

Prediksi Kemampuan Siswa Dalam Bersaing di Dunia Kerja Menggunakan Perbandingan Algoritma *Naïve Bayes* Dan *K-Nearest Neighbor*

Andi Purnomo^{*1}, Ahmad Sururi²,

^{1,2} Program Studi Magister Ilmu Komputer Universitas Budi Luhur Jakarta
E-mail: ^{*1}andipurnomo999@gmail.com, ²ahmadsururi71@gmail.com

Abstrak

SMK Yadika 5 Pondok Aren setiap tahunnya meluluskan + 200 siswa/i nya. Semakin ketatnya persaingan dalam dunia kerja maka kualitas pendidikan pada siswa/i SMK Yadika 5 Pondok Aren adalah salah satu tujuan dari sekolah. Kurangnya daya serap lulusan SMK Yadika 5 Pondok Aren di dunia kerja menjadi salah satu persoalan penting yang harus segera diperbaiki, belum adanya sebuah pola dalam memprediksi siswa yang belum mampu memenuhi beberapa faktor dalam perekrutan karyawan di sebuah perusahaan maupun industri. Oleh karena itu, perlu adanya sebuah sistem untuk menemukan sebuah pola dalam melakukan prediksi siswa/i SMK Yadika dalam bersaing di dunia kerja. Salah satu Teknik pengolahan data yang dapat digunakan adalah Data Mining. Dalam pengolahan Data Mining ini peneliti menggunakan metode klasifikasi dengan algoritma *Naïve Bayes* dan membandingkan dengan algoritma *K-Nearest Neighbor* serta melakukan pengujian pada nilai precision, nilai recall, nilai akurasi. Berdasarkan pengujian yang telah dilakukan dalam penelitian ini mendapati hasil bahwa Algoritma *Naïve Bayes* memiliki Akurasi 97.66 % , Nilai Precision 100% , dan Nilai Recall 97.59% sedangkan Algoritma *K-Nearest Neighbor* mendapati hasil Nilai Akurasi 98.22 % , Nilai Precision 99.38% , Nilai Recall 98.77 % . Dari hasil yang didapatkan maka algoritma *Naïve Bayes* dan *K-Nearest Neighbor* dapat diterapkan serta dijadikan suatu pola baru dalam memprediksi siswa dalam memenuhi beberapa faktor dalam perekrutan karyawan di sebuah perusahaan maupun industri.

Keywords— Data Mining, Algoritma *Naïve Bayes*, Algoritma *K-Nearest neighbor*

Abstract

*SMK Yadika 5 Pondok Aren annually graduates + 200 students. The more intense competition in the world of work, the quality of education for students of SMK Yadika 5 Pondok Aren is one of the goals of the school. The lack of absorption of graduates of SMK Yadika 5 Pondok Aren in the world of work is one of the important problems that must be corrected immediately, there is no pattern in predicting students who have not been able to overcome several factors in employees in an industrial company. Therefore, it is necessary to have a system to find a pattern in predicting Yadika Vocational High School students in competing in the world of work. One of the data processing techniques that can be used is Data Mining. In processing Data Mining, the researcher uses a classification method with the *Naïve Bayes* algorithm and compares it with the *K-Nearest Neighbor* algorithm and tests the precision value, recall value, and accuracy value. Based on the tests that have been carried out in this study found the results that the *Naïve Bayes* Algorithm has 97.66% Accuracy, 100% Precision Value, and 97.59% Recall Value while the *K-Nearest Neighbor* Algorithm gets 98.22% Accuracy Value, Precision Value 99.38% , Recall value 98.77%. From the results obtained, the *Naïve Bayes* and *K-Nearest Neighbor* algorithms can be applied and a new pattern is applied in predicting students in fulfilling several factors in the achievement of employees in a company or industry.*

Keywords— Data Mining, Algoritma *Naïve Bayes*, Algoritma *K-Nearest neighbor*

1. PENDAHULUAN

Perkembangan teknologi yang pesat di era modern ini sudah banyak memberikan manfaat dalam berbagai aspek, khususnya kontribusi untuk mempermudah pekerjaan manusia dalam bentuk komputerisasi. Dengan adanya perkembangan teknologi, sesuatu hal yang belum pernah terjadi akan sangat mudah untuk di prediksi dengan teknologi meskipun tingkat keakuratannya tidak mencapai 100%.

Sekolah Menengah Kejuruan (SMK) merupakan salah satu lembaga pendidikan yang berupaya menciptakan sumber daya manusia yang memiliki kemampuan, keterampilan, dan keahlian sehingga lulusannya dapat mengembangkan keterampilannya agar dapat terjun dalam dunia kerja maupun untuk melanjutkan ke jenjang pendidikan yang lebih tinggi. SMK Yadika 5 Pondok Aren merupakan salah satu instansi Pendidikan Sekolah Menengah Kejuruan (SMK) yang berkawasan di Kecamatan Pondok Aren, yang setiap tahunnya SMK Yadika 5 Pondok Aren meluluskan + 200 siswa/i nya.

Terdapat beberapa faktor pada saat mengikuti perekrutan di perusahaan tersebut yang menjadi standar poin penilaian bagi calon karyawan dengan Pendidikan terakhir SMK/Sederajat, diantaranya ialah nilai rata-rata Ujian Kompetensi Kejuruan, Nilai Praktek Kerja Lapangan, disiplin, tanggung jawab, sikap, dan kemampuan komunikasi dalam team. Untuk memastikan bahwa siswa/i SMK Yadika 5 mendapatkan kesempatan bekerja setelah mereka lulus, maka SMK Yadika 5 dapat melakukan pengumpulan data berdasarkan hasil nilai mata pelajaran, nilai ujian akhir, nilai praktek kerja lapangan dan nilai perilaku keseharian dari setiap siswa/i mulai dari kelas 1 hingga kelas 3, dimana hasil dari proses pembelajaran tersebut setiap tahunnya akan diinput oleh wali kelas dari setiap masing-masing kelas, maka dengan data tersebut SMK Yadika 5 dapat melihat perkembangan dari siswa/i setiap tahunnya. Terdapat beberapa metode yang dapat digunakan untuk memprediksi kemampuan siswa/i SMK Yadika 5, salah satunya ialah penggunaan metode klasifikasi untuk mendapatkan sebuah informasi untuk pihak sekolah agar dapat mengarahkan, membimbing, serta memberikan motivasi kepada siswa/i SMK Yadika 5 yang belum memenuhi beberapa faktor dalam perekrutan karyawan di sebuah perusahaan maupun industri, agar mereka dapat bersaing dengan calon karyawan yang lainnya. Dalam penelitian ini peneliti mengidentifikasi permasalahan yang terjadi ialah belum adanya suatu pola dalam memprediksi siswa yang belum mampu memenuhi beberapa faktor dalam perekrutan karyawan di sebuah perusahaan maupun industri berdasarkan nilai mata pelajaran hingga nilai perilaku keseharian. Dari permasalahan tersebut perlu adanya sebuah solusi untuk menemukan sebuah pola dalam melakukan prediksi siswa/i SMK Yadika dalam bersaing di dunia kerja, salah satu teknik pengolahan data yang dapat digunakan adalah Data Mining. Dalam pengolahan Data Mining ini peneliti menggunakan metode klasifikasi dengan algoritma Naïve Bayes dan membandingkan dengan algoritma K-Nearest Neighbor agar dapat mengetahui serta mengukur akurasi yang dihasilkan oleh algoritma tersebut.

Data Mining merupakan suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data mining adalah proses menganalisa data dari perspektif yang berbeda dan menyimpulkannya menjadi informasi-informasi penting yang dapat dipakai untuk meningkatkan keuntungan, memperkecil biaya pengeluaran, atau bahkan keduanya. Secara teknis, data mining dapat disebut sebagai proses untuk menemukan korelasi atau pola dari ratusan atau ribuan field dari sebuah relasional database yang besar. data mining juga merupakan proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data. Informasi yang dihasilkan diperoleh dengan cara mengekstraksi dan mengenali pola yang penting atau menarik dari data yang terdapat dalam basis data (Nofriansyah et al,2016). Data mining berisi pencarian trend atau pola yang diinginkan dalam database besar untuk membantu pengambilan keputusan di waktu yang akan datang. Pola-pola ini dikenali oleh perangkat tertentu yang dapat memberikan suatu analisa data yang berguna dan berwawasan yang kemudian dapat dipelajari dengan lebih teliti, yang mungkin saja menggunakan perangkat pendukung keputusan lainnya (Mustafa et al, 2017).

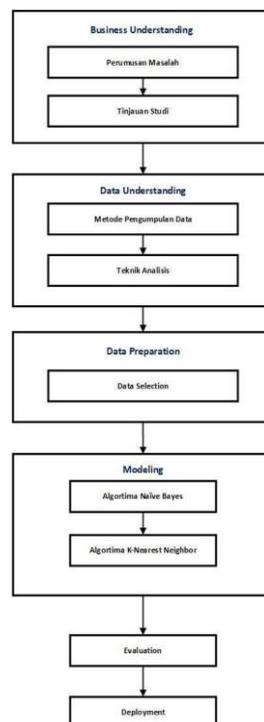
Algoritma Naïve Bayes merupakan sebuah metode klasifikasi yang berdasar pada teorema Bayes. Metode pengklasifikasian ini menggunakan metode probabilitas dan statistik, Ciri utama dari Naive Bayes ini adalah asumsi yang sangat kuat akan independensi dari masing-masing kondisi atau kejadian (Mustafa et al, 2018). Keuntungan penggunaan algoritma Naïve Bayes adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (Training Data) yang kecil untuk menentukan estimasi parameter yang dibutuhkan dalam proses pengklasifikasian(Saleh, 2015). Naive bayes termasuk ke

dalam pembelajaran supervised, sehingga pada tahapan pembelajaran dibutuhkan data awal berupa data pelatihan untuk dapat mengambil keputusan. Pada tahapan pengklasifikasian akan dihitung nilai probabilitas dari masing-masing label kelas yang ada terhadap masukan yang diberikan. Label kelas yang memiliki nilai probabilitas paling besar yang akan dijadikan label kelas data masukan tersebut. Naive bayes merupakan perhitungan teorema bayes yang paling sederhana, karena mampu mengurangi kompleksitas komputasi menjadi multiplikasi sederhana dari probabilitas. Selain itu, algoritma naive bayes juga mampu menangani set data yang memiliki banyak atribut (Sartika et al, 2017). Algoritma Naive Bayes memberikan suatu cara mengkombinasikan peluang terdahulu dengan syarat kemungkinan menjadi sebuah formula yang dapat digunakan untuk menghitung peluang dari setiap kemungkinan yang terjadi.

Pada penelitian sebelumnya Kusuma (2019) penggunaan Algoritma Naïve Bayes pada prediksi kemampuan lulusan SMK Buddhi Tangerang dapat memberikan hasil yang baik, karena memiliki nilai akurasi sebesar 98% dan juga pada penelitian Khoirunnisa (2018) mengungkapkan bahwa tujuan dari metode klasifikasi data mining untuk meningkatkan keandalan hasil yang diperoleh dari data. Serta penggunaan Confusion Matrix sebagai evaluasi untuk mendapatkan nilai akurasi, recall, dan precision yang baik. Oleh Karena itu, diharapkan dengan penerapan metode klasifikasi data mining dengan Algoritma Naïve Bayes dan Algoritma K-Nearest Neighbor dapat memprediksi kemampuan siswa/i SMK Yadika 5 dalam memenuhi beberapa faktor perekrutan karyawan di sebuah perusahaan maupun industri, serta menghasilkan nilai akurasi yang baik agar dapat menjadi acuan dalam meningkatkan kualitas pembelajaran SMK Yadika 5 agar menjadi lebih baik. Dan dapat diterapkan serta dijadikan suatu pola baru dalam memprediksi siswa dalam memenuhi beberapa faktor dalam perekrutan karyawan di sebuah perusahaan maupun industri agar dapat memastikan siswa/i SMK Yadika 5 mendapatkan kesempatan bekerja setelah mereka lulus.

2. METODE PENELITIAN

Pada penelitian ini menggunakan metodologi CRIPS-DM untuk melakukan Analisa dan mengelola data. Tahapan penelitian ini dibagi ke dalam beberapa tahap yang dapat di lihat pada gambar berikut.



Gambar 1 Metode Penelitian

1. Perumusan Masalah

Dalam penelitian ini peneliti mengidentifikasi permasalahan yang terjadi ialah belum adanya suatu pola dalam memprediksi siswa yang belum mampu memenuhi beberapa faktor dalam perekrutan karyawan di sebuah perusahaan maupun industri berdasarkan nilai mata pelajaran hingga nilai perilaku keseharian. Dari permasalahan tersebut perlu adanya sebuah solusi untuk menemukan sebuah pola dalam melakukan prediksi siswa/i SMK Yadika dalam bersaing di dunia kerja, salah satu teknik pengolahan data yang dapat digunakan ialah *Data Mining*. Dalam pengolahan Data Mining ini peneliti menggunakan metode klasifikasi dengan algoritma Naïve Bayes dan membandingkan dengan algoritma *K-Nearest Neighbor* agar dapat mengetahui serta mengukur akurasi yang dihasilkan oleh algoritma tersebut.

2. Tinjauan Studi

Tinjauan Studi ini dilakukan untuk mendapatkan teori mengenai metode klasifikasi *data mining*, algoritma *Naïve bayes*, Algoritma *K-Nearest Neighbor* yang terkait dengan rumusan permasalahan. Metode yang dilakukan penulis adalah membaca berbagai jurnal ilmiah terkait dengan metode klasifikasi *Data Mining*, Algoritma *K-Nearest Neighbor* dan *algoritma Naïve Bayes*. Hasil dari langkah ini adalah literatur yang terkait dengan perumusan masalah.

3. Pengumpulan Data

Untuk mendapatkan sebuah informasi, maka tentu saja dibutuhkan sumber data yang akurat dalam penyusunannya. Untuk itu penulis melakukan pengumpulan data dengan mengambil langsung dari SMK Yadika 5 Pondok Aren. Data ini terdiri dari hasil penilaian siswa/i SMK Yadika 5 Pondok Aren.

No	Tahun Ajaran	Keterangan	Jumlah
1	2016/2017	Alumni	177
2	2017/2018	Alumni	176
3	2018/2019	Alumni	194
4	2019/2020	Alumni	151
5	2020/2021	Kelas X	171
		Kelas XI	157
		Kelas XII	169
Total			1195

Gambar 2 Dataset Penelitian

Kemudian dilakukan Teknik analisis data yang bertujuan untuk mengatasi masalah yang ada berdasarkan data yang diperoleh. Analisis data yang dilakukan yaitu mengumpulkan data kuantitatif. Data kuantitatif tersebut yang nantinya akan diolah dan digunakan dalam penelitian ini.

4. Data Selection

Proses *data selection* ini dilakukan dengan tujuan memilih atribut yang relevan pada data untuk dapat dilakukan suatu analisis data. Pemilihan atribut tersebut di dasari oleh data yang diterima penulis dari SMK Yadika 5 Pondok Aren yang terdapat beberapa atribut (*variable*) yang berpengaruh kepada kemampuan siswa/i SMK Yadika 5 dalam bersaing di dunia kerja. Berikut atribut-atribut yang akan digunakan antara lain.

No	Atribut	Keterangan
1	JK	Jenis Kelamin
2	JUR	Jurusan
3	UKK	Nilai Ujian Kenaikan kelas
4	PROD	Nilai Mata Pelajaran Produktif
5	PKL	Nilai Praktek Kerja Lapangan
6	TJ	Nilai Tanggung Jawab
7	DIS	Nilai Disiplin
8	KOM	Nilai Komunikasi
9	PRI	Nilai Prilaku
10	KM	Nilai Keterampilan
11	KET	Keterangan Mampu atau Belum Mampu

Gambar 3 Atribut yang Digunakan

5. Algoritma Naive Bayes

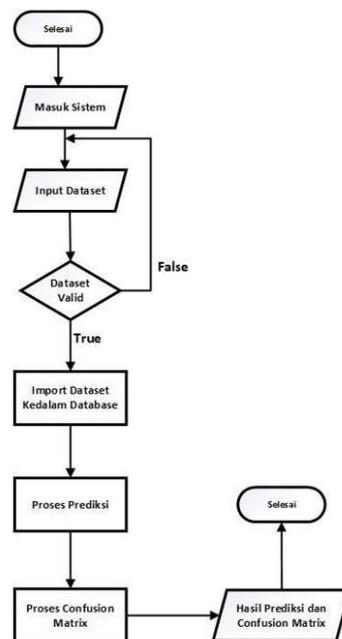
Algoritma Naïve Bayes merupakan sebuah metode klasifikasi yang berdasar pada *teorema Bayes*. Metode pengklasifikasian ini menggunakan metode probabilitas dan statistik, Ciri utama dari *Naive Bayes* ini adalah asumsi yang sangat kuat akan independensi dari masing-masing kondisi atau kejadian(Mustafa et al, 2018).

6. Algoritma K-Nearest Neighbor

K-Nearest Neighbor (K-NN) merupakan sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. K-NN termasuk algoritma *supervised learning* dimana hasil dari *query instance* yang baru diklasifikasi berdasarkan mayoritas dari kategori pada K-NN(Syaliman et al,2017). *K-Nearest Neighbor* melakukan pendekatan untuk mencari kasus dengan menghitung kedekatan antara kasus baru dengan kasus lama, yaitu berdasarkan pada pencocokan bobot dari sejumlah fitur yang ada(Umaidah dan Purwantoro,2019).

7. Evaluation

Pada tahapan ini merupakan penerapan dari penelitian menggunakan metode klasifikasi *data mining* menggunakan *Algoritma Naïve Bayes* dan *Algoritma K-Nearest Neighbor* yang di implementasikan pada sebuah sistem. Sistem ini dimulai dengan masukan berupa *data training* dan *data uji* dari *user* yang kemudian akan masuk ke dalam *database* sistem, kemudian akan dilakukan proses prediksi dengan metode klasifikasi *data mining* menggunakan algoritma *Naïve Bayes* dan *Algoritma K-Nearest Neighbor*, Setelah proses prediksi selesai maka akan muncul hasil dari *confusion matrix* berupa akurasi, *recall*, *precision* dan hasil dari prediksinya.



Gambar 4 Diagram Alur Perancangan

8. Deployment

Sistem ini terdiri dari beberapa proses utama yaitu *Data cleaning*, *Data Selection*, perhitungan Standar deviasi, perhitungan *Mean*, perhitungan probabilitas, perhitungan *Densitass Gaus*, Perhitungan *Manhattan Distance* dan melakukan pengujian menggunakan *Confusion Matrix*

3. HASIL DAN PEMBAHASAN

Lokasi penelitian ini dilakukan di SMK Yadika 5 Pondok aren. Tujuan dari penelitian ini adalah untuk memprediksi kemampuan siswa dan siswi SMK Yadika 5 Pondok aren falam bersaing di dunia

kerja menggunakan perbandingan algoritma naïve bayes dan k-nearest. Dalam penelitian ada beberapa tahapan sebagai berikut.

1. Pengumpulan Data

Data yang dimasukan ke dalam *database* akan digunakan sebagai objek utama dalam penelitian ini. Sebagai sumber informasi yang akan dilakukan proses prediksi terhadap kemampuan siswa/I SMK Yadika 5 Pondok Aren menggunakan Algoritma *Naïve Bayes* dan Algoritma *K-Nearest Neighbor* setelah itu akan dilakukan pengujian menggunakan *Confusion Matrix* untuk menentukan nilai Akurasi, *Recall*, *Precision*. Data yang penulis gunakan sebagai data model dan data uji ialah data berdasarkan gambar yang dapat dilihat pada gambar 5.

DATA							
Jumlah data : 171 Data							
NIS	NAMA	JENIS KELAMIN	JURUSAN	UKK	NILAI PRODUKTIF	PKL	TANGGUNG
41200001	ABELIA SYAHLA TALITHA	P	AK	79.0	79.0	88.0	83.0
41200002	ACHMAD ZAIDAN	L	TKJ	85.0	83.0	82.0	82.0
41200003	ADE IHAH SOLIHAH	L	AP	78.0	84.0	85.0	77.0
41200004	ADIMAS RESTU RAMADHAN	L	TKJ	79.0	77.0	82.0	77.0
41200005	ADINDA VANIA KHAYANTI	P	AP	76.0	77.0	81.0	81.0
41200006	AISYAH ALUSMAYANTI	P	AP	75.0	73.0	85.0	75.0
41200007	AJENG PRAMESWARI	P	AK	78.0	80.0	86.0	82.0
41200008	ALLESANDRO BAGINTA C...	L	TKJ	83.0	80.0	82.0	75.0
41200009	ALSYIRA SUGANTI METIARI	P	AP	80.0	79.0	83.0	86.0
41200010	AMBANGRIJANI	P	AK	79.0	76.0	87.0	80.0
41200011	ANGGUN NUR ARIFIANI	P	AP	77.0	82.0	84.0	78.0
41200012	ANISA ZAHWA AZZAHROO	P	AP	79.0	82.0	86.0	84.0
41200013	ANJANI KIRANA LARASATI	P	AP	76.0	75.0	84.0	79.0
41200014	ANINDYA NABELLA SARI	P	AP	75.0	76.0	83.0	80.0
41200015	ASLI SETIANINGSIH	P	AK	76.0	75.0	85.0	75.0
41200016	ATIKA RIZKI APRILIYANTI	P	AP	75.0	75.0	81.0	75.0
41200017	AUDIA CAHYANI	P	AK	76.0	75.0	87.0	75.0
41200018	AULIA CAVITA AZHAR	P	AP	75.0	76.0	83.0	78.0
41200019	AULIA RETNO WILANDARI	P	AK	78.0	81.0	88.0	82.0
41200020	AYU PRAVESTI	P	AP	75.0	75.0	80.0	75.0
41200021	AZZAHRA DHEVA NIRMALA	P	AP	69.0	79.0	84.0	77.0
41200022	AZZAHRA KAYLA SAL SABILA	P	AK	78.0	86.0	86.0	75.0
41200023	BAYU BAGUS SATRIYO	L	AK	76.0	76.0	86.0	75.0
41200024	BIMO NALFAL ADILIAN	L	TKJ	80.0	81.0	83.0	76.0
41200025	DANU SISDA RAFFYANDE	L	TKJ	74.0	80.0	84.0	75.0
41200026	DARREN JACQUELYN HER...	L	AK	79.0	80.0	86.0	79.0
41200027	DEA NUR FEBRIANI	P	AK	78.0	79.0	85.0	76.0
41200028	DESTA MULYANA PUTRI	P	AK	79.0	80.0	84.0	81.0
41200029	DIAN WILANDARI	P	AP	77.0	79.0	80.0	83.0
41200030	DIMAS SEPTIawan	L	TKJ	73.0	75.0	88.0	79.0
41200031	DINA FARHURROCHMAH	P	AK	81.0	90.0	86.0	80.0
41200032	DINDA AMALLIA	P	AP	75.0	68.0	87.0	74.0
41200033	DINI NURHANE	P	AP	80.0	80.0	84.0	85.0
41200034	EBENEZER BANY SETYA...	L	TKJ	59.0	75.0	85.0	75.0
41200035	EKA SABRINA LESTARI	P	AP	74.0	72.0	85.0	76.0
41200036	ELISA DWI RAHMAWATI	P	TKJ	75.0	75.0	86.0	75.0
41200037	FANNY RENDU FAUZAN	P	AP	76.0	76.0	85.0	69.0
41200038	FEBRIANA SIMARAKATA	P	AP	69.0	71.0	81.0	78.0
41200039	FTONA DEVINDO	P	AK	78.0	78.0	85.0	78.0
41200040	FITRIA RAHMA RAHAYU	P	AP	73.0	77.0	80.0	77.0
41200041	FRANKO TRI AL EYANGFER. VO...	L	TKJ	80.0	76.0	81.0	78.0

Gambar 5. Tampilan Dataset

DATASET	
TAHUN AJARAN	JUMLAH DATA
2016 / 2017 Alumni	177
2017 / 2018 Alumni	176
2018 / 2019 Alumni	194
2019 / 2020 Alumni	151
2020 / 2021 X	171
2020 / 2021 XI	157
2020 / 2021 XII	169

Jumlah Dataset : 7 Dataset

Jumlah Data : 1195 Data

Delete

Gambar 6. Informasi Jumlah Dataset

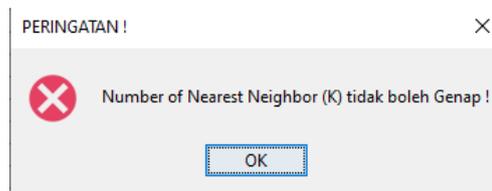
Jumlah *Dataset* yang telah dimasukan kedalam *database* akan digunakan sebagai data model dan data uji. Dimana jumlah data model harus lebih banyak dibandingkan data uji agar dapat dilakukan proses prediksi kemampuan siswa/I dalam bersaing di dunia kerja. untuk jumlah *Dataset* yang sudah di masukan ke dalam *database* dapat dilihat pada gambar 6.

2. Pengujian Sistem

Dalam tahapan ini, Sistem akan dijalankan dan diuji cobakan untuk mengetahui apakah sistem berjalan sesuai dengan hasil Analisa dan tujuan. Untuk mengetahui kemampuan sistem Prediksi kemampuan siswa/I dalam bersaing di dunia kerja yang telah dibangun dalam tesis ini, maka akan dilakukan pengujian dengan mengukur kualitas dari nilai akurasi, *precision* dan *recall* dari algoritma *Naïve Bayes* dan Algoritma *K-Nearest Neighbor*. Manakah di antara kedua Algoritma tersebut yang mendapatkan nilai akurasi yang paling baik. Pada gambar 3 adalah jumlah dataset yang akan digunakan untuk melakukan pengujian dari Algoritma *Naïve Bayes* dan Algoritma *K-Nearest Neighbor*.

3. Klasifikasi Algoritma *K-Nearest Neighbor*

Klasifikasi menggunakan Algoritma *K-Nearest Neighbor* berfungsi untuk memprediksi kemampuan siswa/I dalam bersaing didunia kerja menggunakan *dataset* yang belum memiliki keterangan mampu dan belum mampu. data model yang digunakan dalam pengujian ini merupakan *dataset* terdapat pada *database*, serta *dataset* uji yang digunakan dalam pengujian ini menggunakan 497 data yang berasal dari tahun ajaran 2020/2021 Kelas X, Kelas XI, dan Kelas XII dengan menggunakan *Number of Nearest Neighbor (K) = 3* dan *Number of Nearest Neighbor (K) = 5*. *Number of Nearest Neighbor (K)* yang dapat digunakan ialah nilai dari angka ganjil apabila yang diinputkan selain dari nilai angka ganjil maka sistem akan menampilkan sebuah *messages box* yang dapat di lihat pada gambar 7.



Gambar 7 Tampilan Messages Box

4. Klasifikasi *Number of Nearest Neighbor (K) = 3*

Hasil dari pengklasifikasian menggunakan algoritma *K-Nearest Neighbor* Dengan *Number of Nearest Neighbor (K) = 3* pada tahun ajaran 2020/2021 Kelas XII mendapati hasil siswa yang di nyatakan mampu terdapat 162 siswa Hasil dari pengklasifikasian dapat di *save* untuk di masukan kedalam *database* serta akan di lakukan pengujian dengan *data* model untuk dapat melihat nilai dari akurasi, *Precision* dan *Recall* yang dihasilkan.

ID	Nama	Jenis Kelamin	Umur	Nilai	Nilai	Nilai	Nilai	Tanggung Jawab	Disiplin	Atensi	Relasi	Keterampilan	Kemampuan
010001	AKEL WINDOYO	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010002	ABDULLAH SYAHRI	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010003	AMIRUDDIN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010004	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010005	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010006	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010007	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010008	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010009	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010010	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010011	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010012	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010013	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010014	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010015	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010016	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010017	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010018	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010019	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010020	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010021	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010022	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010023	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010024	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010025	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010026	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010027	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010028	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010029	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010030	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010031	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010032	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010033	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010034	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010035	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010036	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010037	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010038	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010039	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010040	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010041	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010042	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010043	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010044	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010045	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0
010046	ALYAN ALYAN	L	17	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0	80,0

Gambar 8 Hasil Klasifikasi Kelas K=3 XII

5. Klasifikasi *Number of Nearest Neighbor (K) = 5*

Hasil dari pengklasifikasian menggunakan algoritma *K-Nearest Neighbor* Dengan *Number of Nearest Neighbor (K) = 5* pada tahun ajaran 2020/2021 Kelas XII mendapati hasil siswa yang di nyatakan mampu terdapat 166 siswa. Hasil dari pengklasifikasian dapat di *save* untuk di masukan kedalam *database* serta akan di lakukan pengujian dengan *data* model untuk dapat melihat nilai dari akurasi, *Precision* dan *Recall* yang dihasilkan.

Detail Klasifikasi
 Hasil Klasifikasi Komparasi Siswa pada Tahun Ajaran 2020/2021 dengan Parameter K = 5 sebagai berikut:

Siswa yang Mampu : 166 Siswa
 Siswa yang Belum Mampu : 2 Siswa

Gambar 9 Hasil Klasifikasi K= 5 Kelas XII

6. Pengujian Algoritma K-Nearest Neighbor

Pengujian menggunakan Algoritma K-Nearest Neighbor Menggunakan 521 Data Model dan Dataset Uji menggunakan tahun ajaran 2020/2021 Kelas ahun ajaran Kelas XII dengan total data 169 dapat dilihat pada gambar 10 .

Gambar 10 Dataset Model dan Dataset Uji Algoritma K-Nearest Neighbor

Keseluruhan data uji sebelumnya sudah dilakukan proses klasifikasi menggunakan parameter *number of Nearest Neighbor* (K) = 3 yang dapat dilihat pada gambar 5, dan gambar 6, *number of Nearest Neighbor* (K) = 5 yang dapat di lihat pada gambar 8.

7. Pengujian Number of Nearest Neighbor (K) = 3 Kelas XII

Hasil dari pengujian menggunakan algoritma K-Nearest Neighbor menggunakan *number of Nearest Neighbor* (K) = 3 Pada Kelas XII dengan 521 data model dan 169 data uji mendapatkan hasil nilai aktual mampu yang di prediksikan mampu (TP) sebanyak 160 Siswa, nilai aktual mampu yang di prediksikan belum mampu (FN) sebanyak 2 siswa, Sementara nilai actual belum mampu yang di prediksikan mampu (FP) sebanyak 1 siswa dan nilai actual belum mampu yang di prediksikan belum mampu (TN) sebanyak 6 siswa yang dapat di lihat pada gambar 11.

HASIL EVALUASI			
NIS	NAMA	TARGET CLASS	HASIL EVALUASI
41180001	ADIL WIBOWO	MAMPU	MAMPU
41180002	ADISTI AYU PRADITA	MAMPU	MAMPU
41180003	AFRI SURYADI	MAMPU	MAMPU
41180004	AGUS PRASETYO	MAMPU	MAMPU
41180005	AGUS YULYANTO	MAMPU	MAMPU
41180006	AHMAD AFRANSYAH	BELUM MAMPU	BELUM MAMPU
41180007	AHMAD FAJAR	MAMPU	MAMPU
41180008	AHMAD REGI NUR	MAMPU	MAMPU
41180009	AISYAH APRILIA	MAMPU	MAMPU
41180010	ALIF ILHAM PRAYUDA	MAMPU	MAMPU
41180011	ALIF MUHAMAD ALFARIZI	MAMPU	MAMPU
41180012	ALFIA SABRINA AULIA	MAMPU	MAMPU
41180013	AMRILAH MUZAKI	MAMPU	MAMPU
41180014	ANANG SULISTIAWAN	MAMPU	MAMPU
41180015	ANDINI EKA SAPUTRI NINGSIH	MAMPU	MAMPU
41180016	ARLA AMELIA PRIYONO	MAMPU	MAMPU
41180017	AYU MIRANDA	MAMPU	MAMPU
41180018	AZAHRA APRILIA	MAMPU	MAMPU
41180019	AZHRA ALIFIA	MAMPU	MAMPU

CONFUSION MATRIX			
	PREDIKSI MAMPU	PREDIKSI BELUM MAMPU	CLASS RECALL
ACTUAL MAMPU	160	2	98.77 %
ACTUAL BELUM MAMPU	1	6	85.71 %
CLASS PRECISION	99.38 %	75 %	

***ACCURACY**
 Overall Accuracy is 98.22 %

Gambar 11 Hasil Pengujian K=3 Kelas XII

Dari hasil *confusion matrix* yang didapatkan pada gambar 11 maka peneliti dapat menghitung nilai dari akurasi, *Precision*, dan *Recall* sebagai berikut :

$$\begin{aligned} \text{Precision} &= TP / (TP + FP) \\ &= 160 / (160 + 1) \\ &= 160 / 161 = 0.9938 = 99.38 \% \\ \text{Recall} &= TP / (TP + FN) \\ &= 160 / (160 + 2) \\ &= 160 / 162 = 0.9877 = 98.77 \% \\ \text{Akurasi} &= TP + TN / (TP + TN + FP + FN) \\ &= 160 + 6 / (160 + 6 + 1 + 2) \\ &= 166 / 169 = 0.9822 = 98.22 \% \end{aligned}$$

8. Pengujian Number of Nearest Neighbor (K) = 5 Kelas XII

Hasil dari pengujian menggunakan algoritma K-Nearest Neighbor menggunakan *number of Nearest Neighbor* (K) = 5 Pada Kelas XII dengan 521 data model dan 169 data uji mendapatkan hasil nilai aktual mampu yang di prediksikan mampu (TP) sebanyak 165 Siswa, nilai aktual mampu yang di prediksikan belum mampu (FN) sebanyak 1 siswa, Sementara nilai actual belum mampu yang di prediksikan mampu (FP) sebanyak 1 siswa dan nilai actual belum mampu yang di prediksikan belum mampu (TN) sebanyak 42 siswa yang dapat di lihat pada gambar 12.

HASIL EVALUASI			
NIS	NAMA	TARGET CLASS	HASIL EVALUASI
41180001	ADIL WIBOWO	MAMPU	MAMPU
41180002	ADISTI AYU PRADITA	MAMPU	MAMPU
41180003	AFRI SURYADI	MAMPU	MAMPU
41180004	AGUS PRASETYO	MAMPU	MAMPU
41180005	AGUS YULYANTO	MAMPU	MAMPU
41180006	AHMAD AFRANSYAH	MAMPU	MAMPU
41180007	AHMAD FAJAR	MAMPU	BELUM MAMPU
41180008	AHMAD REGI NUR	MAMPU	MAMPU
41180009	AISYAH APRILIA	MAMPU	MAMPU
41180010	ALIF ILHAM PRAYUDA	MAMPU	MAMPU
41180011	ALIF MUHAMAD ALFARIZI	MAMPU	MAMPU
41180012	ALFIA SABRINA AULIA	MAMPU	MAMPU
41180013	AMRILAH MUZAKI	MAMPU	MAMPU
41180014	ANANG SULISTIAWAN	MAMPU	MAMPU
41180015	ANDINI EKA SAPUTRI NINGSIH	MAMPU	MAMPU
41180016	ARLA AMELIA PRIYONO	MAMPU	MAMPU
41180017	AYU MIRANDA	MAMPU	MAMPU
41180018	AZAHRA APRILIA	MAMPU	MAMPU
41180019	AZHRA ALIFIA	MAMPU	MAMPU

CONFUSION MATRIX			
	PREDIKSI MAMPU	PREDIKSI BELUM MAMPU	CLASS RECALL
ACTUAL MAMPU	165	1	99.4 %
ACTUAL BELUM MAMPU	1	2	66.67 %
CLASS PRECISION	99.4 %	66.67 %	

***ACCURACY**
 Overall Accuracy is 98.82 %

Gambar 12 Hasil Pengujian K=5 Kelas XII

Dari hasil *confusion matrix* yang didapatkan pada gambar 12 maka peneliti dapat menghitung nilai dari akurasi, *Precision*, dan *Recall* sebagai berikut :

$$\begin{aligned} \text{Precision} &= TP / (TP + FP) \\ &= 165 / (165 + 1) \\ &= 165 / 166 = 0.994 = 99.4 \% \end{aligned}$$

$$\text{Recall} = TP / (TP + FN)$$

$$= 165 / (165 + 1)$$

$$= 165 / 166 = 0.994 = 99.4 \%$$

$$\text{Akurasi} = TP + TN / (TP + TN + FP + FN)$$

$$= 165 + 2 / (165 + 2 + 1 + 1)$$

$$= 167 / 169 = 0.9982 = 99.82 \%$$

9. Klasifikasi Algoritma Naïve Bayes

Klasifikasi menggunakan Algoritma *Naïve bayes* berfungsi untuk memprediksi kemampuan siswa/I dalam bersaing didunia kerja menggunakan *dataset* yang belum memiliki keterangan mampu dan belum mampu. data model yang digunakan dalam pengujian ini merupakan *dataset* terdapat pada *database*, serta *dataset* uji yang digunakan dalam pengujian ini menggunakan 497 data yang berasal dari tahun ajaran 2020/2021 Kelas Kelas XII. hasil dari klasifikasi Algoritma *Naïve Bayes* dapat di lihat pada gambar berikut.

Gambar 13 Hasil Klasifikasi Algoritma Naïve Bayes Kelas XII

Hasil dari pengklasifikasian Kelas XII menggunakan algoritma *Naïve bayes* mendapati hasil siswa yang di nyatakan mampu terdapat 163 siswa dan yang di nyatakan belum mampu terdapat 6 Siswa dapat dilihat pada gambar 14. Hasil dari pengklasifikasian dapat di *save* untuk di masukan kedalam database serta akan di lakukan pengujian dengan data model untuk dapat melihat nilai dari akurasi, *Precision* dan *Recall* yang di hasilkan.

10. Pengujian Algoritma Naïve Bayes

Pengujian menggunakan Algoritma *Naïve Bayes* Menggunakan 521 Data Model dan Data Uji menggunakan tahun ajaran 2020/2021 Kelas XII dengan total 169 data.

NIS	NAMA	TARGET CLASS	HASIL NAIVE BAYES
41200001	ABELIA SYAHLA TALITHA	MAMPU	MAMPU
41200002	ACHMAD ZAIDAN	MAMPU	MAMPU
41200003	ADE IHAH SOLIHAH	BELUM MAMPU	BELUM MAMPU
41200004	ADIMAS RESTU RAMAD...	MAMPU	MAMPU
41200005	ADINDA VANIA KINANTI	MAMPU	MAMPU
41200006	AISYAH ALLISMAYANTI	MAMPU	MAMPU
41200007	AJENG PRAMESWARI	MAMPU	MAMPU
41200008	ALLESANDRO BAGINTA ...	MAMPU	MAMPU
41200009	ALSYIRA SUSANTI METI...	MAMPU	MAMPU
41200010	AMBARSRIRANI	MAMPU	MAMPU
41200011	ANGGUN NUR ARIFIANI	MAMPU	BELUM MAMPU
41200012	ANISA ZAHWA AZZAHROO	MAMPU	MAMPU
41200013	ANJANI KIRANA LARASATI	MAMPU	MAMPU
41200014	ANNADIRA NABILLA SARI	MAMPU	MAMPU
41200015	ASRI SETIANINGSIH	MAMPU	MAMPU
41200016	ATIKA RIZKI APRILIYANTI	MAMPU	MAMPU
41200017	AUDIA CAHYANI	MAMPU	MAMPU
41200018	AULIA CAVITA AZHAR	MAMPU	MAMPU
41200019	AULIA RETNO WULANDARI	MAMPU	MAMPU

CONFUSION MATRIX			
	PREDIKSI MAMPU	PREDIKSI BELUM MAMPU	CLASS RECALL
ACTUAL MAMPU	162	4	97.59 %
ACTUAL BELUM MAMPU	0	5	100 %
CLASS PRECISION	100 %	55.56 %	

***ACCURACY**
Overall Accuracy is 97.66 %

Gambar 14 Dataset Model dan Dataset Uji Algoritma Naive Bayes

11. Pengujian Algoritma Naïve Bayes Kelas XII

Hasil dari pengujian menggunakan algoritma Naïve Bayes pada Kelas XII dengan data model 521 data dan data uji menggunakan 169 data mendapati hasil nilai aktual mampu yang di prediksikan mampu (TP) sebanyak 156 Siswa, nilai aktual mampu yang di prediksikan belum mampu (FN) sebanyak 7 siswa, Sementara nilai aktual belum mampu yang di prediksikan mampu (FP) sebanyak 2 siswa dan nilai aktual belum mampu yang di prediksikan belum mampu (TN) sebanyak 4 siswa yang dapat di lihat pada gambar 15.

HASIL EVALUASI			
NIS	NAMA	TARGET CLASS	HASIL NAIVE BAYES
#1180001	ADIL WIBOWO	MAMPU	MAMPU
#1180002	ADISTI AYU PRADITA	MAMPU	BELUM MAMPU
#1180003	AFFRI SURYADI	BELUM MAMPU	BELUM MAMPU
#1180004	AGFUS PRASETYO	MAMPU	MAMPU
#1180005	AGUS YULIYANTO	MAMPU	MAMPU
#1180006	AHMAD AFRIANSYAH	MAMPU	MAMPU
#1180007	AHMAD FAJAR	MAMPU	BELUM MAMPU
#1180008	AHMAD REGI NUR	MAMPU	MAMPU
#1180009	AISYAH APRILIA	MAMPU	MAMPU
#1180010	ALIF ILHAM PRAYUDA	MAMPU	MAMPU
#1180011	ALIF MUHAMAD ALFARIZI	MAMPU	MAMPU
#1180012	ALIFIA SABRINA AULIA	MAMPU	MAMPU
#1180013	AMRILAH MUJAKI	MAMPU	BELUM MAMPU
#1180014	ANANG SULISTIAWAN	MAMPU	MAMPU
#1180015	ANDINI EKA SAPUTRI NI...	MAMPU	MAMPU
#1180016	ARLA AMELIA PRIYONO	MAMPU	MAMPU
#1180017	AYU MIRANDA	MAMPU	MAMPU
#1180018	AZAHRA APRILIA	MAMPU	MAMPU
#1180019	AZHRA ALIFIA	BELUM MAMPU	BELUM MAMPU

CONFUSION MATRIX			
	PREDIKSI MAMPU	PREDIKSI BELUM MAMPU	CLASS RECALL
ACTUAL MAMPU	156	7	95.71 %
ACTUAL BELUM MAMPU	2	4	66.67 %
CLASS PRECISION	98.73 %	36.36 %	

***ACCURACY**
 Overall Accuracy is 94.67 %

Gambar 15 Hasil Hasil Pengujian Algoritma Naïve Bayes Kelas XII

Dari hasil *confusion matrix* yang didapatkan pada gambar 15 maka peneliti dapat menghitung nilai dari akurasi, *Precision*, dan *Recall* sebagai berikut :

$$\begin{aligned} \text{Precision} &= TP / (TP + FP) \\ &= 156 / (152 + 2) \\ &= 152 / 154 = 0.9873 = 98.73 \% \end{aligned}$$

$$\begin{aligned} \text{Recall} &= TP / (TP + FN) \\ &= 156 / (156 + 7) \\ &= 156 / 163 = 0.9571 = 95.71 \% \end{aligned}$$

$$\begin{aligned} \text{Akurasi} &= TP + TN / (TP + TN + FP + FN) \\ &= 156 + 4 / (156 + 4 + 2 + 7) \\ &= 160 / 169 = 0.9467 = 94.67 \% \end{aligned}$$

4. KESIMPULAN

Berdasarkan hasil pengujian yang telah dijabarkan diatas Tujuan dari penelitian ini adalah untuk memprediksi kemampuan siswa dan siswi SMK Yadika 5 Pondok aren falam bersaing di dunia kerja menggunakan perbandingan algoritma naïve bayes dan k-nearest. Berdasarkan pengujian yang telah dilakukan dalam penelitian ini mendapati hasil bahwa *Algoritma Naïve Bayes* memiliki Akurasi 97.66 % , Nilai *Precision* 100% , dan Nilai *Recall* 97.59% sedangkan *Algoritma K-Nearest Neighbor* mendapati hasil Nilai Akurasi 98.22 % , Nilai *Precision* 99.38% , Nilai *Recall* 98.77 %.

DAFTAR PUSTAKA

- [1] Buani, D.C. P. (2016). Optimasi Algoritma Naive Bayes dengan menggunakan Algoritma Genetika untuk prediksi Kesuburan (*Fertility*). *Jurnal Evolusi*, 4(1), pp. 54-63.
- [2] Ginantra, N., Wijaya, F., dkk. 2021. Data Mining dan Penerapan Algoritma. Medan : Kita Menulis
- [3] Khoirunnisa, Susanti, L., Rokmah, I. T., and Stianingsih, L. (2020). Prediksi Siswa SMK Al-Hidayah Yang Masuk Perguruan Tinggi Dengan Metode Klasifikasi. *Jurnal Informatika*. 8(1), pp. 26-33.
- [4] Kurniawan, E. D., and Mufti. (2018). Implementasi Algoritma K-Nearest Neighbor Dengan Metode Klasifikasi dan Pengukuran Jarak Manhattan Distance Untuk Prediksi Kelulusan UN Berdasarkan Hasil Nilai Tryout Berbasis Java Desktop Pada SMA Harapan Jaya 2. *Jurnal Teknik Informatika*. 1(1), pp. 76-81.
- [5] Kurniawan. (2018). Perbandingan Algoritma Naive Bayes dan C.45 Dalam Klasifikasi Data Mining. *Jurnal Teknologi Informasi dan Ilmu Komputer*. 5(4), pp. 455-464.
- [6] Kusuma, L. W. (2019). Prediksi Kemampuan Lulusan SMK Untuk Dapat Bersaing di Dunia Kerja Dengan Menggunakan Naive Bayes: Studi Kasus SMK Buddhi Tangerang. *Jurnal Algor*. 1(1), pp. 56-63.
- [7] Maburur, A. G., and Lubis, R. (2012). Penerapan Data Mining Untuk Memprediksi Kriteria Nasabah Kredit. *Jurnal Komputer dan Informatika*. 1(1), pp. 53-57.
- [8] Malik, A. M., and Sibarani, A. J. P. (2018). Aplikasi Prediksi Kelulusan Ujian Nasional Menggunakan Algoritma K-Nearest Neighbor Dengan Pengukuran Jarak Manhattan Distance. *Jurnal Teknik Informatika*. 1(2), pp. 829-835.
- [9] Mujiasih, S. (2011). Pemanfaatan Data Mining Untuk Prakiraan Cuaca. *Jurnal Meteorologi dan Geofisika*. 12(2), pp. 189-195.
- [10] Munanda, W., A.(2020). Klasifikasi Peminatan Siswa Dengan Metode *K-Nearest Neighbor* (Studi Kasus SMP Negeri 1 Pabuaran) Diss. University of Technology Yogyakarta.
- [11] Mustafa, M. S., Ramadhan. M. R., and Thenata, A. P. (2017). Implementasi Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier. *Citec Journal*. 4(2), pp. 151-162.
- [12] Nofriansyah, D., Erwansyah, K., and Ramadhan, M. (2016). Penerapan Data Mining Dengan Algoritma Naive Bayes Classifier Untuk Mengetahui Minat Beli Pelanggan Terhadap Kartu Internet XL. *Jurnal Saintikom*. 15(2), pp. 81-92.
- [13] Ridwan, M., Suyono, H., and Sarosa, M. (2013). Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier. *Jurnal EECCIS*. 7(1), pp. 59-64.
- [14] Saleh, A. (2015). Implementasi Metode Klasifikasi Naive Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga. *Citec Journal*. 2(3), pp. 207-217.