

## Klasifikasi Akreditasi Sekolah Di Kota Jambi Menggunakan Metode Naïve Bayes Di Bpmp Jambi

Agnes Febrin Sitanggang<sup>1</sup>, Errisya Rasyawir<sup>3</sup>, Lies Aryani<sup>3</sup>

<sup>1</sup> Fakultas Ilmu Komputer, Teknik Informatika, Universitas Dinamika Bangsa Jambi, Indonesia

Email: [agneszsitanggang@gmail.com](mailto:agneszsitanggang@gmail.com), [errissya.rasyawir@gmail.com](mailto:errissya.rasyawir@gmail.com), [liesaryani6@gmail.com](mailto:liesaryani6@gmail.com)

Penulis Korespondensi: Errisya Rasyawir, S.Kom, MT, Email Penulis Korespondensi: [errissya.rasyawir@gmail.com](mailto:errissya.rasyawir@gmail.com)

Artikel Info :

Artikel History :

Submitted : 08-03-2024

Accepted : 12-03-2024

Published : 30-04-2024

**Kata Kunci :**

Data Mining,  
Klasifikasi, Akreditasi,  
Algoritma Naïve  
Bayes, WEKA

**Abstrak**– Akreditasi adalah pemberian predikat kepada sebuah sekolah dengan tujuan untuk meningkatkan dan juga memperoleh gambaran pengembangan, peningkatan mutu serta menentuka suatu sekolah sebagai penyelenggara Pendidikan. Akreditasi sekolah merupakan salah satu factor yang penting dalam peningkatan mutu Pendidikan. Penilaian yang dilakukan juga diselenggarakan secara sistematis dan juga komprehensif dengan melewati evaluasi oleh Lembaga yang berwenang .Sekolah akan mendapatkan predikat akreditasi jika telah memenuhi standar-standar yang telah ditentukan.Klasfikasi merupakan salah satu teknik data mining yang bisa digunakan dalam proses klasifikasi akreditasi sekolah. Pengujian dilakukan dengan dilakukan dengan 2 options test yaitu Use Training Set dan 10 Fold Validation. Pengujian data dengan persentase terbaik yaitu 99% dengan menggunakan Use Training Set. Maka dari itu algoritma Naïve Bayes merupakan salah satu algoritma yang cukup efektif dan baik dalam perhitungan maupun hsil akhir.

**Keywords:**

Data Mining,  
Classification,  
Accreditation, Naïve  
Bayes Algorithm, WEKA

**Abstract**– Accreditation is the awarding of a predicate to a school with the aim of improving and also obtaining an overview of development, quality improvement and determining a school as an education provider. School accreditation is an important factor in improving the quality of education. Assessments are also carried out systematically and comprehensively by passing evaluations by authorized institutions. Schools will get an accreditation title if they meet predetermined standards. Classification is one of the data mining techniques that can be used in the process of classifying school accreditation. The test was carried out by using 2 options tests, namely the Use Training Set and 10 Fold Validation. Data testing with the best percentage is 99% using the Use Training Set. Therefore the Naïve Bayes algorithm is an algorithm that is quite effective and good in calculations and final results.

## 1. PENDAHULUAN

### 1.1 Latar Belakang Masalah

Klasifikasi *data mining* merupakan salah satu cara penempatan objek-objek ke salah satu dari beberapa kategori yang telah ditetapkan sebelumnya. Klasifikasi banyaknya digunakan untuk memprediksi kelas pada suatu label tertentu, yaitu dengan mengklasifikasi data (membangun model) berdasarkan training set dan nilai-nilai (label kelas) dalam mengklasifikasikan atribut tertentu dan menggunakannya dalam mengklasifikasi data yang baru[1]. Untuk dapat membandingkan serta memetakan mutu dari setiap satuan pendidikan, perlu dilakukan akreditasi bagi setiap lembaga dan program pendidikan. Proses ini dilakukan secara berkala dan terbuka dengan tujuan membantu dan memberdayakan satuan pendidikan agar mampu mengembangkan sumber daya dalam mencapai tujuan pendidikan nasional [2]. Setelah melakukan penelitian tersebut maka data akan dimasukkan ke file yang berbentuk kertas yang akan diinput oleh petugasnya. Data dengan skala besar tersebut butuh waktu lama dan panjang dalam melakukan pengelompokan data. Data yang sudah ada akan dikelompokkan kembali dan juga akan dihitung total pengelompokannya. Maka dari itu penerapan data mining ini dapat membantu instansi dalam melakukan pengelompokan dengan rumusan dan juga penerapannya. Karena banyaknya data dan juga variable penentu, tentunya akan menyulitkan tim penilai dalam melakukan penilalain dan tidak bisa dihindari jika terjadi kesalahan (human error) dalam proses penilaian. Prediksi tersbeut dapat diketahui dengan proses pendekatan matematik sehingga diketahui polanya, proses tersebut dinamakan data mining. Klasifikasi Bayes merupakan klasifikasi secara statistik, klasifikasi

ini dapat memprediksi peluang keanggotaan kelas seperti probabilitas suatu tupel merupakan milik kelas tertentu. Pada Naïve Bayes berasumsi bahwa efek dari nilai atribut pada kelas tertentu independen dari nilai-nilai dari atribut lainnya [3]. Akreditasi sekolah juga diartikan sebagai proses penilaian secara komprehensif terhadap kelayakan dan kinerja satuan dan/atau program pendidikan, yang dilakukan sebagai bentuk akuntabilitas publik. Akuntabilitas publik ini merupakan suatu bentuk pertanggungjawaban sekolah kepada publik, apakah layanan yang dilaksanakan dan diberikan oleh sekolah telah memenuhi harapan atau keinginan masyarakat [4].

Data mining adalah sebuah ilmu yang mempelajari alur kerja pengalihan data atributnya yang saling berkaitan untuk menemukan pengetahuan atau menemukan sebuah pola dari suatu data yang besar[5]. *Naïve Bayes* merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Algoritma menggunakan teorema Bayes dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas. Definisi lain mengatakan *Naïve Bayes* merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya [6]. Proses klasifikasi adalah metode untuk menemukan model atau fungsi yang dapat menjelaskan atau membedakan konsep kelas data, dengan tujuan untuk memprediksi kelas dari suatu objek yang tidak diketahui labelnya[7].

## 1.2 Penelitian Sejenis

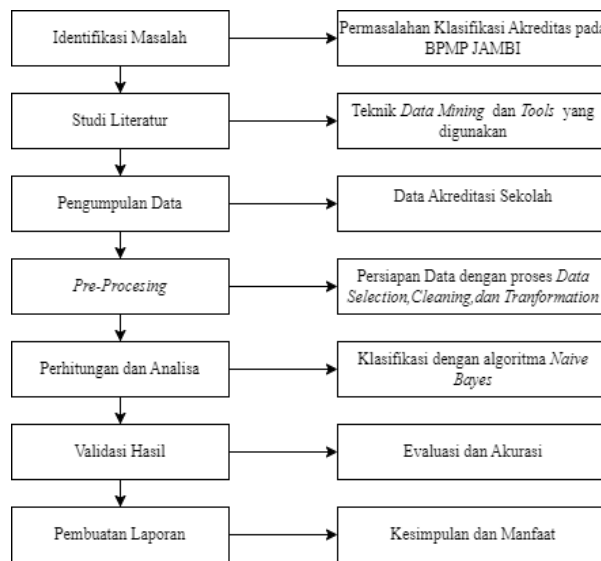
Sebagai pembanding perlu adanya penelitian sejenis untuk meninjau literatur antara penelitian sebelumnya, yaitu:

1. Penelitian yang dilakukan oleh Dodi Fernando Tambunan mengenai Klasifikasi Akreditasi SMA di Pulau Sumatera Menggunakan metode Naïve Bayes dengan menghasilkan nilai akurasi confusion matrix sebesar 94,165% [8].
2. Penelitian yang dilakukan oleh Pusphita Anna Octaviani, Yuciana Wilandari, dan Dwi Ispriyant mengenai Penerapan Metode Klasifikasi Support Vector Machine (SVM). Pada Data Akreditasi Sekolah Dasar (SD) di Kabupaten Magelang Pada pengujian dengan data testing sebanyak 82 data, akurasi klasifikasi yang didapat yaitu sebesar 93.902% menggunakan kernel Gaussian Radial Basic Function (RBF). Sedangkan menggunakan fungsi kernel Polynomial akurasi klasifikasinya adalah sebesar 92.683 % [9].
3. Penelitian yang dilakukan oleh Afifiudin Khadamsyah Ispriyant mengenai Sistem Pendukung Keputusan Penetapan Peserta Sertifikasi Guru Menggunakan Naïve Bayes diperoleh rating pendukung keputusan penetapan peserta sertifikasi guru menggunakan Naïve Bayes dihitung berdasarkan kecocokan dalam 22 daftar peserta yang dinyatakan lulus sebesar 95% [10].
4. Penelitian yang dilakukan oleh Nurul Alfiah mengenai Klasifikasi Penerima Bantuan Sosial Program Keluarga Harapan Menggunakan Metode Naive Bayes penelitian ini membahas tentang komparasi antara algoritma dan Naive Bayes untuk mengklasifikasikan. Tingkat kemiskinan dengan 14 atribut dan jumlah data yang telah di Cleaning dengan Metode Naive Bayes [11].
5. Penelitian yang dilakukan oleh Fitria Septianingrum, Agung Susilo Yuda Irawan mengenai Metode Seleksi Fitur Untuk Klasifikasi Sentimen Menggunakan Algoritma Naive Bayes: Sebuah Literature Review sistem klasifikasi adalah 23 reaksi benar dan 2 reaksi salah. Berdasarkan hasil dengan menggunakan metode Confusion Matrix menunjukkan bahwa nilai akurasi mencapai 92%, nilai presisi mencapai 100%, nilai recall mencapai 91% [12].

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

Kerangka kerja penelitian sebuah tahapan-tahapan yang dilakukan selama mengerjakan penelitian dan disusun secara sistematis. Kerangka kerja penelitian dibuat agar mempermudah dalam pencapaian hasil penelitian, dapat menyelesaikan penelitian tepat waktu, dan berjalan sesuai dengan yang diharapkan. Adapun kerangka kerja penelitian yang digunakan adalah sebagai berikut:



**Gambar 2.1 Tahap Penelitian**

Berdasarkan kerangka kerja penelitian yang telah digambarkan di atas, maka dapat diuraikan pembahasan masing-masing tahap dalam penelitian adalah sebagai berikut:

#### 1. Identifikasi Masalah

Pada tahap ini penulis melakukan identifikasi masalah mengenai klasifikasi akreditasi sekolah di BPMP Provinsi Jambi. Masih dapat banyak kekliruan dalam pengklasifikasian data akreditasi sekolah. Sehingga penulis dapat menentukan rencana kerja serta menentukan data apa yang akan dibutuhkan dalam penelitian ini.

#### 2. Studi Literatur

Pada tahap ini penulis melakukan pencarian terhadap landasan-landasan teori yang diperoleh dari berbagai sumber seperti buku, jurnal ilmiah, dan lain sebagainya yang dapat dijadikan sebagai referensi yang baik dan relevan yang dapat dijadikan sebagai acuan. Studi literatur ini bertujuan untuk mendapatkan landasan teoritis mengenai permasalahan yang akan diteliti seperti teknik data mining. Proses klasifikasi menggunakan algoritma Naïve Bayes. Cara ini digunakan agar peneliti dapat memahami permasalahan yang diteliti dengan benar sesuai dengan pembahasan yang dilakukan.

#### 3. Pengumpulan Data

Pada tahap ini penulis melakukan pencarian dan pengumpulan data informasi yang dibutuhkan dalam penelitian. Metode pengumpulan data yang digunakan peneliti dalam penelitian ini dengan melakukan wawancara secara langsung kepada pihak terkait sehingga peneliti dapat mengambil dan juga mengumpulkan informasi yang diperlukan.

#### 4. Pre-Processing

Tahapan preprocessing akan dilakukan tahapan sebagai berikut:

##### a. Data Selection

Pemilihan (seleksi) data baru sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam Data mining dimulai. Penyeleksian ini berupa penghilangan atribut yang tidak perlu ataupun tidak dipakai.

##### b. Data Cleaning

Pada tahapan preprocessing terdapat tahapan data cleaning dimana pada tahapan ini melakukan pembersihan data antara lain data yang tidak dipakai dan data yang kosong. Cleaning data dilakukan untuk memisahkan dan juga membersihkan noise data yang tidak konsisten pada data akreditasi sekolah.

c. Data Transformation

Pada tahapan ini dilakukan pengurangan dimensi atau feature yang tidak dipakai atau tidak valid. Pada feature selection yang digunakan pada penelitian ini adalah information gain.

5. Perhitungan dan Analisis

Pada tahap ini dilakukan proses klasifikasi akreditasi sekolah menggunakan algoritma naïve bayes dan menganalisis data akreditasi BPMP Provinsi.Jambi yang telah dipilih sehingga data tersebut sesuai untuk proses dalam Data mining. Hasil yang didapatkan berupa adalah akreditas A, B, dan C suatu sekolah mendapatkan akreditasi.

6. Validasi Hasil

Validasi hasil menunjukkan kedekatan hasil pengukuran dengan nilai akurasi.

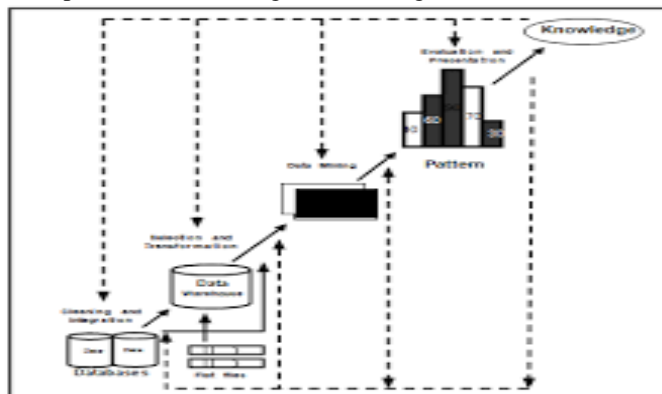
7. Pembuatan Laporan

Hasil dari analisis akan dinilai keakuratan data model yang sudah didapatkan dengan perhitungnan naïve bayes dan setelah semua tahapan penelitian dilakukan, kama akan dibuat laporan sebagai dokumentasi penellitian agar dapat dimanfaatkan pada waktu yang akan datang baik oleh peneliti sendiri maupun peneliti selanjutnya.

## 2.2 Tahapan Data Mining

Klasifikasi merupakan sebuah proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep dan kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang kelasnya tidak diketahui. Data training adalah data yang siap untuk di-mining yang telah melewati data preprocessing. Sedangkan data testing adalah data yang digunakan untuk menguji rule klasifikasi yang diperoleh dari data training[13].

Adapun tahapan-tahapan dari data mining adalah sebagai berikut[14]:



Gambar 2.2 Tahapan Data Mining

1. Pembersihan data (data cleaning)

Sebelum proses data mining dapat dilaksanakan, perlu dilakukan proses cleaning pada data yang menjadi fokus KDD. Proses cleaning mencakup antara lain membuang 4 duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi). Juga dilakukan proses enrichment, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal.

2. Integrasi data (data integration)

Integrasi data merupakan penggabungan data dari berbagai database ke dalam satu database baru. Tidak jarang data yang diperlukan untuk data mining tidak hanya berasal dari satu database tetapi juga berasal

dari beberapa database atau file teks. Integrasi data dilakukan pada atribut-atribut yang mengidentifikasi entitas-entitas yang unik seperti atribut nama, jenis produk, nomor pelanggan dan lainnya. Integrasi data perlu dilakukan secara cermat karena kesalahan pada integrasi data bisa menghasilkan hasil yang menyimpang dan bahkan menyesatkan pengambilan aksi nantinya. Sebagai contoh bila integrasi data berdasarkan jenis produk ternyata menggabungkan produk dari kategori yang berbeda maka akan didapatkan korelasi antar produk yang sebenarnya tidak ada.

### 3 Seleksi Data (Data Selection)

Data yang ada pada database sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari database. Sebagai contoh, sebuah kasus yang meneliti faktor kecenderungan orang membeli dalam kasus market basket analysis, tidak perlu mengambil nama pelanggan, cukup dengan id pelanggan saja.

### 4 Transformasi data (Data Transformation)

Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam data mining. Beberapa metode data mining membutuhkan format data yang khusus sebelum bisa diaplikasikan. Sebagai contoh beberapa metode standar seperti analisis asosiasi dan clustering hanya bisa menerima input data kategorikal. Karenanya data berupa angka numerik yang berlanjut perlu dibagi-bagi menjadi beberapa interval. Proses ini sering disebut transformasi data.

### 5 Proses mining

Sebuah proses yang paling utama pada saat metode diterapkan untuk mencari pengetahuan tersembunyi dan berharga dari data.

### 6 Evaluasi pola (pattern evaluation)

Untuk mengidentifikasi pola-pola menarik kedalam knowledge based yang ditemukan. Dalam tahap ini hasil dari teknik data mining berupa pola-pola yang khas maupun model prediksi dievaluasi untuk menilai apakah hipotesa yang ada memang tercapai.

### 7 Presentasi pengetahuan (knowledge presentation)

Merupakan penyajian dan visualisasi pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna. Tahap terakhir dari proses data mining adalah bagaimana memformulasikan keputusan atau aksi dari hasil analisis yang didapat. Ada kalanya hal ini harus melibatkan orang-orang yang tidak memahami data mining. Karenanya presentasi hasil data mining dalam bentuk pengetahuan yang bisa dipahami semua orang adalah satu tahapan yang diperlukan dalam proses data mining. Dalam presentasi ini, visualisasi juga bisa membantu mengkomunikasikan hasil data mining.

Berdasarkan pengertian diatas dapat disimpulkan bahwa data mining merupakan proses ekstraksi atau pengalihan data yang kita peroleh menjadi informasi dengan menggunakan metode statistika, matematika, kecerdasan buatan (Artificial Intelligence), dan machine learning untuk mengidentifikasi informasi yang bermanfaat dalam pengambilan keputusan.

## 2.3 ALGORITMA NAÏVE BAYES

*Naïve Bayes* merupakan teknik prediksi berbasis probabilitas sederhana yang berdasar pada penerapan aturan bayes dengan asumsi ketidaktergantungan yang kuat. Selain itu *Naïve Bayes* juga dapat menganalisis variabel-variabel yang paling mempengaruhinya dalam bentuk peluang [15].

Persamaan dari teorema Bayes adalah :

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

Di mana :

- X : Data dengan class yang belum diketahui
- H : Hipotesis data X merupakan suatu class spesifik
- P(H|X) : Probabilitas hipotesis H berdasarkan kondisi x (posteriori prob.)
- P(H) : Probabilitas hipotesis H (prior prob.)
- P(X|H) : Probabilitas X berdasarkan kondisi tersebut
- P(X) : Probabilitas dari X [16]

## 2.4 WEKA

WEKA merupakan *software* terintegrasi yang berisi implementasi dari metode-metode *data mining* WEKA dikembangkan oleh Universitas Waikato, Selandia Baru menggunakan Bahasa pemrograman *Java*. Oleh karena itu, WEKA merupakan singkatan dari *Waikato Environment For Knowledge Analysis*. Dengan mengadopsi konsep *open source software* menjadikan WEKA dapat digunakan dan dimodifikasi siapapun secara gratis [17].

## 3. HASIL DAN PEMBAHASAN

### 3.1 Representasi Data

Pada penelitian ini terdapat 350 data sekolah jenjang SD,SMP,SMA dan SMK dan 11 atribut. Berikut merupakan profil data penelitian.

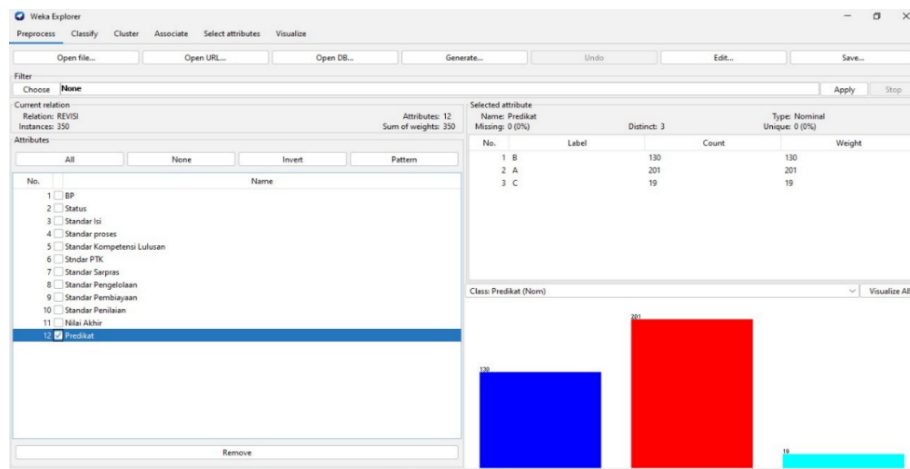
Tabel 3.1 Data Akreditasi Sekolah

NO.	Nama Sekolah	NPSN	BP	Status	Standar Isi	Standar proses	Standar Kompetensi Lulusar	Standar PTK	Standar Sarpras	Standar Pengelolaan	Standar Pembiayaan	Standar Penilaian	Nilai Akhir	Predikat
1	SD NEGERI 0201V JAMBI	10504894	SD	Negeri	BAIK	BAIK	BAIK	BAIK	CUKUP	BAIK	BAIK	BAIK	BAIK	B
2	SD NEGERI 0401V JAMBI	10504909	SD	Negeri	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	BAIK	SANGAT BAIK	BAIK	SANGAT BAIK	SANGAT BAIK	A
3	SD NEGERI 0501V JAMBI	10504823	SD	Negeri	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	A
4	SD NEGERI 0701V JAMBI	10504808	SD	Negeri	SANGAT BAIK	SANGAT BAIK	BAIK	BAIK	CUKUP	BAIK	SANGAT BAIK	SANGAT BAIK	BAIK	B
5	SD NEGERI 1271V JAMBI	10504509	SD	Negeri	BAIK	BAIK	BAIK	BAIK	BAIK	BAIK	SANGAT BAIK	BAIK	BAIK	B
6	SD ISLAM ASAD JAMBI	10504472	SD	Swasta	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	A
7	SMP NEGERI 003 JAMBI	10504638	SMP	Negeri	BAIK	BAIK	CUKUP	BAIK	BAIK	SANGAT BAIK	SANGAT BAIK	BAIK	BAIK	B
8	SMAN 7 KOTA JAMBI	10504583	SMA	Negeri	BAIK	CUKUP	SANGAT BAIK	BAIK	KURANG	CUKUP	SANGAT BAIK	BAIK	BAIK	B
9	SMK NEGERI 5 KOTA JAMBI	10505934	SMK	Negeri	CUKUP	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	CUKUP	SANGAT BAIK	BAIK	SANGAT BAIK	BAIK	B
10	SD NEGERI 0251V JAMBI	10504918	SD	Negeri	SANGAT BAIK	SANGAT BAIK	BAIK	SANGAT BAIK	BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	A
11	SD NEGERI 0361V JAMBI	10504914	SD	Negeri	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	A
12	SD NEGERI 0531V JAMBI	10504807	SD	Negeri	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	BAIK	SANGAT BAIK	SANGAT BAIK	BAIK	SANGAT BAIK	A
13	SD NEGERI 0631V JAMBI	10504831	SD	Negeri	SANGAT BAIK	SANGAT BAIK	BAIK	BAIK	CUKUP	BAIK	SANGAT BAIK	BAIK	BAIK	B
14	SD NEGERI 0641V JAMBI	10504832	SD	Negeri	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	A
15	SD NEGERI 0651V JAMBI	10504833	SD	Negeri	CUKUP	BAIK	BAIK	CUKUP	BAIK	CUKUP	CUKUP	CUKUP	CUKUP	C
16	SD NEGERI 0961V JAMBI	10504864	SD	Negeri	BAIK	BAIK	BAIK	BAIK	CUKUP	BAIK	SANGAT BAIK	BAIK	BAIK	B
17	SD NEGERI 0971V JAMBI	10504852	SD	Negeri	BAIK	SANGAT BAIK	BAIK	BAIK	BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	BAIK	B
18	SD NEGERI 0981V JAMBI	10504851	SD	Negeri	SANGAT BAIK	SANGAT BAIK	CUKUP	BAIK	BAIK	SANGAT BAIK	BAIK	BAIK	BAIK	B
19	SD NEGERI 0991V JAMBI	10504850	SD	Negeri	BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	A
20	SD NEGERI 1041V JAMBI	10504842	SD	Negeri	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	BAIK	BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	A
21	SD NEGERI 1321V JAMBI	10504497	SD	Negeri	CUKUP	BAIK	BAIK	BAIK	CUKUP	BAIK	SANGAT BAIK	BAIK	BAIK	B
22	SD NEGERI 1401V JAMBI	10504524	SD	Negeri	BAIK	BAIK	SANGAT BAIK	SANGAT BAIK	BAIK	SANGAT BAIK	SANGAT BAIK	BAIK	BAIK	B
23	SD NEGERI 1461V JAMBI	10504541	SD	Negeri	SANGAT BAIK	BAIK	SANGAT BAIK	BAIK	BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	A
24	SD NEGERI 1481V JAMBI	10504543	SD	Negeri	BAIK	BAIK	BAIK	KURANG	CUKUP	SANGAT BAIK	SANGAT BAIK	BAIK	BAIK	B
25	SD NEGERI 1761V JAMBI	10504450	SD	Negeri	BAIK	BAIK	KURANG	CUKUP	BAIK	KURANG	CUKUP	CUKUP	CUKUP	C
26	SD NEGERI 2001V JAMBI	10504464	SD	Negeri	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	BAIK	BAIK	BAIK	SANGAT BAIK	SANGAT BAIK	A
27	SD NEGERI 2061V JAMBI	10504481	SD	Negeri	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	BAIK	BAIK	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	A

### 3.2 Data Cleaning

Sebelum proses data mining dapat dilaksanakan perlu dilakukan tahap preprocessing. Adapun tujuan dari pre-processing data yaitu tahapan untuk menghilangkan beberapa permasalahan yang bisa mengganggu saat pemrosesan data. Data preprocessing merupakan teknik paling awal sebelum melakukan data mining. Pada tahap praproses, data yang akan dijadikan data latih dianggap sebagai data mentah ,yang mana pada tahap ini dilakukan pengecekan apakah telah lengkap atau belum. Yang mana jika belum lengkap maka akan dilakukan Replace Missing Value dan dilakukan data cleaning. Dan pada data yang digunakan peneliti tidak adanya missing value.





Gambar 3.1 Remove Missing Value

### 3.3 Hasil Klasifikasi Perhitungan Data

1. Hasil Klasifikasi *Naïve Bayes* menggunakan WEKA (*Use Training Set*), Berikut ini adalah test dengan menggunakan Training Set yang akan menggunakan data training.

Tabel 3.2 Hasi Klasifikasi *Naïve Bayes* dengan WEKA (*Use Training Set*)

TP ATE	FP ATE	PRECISION	RECALL	F-MEASURE	MCC	ROC REA	PRC REA	CLASS
0.985	0.005	0.992	0.985	0.989	0.982	0.99	0.992	B
0.995	0.013	0.99	0.995	0.993	0.983	0.991	0.987	A
1	0	1	1	1	1	1	1	C
0.991	0.009	0.991	0.991	0.991	0.983	0.991	0.99	W.AVG

2. Hasil Klasifikasi *Naïve Bayes* Menggunakan WEKA 10 Cross-Validation) Test 5 Cross-Validation adalah pengujian data dimana data training dibagi menjadi k subset (sub himpunan)

Tabel 3.3 Confusion Matrix 10 Fold Cross Validation Menggunakan Tools WEKA

TP ATE	FP ATE	PRECISION	RECALL	F-MEASURE	MCC	ROC REA	PRC REA	CLASS
0.985	0.009	0.985	0.985	0.985	0.976	0.988	0.991	B
0.990	0.013	0.990	0.990	0.990	0.977	0.990	0.985	A
1	0	1	1	1	1	1	1	C
0.989	0.011	0.989	0.989	0.989	0.978	0.990	0.988	W.AVG

3. Berikutnya merupakan hasil dari analisis klasifikasi dengan menggunakan Metode *Naïve Bayes* padatools WEKA. Pada penggunaan tools WEKA penulis menggunakan 3 options yaitu Use Training Set.

Tabel 3.4 *Summary Use Training Set* Klasifikasi Naive Bayes Dengan Tools WEKA

Mode EVALUASI	Akurasi	Jumlah	Persentase
<i>Use Training Set</i>	<i>Correctly Classified Instance</i>	347	99,14%
	<i>Incorrectly Classified Instance</i>	3	0,8571%
<i>10 Fold Cross Validation</i>	<i>Correctly Classified Instance</i>	346	98,85%
	<i>Incorrectly Classified Instance</i>	4	1,1429%
Perhitungan Excel		350	98%

4. Tabel Perbandingan Evaluasi Akurasi

Tabel 3.5 Perbandingan Evaluasi Akurasi

<i>Data Summary</i>	<i>Nilai</i>	<i>Persentase</i>
<i>Correctly Classified Instance</i>	347	99.14%
<i>Incorrectly Classified Instance</i>	3	0.8571%
<i>Kappa Statistic</i>	0.9839	
<i>Mean Absolute Error</i>	0.0102	
<i>Root Mean Square Error</i>	0.0758	
<i>Relative Absolute Error</i>	2.8899%	
<i>Root Relative Squared Error</i>	18.0291%	
<i>Total Number of Instance</i>	350	

## 4.KESIMPULAN

Penelitian ini menggunakan data akreditasi sekolah Kota Jambi. Atribut yang di pakai yaitu Bentuk Pendidikan, Status, Standar Kompetensi, Standar PTK, Standar Sarpras, Standar Pengelolaan, Standar Pembiayaan, Standar Penilaian, dan Nilai Akhir. Hasil klasifikasi Siswa Penerima Beasiswa menggunakan algoritma naïve bayes pada tool Rapidminer diperoleh sebanyak 201 pada kelas “A”,130 pada kelas “B” dan 19 pada kelas “C”, sedangkan pada Excel diperoleh sebanyak 201 pada kelas “A”,130 pada kelas “B” dan 19 pada kelas “C”.

.Persentasi hasil Akurasi klasifikasi naïve bayes dengan menggunakan 3 test options. Menunjukkan nilai akurasi terbesar yaitu pada test options Use Training Set dengan hasil persentasi akurasi sebesar 99.14%, Pada 10 Fold Cross Validation dengan hasil persentasi akurasi sebesar 98,85%, dan Pada Excel diperoleh hasil persentasi akurasi 99%. Dari Persentasi diatas dengan Acuraccy 99% Sangat Baik.



## REFERENCES

- [1] M. S. Mustafa, M. R. Ramadhan, and A. P. Thenata, "Implementasi Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier," *Creat. Inf. Technol. J.*, vol. 4, no. 2, p. 151, 2018, doi: 10.24076/citec.2017v4i2.106.
- [2] P. N. Indriyawati, "Pengaruh Akreditasi Sekolah Terhadap Kualitas Lulusan Sma Se Kota Jambi Tahun 2017," ... *Akreditasi Sekol. Terhadap Kualitas ...*, pp. 1–9, 2017, [Online]. Available: <https://repository.unja.ac.id/5217/%0Ahttps://repository.unja.ac.id/5217/1/ARTIKEL.pdf>
- [3] H. F. Putro, R. T. Vuldari, and W. L. Y. Saptomo, "Penerapan Metode Naive Bayes Untuk Klasifikasi Pelanggan," *J. Teknol. Inf. dan Komun.*, vol. 8, no. 2, 2020, doi: 10.30646/tikomsin.v8i2.500.
- [4] A. A. R. Awaludin, "Akreditasi Sekolah sebagai Suatu Upaya Penjaminan Mutu Pendidikan di Indonesia," *SAP (Susunan Artik. Pendidikan)*, vol. 2, no. 1, pp. 12–21, 2017, doi: 10.30998/sap.v2i1.1156.
- [5] A. Ridwan, "Penerapan Algoritma Naive Bayes Untuk Klasifikasi Penyakit Diabetes Mellitus," *J. SISKOM-KB (Sistem Komput. dan Kecerdasan Buatan)*, vol. 4, no. 1, pp. 15–21, 2020, doi: 10.47970/siskom-kb.v4i1.169.
- [6] Darma Juang, "Analisis Spam Dengan Menggunakan Naive Bayes," *J. Teknovasi*, vol. 03, no. 1998, pp. 51–57, 2016, [Online]. Available: <http://www.ejurnal.plm.ac.id/index.php/Teknovasi/article/view/69%0Ahttps://www.ejurnal.plm.ac.id/index.php/Teknovasi/article/download/69/58>
- [7] G. A. Rosso, "Milton," *William Blake Context*, no. September, pp. 184–191, 2019, doi: 10.1017/9781316534946.021.
- [8] D. F. Tambunan, "Bayes Classification of High School Accreditation in Sumatera Island Using Naive Bayes Method Faculty of Science and Technology," p. 91, 2020.
- [9] P. A. Octaviani, Yuciana Wilandari, and D. Ispriyanti, "Penerapan Metode Klasifikasi Support Vector Machine (SVM) pada Data Akreditasi Sekolah Dasar (SD) di Kabupaten Magelang," *J. Gaussian*, vol. 3, no. 8, pp. 811–820, 2014, [Online]. Available: [http://download.portalgaruda.org/article.php?article=286497&val=4706&title=PENERAPAN METODE KLASIFIKASI SUPPORT VECTOR MACHINE \(SVM\) PADA DATA AKREDITASI SEKOLAH DASAR \(SD\) DI KABUPATEN MAGELANG](http://download.portalgaruda.org/article.php?article=286497&val=4706&title=PENERAPAN METODE KLASIFIKASI SUPPORT VECTOR MACHINE (SVM) PADA DATA AKREDITASI SEKOLAH DASAR (SD) DI KABUPATEN MAGELANG)
- [10] A. Kadham, "Sistem Pendukung Keputusan Penetapan Peserta Sertifikasi Guru Menggunakan Naive Bayes," *Tek. Inform.*, p. 95, 2013, [Online]. Available: [file:///C:/Users/Icha\\_Tahir/Documents/SEMESTER 6/PROPOSAL/icha/07650074.pdf](file:///C:/Users/Icha_Tahir/Documents/SEMESTER 6/PROPOSAL/icha/07650074.pdf)
- [11] N. Alfiah, "Klasifikasi Penerima Bantuan Sosial Program Keluarga Harapan Menggunakan Metode Naive Bayes," *Respati*, vol. 16, no. 1, p. 32, 2021, doi: 10.35842/jtir.v16i1.386.
- [12] F. Septianingrum and A. S. Y. Irawan, "Metode Seleksi Fitur Untuk Klasifikasi Sentimen Menggunakan Algoritma Naive Bayes: Sebuah Literature Review," *J. Media Inform. Budidarma*, vol. 5, no. 3, p. 799, 2021, doi: 10.30865/mib.v5i3.2983.
- [13] I. Budiman, Muliadi, and R. Ramadina, "Penerapan Fungsi Data Mining Klasifikasi untuk Prediksi Masa Studi Mahasiswa Tepat Waktu pada Sistem Informasi Akademik Perguruan Tinggi," *J. Jupiter*, vol. 7, no. 1, pp. 39–50, 2015.
- [14] J. Eska, "Penerapan Data Mining Untuk Prekdiksi Penjualan Wallpaper Menggunakan Algoritma C4.5 STMIK Royal Ksieran," *JURTEKSI (Jurnal Teknol. dan Sist. Informasi)*, vol. 2, pp. 9–13, 2016.
- [15] R. Retnosari, "Analisa Kelayakan Kredit Usaha Mikro Berjalan pada Perbankan dengan Metode Naive Bayes," *PROSISKO J. Pengemb. Ris. dan Obs. Sist. Komput.*, vol. 8, no. 1, pp. 53–59, 2021, [Online]. Available: <https://e-jurnal.lppmunsera.org/index.php/PROSISKO/article/view/2848>
- [16] D. P. Utomo and M. Mesran, "Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung," *J. Media Inform. Budidarma*, vol. 4, no. 2, p. 437, 2020, doi: 10.30865/mib.v4i2.2080.
- [17] S. Adinugroho and Y. A. Sari, *Implementasi Data Mining Menggunakan WEKA*. 2018.