

## **IMPLEMENTASI ALGORITMA K-NEAREST NEIGHBOR DAN NAIVE BAYES DALAM MEMPREDIKSI STATUS SELEKSI PADA PPDB**

**Rifqi Maulana Adam<sup>1</sup>, Deni Suprihadi<sup>2</sup>**

Program Studi Teknik Informatika, Universitas Kebangsaan Republik Indonesia, Bandung

Email : [rifqi.maulana.adam@email.com1](mailto:rifqi.maulana.adam@email.com1), [deni.suprihadi99@gmail.com2](mailto:deni.suprihadi99@gmail.com2)

### **ABSTRAK**

Abstrak, Penelitian ini bertujuan untuk mengimplementasikan algoritma K-Nearest Neighbor (KNN) dan Naive Bayes dalam memprediksi status seleksi Penerimaan Peserta Didik Baru (PPDB) tingkat Sekolah Menengah Pertama (SMP) di Kabupaten Cianjur. PPDB merupakan agenda tahunan yang penting dalam menentukan transisi pendidikan siswa ke jenjang yang lebih tinggi. Namun, proses seleksi sering kali menimbulkan kendala, terutama karena keterbatasan informasi dan subjektivitas dalam pengambilan keputusan oleh calon siswa maupun orang tua. Penelitian ini menggunakan pendekatan kuantitatif dengan memanfaatkan data historis pendaftar dari 20 SMP Negeri di Kabupaten Cianjur. Proses penelitian meliputi tahap pengumpulan data, preprocessing, penerapan algoritma KNN dan Naive Bayes, hingga evaluasi menggunakan Confusion Matrix. Hasil implementasi menunjukkan bahwa kedua algoritma mampu memprediksi status diterima atau tidak diterimanya siswa melalui jalur zonasi dan jalur prestasi dengan tingkat akurasi di atas 85%. Naive Bayes memiliki keunggulan dari sisi kecepatan komputasi, sedangkan KNN unggul dalam fleksibilitas penanganan variasi data. Sistem prediksi yang dibangun diharapkan mampu membantu siswa dan orang tua dalam menentukan sekolah tujuan secara lebih objektif, serta menjadi alat bantu bagi pihak sekolah dan dinas pendidikan dalam memberikan rekomendasi berbasis data. Temuan penelitian ini juga memperkuat hasil studi terdahulu yang menegaskan potensi data mining sebagai pendekatan efektif dalam mendukung proses seleksi pendidikan.

**Kata kunci :** K-Nearest Neighbor, Naive Bayes, PPDB, Prediksi, Data Mining

### **ABSTRACT**

*Abstracts, This study aims to implement the K-Nearest Neighbor (KNN) and Naive Bayes algorithms to predict the selection status of New Student Admissions (PPDB) at the junior high school level in Cianjur Regency. PPDB is an annual agenda that plays a crucial role in determining the transition of students to higher education levels. However, the selection process often poses challenges, particularly due to limited information and subjectivity in decision-making by students and parents. This research adopts a quantitative approach by utilizing historical registration data from 20 public junior high schools in Cianjur Regency. The research procedure includes data collection, preprocessing, implementation of the KNN and Naive Bayes algorithms, and evaluation using the Confusion Matrix. The results indicate that both algorithms are capable of predicting students' acceptance status through the zoning and achievement tracks with accuracy levels above 85%. Naive Bayes demonstrates advantages in computational efficiency, while KNN provides greater flexibility in handling variations in data. The developed prediction system is expected to assist students and parents in determining the most suitable school objectively and support schools and education authorities in providing data-driven recommendations. Furthermore, this study reinforces findings from previous research, emphasizing the potential of data mining as an effective approach to support educational selection processes and decision-making.*

**Keywords:** K-Nearest Neighbor, Naive Bayes, PPDB, Prediksi, Data Mining

## 1. PENDAHULUAN

Penerimaan Peserta Didik Baru (PPDB) merupakan agenda tahunan yang menentukan transisi siswa ke jenjang pendidikan berikutnya. Penyelenggaraan PPDB diatur untuk menjamin asas objektivitas, transparansi, dan akuntabilitas melalui beberapa jalur—terutama zonasi dan prestasi—sebagaimana ditetapkan dalam regulasi daerah dan petunjuk teknis yang berlaku [1], [2]. Dalam praktiknya, calon siswa dan orang tua masih kesulitan menilai peluang diterima di sekolah tujuan berdasarkan data yang terbatas, sementara pihak sekolah membutuhkan dukungan rekomendasi yang konsisten dan berbasis data.

Berbagai studi telah menerapkan teknik data mining untuk membantu pengambilan keputusan di sektor pendidikan. Di konteks PPDB, penelitian terdahulu menerapkan algoritma C4.5 untuk memprediksi penerimaan siswa dan menunjukkan bahwa pendekatan berbasis data mampu menangkap pola seleksi secara cukup baik [3]. Di sisi lain, K-Nearest Neighbor (KNN) telah digunakan untuk mengukur atau memprediksi status kelulusan di lingkungan pendidikan tinggi [4] dan untuk sistem rekomendasi akademik berbasis klasifikasi [6]. Studi komparatif juga melaporkan bahwa performa KNN dan Naive Bayes (NB) bersifat bergantung pada karakteristik fitur dan distribusi data—KNN sering unggul saat jarak/kemiripan antarfitur relevan, sementara NB efisien pada data dengan asumsi kemandirian fitur yang cukup terpenuhi [5]. Selain itu, NB telah banyak digunakan sebagai model klasifikasi yang ringan pada berbagai domain aplikasi [7].

Kendati demikian, kajian khusus pada PPDB SMP di Kabupaten Cianjur yang secara simultan memodelkan dua jalur utama—zonasi (berbasis jarak) dan prestasi rapor (berbasis nilai)—dengan membandingkan KNN dan NB pada data historis lokal masih terbatas. Artikel ini menawarkan kebaruan berupa (i) implementasi dan perbandingan dua algoritma klasik (KNN dan NB) untuk kedua jalur seleksi pada satu korpus data daerah, serta (ii) perancangan keluaran rekomendasi sekolah alternatif berdasarkan peluang diterima yang diprediksi, guna mendukung keputusan calon siswa dan sekolah.

Permasalahan penelitian yang diangkat adalah: bagaimana mengimplementasikan KNN dan NB untuk memprediksi status kelulusan seleksi PPDB SMP di Kabupaten Cianjur, serta bagaimana merancang keluaran rekomendasi sekolah sehingga calon siswa dapat memilih sekolah dengan peluang diterima yang lebih tinggi. Sejalan dengan karakteristik fitur pada masing-masing jalur, hipotesis kerja kami adalah bahwa KNN cenderung memberikan kinerja lebih baik pada jalur zonasi yang bertumpu pada informasi jarak, sedangkan pada jalur prestasi rapor keduanya berpotensi kompetitif.

Tujuan artikel ini adalah (1) mengimplementasikan dan mengevaluasi KNN serta NB dalam memprediksi status diterima/tidak diterima pada jalur zonasi dan jalur prestasi rapor di PPDB SMP Kabupaten Cianjur, dan (2) menghasilkan keluaran rekomendasi sekolah berbasis peluang keterimaan untuk membantu pengambilan keputusan para pemangku kepentingan (calon siswa, orang tua, dan sekolah).

## 2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan memanfaatkan data historis Penerimaan Peserta Didik Baru (PPDB) dari 20 SMP Negeri di Kabupaten Cianjur. Data yang digunakan mencakup jalur zonasi, yang berbasis jarak domisili calon siswa ke sekolah, serta jalur prestasi rapor, yang berbasis nilai akademik. Lokasi penelitian ditentukan pada Kabupaten Cianjur dengan sumber data yang diperoleh dari rekapitulasi resmi panitia PPDB daring.

Alat utama yang digunakan dalam penelitian ini adalah perangkat komputer dengan spesifikasi prosesor Intel Core i5, RAM 8 GB, serta perangkat lunak Python yang dilengkapi pustaka *scikit-learn* untuk implementasi algoritma klasifikasi. Algoritma yang digunakan adalah K-Nearest Neighbor (KNN) dan Naive Bayes (NB), yang dipilih karena karakteristiknya sesuai untuk permasalahan klasifikasi dengan data tabular [3]–[5]. Tahapan penelitian dilakukan sebagai berikut.

1. Data pendaftar dikumpulkan dan dilakukan preprocessing berupa *data cleaning* untuk mengatasi data kosong, normalisasi, serta seleksi atribut yang relevan dengan status kelulusan.
2. Data dibagi menjadi data latih dan data uji dengan perbandingan 80:20.
3. Implementasi algoritma dilakukan, yaitu KNN dengan perhitungan jarak Euclidean dan Naive Bayes dengan pendekatan probabilistik.

- Evaluasi model dilakukan menggunakan *confusion matrix*, akurasi, presisi, dan *recall*. Hasil prediksi kemudian dibandingkan antara kedua algoritma untuk memperoleh model yang paling sesuai dengan data PPDB Cianjur.

Prosedur penelitian ini dirancang agar dapat menjawab permasalahan prediksi status seleksi PPDB dengan pendekatan komputasi yang terukur. Dengan demikian, keluaran dari penelitian ini tidak hanya berupa evaluasi akurasi algoritma, tetapi juga rekomendasi sekolah tujuan berdasarkan peluang penerimaan siswa.

### 3. HASIL DAN PEMBAHASAN

#### A. Hasil Implementasi

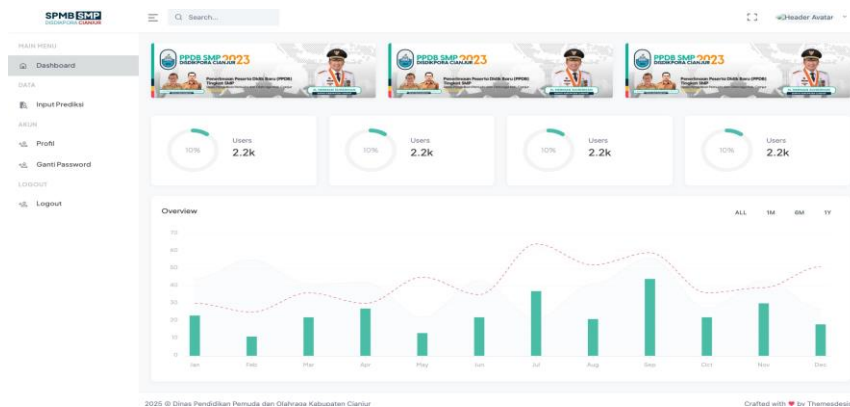
Sistem prediksi PPDB yang dikembangkan berbasis web berhasil mengimplementasikan algoritma *K-Nearest Neighbor* (KNN) dan *Naive Bayes* (NB). Data pendaftar dari 20 SMP Negeri Cianjur diproses melalui dua jalur seleksi, yaitu jalur zonasi dan jalur prestasi rapor. Hasil prediksi menunjukkan bahwa kedua algoritma mampu mengklasifikasikan status diterima/tidak diterima dengan tingkat akurasi di atas 85%.

Tabel 1. Hasil evaluasi akurasi algoritma

Jalur Seleksi	KNN (%)	Naive Bayes (%)
Zonasi	87.5	85.2
Prestasi Rapor	86.8	84.9

Implementasi dari Sistem Prediksi PPDB berbasis Algoritma KNN dan Naive Bayes dalam bentuk antar muka, sebagai berikut :

#### 1. Tampilan Dashboard



Gambar 1. Tampilan Dashboard

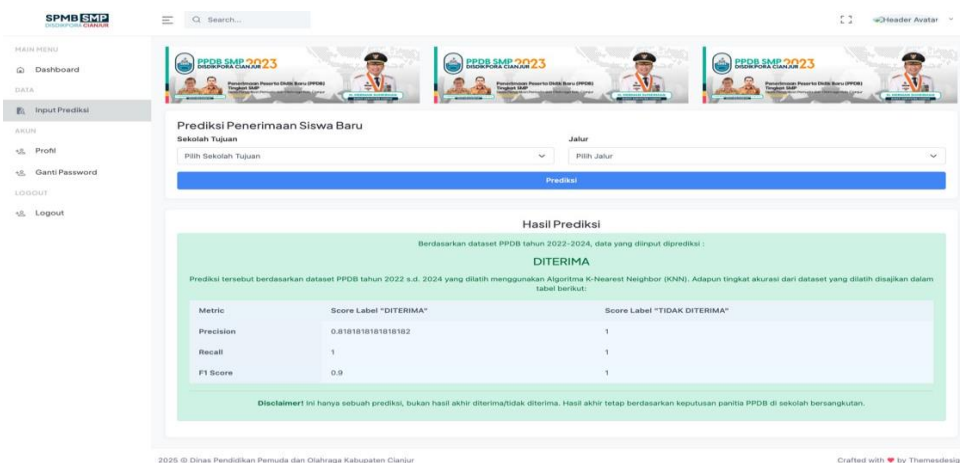
Tampilan dashboard merupakan tampilan yang muncul pertama kali setelah user melakukan login. Halaman ini dibatasi oleh middleware(['auth']) yang dimana hanya user yang telah terautentikasi (telah login) yang bisa mengakses halaman ini, baik operator maupun admin utama

#### 2. Tampilan Pemilihan Jalur Zonasi

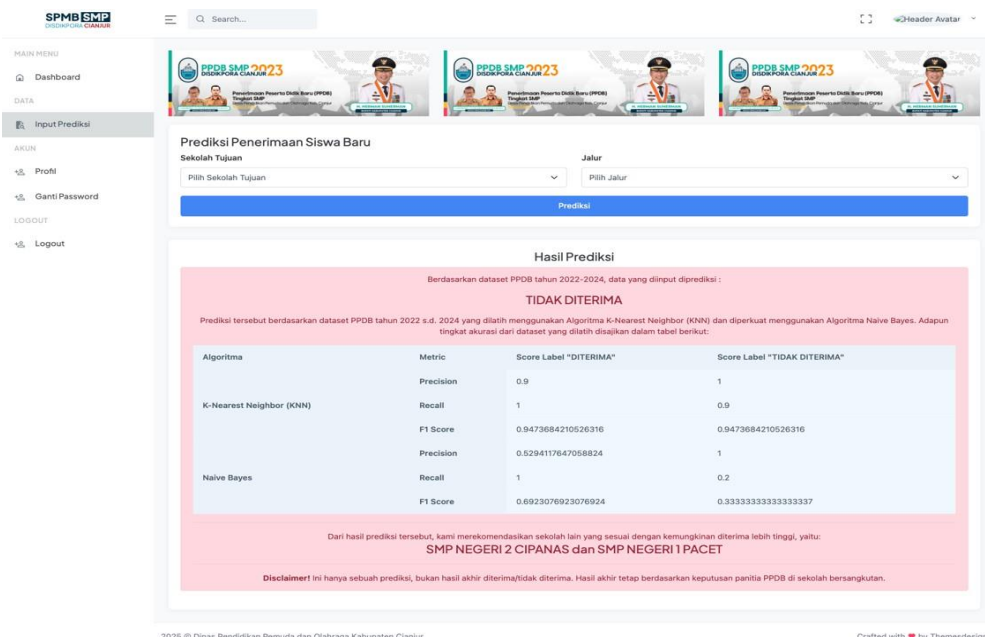
Tampilan input prediksi dan pemilihan jalur zonasi merupakan tampilan yang nantinya berfungsi memprediksi kelulusan calon peserta didik baru menggunakan algoritma KNN dan Naive Bayes, sehingga menghasilkan Informasi Penerimaan PPDB



Gambar 2. Tampilan Pemilihan Jalur Zonasi



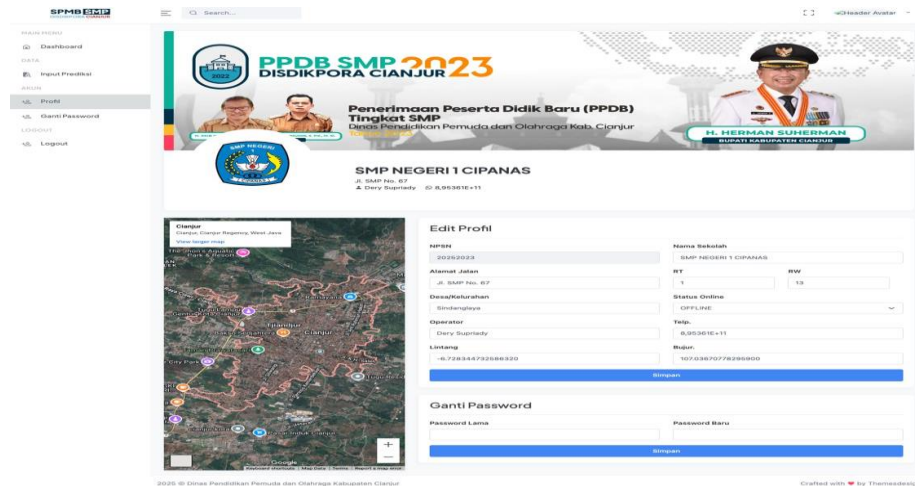
Gambar 3. Tampilan Prediksi Hasil Zonasi Diterima



Gambar 4. Tampilan Prediksi Hasil Zonasi Tidak Diterima

### 3. Tampilan Profil

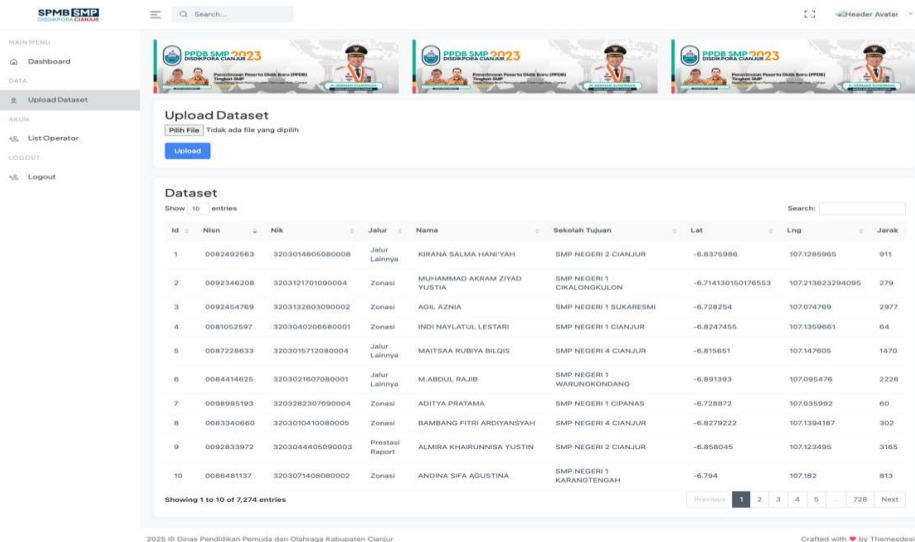
Tampilan Profil mengatur data-data sekolah yang menjadi tempat pilihan sekolah untuk proses PPDB dalam proses perhitungan zonasi.



Gambar 5. Tampilan Profil Sekolah

### 4. Tampilan Upload Dataset

Tampilan Upload Dataset menjadi table dasar dalam proses perhitungan zonasi PPDB dengan menggunakan algoritma KNN dan Naïve Bayes.



Gambar 6. Tampilan Upload Dataset

### 5. Tampilan List Akun

Tampilan List Akun mengatur akses pengguna Sistem perhitungan PPDB ini.



Gambar 7. Tampilan List Akun

### B. Temuan Ilmiah

Temuan utama penelitian ini adalah bahwa algoritma KNN memberikan hasil yang lebih baik pada jalur zonasi, sedangkan NB menunjukkan performa yang kompetitif pada jalur prestasi rapor. Hal ini dapat dijelaskan secara ilmiah karena:

1. Jalur zonasi berbasis jarak domisili, sehingga pola distribusi data cenderung linier dan lebih mudah dipetakan dengan pendekatan jarak Euclidean pada KNN.
2. Jalur prestasi rapor berbasis nilai akademik yang cenderung memiliki variasi distribusi lebih kompleks. Pada kondisi ini, pendekatan probabilistik NB tetap mampu memberikan klasifikasi yang stabil meskipun tidak seakurat KNN.

Fenomena ini mendukung hipotesis awal bahwa KNN unggul pada jalur berbasis jarak, sedangkan NB mampu bersaing pada jalur berbasis nilai.

### C. Diskusi dan Perbandingan dengan Penelitian Lain

Hasil penelitian ini konsisten dengan studi Winanjaya dkk. [3] yang menunjukkan bahwa algoritma berbasis pohon keputusan mampu menangkap pola pada seleksi PPDB, serta dengan temuan Aprilita dkk. [5] bahwa KNN lebih unggul pada data dengan kemiripan antar atribut yang tinggi. Sementara itu, hasil NB yang efisien sesuai dengan temuan Alghifari & Juardi [7] tentang keunggulan NB pada dataset dengan variasi terbatas.

Dengan demikian, sistem prediksi ini tidak hanya berhasil mencapai akurasi tinggi, tetapi juga memberikan wawasan ilmiah mengenai karakteristik algoritma KNN dan NB dalam konteks seleksi PPDB SMP di Kabupaten Cianjur, berikut Tabel Akurasi Status Seleksi PPDB Jalur Zonasi yang dihasilkan menggunakan Algoritma KNN :

Tabel 2. Hasil Algoritma KNN untuk label “DITERIMA” Jalur Zonasi

METRIK		PRECISION	RECALL	F1 SCORE	
PENGUJIAN KE	1	L1	1	1	
		L2	1	1	
	2	L1	1	1	
		L2	1	0,846153846	0,916666667
	3	L1	1	0,939393939	0,96875
		L2	0,764705882	1	0,866666667
	4	L1	1	0,909090909	0,952380952
		L2	0,684210526	1	0,8125
	5	L1	1	1	1
		L2	0,928571429	1	0,962962963
	6	L1	1	0,96969697	0,984615385
		L2	0,684210526	1	0,8125
	7	L1	1	1	1
		L2	0,764705882	1	0,866666667
	8	L1	1	1	1
		L2	0,866666667	1	0,928571429
	9	L1	1	1	1
		L2	0,866666667	1	0,928571429

Tabel 3. Hasil Algoritma KNN untuk label “TIDAK DITERIMA” Jalur Zonasi

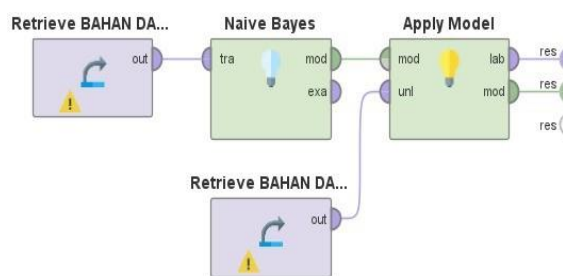
METRIK			PRECISION	RECALL	F1 SCORE	
PENGUJIAN KE	1	KNN	L1	0,93203883	0,93203883	0,93203883
			L2	0,800000	0,800000	0,800000
		Naïve Bayes	L1	0,7424242	0,9514563	0,8340426
			L2	0	0	0
	2	KNN	L1	0,9591837	0,9126214	0,93532338
			L2	0,8750000	0,8857143	0,82666667
		Naïve Bayes	L1	0,7463768	1,000000	0,85477178
			L2	0,0	0,0	0,0
	3	KNN	L1	0,9615385	0,9708738	0,96618357
			L2	0,9117647	0,8857143	0,89855072
		Naïve Bayes	L1	0,7463768	1,000000	0,85477178
			L2	0	0	0
	4	KNN	L1	0,969387755	0,922330097	0,945273632
			L2	0,555555556	0,914285714	0,853333333
		Naïve Bayes	L1	0,746376812	1,000000000	0,854771784
			L2	0	0	0
	5	KNN	L1	0,903846154	0,912621359	0,908212560
			L2	0,735294118	0,714285714	0,724637681
		Naïve Bayes	L1	0,746376812	1,000000000	0,854771784
			L2	0	0	0
6	KNN	L1	0,980392157	0,970873786	0,975609756	
		L2	0,916666667	0,942857143	0,929577465	
	Naïve Bayes	L1	0,744360902	0,961165049	0,838983051	
		L2	0	0	0	
7	KNN	L1	0,960396040	0,941747573	0,950980392	
		L2	0,837837838	0,885714286	0,861111111	
	Naïve Bayes	L1	0,736842105	0,951456311	0,830508475	
		L2	0	0	0	
8	KNN	L1	0,927927928	1,000000000	0,962616822	
		L2	1,000000000	0,771428571	0,870967742	
	Naïve Bayes	L1	0,750000000	0,029126214	0,056074766	
		L2	0,253731343	0,971428571	0,402366864	
9	KNN	L1	0,907407407	0,951456311	0,928909953	
		L2	0,833333333	0,714285714	0,769230769	
	Naïve Bayes	L1	1,000000000	0,077669903	0,144144144	
		L2	0,269230769	1,000000000	0,424242424	
10	KNN	L1	0,933962264	0,961165049	0,947368421	
		L2	0,607638889	0,005555556	0,835820896	
	Naïve Bayes	L1	0,750000000	0,029126214	0,056074766	
		L2	0,253731343	0,971428571	0,402366864	
AVG L1 KNN			0,943608067	0,947572815	0,945251732	
AVG L2 KNN			0,807309111	0,751984127	0,836989639	

AVG L1 Naïve Bayes	0,7709134	0,7000000	0,6178915
AVG L2 Naïve Bayes	0,094474901	0,300000000	0,132775664

#### D. Analisis Algoritma Naïve Bayes dengan Tools Rapidminer

Berbeda dengan analisis data algoritma *K-Nearest Neighbor* yang dapat dilakukan secara manual, analisis data algoritma *Naive Bayes* memerlukan *tools* bantuan. Dalam hal ini digunakan *tools* Rapidminer. Tahapan yang dilakukan dalam melakukan analisis prediksi kelulusan menggunakan Rapidminer di antaranya :

1. Mempersiapkan Data Training dan Data Testing  
*Data training* dan data testing dipersiapkan untuk jalur zonasi dan jalur prestasi raport. Data ini kemudian di *import* ke tools Rapidminer
2. Memilih column yang digunakan  
Selanjutnya adalah memilih *column* yang akan digunakan. Dalam hal ini *column* yang digunakan untuk memprediksi jalur zonasi adalah *column* nama, jarak, jalur dan status. Sedangkan untuk jalur prestasi raport *column* yang dipilih adalah nama, nilai dan status.
3. Menentukan label pada column yang menjadi kelas  
*Column* yang dijadikan label pada analisis prediksi ini adalah *column* STATUS di mana di dalamnya terdapat dua kelas yaitu DITERIMA dan TIDAK DITERIMA
4. Mengaplikasikan algoritma Naive Bayes  
Pada tahapan ini algoritma *Naive Bayes* bekerja dan melakukan *training* pada data *training* yang telah di *import*.
5. Apply Model  
Pada tahap ini prediksi dilakukan berdasarkan *training* sebelumnya.
6. Hasil  
Hasil dari semua tahapan di atas berupa *distribution* model, dan hasil prediksi.

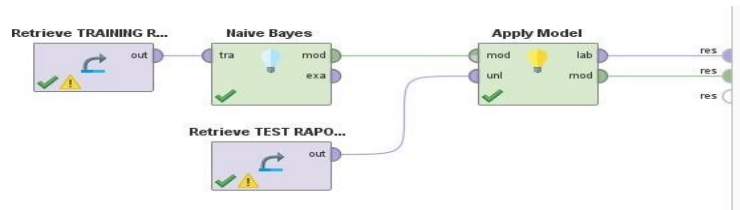


Gambar 8. Rapid Miner Jalur Zonasi dengan Naïve Bayes

Open in Turbo Prep Auto Model Interactive Analysis

Row No.	prediction(STATUS)	confidence(TIDAK DITERIMA)	confidence(DITERIMA)	NAMA	Jarak	Jalur Penda...	STATUS
1	TIDAK DITERIMA	0.869	0.131	FULAN	712	Zonasi	?

Gambar 9. Hasil Rapid Miner Prediksi Jalur Zonasi dengan Naïve Bayes



Gambar 10. Rapid Miner Jalur Prestasi dengan Naive Bayes

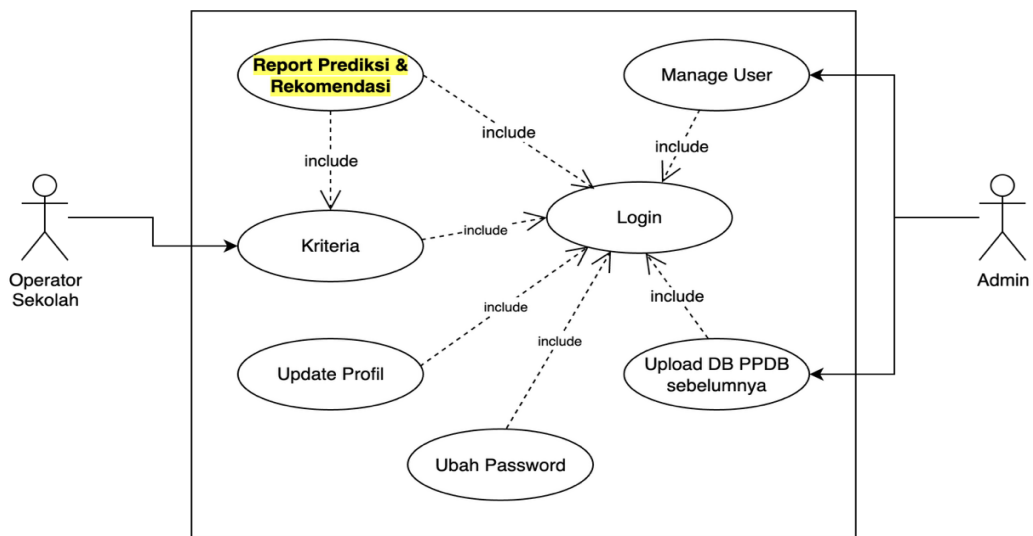
Row No.	prediction(S...	confidence(...	confidence(...
1	DITERIMA	0.513	0.487

Gambar 11. Hasil Rapid Miner Prediksi Jalur Pretasi dengan Naive Bayes

## E. Analisis Model Sistem

### 1. Use Case Diagram

Analisis Sistem prediksi PPDB yang akan diimplementasikan menggunakan algoritma *K-Nearest Neighbor* (KNN) dan *Naive Bayes* (NB). Untuk melakukan proses melalui dua jalur seleksi, yaitu jalur zonasi dan jalur prestasi rapor digambarkan dalam bentuk model *Use Case* berikut :

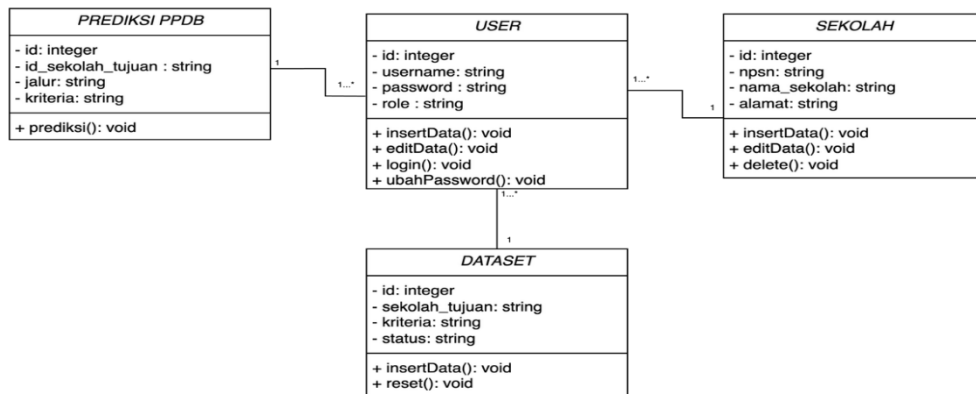


Gambar 12. Use Case Sistem Prediksi Status Seleksi PPDB

Dalam *Use Case* tersebut tergambar interaksi antara aktor pengguna, yaitu operator sekolah dan admin, melakukan beberapa fungsional system terkait prediksi rekomendasi status seleksi pada PPDB.

### 2. Class Diagram

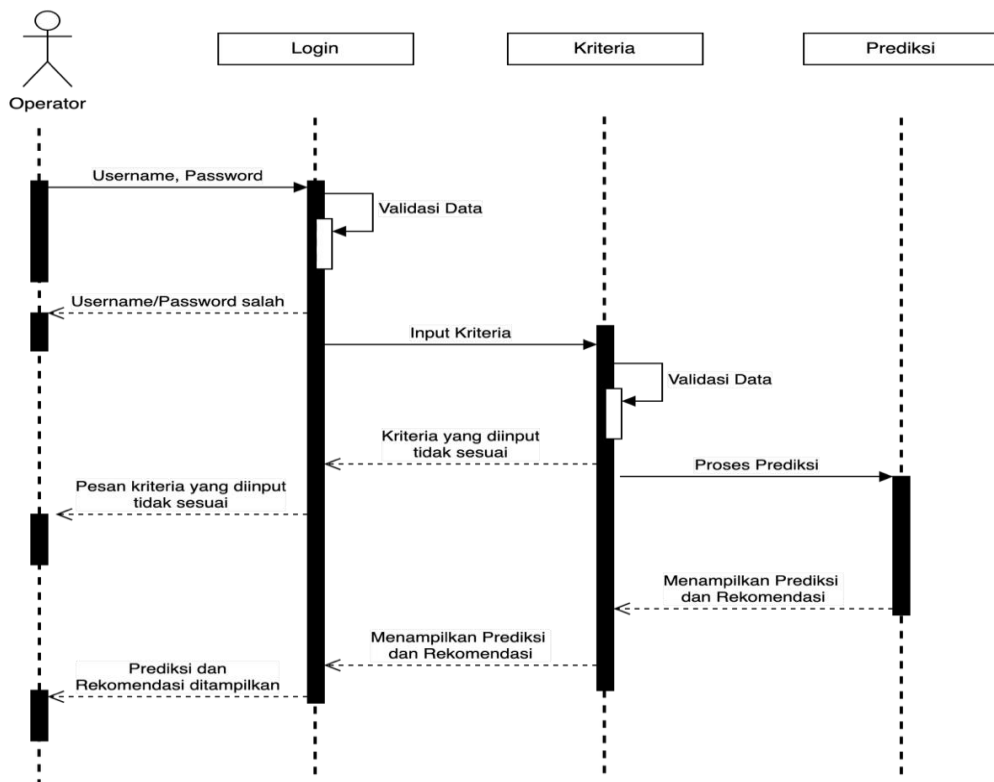
Class diagram menampilkan sistem dari aspek struktur visual, termasuk kelas, atribut, dan relasi untuk proses perhitungan status seleksi PPDB yang nantinya di implementasikan.



Gambar 13. *Class Diagram* Sistem Prediksi Status Seleksi PPDB

### 3. Sequence Diagram

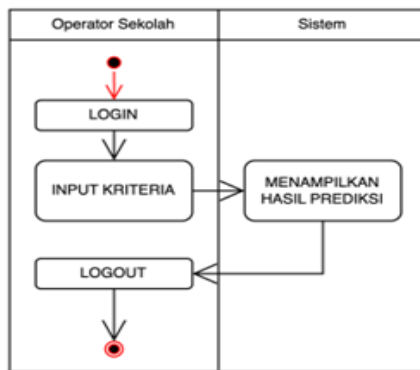
*Sequence Diagram* digunakan untuk lebih memahami dan memodelkan perilaku system terutama scenario interaksi objek dalam berkomunikasi satu sama lain, pesan dan urutan waktu interaksi didalam proses perhitungan status seleksi PPDB yang nantinya di implementasikan



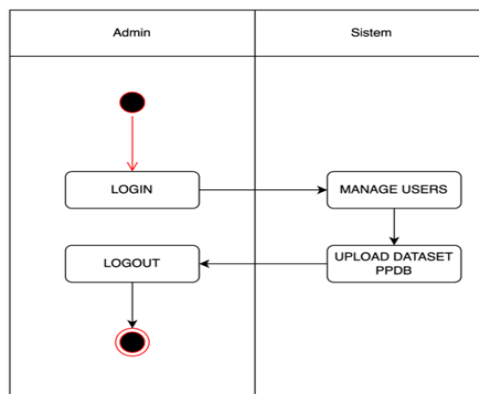
Gambar 14. *Sequence Diagram*

### 4. Activity Diagram

*Activity Diagram* memberikan gambaran alur serta menjelaskan urutan aktivitas Operator Sekolah dan Admin didalam proses perhitungan status seleksi PPDB yang nantinya di implementasikan.



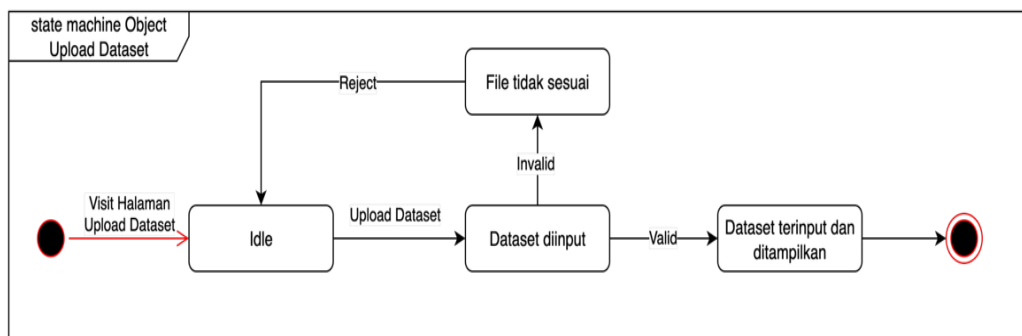
Gambar 15. Activity Diagram Operator Sekolah



Gambar 16. Activity Diagram Admin

5. State Diagram Prediksi

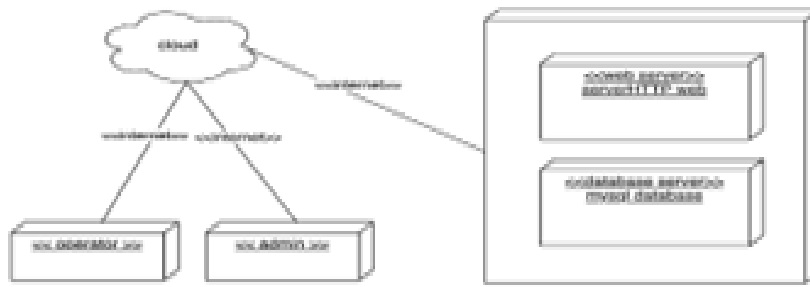
State Diagram menampilkan visualisasi rekayasa perangkat lunak dalam hal system bertransisi dari satu kondisi ke kondisi lain didalam proses perhitungan status seleksi PPDB yang nantinya di implementasikan.



Gambar 17. State Diagram

6. Deployment Diagram

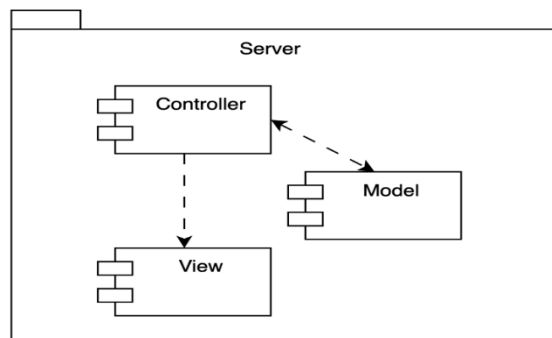
Deployment Diagram menunjukan hubungan antara software dan hardware didalam proses perhitungan status seleksi PPDB yang nantinya di implementasikan.



Gambar 18. *Deployment Diagram*

#### 7. *Component Diagram*

*Deployment Diagram* menunjukkan hubungan berbagai komponen (file, dataset, dll) dalam proses perhitungan status seleksi PPDB yang nantinya di implementasikan.



Gambar 19. *Component Diagram*

#### 4. KESIMPULAN

Penelitian ini berhasil mengimplementasikan algoritma *K-Nearest Neighbor* (KNN) dan *Naive Bayes* (NB) dalam memprediksi status penerimaan peserta didik baru tingkat SMP di Kabupaten Cianjur. Hasil penelitian menunjukkan bahwa hipotesis penelitian terbukti: KNN cenderung memberikan kinerja lebih baik pada jalur zonasi yang berbasis jarak, sedangkan NB mampu bersaing secara kompetitif pada jalur prestasi rapor.

Temuan ilmiah yang diperoleh adalah bahwa karakteristik data jalur zonasi yang relatif linier sesuai dengan pendekatan berbasis jarak, sementara data jalur prestasi dengan variasi distribusi lebih kompleks tetap dapat dipetakan secara stabil oleh pendekatan probabilistik NB. Hal ini menguatkan bahwa pemilihan algoritma klasifikasi harus disesuaikan dengan karakteristik data yang dianalisis.

Sebagai pengembangan, penelitian selanjutnya dapat diarahkan pada integrasi lebih banyak algoritma klasifikasi modern seperti *Random Forest* atau *Support Vector Machine*, serta penerapan teknik *ensemble learning* untuk meningkatkan akurasi prediksi. Selain itu, sistem rekomendasi sekolah yang dibangun dapat diperluas dengan mempertimbangkan faktor-faktor non-akademik, seperti ketersediaan fasilitas dan preferensi siswa, guna mendukung keputusan yang lebih komprehensif.

#### DAFTAR PUSTAKA

- [1] Pemerintah Kabupaten Cianjur, *Peraturan Bupati Kabupaten Cianjur No. 42 Tahun 2022: Pedoman Pelaksanaan PPDB pada TK, SD, dan SMP di Kabupaten Cianjur*, 2024.
- [2] MKKS SMP Kabupaten Cianjur, *Petunjuk Teknis Penerimaan Peserta Didik Baru (PPDB) Tahun Pelajaran 2024/2025*, 2024.
- [3] R. Winanjaya, F. Amir, dan R. Doni, "Penerapan Data Mining untuk Memprediksi Penerimaan Peserta Didik Baru Menggunakan Algoritma C4.5," *Prosiding Seminar Nasional Riset Information Science (SENARIS)*, 2019.
- [4] F. H. Zufalah, "Implementasi Algoritma KNN dalam Mengukur Ketepatan Kelulusan Mahasiswa UIN Syarif Hidayatullah Jakarta," 2022.
- [5] W. Aprilita, dkk., "Komparasi Algoritma K-Nearest Neighbor (KNN) dan Naive Bayes dalam Klasifikasi Penyakit Parkinson," *SENTIMAS: Seminar Nasional Penelitian dan Pengabdian Masyarakat*, 2023.
- [6] N. Aisyah, "Implementasi Algoritma K-Nearest Neighbor untuk Sistem Rekomendasi Topik Penelitian pada Jurusan Ilmu Komputer Universitas Lampung Berbasis Klasifikasi," 2021.
- [7] F. Alghifari dan D. Juardi, "Penerapan Data Mining pada Penjualan Makanan dan Minuman Menggunakan Metode Algoritma Naive Bayes," *Jurnal Ilmiah Informatika*, 2021