

Implementasi Algoritma Naive Bayes Untuk Memprediksi Predikat Ketuntasan Belajar Siswa Pasca Pandemi Covid 19

Muhammad Saiful^{1*}, Syamsuddin², Moh. Farid Wajdi³

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

³Program Studi Teknik Informatika, Universitas Hamzanwadi

*saipulslb@gmail.com

Abstrak

Selama masa pandemi Covid-19 SMA Negeri 3 Selong mengubah kegiatan pembelajaran dari yang semula melalui tatap muka secara langsung namun saat ini pembelajaran di alihkan ke Sistem Pembelajaran Daring (SPADA) dengan menggunakan beberapa platform yang ada. Dilihat dari tingkat kemajemukan pola berpikir siswa selama pelaksanaan belajar daring ada banyak problem yang muncul salah satunya adalah, tidak stabilnya jaringan internet, perangkat media hendpone yang dimiliki siswa bervariasi dan masih kurangnya pengetahuan siswa dalam menggunakan platform daring. Tujuan dalam penelitian ini adalah untuk mengetahui indikator masalah dalam predikat ketuntasan belajar siswa kelas XII SMAN 3 Selong selama pasca pandemic covid 19. Metode yang digunakan dalam menyelesaikan permasalahan ini yaitu dengan algoritma Naive Bayes. Naive Bayes merupakan salah satu metode pada probabilistic reasoning. Dan kedepannya hasil penelitian ini diharapkan mampu memberikan solusi yang tepat dalam penyelesaian problem dalam pembelajaran via daring.

Kata Kunci : Data Mining ,Algoritma Naive Bayes, Belajar Daring

Abstract

During the Covid-19 pandemic, SMA Negeri 3 Selong changed learning activities from what was originally face-to-face, but currently learning is being transferred to the Online Learning System (SPADA) using several existing platforms. Judging from the level of plurality of students' thinking patterns during the implementation of online learning there are many problems that arise, one of which is the instability of the internet network, the various hendpone media devices owned by students and the lack of student knowledge in using online platforms. The purpose of this study was to determine the indicator of the problem in the predicate of student learning completeness of class XII SMAN 3 Selong during the post-COVID-19 pandemic. The method used to solve this problem is the Naive Bayes algorithm. Naive Bayes is a method of probabilistic reasoning. And in the future the results of this study are expected to be able to provide the right solution in solving problems in online learning.

Keywords: Data Mining, Naive Bayes Algorithm, Learn Online

1. Pendahuluan

Pembelajaran secara daring menuntut tenaga pendidik untuk mampu berinovasi dalam melaksanakan pembelajaran. Pembelajaran seharusnya dirancang agar membawa peserta didik ke pengenalan nilai secara kognitif, penghayatan nilai secara afektif, dan akhirnya

ke pengamalan nilai secara nyata. Dengan demikian semua yang dipelajari harus bermuatan pendidikan karakter. Misi utama pendidikan tidak sekadar membuat peserta didik pintar dari segi intelektual namun juga berkarakter baik. Misi tersebut tetap harus dijalankan apapun metode pembelajaran yang

digunakan baik secara konvensional maupun pembelajaran daring. Pembelajaran secara daring diimplementasikan dengan beragam cara oleh siswa di tengah penutupan sekolah selama kurang lebih beberapa bulan berjalan untuk mengantisipasi virus corona. Namun implementasi itu dinilai tidak maksimal dan menunjukkan masih ada ketidaksiapan di kalangan siswa untuk beradaptasi di iklim digital sekarang ini, karena faktornya bisa bersumber pada pengetahuan yang dimiliki siswa cukup beragam pengetahuan mereka tentang bagaimana menggunakan media online (internet), kemudian mengaplikasikan platform untuk belajar daring, seperti google class room, whatsapp, zoom meeting, dan lain sebagainya masih belum maksimal. Adanya Pandemi ini 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. Namun apakah penerapan Sistem Pembelajaran Daring saat ini di tengah Pandemi Covid-19 telah efektif? Nah pertanyaan itu ada banyak yang muncul disemua elemen pendidikan baik dari tingkat SD, SMP, SMA/SMK, bahkan di tingkat Perguruan Tinggi pun dalam penerapan pembelajaran daring masih banyak problem

yang dihadapi oleh guru sebagai tenaga pendidik. Seperti yang dirasakan beberapa guru di SMA Negeri 3 Selong dalam penerapan pembelajaran daring mengalami beberapa problem atau kendala, diantaranya ketika menyampaikan pembelajaran melalui sistem daring, maka peserta didik harus mengakses internet, yang tentunya memerlukan jaringan internet yang stabil untuk menerimanya, namun tidak semua peserta didik dapat mengakses jaringan internet yang stabil di tempat tinggalnya. Peserta didik yang tidak dapat mengakses internet tentunya terkendala dan harus keluar rumah untuk bisa mengakses internet, ketika peserta didik keluar dari rumah tentunya ini tidak sejalan lagi dengan prinsip penerapan Sistem Pembelajaran Daring (SPADA) di tengah Pandemi ini. Kendala berikutnya lagi soal pemberian tugas yang dimana pendidik tak henti-hentinya memberikan tugas online terhadap peserta didik yang tentunya berdampak pada ketahanan tubuh dan pola berfikir siswa yang bisa saja menjadi stress, padahal di tengah pandemi ini para Tim Medis Kesehatan dan Pemerintah menganjurkan kita untuk meningkatkan imunitas tubuh dan menjaga kesehatan. Masih banyak permasalahan yang muncul di lapangan yang perlu dituntaskan supaya tercapai tujuan pendidikan dengan baik. 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. Oleh karena itu, proses belajar mengajar yang terjadi dari rumah tetap dibutuhkan keaktifan, kreativitas dan inovasi dari seorang guru supaya pembelajaran tetap menarik, menyenangkan dan mudah dicerna. Dengan demikian aktifitas belajar berjalan dengan baik. Penelitian ini dilakukan di SMA Negeri 3 Selong. Berdasarkan observasi yang telah dilakukan selama menggunakan system pembelajaran daring masih banyak problem yang muncul sehingga efektifitas pembelajaran berlangsung tidak maksimal dan masih kurang dan tidak sesuai harapan sehingga akan berpengaruh pada predikat ketuntasan siswa selama belajar daring. Maka dengan menggunakan *data mining* algoritma naïve bayes, hasil belajar daring yang dilakukan oleh guru tercatat dan akan diimplementasikan dengan menggunakan algoritma naïve bayes untuk memprediksi predikat apakah siswa dianggap tuntas atau tidak tuntas selama proses belajar daring berlangsung. Adapun atribut dari data yang akan dikelola meliputi data nama siswa, jumlah pertemuan, presentasi kehadiran, tugas, ujian, KKM dan ketuntasan. Data-data tersebut

belum termanfaatkan dengan baik sebagai bahan pertimbangan, kajian dan penelitian untuk membuat suatu informasi yang berguna bagi sekolah. Berdasarkan uraian di atas, penelitian ini menggunakan algoritma naïve bayes untuk memprediksi tingkat ketuntasan belajar siswa selama pasca pandemic covid 19.[1] Berdasarkan uraian dan hasil dari penjelasan diatas maka penulis mengusulkan tema dengan judul : “Implementasi Algoritma Naïve Bayes Untuk Memprediksi Predikat Ketuntasan Siswa Selama Belajar Daring Pasca Pandemi Covid. Dengan tujuan untuk memprediksi seberapa besar akurasi tingkat predikat ketuntasan siswa selama belajar daring siswa pasca pandemic covid 19.

2. Tinjauan Pustaka.

2.1. Penelitian Terkait.

Beberapa penelitian terkait untuk penggunaan *Data Mining* metode klasifikasi dengan menggunakan Algoritma Naïve Bayes dalam mencari pola untuk menentukan yang terbaik atau untuk memprediksi yang akan terjadi kedepannya, diantaranya adalah :

- 1). Penelitian yang dilakukan oleh Muhammad Saiful, Suhartini, Nurhidayati yang berjudul “Klasifikasi Kinerja Dosen STT Hamzanwadi Menggunakan *Naive Bayes* Berbasis PSO” (2016) didapatkan kesimpulan Berdasarkan hasil penelitian yang telah dilakukan penulis

terhadap 736 data responden dalam memprediksi tingkat kinerja dosen maka dilakukan pengujian model dengan menggunakan *Naive Bayes* dan *Naive Bayes* berbasis *Particle Swarm Optimization*. Model yang dihasilkan diuji untuk mendapatkan nilai *accuracy* dari setiap algoritma sehingga didapat pengujian dengan menggunakan *Naive Bayes* dan setelah dilakukan pengujian dengan *tools rapid miner* didapat nilai *accuracy* adalah (98,37%). Sedangkan pengujian dengan menggunakan *Naive bayes* berbasis *Particle Swarm Optimization* dilakukan seleksi atribut dan penyesuaian pada parameter *Population Size*, *Max number of generation*, dan *number of validation*. Dari 26 variabel prediktor dilakukan seleksi atribut sehingga menghasilkan terpilihnya 22 atribut yang digunakan. didapatkan nilai *accuracy* (98,78%). Maka dapat disimpulkan pengujian data kinerja dosen menggunakan *Naive Bayes* dan penerapan *Particle Swarm Optimization* dalam pemilihan atribut didapat bahwa metode tersebut lebih akurat dalam penentuan Prediksi kinerja dosen STT Hamzanwadi dibandingkan dengan metode *NAive Bayes* tunggal, ditandai dengan peningkatan nilai akurasi sebesar (0,41%), dengan nilai tersebut masuk kedalam klasifikasi 0.90 - 1.00 = klasifikasi sangat baik (*excellent classification*).[2]

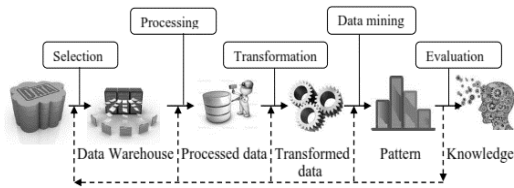
2.2. Landasan Teori.

1. Belajar Daring

Pelaksanaan Pembelajaran daring melalui 3 tahapan yaitu persiapan, pelaksanaan dan penilaian. Adapun rangkaian kegiatan persiapan, pelaksanaan dan penilaian pembelajaran daring yang dilakukan oleh guru dan siswa adalah sebagai berikut [3]:1. Persiapan belajar daring, 2. Pelaksanaan pembelajaran daring, 3. Penilaian hasil belajar daring.

2. Data Mining.

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* besar. Dalam data mining terdapat dua pendekatan metode pelatihan, yaitu [4]: *Unsupervised learning*, metode ini diterapkan tanpa adanya Latihan (*training*) dan tanpa ada guru (*teacher*). Guru di sini adalah label dari data. *Supervised learning*, yaitu metode belajar dengan adanya latihan dan pelatih. Dalam pendekatan ini, untuk menemukan fungsi keputusan, fungsi pemisah atau fungsi regresi, digunakan beberapa contoh data yang mempunyai output atau label selama proses training. Tahap-tahap tersebut bersifat interaktif di mana pemakai terlibat langsung atau dengan perantaraan knowledge base [5].



Gambar. 1 : Tahapan - tahapan Data Mining

3. Algoritma Naïve Bayes.

Naïve Bayes merupakan sebuah metode klasifikasi yang berakar pada teorema Bayes. Metode pengklasifikasian dengan menggunakan metode probabilitas dan statistik yg dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes. Ciri utama dr Naïve Bayes Classifier ini adalah asumsi yg sangat kuat (naïf) akan independensi dari masing-masing kondisi atau kejadian [6]. Teorema tersebut dikombinasikan dengan *Naive* dimana diasumsikan kondisi antar atribut saling bebas. Klasifikasi *Naive Bayes* diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lainnya [7].

Persamaan dari teorema Bayes adalah :

$$P(H|X) = P(X|H) \cdot P(H) / P(X)$$

Keterangan :

X : Data dengan class yang belum diketahui

H : Hipotesis data X merupakan suatu class spesifik

$P(H|X)$: Probabilitas hipotesis H berdasar kondisi X (posteriori probability)

$P(H)$: Probabilitas hipotesis H (prior probability)

$P(X|H)$: Probabilitas X berdasarkan kondisi pada hipotesis H

$P(X)$: Probabilitas X

4. Klasifikasi.

Data Mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu : Deskripsi, Estimasi, Prediksi, Klasifikasi, Pengklusteran, dan Asosiasi. Klasifikasi merupakan bagian dari algoritma data mining, klasifikasi ini adalah algoritma yang menggunakan data dengan target (*class/label*) yang berupa nilai kategorikal/nominal. Menurut Gorunescu [4] proses klasifikasi didasarkan pada empat komponen mendasar, yaitu: 1. Kelas (*Class*), 2. Prediktor (*Predictor*), 3. Pelatihan dataset (*Training dataset*), 4. Data set pengujian (*Testing Dataset*)

5. Confusion Matrix

Confusion Matrix adalah *tool* yang digunakan untuk evaluasi model klasifikasi untuk memperkirakan objek yang benar atau salah. Sebuah matrix dari prediksi yang akan dibandingkan dengan kelas yang asli dari inputan atau dengan kata lain berisi informasi nilai actual dan prediksi pada klasifikasi [2].

Rumus untuk menghitung tingkat akurasi pada matriks adalah :

$$\text{Akurasi} = \frac{TP + TN}{TP + FP + TN + FN} = \frac{a + d}{a + b + c + d} \times 100\%$$

3. Metoden Penelitian.

1. Metode Pengumpulan data awal.

Dalam penyusunan penelitian ini metode yang digunakan dalam proses pengumpulan data antara lain : a. Studi Pustaka Studi pustaka dilakukan dengan melakukan pengumpulan informasi yang sesuai dengan topik atau masalah yang sedang diteliti. Informasi didapatkan melalui buku ilmiah, jurnal penelitian, tesis dan disertasi serta sumber – sumber tercetak maupun elektronik. b. Dokumentasi Pengumpulan data dilakukan dengan cara pengumpulan dokumen tertulis maupun dokumen elektronik dari lembaga / institusi. Dokumen ini diperlukan guna mendukung kelengkapan penelitian. c. Wawancara Wawancara dilakukan secara langsung dengan kepala cabang sebagai pihak yang terkait secara langsung dengan penelitian yang sedang dilakukan dengan cara mengajukan pertanyaan sesuai kebutuhan. Informan dapat menyampaikan informasi sebagaimana yang diharapkan oleh peneliti [8][9][10]

2. Seleksi data

Seleksi data merupakan pemilihan data yang akan digunakan dalam proses algoritma Naïve Bayes. Tujuan dari seleksi data adalah mendapatkan himpunan data target, pemilihan himpunan data atau memfokuskan data pada

subset variable atau sample data dimana penemuan akan dilakukan.

3. Eksperimen

Dalam melakukan eksperimen, penulis melakukan beberapa tahap agar mendapatkan hasil yang maksimal, di antaranya :1. Tahap Pemahaman Belajar Daring di masa Pandemi Covid 19, 2. Tahap Pemahaman (*understanding*) Data, 3. Tahap Persiapan (*preparation*) data, 4. Tahap modeling, 5. Tahap Evaluation, 6. Tahap Deployment [10].

4. Hasil dan Pembahasan.

1. Pengumpulan data awal.

Pada tahap awal, data di peroleh dari pengumpulan data yang didapatkan dari hasil belajar daring siswa SMA Negeri 3 Selong selama masa pandemic covid 19 TP. 2019/2020 berupa data private. Data yang di peroleh sebanyak 174 sample data.

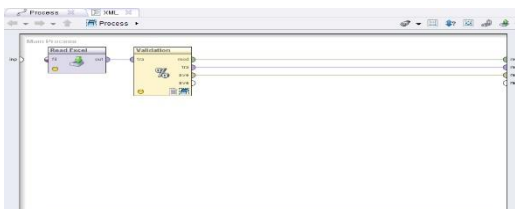
Dan Adapun beberapa atribut yang digunakan berjumlah 12 atribut yang terdiri dari : Nomor, Nama siswa, Hadir, Tidak, Jumlah Pertemuan, Presentasi Kehadiran, Tugas, MID, UAS, Total, KKM dan Predikat. Berikut adalah beberapa gambaran dari sample yang digunakan dalam penelitian ini :

No	NAMA SISWA	HA DIR	TIDA K	JUMLAH PERTEM UAN	PRESENT ASE KEHADIR AN (%)	TUGAS	MD	UA S	TOT AL	KKII	PREDIKAT
1	ADITYA DWI PRIMA NANCA	6	4	10	0,6	80	89	85	77	75	TUNTAS
2	AQUSTIAN PRIMARYA PUTRA	7	3	10	0,7	80	88	90	79	75	TUNTAS
3	ALANG ARTHA IWANA	8	2	10	0,8	80	87	90	79	75	TUNTAS
4	ARCIAN LAUJHARI PRATAMA	9	1	10	0,9	80	80	90	77	75	TUNTAS
5	ARIPITA RAMADHAN	7	3	10	0,7	80	87	90	79	75	TUNTAS
6	BAEDHOWI NAUFAL KURNIA	8	2	10	0,8	80	78	90	76	75	TUNTAS
7	BACUS ROFIQUL AKBAR	5	5	10	0,5	80	80	80	72	75	TIDAK TUNTAS
8	BAHTIAR EFFENDY	4	6	10	0,4	80	78	78	70	75	TIDAK TUNTAS
9	BAIQ INTAN PURNAMAWATI	7	3	10	0,7	88	88	80	76	75	TUNTAS
10	ELYANA INTAN LESTARI	9	1	10	0,9	80	90	90	80	75	TUNTAS
11	FATHUL HADI	8	2	10	0,8	80	88	78	77	75	TUNTAS
12	GILANG JATI SIWARA	9	1	10	0,9	80	90	89	80	75	TUNTAS
13	ILHAM ABDIANSYAH	7	3	10	0,7	80	88	87	78	75	TUNTAS
14	M. GHOZALI MARI	8	2	10	0,8	80	98	78	77	75	TUNTAS
15	M. JIHAT	7	3	10	0,7	88	87	80	75	75	TUNTAS
16	M. KHARUL AZMI	9	1	10	0,9	89	90	90	81	75	TUNTAS
17	MAULANA BAGINDA PUTRA	8	2	10	0,8	80	90	80	76	75	TUNTAS

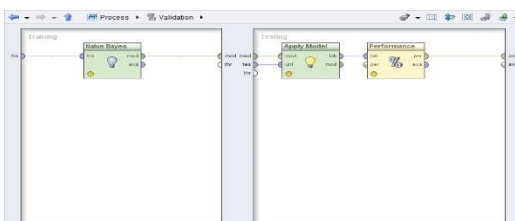
Gambar. 2 : Data Hasil Ketuntasan Siswa Pasca Pandemi Covid 19

2. Pengolahan data dengan Algoritma Naïve Bayes.

Pengolahan data menggunakan algoritma Naïve Bayes menggunakan tools rapidminer ditunjukkan oleh gambar berikut:

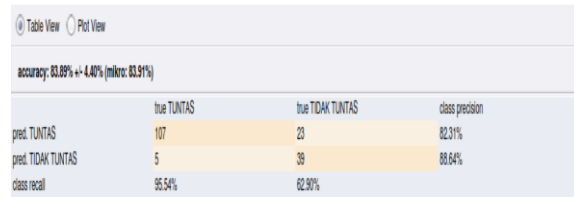


Gambar. 3 : Tampilan Pengolahan menggunakan x-validation



Gambar. 4 : Tampilan Area Kerja RapidMiner dengan Naïve Bayes

Diperoleh hasil akurasi yang diujikan adalah 83,89%, berikut gambar hasil accuracy yang telah diujikan.



	true TUNTAS	true TIDAK TUNTAS	class precision
pred. TUNTAS	107	23	82,31%
pred. TIDAK TUNTAS	5	39	88,64%
class recall	95,54%	62,90%	

Gambar. 5 : Hasil accuracy dengan K-Fold Validation 4

Dari gambar diatas dapat dijelaskan bahwa pengujian yang dilakukan dengan menerapkan nilai K-Fold Validation 4 menghasilkan Jumlah :

- Jumlah *True Positive (TP)* = 107 Artinya bahwa jumlah data ini sebagai data (Tuntas) sesuai dengan prediksi yang dilakukan yaitu (Tuntas).
- False Negative (FN)* = 23 artinya data diprediksi sebagai data (Tuntas) tetapi ternyata (Tidak Tuntas)
- True Negative (TN)*= 39 artinya data diprediksi sebagai (Tidak Tuntas) sesuai dengan prediksi (Tidak Tuntas), dan
- False Positive (FP)* = 5 artinya data diprediksi sebagai (Tidak Tuntas) ternyata masuk dalam klasifikasi (Tuntas).

Tingkat akurasi yang dihasilkan dengan menggunakan algoritma *Naive Bayes* adalah sebesar 83,89 %, kemudian setelah diidentifikasi *True Positive (TP)*, *False Negative (FN)*, *True Negative (TN)*, dan *False Positive (FP)* maka dapat dihitung nilai *accuracy*, *sensitivity*, *specificity*, *ppv*, dan *npv* pada persamaan dibawah ini:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$= \frac{107 + 39}{107 + 39 + 5 + 23}$$

$$= 83,89\%$$

$$sensitivity = \frac{TP}{FP + FN} = \frac{107}{5 + 23}$$

$$= 82,31\%$$

$$Specificity = \frac{TN}{TN + FP} = \frac{39}{39 + 5}$$

$$= 88,64\%$$

$$PPV = \frac{TP}{TP + FP} = \frac{107}{107 + 5}$$

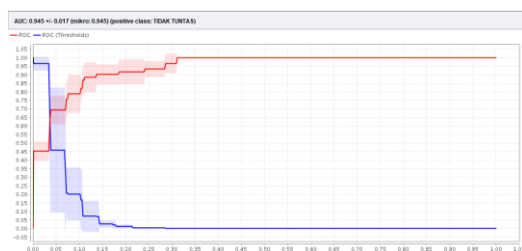
$$= 95,54\%$$

$$NPV = \frac{TN}{TN + FN} = \frac{39}{39 + 23}$$

$$= 62,90\%$$

Hasil perhitungan penggunaan rumus diatas dapat dilihat dalam tabel berikut :

NAIVE BAYES	
	Nilai (%)
Accuracy	83,89%
Sensitivity	82,31%
Specificity	88,64%
PPV	95,54%
NPV	62,90%



Gambar. 6 : Nilai AUC dalam Grafik ROC dengan K-Fold Validition 4

Pada gambar menjelaskan Grafik ROC dengan nilai AUC (Area Under Curve) menghasilkan nilai sebesar 0.945 dengan tingkat diagnose nilai akurasi *good classification*.

3. Evaluasi dan Hasil Validasi.

Dari uji coba menggunakan *K-Fold Validation* 2 sampai 10. Mengapa harus sampai 10 kali pengujian?, ini karena dari para ahli serta berdasarkan percobaan yang ekstensif dan pembuktian teoritis menunjukkan bahwa 10 k-fold validation adalah pilihan terbaik untuk mendapatkan hasil validasi yang akurat, bisa saja jika dilakukan pengujian data lebih dari 10 kali akan tetapi tidak akan menjadi efektif dan akurat jika data dibagi terlalu banyak.

Hasil uji coba K-Fold Validation 2 sampai 10.

K-Fold Validation	Accuracy	Precision	Recall	AUC
2	81.61%	87.99%	56.51%	0.933
3	81.61%	87.99%	56.51%	0.933
4	83.89%	88.72%	63.02%	0.945
5	82.74%	88.47%	59.49%	0.937
6	82.18%	89.45%	58.48%	0.942
7	82.19%	90.79%	57.34%	0.945
8	82.20%	88.54%	58.48%	0.946
9	81.64%	89.97%	56.35%	0.933
10	82.12%	90.00%	58.10%	0.938

Berdasarkan table di atas,dapat disimpulkan bahwa nilai dengan hasil terbaik yaitu terdapat pada K-Fold Validation 4 dengan nilai accuracy sebesar 83.89%, Precision 88.72%.Recall 63.02%,dan nilai Area Under Curve (AUC) adalah 0.945 dengan predikat Good Clasification.

Nilai akurasi

	K-Fold Validation 4	K-Fold Validation 5
Accuracy	83.89%	82.74%

Berdasarkan hasil uji coba maka didapatkan *accuracy* tertinggi pada dua *K-Fold Validation* yaitu pada *K-Fold Validation* 4 dan *K-Fold*

Validation 5. Dimana didapatkan hasil *accuracy* sebesar 83.89% pada *K-Fold Validation 4* dan *accuaracy* sebesar 82.74% pada *K-Fold Validation 5*. Dan selisih akurasi keduanya adalah 1% yang menunjukkan bahwa terjadi peningkatan 0.80 % pada rentang *k-fold validation 4* dan 5. Peningkatan pada setiap uji coba tidaklah selalu sama ini dipengaruhi oleh seberapa banyak pembagian data yang dilakukan.

5. Kesimpulan.

Berdasarkan implementasi dan pembahasan dalam penelitian dilakukan untuk memprediksi ketuntasan belajar Siswa SMA Negeri 3 Selong, selama pandemic pasca covid 19, maka dapat disimpulkan bahwa dalam menentukan data tuntas dan tidak tuntas dapat diprediksi dan dievaluasi dengan memanfaatkan teknik *data mining* menggunakan algoritma *Naïve Bayes*. Algoritma ini mampu menganalisa pola data tuntas dan tidak tuntas, dalam mengambil kebijakan untuk melengkapi data siswa yang di katagorikan tuntas dan tidak tuntas dengan memanfaatkan teknik data mining, dimana dilakukan dengan cara melakukan menampilkan *output* berupa nilai akurasi terbaik dan akurat. 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.

6. Daftar Pustaka.

- [1] H. Ouassou *et al.*, "The Pathogenesis of Coronavirus Disease 2019 (COVID-19): Evaluation and Prevention," *J. Immunol. Res.*, vol. 2020, no. July, 2020, doi: 10.1155/2020/1357983.
- [2] M. S. Muhammad Saiful, "Algoritma Support Vector Machines (SVM) Berbasis Particle Swarm Optimization (PSO) Untuk Klasifikasi Kinerja Dosen STT Hamzanwadi."
- [3] N. Hayati, "Metode Pembelajaran Daring," *E-learning Yang Ef. Bali Jur. Ilmu Pendidik.* ..., no. April, 2020.
- [4] M. S. Alimuddin, "Implementation of the Neural Network (NN) Algorithm in Analysis of Student Class Increment Data Based on Report Card Value Implementation of the Neural Network (NN) Algorithm in Analysis of Student Class Increment Data Based on Report Card

- Value .,” *Conf. Ser.*, vol. 1539, pp. 0–6, 2020, doi: 10.1088/1742-6596/1539/1/012034.
- [5] H. B. Suhatini, “Klasifikasi Algoritma K-Nearest Neighbor Berbasis Particle Swarm Optimization Untuk Kelayakan Bantuan Rehabilitasi Rumah Tidak Layak Huni Pada Desa Lenek Duren Kecamatan Aikmel,” *Klasifikasi Algoritma K-Nearest Neighbor Berbasis Part. Swarm Optim. Untuk Kelayakan Bantu. Rehabil. Rumah Tidak Layak Huni Pada Desa Lenek Duren Kec. Aikmel*, vol. 8, no. 5, p. 55, 2019.
- [6] M. Hamonangan Nasution and M. Iqbal Sofyan, “Penerapan Metode Decision Tree Untuk Memprediksi Prestasi Siswa Kelas XII Dilihat dari Nilai Akhir Semester di SMK Negeri 1 Selong Tahun Pelajaran 2017/2018,” *Infotek J. Inform. dan Teknol.*, vol. 3, no. 1, pp. 58–65, 2020, doi: 10.29408/jit.v3i1.1825.
- [7] B. H. Amri Muliawan Nur, “Komparasi Algoritma SVM Dan SVM Berbasis PSO Dalam Menganalisa Kinerja Guru SMAN 3 Selong,” vol. 2, no. 2, pp. 15–16, 2019.
- [8] N. Muhammad Saiful, Aris Sudianto, “Penerapan Sistem Informasi Tracher Study Untuk Mengetahui Tingkat Kontribusi Perguruan Tinggi,” vol. 2, no. 1, pp. 573–574, 2019.
- [9] A. M. N. MUhammad Saiful, “Application of Expert System with Web-Based Forward Chaining Method in Diagnosing Corn Plant Disease Application of Expert System with Web-Based Forward Chaining Method in Diagnosing Corn Plant Disease,” *Conf. Ser.*, vol. 1539, pp. 0–7, 2020, doi: 10.1088/1742-6596/1539/1/012019.
- [10] I. F. Muhammad Saiful, LM Samsu, “Perancangan Kerangka Crowdsourcing Berbasis Wisdom Of Crowds Untuk Kamus Naskah Lontar (Takepan) Sasak Online,” *Infotek J. Inform. dan Teknol.*, vol. 3, no. 2, pp. 165–173, 2020, doi: 10.29408/jit.v3i2.2300.