki data yang berjumlah 270 data dengan perbandingan nilai rekomendasi layak dan tidak layak sebesar 50:50, sehingga dilakukan percobaan dengan banyak *fold* atau nilai *K* terbaik adalah sembilan dengan setiap *fold* berisi 30 data. Berikut hasil informasi *confusion matrix* salah satu *fold*, berikut Gambar 4 hasil *confusion matrix*.



Gambar 3. Hasil *Confusion Matrix*

Hasil *Confusion Matrix* yang dilakukan sebanyak *9-fold* menghasilkan nilai sebagai berikut, berikut Tabel 4 nilai hasil *9-fold cross validation*.

Tabel 4. Nilai hasil *9-fold Cross Validation*

Fold Ke-	Accuracy	Precision	Recall	F Score
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	0,934	0,941	0,933	0,933
7	1	1	1	1
8	0,934	0,941	0,933	0,933
9	1	1	1	1
<b>Rata-rata</b>	<b>0,985</b>	<b>0,987</b>	<b>0,985</b>	<b>0,985</b>

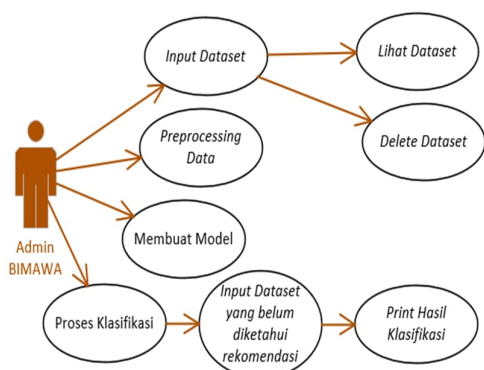
Berdasarkan Tabel 4, nilai hasil *9-fold cross validation* mendapatkan nilai akhir atau rata-rata dari setiap *fold*, dimana nilai *accuracy* 0,985, *precision* 0,987, *recall* 0,985 dan *f score* 0,985 sehingga *Score* tersebut dapat menunjukkan bahwa sistem *data mining* yang dibangun dapat dikatakan baik.

### 3.1. Rancangan Sistem

Sistem yang akan dibangun berdasarkan kebutuhan dari BIMAWA, kebutuhan tersebut dapat dijadikan landasan dari kebutuhan sebuah sistem sehingga sistem dapat dibangun dengan semestinya. Rancangan *use case diagram* dibuat pada penelitian ini,



berikut merupakan *use case diagram* pada Gambar 5.



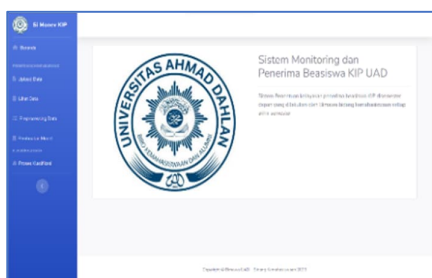
Gambar 4. *Use case Diagram*

Gambar 5 *use case diagram* dapat dijelaskan bahwa user *dari* sistem adalah *Admin BIMAWA* dengan dapat melakukan *input dataset*, *preprocessing*, *membuat model* dan melakukan *klasifikasi dengan dataset yang belum diketahui rekomendasinya*.

### 3.2. Implementasi Sistem

#### a. Halaman Awal Sistem

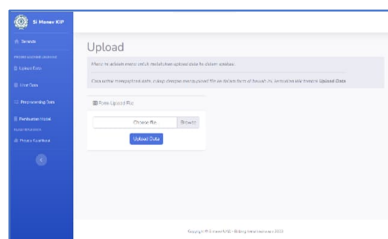
Halaman beranda yaitu tampilan awal dari sistem ini, dimana terdapat sekilas informasi tentang sistem ini. Tampilan halaman awal sistem ditujukan Gambar 6.



Gambar 5. Halaman awal sistem

#### b. Halaman *Upload Dataset*

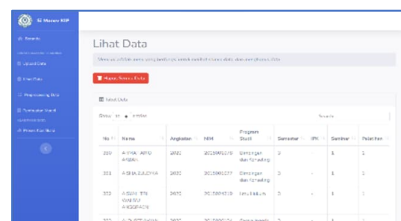
Halaman ini yaitu halaman untuk mengunggah *dataset* dengan format file *.xlsx*. Tampilan halaman *upload dataset* ditujukan Gambar 7.



Gambar 6. Halaman *upload dataset*

#### c. Halaman Lihat Data

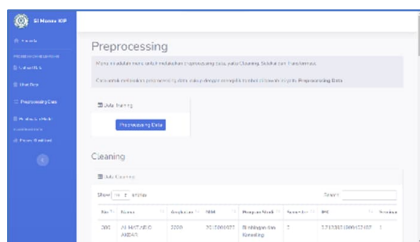
Halaman lihat data akan menampilkan data seluruh isi dari *dataset* yang di *upload* ke dalam sistem, juga terdapat sebuah tombol untuk menghapus seluruh *dataset* yang telah di *upload* ke sistem, berikut merupakan tampilan halaman lihat data ditujukan Gambar 8.



Gambar 7. Halaman lihat data

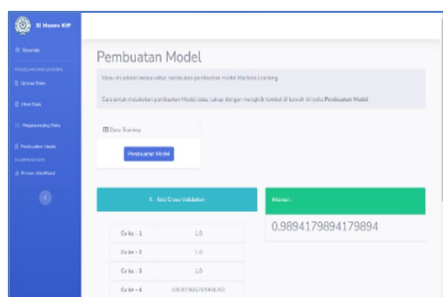
#### d. Halaman *Preprocessing*

Halaman ini akan menampilkan berdasarkan *dataset* yang ada dilakukan proses *preprocessing* seperti *cleaning*, *selection* dan *transform data*. Tampilan halaman ini ditujukan Gambar 9.

Gambar 8. Halaman *preprocessing*

## e. Halaman Pembuatan Model

Pada halaman ini akan terdapat tombol yang digunakan untuk membuat model *machine learning* sehingga model tersebut dapat memprediksi data yang belum diketahui rekomendasinya. Halaman ini juga akan menampilkan nilai akurasi yang didapat berdasarkan *data-set* yang digunakan. Tampilan halaman ini pada ditunjukkan 10.

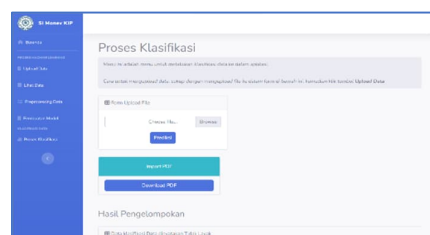


Gambar 9. Halaman pembuatan model

## f. Halaman Proses Klasifikasi

Proses *monitoring* dan evaluasi pada data baru dapat dilakukan di halaman ini. Sistem dapat mengunggah data baru yang belum diketahui rekomendasinya sehingga dapat melakukan proses *monitoring* dan menghasilkan evaluasi berupa prediksi kelayakan dari setiap data.

Hasil dari prediksi ini akan langsung mengelompokkan mahasiswa berdasarkan hasil rekomendasinya yaitu layak dan tidak layak, Tahapan akhir sistem dapat mengunduh hasil klasifikasi menjadi file .pdf. Tampilan halaman ini ditunjukkan Gambar 11.

Gambar 10. Halaman proses Klasifikasi 3.3 *System Usability Scale (SUS)*

Pengujian *usability* digunakan untuk menghitung nilai kelayakan berdasarkan perhitungan SUS sehingga dapat dikategorikan kelayakan sistem. Pengujian ditujukan kepada pengguna dari sistem yaitu pihak BIMAWA dengan menyebarkan sebanyak 5 kuesioner kepada pengelolaan Beasiswa dan penanggung jawab dari pengelolaan Beasiswa KIP. Rata-rata nilai dari responden yaitu 90 atau 90%, sehingga dapat dikategorikan pada kriteria sangat baik dan dapat dinyatakan layak untuk digunakan.

## 4. KESIMPULAN

Berdasarkan penelitian ini dengan judul sistem klasifikasi *data mining* monitoring dan evaluasi

kelayakan penerima beasiswa UAD menggunakan algoritma *Naïve Bayes* dapat diambil beberapa kesimpulan yaitu dapat dibangun sebuah sistem untuk melakukan proses klasifikasi *monitoring* dan evaluasi penerima beasiswa UAD dengan *accuracy score* sebesar 0,985, *precision score* sebesar 0,987, *recall score* sebesar 0,985 dan *F-score* sebesar 0,985. *Score* tersebut dapat menunjukkan bahwa sistem *data mining* yang dibangun dapat dikatakan baik dan pengujian kelayakan sistem menggunakan dengan *score* SUS sebesar 90 sehingga dapat disimpulkan sistem yang dibangun layak/sangat baik untuk digunakan. Adapun saran yang dapat dilakukan oleh peneliti lainnya, bahwa lakukan uji perbandingan dengan metode klasifikasi lainnya dan pada tahap *preprocessing* dilakukan menggunakan teknik lain agar mendapatkan *accuracy score* yang lebih akurat.

## 5. UCAPAN TERIMA KASIH

Alhamdulillah, segala puji bagi Allah yang telah memberikan kelancaran dalam mengerjakan penelitian ini, ucapan terimakasih kepada pihak BIMAWA UAD yang telah mengizinkan dalam penelitian ini,

kepada orang tua, dosen pembimbing skripsi telah membimbing selama penelitian ini dan tidak lupa kepada pihak lainnya yang membantu dalam bentuk apapun.

## DAFTAR PUSTAKA

- Abdullah, L., Tamin, R., & Qashlim, A. A. A. (2021). *Klasifikasi Penerimaan Beasiswa Menggunakan Algoritma Naive Bayes Di Universitas Al Asyariah Mandar Kabupaten Polewali Mandar*. 3(1). doi: 10.35329/jp.v3i1.1399
- Abidin, Z., Fredyatama, Y., Teknik Informasi, P., Tinggi Teknik Pati Jl Raya Pati-Trangkil, S. K., & Jawa Tengah, P. (2021). *Klasifikasi Daun Empon-Empo Menggunakan Metode Gray Level Co-Occurrence Matrix Dan Algoritma K-Nn*. *Jurnal Sains, Teknologi Dan Industri*, 18(02), 261–267. doi: dx.doi.org/10.24014/sitekin.v18i2.12913
- Alita, D., Sari, I., & Rahman Isnain, A. (2021). *Penerapan Naive Bayes Classifier Untuk Pendukung Keputusan Penerima Beasiswa*. *Jdmsi*, 2(1), 702022. doi: doi.org/10.55606/jitek.v1i2.378
- Andika, L. A., Azizah, P. A. N., & Respatiwan, R. (2019). *Analisis Sentimen Masyarakat terhadap Hasil Quick Count Pemilihan Presiden Indonesia 2019 pada Media Sosial Twitter Menggunakan Metode Naive Bayes Classifier*. *Indonesian Journal of Applied Statistics*, 2(1), 34. doi: 10.13057/ijas.v2i1.29998

- Damuri, A., Riyanto, U., Rusdianto, H., & Aminudin, M. (2021). Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako. *Jurnal Riset Komputer*, 8(6), 219–225. doi: 10.30865/jurikom.v8i6.3655
- Hozairi, Anwari, & Alim, S. (2021). Implementasi Orange Data Mining Untuk Klasifikasi Kelulusan Mahasiswa Dengan Model K-Nearest Neighbor, Decision Tree Serta Naive Bayes. *Network Engineering Research Operation*, 6(2), 133. doi: 10.21107/nero.v6i2.237
- Kemdikbud, P. W. (n.d.). *KIP Kuliah Merdeka: Akses Pendidikan Tinggi Semakin Merata dan Berkualitas*. Retrieved from <https://www.kemdikbud.go.id/main/blog/2021/03/kip-kuliah-merdeka-akses-pendidikan-tinggi-semakin-merata-dan-berkualitas>
- Lonang, S., & Normawati, D. (2022). Klasifikasi Status Stunting Pada Balita Menggunakan K-Nearest Neighbor Dengan Feature Selection Backward Elimination. *Jurnal Media Informatika Budidarma*, 6(1), 49. doi: 10.30865/mib.v6i1.3312
- Mutawalli, L., Zaen, M. T. A., & Bagye, W. (2019). Klasifikasi Teks Sosial Media Twitter Menggunakan Support Vector Machine (Studi Kasus Penusukan Wiranto). *Jurnal Informatika Dan Rekayasa Elektronik*, 2(2), 43. doi: 10.36595/jire.v2i2.117
- Nasution, D. A., Khotimah, H. H., & Chamidah, N. (2019). Perbandingan Normalisasi Data untuk Klasifikasi Wine Menggunakan Algoritma K-NN. *Computer Engineering, Science and System Journal*, 4(1), 78. doi: 10.24114/cess.v4i1.11458
- Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 5(2), 697–711. doi: <http://dx.doi.org/10.30645/j-sakti.v5i2.369>
- Noviyanto, H., & Mukti, B. (2021). Implementasi Algoritme Naïve Bayes Untuk Menentukan Kelayakan Calon. *Jurnal Informatika Dan Teknologi Komputer*, 1(2), 7–12. doi: [doi.org/10.55606/jitek.v1i2.378](https://doi.org/10.55606/jitek.v1i2.378)
- Ramadhan, D. W., Soedijono, B., & Pramono, E. (2019). *Pengujian Usability Website Time Excelindo Menggunakan System Usability Scale ( Sus ) ( Studi Kasus : Website Time Excelindo )*. 04, 139–147.
- Shianto, K. A., Gunadi, K., & Setyati, E. (2019). Deteksi Jenis Mobil Menggunakan Metode YOLO Dan Faster R-CNN. *Jurnal Infra*, 7(1), 157–163. Retrieved from <http://publication.petra.ac.id/index.php/teknik-informatika/article/view/8065>
- Sumiah, A., & Mirantika, N. (2020). *Perbandingan Metode K-Nearest Neighbor dan Naive Bayes untuk Rekomendasi Penentuan Mahasiswa Penerima Beasiswa pada Universitas Kuningan*. 6(April). doi: [doi.org/10.25134/buffer.v6i1.2907](https://doi.org/10.25134/buffer.v6i1.2907)
- Suyanto, D. (2019). *Data Mining untuk klasifikasi dan klusterisasi data*.