

Komparasi Algoritma Data Mining Dalam Ketuntasan Belajar Daring Siswa Pada Masa Pandemi Covid 19

Muhammad Saiful^{1*}, Hariman Bahtiar², Amri Muliawan Nur³, Yupi Kuspani Putra⁴

^{1,2,3,4}Program Studi Sistem Informasi, Universitas Hamzanwadi

*saipulslbm@gmail.com

Abstrak

Penelitian ini bertujuan untuk mengidentifikasi tingkat akurasi ketuntasan belajar daring siswa selama masa pandemi covid 19. Sasaran penelitian yang dilakukan fokus pada siswa kelas XI IPA dan IPS. Teknik dalam pengambilan sampel menggunakan metode purposive sampling. Adapun algoritma yang digunakan adalah komparasi naïve bayes dan naïve bayes berbasis particle swarm optimization yang berfungsi untuk menemukan hasil nilai akurasi yang terbaik dalam menganalisa ketuntasan belajar daring siswa selama masa pandemi. Hasil analisis setelah digunakan algoritma naïve bayes mendapat tingkat akurasi sebesar 83.91 %, dan hasil analisis setelah digunakan algoritma naïve bayes berbasis PSO mendapat tingkat akurasi sebesar 91.98 %, sehingga dari hasil eksperimen dan pengujian kedua algoritma tersebut dapat disimpulkan bahwa algoritma naïve bayes berbasis PSO memiliki tingkat akurasi yang lebih tinggi dibanding algoritma naïve bayes. Selisih perbandingan nilai akurasi kedua algoritma tersebut adalah sebesar 8.07%, sehingga dapat disimpulkan untuk algoritma yang cocok digunakan untuk klasifikasi ketuntasan belajar siswa selama pandemi covid 19 adalah Naive Bayes berbasis particle swarm optimization.

Kata Kunci: Pembelajaran Daring, Data Mining, Algoritma Naïve Bayes, Naïve Bayes berbasis PSO

Abstract

This research was conducted at SMA Negeri 3 Selong and became the focus of students in class XI IPA and Social Studies. The sampling technique used purposive sampling method. This study aims to describe the extent to which the level of completeness of students during post-covid-19 pandemic learning with online media. This study uses a classification algorithm that functions to find a model that distinguishes data classes or data concepts, with the specific objective of determining the class of unknown object labels. The method used is the PSO-based Naïve Bayes and Naïve Bayes Comparison Algorithms. The results of this study indicate that the use of online media during online learning using the naïve Bayes algorithm is 83.91%, and the PSO-based naïve Bayes algorithm is 91.98%, from the experimental results and testing of the two algorithms, the results of the confusion matrix and AUC testing can be obtained which can be determined the best accuracy value is the PSO-based Naïve Bayes algorithm. As for the comparison of the results in the form of an accuracy value obtained by the Naïve Bayes Algorithm of 83.91% and the PSO-Based Naïve Bayes Algorithm of 91.98% and the difference in the level of accuracy of 8.07%, so it can be concluded that the algorithm that is suitable for classifying student learning completeness during the covid 19 pandemic is Naive Bayes based on particle swarm optimization.

Keywords: Online Learning, Data Mining, Naïve Bayes Algorithm, Naïve Bayes Berbasis PSO

1. Pendahuluan

Implementasi pembelajaran pada masa pandemi covid19 berdampak pada sistem pembelajaran yang dimana biasa dilakukan dengan tatap muka

langsung, namun saat ini pembelajaran di alihkan ke sistem pembelajaran daring (SPADA) dengan menggunakan beberapa platform yang ada. Seperti halnya SMA Negeri 3 Selong selama

masa pandemi covid 19 pembelajaran secara tatap muka dialihkan kesistem pembelajaran berbasis daring [1]-[2]. Adapun selama proses pembelajaran daring tentu ada banyak problem atau kendala, diantaranya ketika menyampaikan pembelajaran melalui system daring, peserta didik harus mengakses internet, yang tentunya memerlukan jaringan internet yang stabil untuk menerimanya, dan tidak semua peserta didik dapat mengakses jaringan internet yang stabil di tempat tinggalnya. Esensi dari penerapan SPADA untuk meningkatkan pemerataan akses terhadap pembelajaran yang bermutu, belum menjadi efektif di tengah wabah pandemi covid 19. Hal tersebut menguatkan opini bahwa sesungguhnya kehadiran seorang guru jelas tidak dapat digantikan oleh mesin secanggih apapun. Dalam hal ini perannya akan dibutuhkan, meskipun yang kita lihat sarana prasarana dan juga bahan ajar yang tersedia di dunia maya (internet) sangat banyak tak terhitung jumlahnya [3]. Oleh karena itu, proses belajar mengajar yang terjadi dari rumah tetap dibutuhkan keaktifan, kreativitas dan inovasi dari seorang guru sangat berpengaruh terhadap minat belajar siswa, terutama selama masa pandemi covid berlangsung. Penelitian ini dilakukan di SMA Negeri 3 Selong, dengan tujuan untuk mengambil sample data hasil pembelajaran daring selama masa pandemi covid 19 [4]-[5]. Dengan demikian, dari hasil yang diperoleh,

peneliti melakukan pengambilan sampel data hasil pembelajaran daring menggunakan metode purposive sampling. Adapun algoritma yang digunakan adalah naïve bayes dan naïve bayes berbasis particle swarm optimization yang digunakan untuk mencari nilai akurasi yang terbaik dalam menganalisa ketuntasan belajar daring siswa selama masa pandemi. Atribut data yang akan dikelola dalam penelitian ini meliputi data nama siswa, jumlah pertemuan, presentasi kehadiran, tugas, ujian, KKM dan ketuntasan [6]-[7]. Berdasarkan uraian diatas maka penulis mengusulkan tema dengan judul : “Komparasi Algoritma Data Mining Dalam Ketuntasan Belajar Daring Siswa Pada Masa Pandemi Covid 19”. Dan adapun tujuan penulis dalam penelitian ini adalah untuk mengetahui seberapa besar tingkat akurasi data predikat ketuntasan siswa selama belajar daring siswa pasca pandemic covid 19.

2. Tinjauan Pustaka

2.1. Penelitian Terkait

Penelitian Terkait Untuk mendukung penelitian ini, peneliti menggunakan beberapa penelitian sebelumnya, antara lain sebagai berikut :

- Muhammad Saiful, Syamsuddin dan Moh. Farid Wajdi, yang berjudul “Implementasi Algoritma Naive Bayes untuk memprediksi Predikat Ketuntasan Belajar Siswa Pasca Pandemi Covid 19” menjelaskan dalam

penelitian yang dilakukan di SMAN 3 Selong untuk memprediksi ketuntasan belajar Siswa selama pandemik pasca covid 19, maka dapat disimpulkan bahwa setelah dilakukannya pengujian sebanyak 9 kali yaitu dari validation 2 sampai dengan 10 maka didapatkan dua hasil akurasi terbaik dari K-Fold Validation 4 dan K-Fold Validation 5 yang masing-masing memiliki nilai accuracy sebesar 83.89% pada K-Fold Validation 4 dan 82.74% dari K-Fold Validation 5. Dan selisih akurasi keduanya adalah 0.01% yang menunjukkan bahwa terjadi peningkatan 0.80 % pada rentang k-fold validation 4 dan 8 . Peningkatan pada setiap uji coba tidak selalu sama ini dipengaruhi oleh seberapa banyak pembagian data yang dilakukan [1].

- Muhammad Saiful, LM Samsu dan Fathurrahman, yang berjudul "Sistem Deteksi Infeksi COVID-19 Pada Hasil X-Ray Rontgen menggunakan Algoritma Convolutional Neural Network (CNN)" menjelaskan hasil analisis dan pembahasan dalam studi kasus dalam penelitian ini, maka dapat diambil beberapa simpulan sebagai berikut : Dengan menggunakan 25.000 step dan 1.000 evaluation_step untuk melakukan pengujian model4 pada data baru terinfeksi COVID-19 menghasilkan akurasi sebesar 98%. Tingkat akurasi pendeteksian thorax terinfeksi COVID-19 pada model4 menggunakan

algoritma Convolutional Neural Network dapat dinilai bekerja dengan baik dengan beragam tingkat akurasi pada hasil pengujian berkisar 60-99% [8].

- Nurhayati, yahya, fathurrahman, LM Samsu dan Wajizatul Amnia yang berjudul " Implementasi Algoritma Naive Bayes Untuk Klasifikasi Penerima Beasiswa (Studi Kasus Universitas Hamzanwadi)". menjelaskan dengan menggunakan 25.000 step dan 1.000 evaluation_step untuk melakukan pengujian model4 pada data baru terinfeksi covid-19 menghasilkan akurasi sebesar 98%. Tingkat akurasi pendeteksian thorax terinfeksi covid-19 pada model4 menggunakan algoritma Convolutional Neural Network dapat dinilai bekerja dengan baik dengan beragam tingkat akurasi pada hasil pengujian berkisar 60-99% [9].
- Eko Ahadi, Indra Gunawan, Ika Okta Kirana, Dedy Hartama dan Muhammad Ridwan Lubis, yang berjudul "Penentuan keberhasilan pembelajaran daring pada masa pandemi covid-19 menggunakan algoritma C.45 di STIKOM Tunas bangsa", menjelaskan hasil penentuan keberhasilan pembelajaran daring di masa pandemi Covid-19 dengan perhitungan manual dan perhitungan menggunakan RapidMiner menghasilkan 23 rule keberhasilan pembelajaran daring dan terdiri dari 90 data mahasiswa yang

digunakan diperoleh nilai entropy total sebesar 0,996791632 dan atribut tertinggi adalah pemahaman dengan gain sebesar 0,180 [10].

- Irma Agustika, Deddy Hartama, Lin Parlina, Indra Gunawan dan Ika Okta Kirana, yang berjudul “Analisis Keberhasilan Pembelajaran Daring pada Masa Pandemi Covid-19 menggunakan Algoritma C4.5 dan Naive Bayes”, Faktor yang paling dominan dalam meningkatkan prestasi belajar siswa dengan menggunakan Algoritma C4.5 dan Naive Bayes adalah mata pelajaran Adaptif. Didalam mata pelajaran Adaptif terdapat Performance atau tingkat Akurasi untuk Algoritma C4.5 sebanyak 98% dan Algoritma Naive Bayes sebanyak 99%. Setelah dilakukan perbandingan diantara dua algoritma tersebut, yang memiliki tingkat ke Akurasian lebih tinggi yaitu Algoritma Naive Bayes [10]

2.2. Landasan Teori

1. Naive Bayes

Naive Bayes Naive Bayes Classifier adalah algoritma yang efisien dan efektif untuk pembelajaran mesin dan penambangan data. Berbasiskan Naive Bayes Classifier pada kombinasi Bayes Theorem dan atribut independensi anggapan. Naive Bayes Classifier

didasarkan pada asumsi yang disederhanakan nilai-nilai atribut independen secara kondisional, dengan asumsi dari nilai target yang diberikan [11]-[12]. Persamaan dari teorema Bayes adalah : $P(H | X) = P(X | H) \times P(H) / P(X)$, keterangan dari teorema bayes berikut :

- X : Data dengan class yang belum diketahui.
- H : Hipotesis pada data X yang merupakan suatu class khusus.
- $P(H|X)$: Nilai probabilitas pada hipotesis H berdasarkan kondisi X.
- $P(H)$: Nilai probabilitas pada hipotesis H.
- $P(X|H)$: Nilai probabilitas X yang berdasarkan dengan kondisi H.
- $P(X)$: Nilai probabilitas pada X.

2. Teknik klasifikasi Dalam klasifikasi.

Teknik dalam klasifikasi terdapat dua pekerjaan utama yang dilakukan, yaitu pembangunan model sebagai prototype untuk disimpan sebagai memori dan penggunaan model tersebut untuk melakukan pengenalan/ klasifikasi/ prediksi pada suatu objek data lain agar diketahui di kelas mana objek data tersebut dalam model yang sudah disimpannya [5]-[13].

3. Particle Swarm Optimization (PSO).

Particle Swarm Optimization (PSO) sering digunakan dalam penelitian, karena PSO memiliki kesamaan sifat dengan Genetic Algorithm (GA). Keuntungan dari PSO adalah

mudah diterapkan dan ada beberapa parameter untuk menyesuaikan. Sistem PSO diinisiasi oleh sebuah populasi solusi acak dan selanjutnya mencari titik optimum dengan cara meng-update tiap hasil pembangkitan [11]

3. Metode Penelitian

3.1. Metode pengumpulan data

Metode penelitian pada dasarnya merupakan cara ilmiah untuk mendapatkan data dengan tujuan dan kegunaan tertentu [14]-[15]. Berdasarkan hal tersebut terdapat beberapa kata kunci yang diperhatikan yaitu, cara ilmiah, data, tujuan, dan kegunaan. Cara ilmiah berarti kegiatan penelitian yang didasarkan pada ciri-ciri keilmuan, yaitu rasional, empiris, dan sistematis. Rasional merupakan kegiatan penelitian yang dilakukan dengan cara yang masuk akal dan bisa dijangkau oleh pemikiran manusia. Empiris berarti cara yang dilakukan dengan mengamati oleh indera manusia dan orang lain dapat mengamati serta mengetahui cara yang digunakan. Sedangkan sistematis merupakan sebuah proses yang digunakan dalam penelitian dengan langkah-langkah tertentu yang bersifat logis. Adapun teknik yang digunakan adalah :

1) Interview (wawancara)

Wawancara digunakan sebagai teknik pengumpulan data apabila ingin melakukan studi pendahuluan untuk menemukan permasalahan yang harus diteliti, dan juga apabila peneliti ingin

mengetahui hal-hal dari responden yang lebih mendalam dan jumlah respondennya sedikit/kecil.

2) Observasi

Observasi merupakan sebuah teknik pengumpulan data yang mempunyai ciri yang spesifik bila dibandingkan dengan teknik yang lain, yaitu wawancara dan kuisisioner. Kalau wawancara dan kuisisioner selalu berkomunikasi dengan orang, maka observasi tidak terbatas pada orang, tetapi juga obyek-obyek alam yang lain. Teknik pengumpulan data dengan observasi digunakan ketika penelitian berkenaan dengan perilaku manusia, proses kerja, gejala-gejala alam dan bila responden yang diamati tidak terlalu besar.

3.2. Tahapan Penelitian

Adapun tahapan dari penelitian ini meliputi beberapa tahapan diantaranya :

1) Pengumpulan data

Dalam penelitian ini, yang menjadi populasi adalah siswa kelas XI IPA dan IPS SMAN 3 Selong yang berjumlah 175 orang. Bila populasi besar, dan peneliti tidak mungkin mempelajari semua yang ada pada populasi, misalnya karena keterbatasan dana, tenaga dan waktu, maka peneliti dapat menggunakan sampel yang diambil dari populasi [16]. Untuk itu sampel yang diambil dari populasi harus betul-betul representatif (mewakili)".

Table 1: Populasi dalam Penelitian

No	Responden	Populasi
1	IPA	85
2	IPS	90
Jumlah		175

Sumber : Tata usaha SMAN 3 Selong , 2022 (diolah).

Metode sampel yang digunakan dalam penelitian ini adalah teknik propotionate startifed random sampling, dimana metode ini digunakan bila populasi mempunyai anggota atau unsur yang tidak homogen dan berstrata secara proporsional [5]. Rumusnya sebagai berikut :

$$n_i = (N_i \times n) / N$$

Keterangan :

n_i : Sampel siswa ke-1

N_i : Populasi siswa ke-1

N : Populasi siswa keseluruhan

n : Jumlah sampel keseluruhan

Berdasarkan rumus diatas dapat diketahui perhitungan rasio sampel dari siswa IPA dan IPS.

Dapat dilihat pada tabel dibawah ini :

Table 2 : Tabel Responden

No	Responden	Rumus	Jumlah Sampel
1	IPA	$85 \times 175 / 175$	85
2	IPS	$90 / 175 / 175$	90
Jumlah			175

Sumber : Tata usaha SMAN 3 Selong , 2022 (diolah)

2) Pengolahan data awal

Jumlah data awal yang diperoleh dari data siswa kelas XI SMA Negeri 3 Selong sebanyak 170 data dan terdiri dari 11 atribut, namun tidak semua

atribut bisa digunakan karena harus melalui beberapa tahap pengolahan awal data. Untuk mendapatkan data yang berkualitas, beberapa teknik yang dilakukan sebagai berikut [9] :

- Data validation, untuk mengidentifikasi dan menghapus data yang ganjil (*outlier/noise*), data yang tidak konsisten, dan data yang tidak lengkap (*missing value*).
- Data *integration and transformation*, untuk meningkatkan akurasi dan efisiensi algoritma. Data yang digunakan dalam penulisan ini bernilai kategorikal. Data ditransformasikan kedalam *software Rapidminer*.
- Data *size reduction and discritization*, untuk memperoleh data set dengan jumlah atribut dan record yang lebih sedikit tetapi bersifat informative.
- Dataset yang tidak ada nilainya atau kosong dapat dihilangkan, namun datanya tidak ada yang kosong hanya atribut atau variabel yang mengalami perubahan. Sehingga jumlah data yang digunakan sebanyak 174 record dan terdiri dari 11 atribut. Dari data semula menggunakan 30 atribut, kemudian yang digunakan hanya 11 atribut yaitu nama siswa, hadir, jumlah pertemuan, Presentase kehadiran, Tugas, MID, UAS, Total, KKM, Predikat. Sampel data training yang digunakan seperti terlihat pada tabel 3.

NAMA SISWA	HADIR	TIDAK	JUMLAH PERTEMUAN	PRESENTASE KEHADIRAN (%)	TUGAS	MID	UAS	TOTAL	KKM	PREDIKAT
ADITYA DINI PRIMA NANDA	6	4	10	60%	80	88	88	77	75	TUNTAS
AGUSTIAN PRIMARIA PUTRA	7	3	10	70%	80	88	90	79	75	TUNTAS
ALANG ARTHA INARA	8	2	10	80%	80	87	90	79	75	TUNTAS
ARDIAN JAUHARI PRATAMA	9	1	10	90%	80	80	90	77	75	TUNTAS
ARPIA RAMACHAN	7	3	10	70%	80	87	90	79	75	TUNTAS
BAECHOM NAUFAL ALRINA	8	2	10	80%	80	79	90	79	75	TUNTAS
BAQUSAFIQUH ANBAR	5	5	10	50%	80	80	80	72	75	TIDAK TUNTAS
BHANTAR EFFENDY	4	6	10	40%	80	79	79	70	75	TIDAK TUNTAS
BHQ INTAN PURNAMAATI	7	3	10	70%	88	88	80	76	75	TUNTAS
ELVANA INTAN LESTARI	9	1	10	90%	80	90	90	80	75	TUNTAS
FATHUL HADI	8	2	10	80%	80	98	79	77	75	TUNTAS
GILANG ATI SARARA	9	1	10	90%	80	90	89	80	75	TUNTAS
LIHAMA SEDIANSYAH	7	3	10	70%	80	88	87	79	75	TUNTAS
M. GHIZZALI MARI	9	2	10	90%	80	98	79	77	75	TUNTAS

Gambar 1. Tabel Atribut yang digunakan

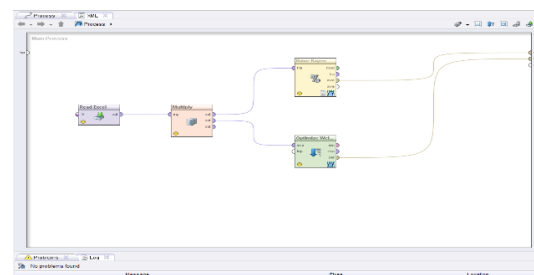
No	Atribut	Keterangan
1.	Nama Siswa	Nama Siswa
2.	Hadir	Kehadiran Siswa
3.	Tidak	Ketidak Hadiran Siswa
4.	Jumlah Pertemuan	Jumlah Pertemuan
5.	Presentase Kehadiran (%)	Presentase Kehadiran
6.	Tugas	Jenis Tugas
7.	MID	Ujian tengah Semester
8.	UAS	Ujian Akhir Semester
9.	Total	Total Jumlah Nilai
10.	KKM	Kriteria Ketuntasan Minimal
11.	Predikat	Tingkatan yang diperoleh

3) Model yang diusulkan

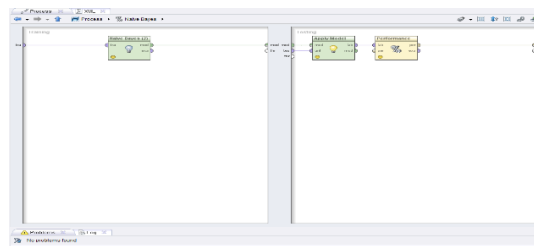
Model yang diusulkan dalam penelitian ini adalah menggunakan algoritma Naive Bayes dan Naive Bayes berbasis Particle swarm optimization. Model ini melakukan pengolahan dataset dan mengukur tingkat akurasi dalam mengklasifikasi ketuntasan belajar siswa menggunakan media daring pasca covid 19 di SMA Negeri 3 Selong [17].

4) Eksperimen dan Pengujian Model

- Pengujian Algoritma Naive Bayes : Pengujian akan dilakukan menggunakan 10 k-fold validation. Data siswa yang digunakan berjumlah 175 record. Algoritma Naive bayes melakukan training terhadap data yang telah diujikan. Hasil dari pengujian model berupa *accuracy*, *AUC*, dan *Confusion Matrix*. Berikut dijelaskan parameter dan operator yang digunakan pada model tersebut :



Gambar 2: Main process pada rapidminer

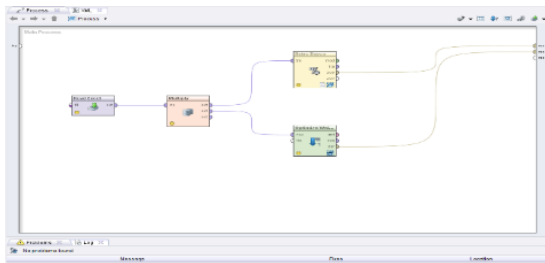


Gambar 3: Pembagian data training dan testing. Selanjutnya adalah perolehan hasil *accuracy* dari algoritma naive bayes menggunakan cross validation 2. Tampilan dapat dilihat pada gambar dibawah ini :

Multiclass Classification Performance - Annotations			
Table View - Plot View			
accuracy: 83.8% +/- 1.15% (macro 83.8%)			
	true TUNTAS	true TIDAK TUNTAS	class precision
pred TUNTAS	100	24	81.82%
pred TIDAK TUNTAS	4	33	89.48%
class recall	98.42%	81.25%	

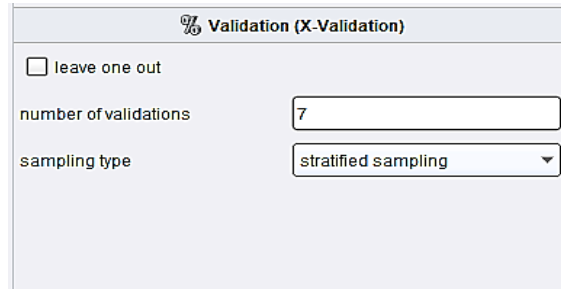
Gambar 4: Hasil accuracy Naïve Bayes menggunakan cross validation 2

- Pengujian Algoritma Naïve Bayes Berbasis PSO: Pengujian akan dilakukan menggunakan 10 k-fold validation. Data siswa yang digunakan berjumlah 175 record. Algoritma Naïve bayes berbasis PSO melakukan training terhadap data yang telah diujikan. Hasil dari pengujian model berupa accuracy, AUC, dan Confusion Matrix. Berikut dijelaskan parameter dan operator yang digunakan pada model tersebut :



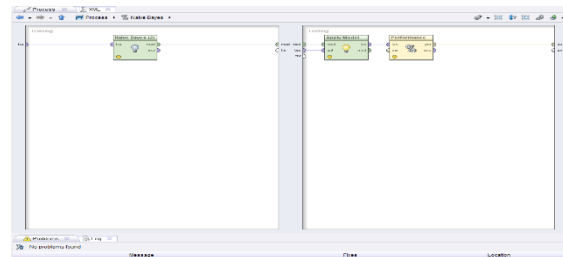
Gambar 5. Main process pada rapidminer

Dalam penelitian ini penulis menggunakan jenis file excel. Validation operator yang digunakan dalam penelitian ini adalah x validation. Operator ini memiliki dua proses, yaitu proses training dan testing. Proses training itu sendiri digunakan untuk mentraining model kemudian model diterapkan ke dalam proses testing dan kinerja model diukur selama fase pengujian. Selanjutnya akan dilakukan pengisian dan pemilihan parameter.



Gambar 6. Parameter x-validation

Number of validation adalah parameter yang menentukan jumlah himpunan bagian yang harus dibagi menjadi beberapa subset yaitu masing-masing subset memiliki jumlah yang sama. Pada number of validation ini penulis memasukkan jumlah data yang akan dibagi menjadi data training dan testing.

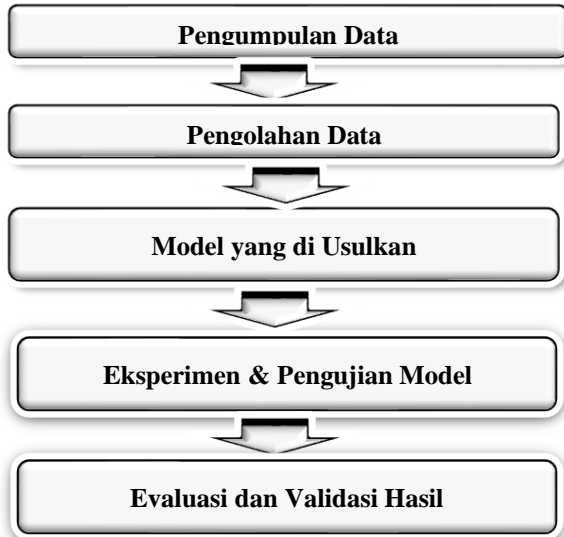


Gambar 7: Pembagian data training dan testing Selanjutnya adalah perolehan hasil accuracy dari algoritma naïve bayes menggunakan cross validation 7. Tampilan dapat dilihat pada gambar dibawah ini :

	True TRAINING	True TEST TRAINING	Class precision
pred TRAINING	100	0	92.00%
pred TEST TRAINING	0	54	91.00%
class recall	94.54%	87.10%	

Gambar 8: Hasil accuracy Naïve Bayes berbasis PSO menggunakan cross validation 7

Berikut tahapan penelitian yang disajikan dalam bentuk gambar sebagai berikut :



Gambar 9: Tahapan Penelitian

3.2. Lokasi Penelitian

Lokasi Penelitian ini dilakukan dengan mengumpulkan sumber data yang di ambil dengan cara melakukan penyebaran angket terhadap siswa SMA Negeri 3 Selong khususnya kelas IX IPA dan IPS.

4. Hasil dan Pembahasan

1) Analisa Dan Perancangan

Hasil dari pengujian algoritma yang sudah dilakukan sebanyak 10 kali pengujian menggunakan 10 k-fold validation, maka didapatkan nilai hasil eksperimen terbaik dapat ditunjukkan pada tabel berikut ini :

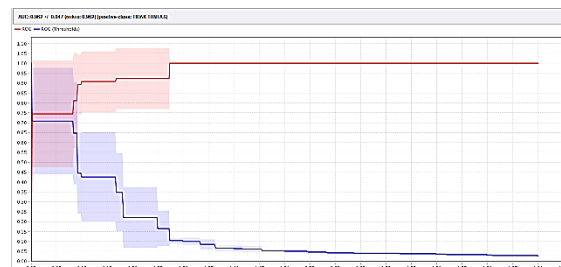
ALGORITMA NAÏVE BAYES				
Validation (%)	Accurasi (%)	Precision (%)	Recall (%)	AUC dan ROC (%)
2	83,91	90,68	61,29	0,933
3	83,33	87,72	61,11	0,938
4	42,77	86,94	59,58	0,943
5	82,29	88,81	64,41	0,943
6	82,18	88,92	57,58	0,938
7	82,19	90,02	57,94	0,945
8	82,79	91,44	59,84	0,951
9	82,08	87,22	58,47	0,935
10	82,19	89,83	58,33	0,950

Gambar 10. Hasil pengujian menggunakan Algoritma Naïve Bayes

Berdasarkan tingkat akurasi yang terbaik dari pengujian menggunakan Algoritma Naïve Bayes sebanyak 10 pengujian menggunakan 10 k-fold validation, maka didapatkan hasil yang terbaik dari pengujian 2 k-fold validation yaitu sebesar 83,91 % dengan nilai AUC 0,933 %, dan dari hasil pengujian yang dilakukan di dapatkan nilai excellent klasifikasi. Hasil pengujian dapat ditunjukkan pada gambar dibawah ini :

accuracy: 0,933% (+ 1,15% pada: 0,939%)			
	true TUNTAS	true TIDAK TUNTAS	class precision
pred TUNTAS	108	24	81,02%
pred TIDAK TUNTAS	4	38	90,48%
class recall	99,45%	91,29%	

Gambar 11 : Nilai Accurasi Algoritma Naïve Bayes



Gambar 12 : Nilai AUC dan Grafik ROC Algoritma Naïve Bayes

Hasil pengujian Algoritma Naïve Bayes Berbasis PSO : Hasil dari pengujian model yang sudah dilakukan sebanyak 10 kali pengujian

menggunakan 10 k-fold validation, maka didapatkan nilai hasil eksperimen terbaik dapat ditunjukkan pada tabel berikut ini :

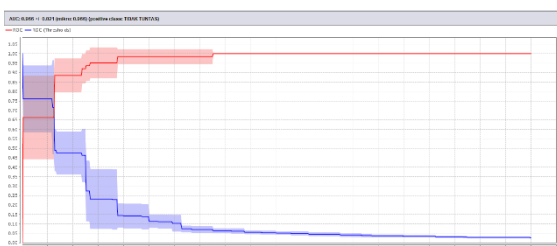
Hasil pengujian menggunakan Algoritma Naïve Bayes Berbasis PSO

ALGORITMA NAÏVE BAYES BERBASIS PSO				
Validation (%)	Accurasi (%)	Precision (%)	Recall (%)	AUC dan ROC (%)
2	89,66	92,73	77,42	0,969
3	89,08	98,25	70,79	0,945
4	91,36	90,06	85,62	0,968
5	90,84	90,84	84,23	0,961
6	91,95	90,14	87,12	0,971
7	91,98	90,63	86,90	0,966
8	90,88	89,64	84,15	0,961
9	89,21	97,78	71,43	0,954
10	91,44	90,40	86,19	0,962

Berdasarkan tingkat akurasi yang terbaik dari pengujian menggunakan Algoritma Naïve Bayes Berbasis PSO sebanyak 10 pengujian menggunakan 10 k-fold validation, maka didapatkan hasil yang terbaik dari pengujian 7 k-fold validation yaitu sebesar 91,98 % dengan nilai AUC dan ROC 0,966 %, dan dari hasil pengujian yang dilakukan di dapatkan nilai excellent klasifikasi. Hasil pengujian dapat ditunjukkan pada gambar dibawah ini :

accuracy 91.98% + 5.00% (nilai 91.98%)			
	true TUNTAS	true TIDAK TUNTAS	class prediction
pred TUNTAS	116	0	92.98%
pred TIDAK TUNTAS	6	54	90.00%
class real	94.64%	87.10%	

Gambar 12: Nilai Accurasi Algoritma Naïve Bayes Berbasis PSO



Gambar 13. Nilai AUC dan Grafik ROC Algoritma Naïve bayes

Perbandingan Algoritma Naïve Bayes Dengan Algoritma Naïve Bayes Berbasis PSO : Setelah dilakukan eksperimen dan pengujian kedua algoritma tersebut maka dapat dilihat hasil pengujian confusion matrix dan AUC yang mana dapat ditentukan nilai akurasi yang terbaik adalah algoritma Naïve Bayes berbasis PSO. Adapun perbandingan hasil yang berupa nilai akurasi dan performa AUC, maka diperoleh data perbandingan sebagai berikut :

Table 3 : Tabel perbandingan Algoritma Naïve Bayes dan Naïve Bayes Berbasis PSO

ALGORITMA	ACCURACY
Naïve Bayes	83.91 %
Naïve Bayes Berbasis PSO	91.98 %
Selisih Tingkat Accuracy	8.07 %

5. Kesimpulan

Berdasarkan hasil eksperimen yang dilakukan algoritma *Naive Bayes* dan *Naive Bayes* berbasis *Particle Swarm Optimization* (PSO) maka dapat disimpulkan algoritma *naive bayes* berbasis PSO menghasilkan nilai akurasi yang lebih tinggi yaitu 91.98% dibandingkan dengan algoritma *Naive Bayes* sebesar 83.91%, dan selisih tingkat akurasi yang didapatkan antara kedua algoritma tersebut adalah 8.07%. Kemudian untuk evaluasi

menggunakan ROC Curve untuk Algoritma Naïve Bayes yaitu , untuk model algoritma *Naive Bayes* nilai AUC adalah 0.933 dengan tingkat diagnosa *Poor Classification*, dan untuk model algoritma *Naïve Bayes* berbasis *particle swarm optimization* nilai AUC adalah 0.966 dengan tingkat diagnosa *Fair Classification*. Dari evaluasi ROC curve tersebut terlihat bahwa model *Naive Bayes* berbasis *particle swarm optimization* lebih tinggi jika dibandingkan dengan algoritma *Naive Bayes*. Dari hasil nilai AUC tersebut didapatkan selisih antara kedua model yaitu 0.033. Sehingga dapat disimpulkan bahwa algoritma *Naive Bayes* berbasis *particle swarm optimization* lebih akurat dalam mengklasikasi tingkat ketuntasan belajar siswa

6. Daftar Pustaka

- [1] M. F. W. Muhammad Saiful, Syamsuddin, "Infotek : Jurnal Informatika dan Teknologi," *Implementasi Algoritma Naive Bayes Untuk Memprediksi Predikat Ketuntasan Belajar Siswa Pasca Pandemi Covid 19*, vol. 4, no. 1, pp. 29–38, 2021.
- [2] R. Assya'bani and M. Majdi, "Pengembangan Model Pembelajaran Pasca Covid-19 Berdasarkan Pembelajaran Abad 21," *Al Qalam J. Ilm. Keagamaan dan Kemasyarakatan*, vol. 16, no. 2, p. 555, 2022, doi: 10.35931/aq.v16i2.903.
- [3] M. Saiful and N. Amalia, "Rancang Bangun Aplikasi Sistem Informasi Peserta Didik Baru menggunakan PHP dan Mysql "Sman 3 Selong," *Infotek J. Inform. dan Teknol.*, vol. 1, no. 1, pp. 18–31, 2018, doi: 10.29408/jit.v1i1.891.
- [4] A. Habiba, R. R. Isnanto, and J. E. Suseno, "Pemilihan Fitur Chi Square Pada Algoritma Naïve Bayes dan Pengaruhnya Terhadap Analisis Sentimen Masyarakat Indonesia Tentang Pembelajaran Tatap Muka Pada Masa Pandemi Covid-19," vol. 12, no. 1, pp. 111–116, 2023.
- [5] M. Saiful and A. Muliawan Nur, "Application of Expert System with Web-Based Forward Chaining Method in Diagnosing Corn Plant Disease," *J. Phys. Conf. Ser.*, vol. 1539, no. 1, pp. 0–7, 2020, doi: 10.1088/1742-6596/1539/1/012019.
- [6] I. A. Sihombing, D. Hartama, I. Parlina, I. Gunawan, and I. O. Kirana, "Analisis Keberhasilan Pembelajaran Daring pada Masa Pandemi Covid-19 menggunakan Algoritma C4.5 dan Naive Bayes," *JUKI J. Komput. dan Inform.*, vol. 3, no. 2, pp. 89–96, 2021, doi: 10.53842/juki.v3i2.68.
- [7] Y. Yahya and H. Bahtiar, "Pengaruh Pertumbuhan Ekonomi Terhadap Tingkat Kesejahteraan Masyarakat Kecamatan Selong Kabupaten Lombok Timur – Nusa Tenggara Barat Menggunakan Algoritma Naive Bayes," *Infotek J. Inform. dan Teknol.*, vol. 4, no. 1, pp. 20–28, 2021, doi: 10.29408/jit.v4i1.2981.
- [8] M. Saiful, L. M. Samsu, and F. Fathurrahman, "Sistem Deteksi Infeksi COVID-19 Pada Hasil X-Ray Rontgen menggunakan Algoritma Convolutional Neural Network (CNN)," *Infotek J. Inform. dan Teknol.*, vol. 4, no. 2, pp. 217–227, 2021, doi: 10.29408/jit.v4i2.3582.
- [9] N. I. Nurhidayati, Y. Yahya, F. Fathurrahman, L. . Samsu, and W. Amnia, "Implementasi Algoritma Naive Bayes Untuk Klasifikasi Penerima Beasiswa (Studi Kasus Universitas Hamzanwadi)," *Infotek J. Inform. dan Teknol.*, vol. 6, no. 1, pp. 177–188, 2023, doi: 10.29408/jit.v6i1.7529.
- [10] E. Ahadi, I. Gunawan, I. O. Kirana, D. Hartama, and M. R. Lubis, "Penentuan Keberhasilan Pembelajaran Daring Pada Masa Pandemi Covid-19 dengan

- Menggunakan Algoritma C4.5 di Stikom Tunas Bangsa,” *J. Komput. dan Inform.*, vol. 10, no. 1, pp. 78–85, 2022, doi: 10.35508/jjicon.v10i1.6446.
- [11] L. Out, *Proceeding Icete 2016*, no. October. 2016.
- [12] Zhou, Yang, and Wang, “Klasifikasi Penentuan Penerima Program Keluarga Harapan (PKH) Menggunakan Algoritma Support Vector Machine (Svm) Pada Kantor Dinas Sosial Lombok Timur, vol. 21, no. 1, pp. 1–9, 2020.
- [13] Alimuddin and M. Saiful, “Implementation of the Neural Network (NN) Algorithm in Analysis of Student Class Increment Data Based on Report Card Value,” *J. Phys. Conf. Ser.*, vol. 1539, no. 1, pp. 0–6, 2020, doi: 10.1088/1742-6596/1539/1/012034.
- [14] M. Saiful, L. M. Samsu, and I. Fathurrahman, “Perancangan Kerangka Crowdsourcing Berbasis Wisdom Of Crowds Untuk Kamus Naskah Lontar (Takepan) Sasak Online,” vol. 3, no. 2, pp. 165–173, 2020.
- [15] S. Sukardi, “Actualizing Strategy for Social Reconstruction-Oriented Learning Management: The Application on the Craftmanship and Entrepreneurship Subject at Senior High School Level,” no. October 2016, 2017.
- [16] Samsinar, “Mobile Learning :Inovasi pembelajaran Di Masa Pandemi Covid-19,” *Al Gurfah*, vol. Volume 1, no. 1, pp. 41–57, 2020, [Online].
- [17] Yupi Kuspani Putra, “Perbandingan Algoritma Naive Bayes dan Naive Bayes Berbasis PSO untuk Analisis Kredit pada PT. BPR Syariah Paokmotong,” *Ayan*, vol. 8, no. 5, p. 55, 2019.