

IMPLEMENTATION OF WEB-BASED NAIVE BAYES ALGORITHM FOR DETERMINING DEPARTMENTS AT SMK 10 MUHAMMADIYAH KISARAN

Nurlaili Sabila^{*1}, Herman Saputra², Muthia Dewi^{*3}

^{1,3}Sistem Informasi, Sekolah Tinggi Manajemen Informatika dan Komputer Royal, Indonesia

²Manajemen Informatika, Sekolah Tinggi Manajemen Informatika dan Komputer Royal, Indonesia

Email: nurlailisabila46@gmail.com

(Naskah masuk: 11 September 2022, Revisi : 14 September 2022, diterbitkan: 26 Desember 2022)

Abstract

Determination of majors is very important for the convenience of prospective students in the process and continuation of education so that they do not experience difficulties in the teaching and learning process in the future. SMK 10 Muhammadiyah Kisaran is one of the private vocational schools in Asahan that provides 3 majors including Audio Video Engineering (TAV), Computer and Network Engineering (TKJ), and Motorcycle Engineering and Business (TBSM). SMK 10 Muhammadiyah Kisaran does not yet have a special system for selecting majors so that prospective students are welcome to choose majors according to their own wishes, not a few students find it difficult because the students themselves do not understand their abilities. so that it's not uncommon for students to choose majors in a random way or follow their friends' choices. Therefore we need a system that can help prospective students in selecting majors that match their interests and talents and reduce mistakes in choosing majors. The technique used for the classification data mining model in this study is the Naïve Bayes Algorithm. The dataset that will be used as training data and test data is data for new students for the 2021/2022 school year, to be precise, for class X SMK 10 Muhammadiyah Kisaran obtained from the results of documentation and questionnaires. The criteria used were school origin, gender, interests, major, influence of friends, parental suggestions, math scores, English grades, and science grades. The results of the classification modeling with the Naïve Bayes Algorithm produce an accuracy value of 89%.

Keywords: Data Mining, Determining, Implementation, Naïve Bayes, Vocational High School.

IMPLEMENTASI ALGORITMA NAIVE BAYES UNTUK PENENTUAN JURUSAN BERBASIS WEB DI SMK 10 MUHAMMADIYAH

Abstrak

Penentuan jurusan sangat penting untuk kenyamanan calon siswa dalam proses dan kelanjutan pendidikan agar tidak mengalami kesulitan dalam proses belajar mengajar dikemudian hari. SMK 10 Muhammadiyah Kisaran merupakan salah satu sekolah kejuruan swasta di Asahan yang menyediakan 3 jurusan diantaranya Teknik Audio Video (TAV), Teknik Komputer dan Jaringan (TKJ), dan Teknik dan Bisnis Sepeda Motor (TBSM). SMK 10 Muhammadiyah Kisaran belum mempunyai sistem khusus untuk pemilihan jurusan sehingga calon siswa dipersilahkan untuk memilih jurusan sesuai keinginan sendiri, tak sedikit siswa yang merasa kesulitan akan hal itu dikarenakan siswa itu sendiri juga belum memahami tentang kemampuan yang mereka miliki sehingga tak jarang para siswa akan menentukan jurusan dengan cara asal asalan atau mengikuti pilihan temannya. Maka dari itu dibutuhkan sebuah sistem yang dapat membantu calon siswa dalam pemilihan jurusan yg sesuai dengan minat dan bakat mereka serta mengurangi kekeliruan dalam pemilihan jurusan. Teknik yang digunakan untuk model data mining klasifikasi pada penelitian ini adalah Algoritma Naïve Bayes. Dataset yang akan digunakan untuk menjadi data latih dan data uji adalah data siswa baru tahun ajaran 2021/2022 tepatnya kelas X SMK 10 Muhammadiyah Kisaran yang diperoleh dari hasil dokumentasi dan kuisioner. Kriteria yang digunakan adalah asal sekolah, jenis kelamin, minat, Jurusan, Pengaruh teman, Saran orang tua, nilai matematika, nilai bahasa Inggris, dan nilai IPA. Hasil pemodelan klasifikasi dengan Algoritma Naïve Bayes menghasilkan nilai akurasi sebesar 89%.

Kata kunci: Data Mining, Implementasi, Naïve Bayes, Penjurusan, Sekolah Menengah Kejuruan.

1. PENDAHULUAN

Penjurusan siswa merupakan suatu proses penempatan siswa ke dalam jurusan yang sesuai dengan minat dan bakat, sehingga siswa dapat menyerap mata pelajaran yang sesuai dengan kemampuan yang dimilikinya.[1] Dengan mengikuti kurikulum yang berlaku di Indonesia, siswa yang akan mendaftar ke tingkat SMK/SMA akan melakukan pemilihan jurusan sesuai dengan minat dan bakat nya. Pemilihan jurusan ini juga harus dilakukan dengan sebaik mungkin agar tidak terjadi kekeliruan dalam pemilihan jurusan dan proses pembelajaran pun akan lebih terarah sesuai dengan minat dan nilai yang dia punya.[2] Banyak sekali siswa yang mengalami kesulitan dalam memilih jurusan yang sesuai dengan minat dan bakat mereka, dikarenakan siswa itu sendiri juga belum mengetahui secara pasti tentang kemampuan yang mereka miliki sehingga tak jarang para siswa akan menentukan jurusan dengan cara asal alasan atau mengikuti pilihan temannya.[3] Sebuah sekolah tentunya memiliki arsip data yang dapat digunakan untuk pengetahuan dalam proses pengambilan keputusan yang nantinya diharapkan bisa meminimalisir kekeliruan dalam pemilihan jurusan. Dan tentunya juga dibutuhkan sebuah teknik untuk mengolah arsip data tersebut. Dengan teknik *Data Mining*, arsip data tadi akan dikelola sedemikian rupa untuk menentukan sebuah pola agar dapat menghasilkan sebuah informasi atau hasil prediksi yang akan membantu para calon siswa dalam memilih jurusan yang sesuai dengan minat dan bakat atau kemampuan yang dimilikinya.[4]

Data mining adalah proses pengumpulan atau penggalian informasi dari sebuah data yang berukuran besar melalui proses perhitungan statistika, matematika, maupun *AI (Artificial Intelligence)*, dan biasanya digunakan untuk menemukan pola pola tertentu dari data yang tersebut.[5] Saat ini *data mining* dikatakan sebagai salah satu ilmu yang relatif baru yang memiliki kemungkinan besar dalam proses pengembangan dan diprediksi akan menjadi salah satu yang paling revolusioner pada dekade ini.[6] *Data mining* memiliki banyak metode yang dapat digunakan, salah satunya yaitu klasifikasi *Naive Bayes*. *Naive Bayes* yaitu sebuah pengklasifikasian sederhana yang memanfaatkan perhitungan probabilitas dan statistik.[7] Metode ini ditemukan oleh ilmuan Inggris *Thomas Bayes* untuk memprediksi probabilitas atau kemungkinan yang akan terjadi dimasa depan berdasarkan pengalaman pada masa sebelumnya atau disebut dengan *Teorema Bayes* dengan berasumsikan bahwa ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lainnya.[8] Klasifikasi *Naive Bayes* dianggap tepat untuk digunakan oleh database yang relatif besar karena mudah dan sederhana.

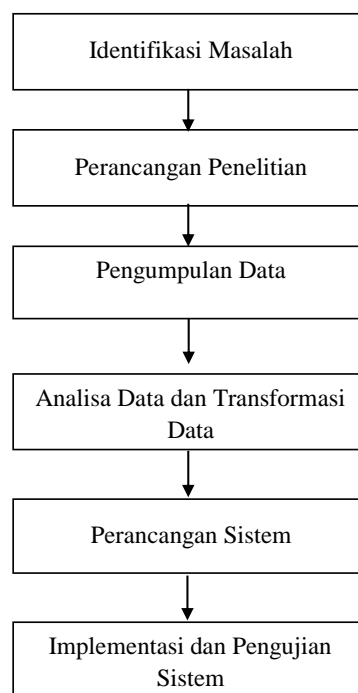
SMK 10 Muhammadiyah Kisaran merupakan salah satu sekolah kejuruan swasta di Asahan yang menyediakan 3 jurusan diantaranya Teknik Audio Vidio (TAV), Teknik Komputer dan Jaringan (TKJ),

dan Teknik dan Bisnis Sepeda Motor (TBSM). [9] Seperti yang kita ketahui, pemilihan jurusan sangatlah penting untuk kenyamanan calon siswa dalam proses dan kelanjutan pendidikan mereka agar tidak mengalami kesulitan dalam proses belajar mengajar dikemudian hari.[10] SMK 10 Muhammadiyah Kisaran belum mempunyai sistem khusus untuk pemilihan jurusan sehingga calon siswa yang akan mendaftar dipersilahkan untuk memilih jurusan sesuai dengan keinginannya sendiri, untuk sebagian calon siswa mungkin sudah mempunyai pilihan karena telah mengetahui perihal minat dan bakat atau kemampuan yang mereka miliki, tetapi tidak sedikit juga calon siswa yang masih bingung untuk menentukan jurusan yang sesuai sehingga setelah menjalani proses belajar mengajar, tak sedikit siswa yang mengeluh salah memilih jurusan karena merasa bahwa pelajarannya tidak sesuai dengan minat mereka dan meminta untuk di pindahkan ke jurusan lain yang dirasa lebih sesuai.

Penelitian ini bertujuan mengimplementasikan sistem pengklasifikasian data berbasis web untuk penentuan jurusan yang sesuai bagi calon siswa dengan menggunakan metode klasifikasi *Naive Bayes* yang diharapkan sistem tersebut dapat meminimalisir kekeliruan calon siswa dalam langkah penentuan jurusan.

2. METODE PENELITIAN

Pada bab ini akan diuraikan metodologi penelitian dan kerangka kerja penelitian. Kerangka kerja ini merupakan langkah-langkah yang akan dilakukan dalam penyelesaian masalah yang akan dibahas. Adapun kerangka kerja penelitian ini dapat dilihat pada gambar 1.



Gambar 1. Kerangka Kerja Penelitian

Berdasarkan kerangka kerja penelitian diatas, uraian dan penjelasan dari masing masing tahapan yaitu sebagai berikut:

1. Identifikasi Masalah

Identifikasi Masalah adalah upaya untuk menjelaskan masalah dan membuat penjelasan yang dapat diukur. Identifikasi ini dilakukan sebagai langkah awal penelitian. Jadi, secara ringkas, identifikasi masalah adalah mendefinisikan masalah penelitian.

2. Perancangan Penelitian

Perancangan Penelitian secara umum dapat diartikan sebagai pembuatan sketsa atau desain penelitian oleh peneliti, atau bisa juga diartikan sebagai kerangka atau proposal penelitian. Jadi sebelum menjalankan sebuah penelitian, peneliti harus membuat proposal penelitian terlebih dahulu.

3. Pengumpulan Data

Pengumpulan data adalah proses mencari, mencatat, dan mengumpulkan semua informasi yang dibutuhkan secara objektif yang mana data tersebut nantinya akan digunakan untuk membantu peneliti dalam mencapai tujuan pada penelitian yang dilakukan.

4. Analisa Data dan Transformasi data

Analisa Data dan Transformasi data adalah proses pemeriksaan, pembersihan, pemodelan, dan perubahan data dengan tujuan agar data menjadi sesuai dengan yang dibutuhkan dan dapat diolah dengan Algoritma *Naive Bayes* untuk menemukan informasi yang berguna, dan membuat suatu kesimpulan.

5. Perancangan Sistem

Perancangan sistem merupakan proses merancang atau mendesain suatu sistem untuk menentukan langkah-langkah operasi dalam proses pengolahan data yang bertujuan untuk memenuhi kebutuhan para pemakai sistem.

6. Implementasi dan Pengujian Sistem

Implementasi dan Pengujian Sistem adalah langkah atau prosedur yang dilakukan untuk menyelesaikan desain sistem yang telah dirancang dan melakukan pengujian agar dapat mengetahui apakah sistem yang dibuat sudah sesuai dengan kebutuhan. Implementasi dan pengujian sistem juga dilakukan untuk mengevaluasi kekurangan yang ada pada sistem yang dibuat.

Data mining merupakan suatu proses penggalian sekumpulan data yang besar untuk mendapatkan sebuah informasi yang tidak diketahui sebelumnya.[11] *Data mining* menggunakan teknik statistik, matematika, kecerdasan buatan dan *machine learning* pada prosesnya dengan cara mengolah data tersebut untuk mendapatkan sebuah informasi yang lebih bermanfaat.[12] Nama lain *data mining* adalah *Knowledge Discovery in Database (KDD)*. *Data mining* adalah salah satu bagian langkah yang penting dalam proses KDD terutama berkaitan dengan penemuan pola-pola dari data yang ditelaah, yang

kemudian diolah menjadi sebuah informasi baru serta lebih bermanfaat.[13]

Bagian atau Teknik *Data Mining* merupakan cara yang akan diterapkannya pada proses penggalian data namun tetap harus disesuaikan dengan berbagai tipe fungsi ataupun tujuan penggunaannya. Adapun beberapa bagian atau teknik *data mining* menurut Herry Derajad Wijaya dan Saruni Dwiasnati[14] yaitu sebagai berikut :

a. *Association*

Teknik *data mining Market basket analysis* (analisa keranjang pasar) atau *association* adalah teknik yang berhubungan dengan pemasaran, mengenali kelakuan melalui kebiasaan kebiasaan yang terjadi sebelumnya.

b. *Classification*

Klasifikasi *data mining* adalah sebuah proses menemukan definisi kesamaan karakteristik dalam suatu kelompok atau kelas (*class*). Klasifikasi *data mining* menjadi salah satu teknik yang paling umum untuk digunakan, bertujuan untuk memperkirakan kelas dari suatu objek yang labelnya belum diketahui.

c. *Clustering*

Clustering data mining adalah nama lain untuk teknik *segmentation*, bertujuan untuk mengelompokkan suatu *class* ke dalam beberapa segmen berdasarkan atribut yang memiliki kemiripan.

Metode klasifikasi *Naive Bayes* adalah metode klasik yang menggunakan konsep statistik probabilitas dalam melakukan klasifikasi data.[11] *Naive Bayes* merupakan metode yang cocok untuk klasifikasi biner dan *multiclass*.

Model *machine learning* yang diterapkan pada program tersebut menggunakan *Teorema Bayes* yang dirumuskan sebagai berikut:[13]

$$P(H | X) = \frac{P(X|H) \times P(H)}{P(X)} \quad (1)$$

Dimana X merupakan ata dengan *class* yang belum diketahui, H adalah Hipotesis data merupakan suatu *class* spesifik, P(H|X) ; Probabilitas Hipotesis H berdasarkan kondisi X (*posterior probabilitas*)

P(H) ; Probabilitas hipotesis H (*prior probabilitas*)

P(X|H) ; Probabilitas X berdasarkan kondisi pada hipotesis H, P(X) ; Probabilitas X

Metode yang digunakan pada penelitian ini adalah metode penelitian kuantitatif. Metode kuantitatif yaitu suatu proses mengumpulkan data berupa angka sebagai alat menganalisis keterangan mengenai apa yang ingin diketahui yang kemudian angka-angka yang didapat akan di analisis dan diolah agar data dapat di proses sesuai dengan prosedur statistik yang berlaku. Tujuan penelitian kuantitatif memiliki beberapa poin, seperti untuk pengembangan model matematis karena penelitian ini tidak sekadar menggunakan teori yang diambil lewat kajian literatur dan teori, namun juga pentingnya membangun hipotesis yang memiliki keterkaitan

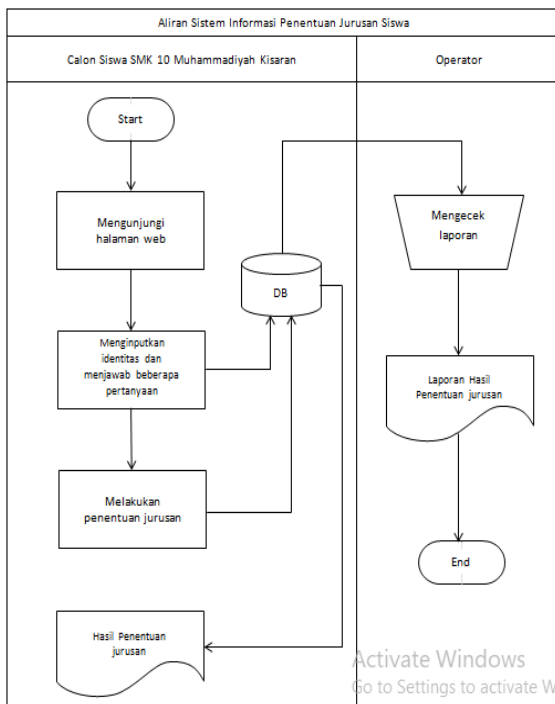
dengan fenomena yang akan diteliti menggunakan metode penelitian ini.

Beberapa Teknik pengumpulan data yang digunakan pada penelitian ini meliputi: (1) Wawancara (2) Kuisisioner (3) Dokumentasi (4) Studi Pustaka.

3. HASIL DAN PEMBAHASAN

3.1. Analisis Sistem

Analisis Sistem adalah suatu teknik atau metode pemecahan masalah dengan cara menguraikan sistem ke dalam komponen-komponen pembentuknya untuk mengetahui bagaimana komponen-komponen tersebut bekerja dan saling berinteraksi satu sama lain guna mencapai tujuan sistem, biasanya analisis sistem dilakukan dalam membuat Design Sistem, dimana komponen-komponen pembentuk sistem digabungkan sehingga membentuk satu kesatuan yang utuh dalam bentuk yang lebih baik lagi. Pada penelitian ini dilakukan klasifikasi dalam penentuan jurusan berbasis web menggunakan metode *Naïve Bayes*.



Gambar 2. Analisis Sistem Yang Diusulkan (ASI Baru).

Sistem penentuan jurusan di SMK 10 Muhammadiyah Kisaran masih menggunakan penjurusan siswa berdasarkan pilihan siswa itu sendiri, sehingga tidak sedikit calon siswa yang akan mendaftar tidak mempunyai gambaran perihal jurusan apa yang harus mereka pilih. Alur dan sistem yang sedang berjalan dalam penentuan jurusan di SMK 10 Muhammadiyah Kisaran diuraikan dengan beberapa pemodelan *Unified Modelling Language* (UML). Dengan adanya Implementasi Algoritma *Naïve Bayes* untuk Penentuan Jurusan Berbasis Web di

SMK 10 Muhammadiyah Kisaran dapat memudahkan calon siswa baru untuk menentukan jurusan yang akan diambil. Adapun analisis proses dari sistem penentuan jurusan berbasis web dapat dilihat pada gambar 2.

3.2. Analisis Masalah

Analisis masalah dalam penelitian ini dilakukan dengan *PIECES Analysis*, yaitu peneliti akan menganalisis data yang sudah dikumpulkan data yang sudah terkumpul guna menentukan apa saja yang diperlukan untuk mengembangkan sistem seperti *performance, information, economy, control, efficiency*, dan *services* yang diperlukan pada sistem yang akan dibangun. Adapun tahapan-tahapan *PIECES Analysis* adalah sebagai berikut:

1. Performance Analysis

Karena penentuan jurusan masih dilakukan secara manual maka dibutuhkan suatu sistem yang dapat membantu penentuan jurusan berbasis web untuk masa yang akan datang agar menjadi lebih efektif dan akurat.

2. Information Analysis

Information Analysis mengenai perancangan sistem penentuan jurusan berbasis web pada SMK 10 Muhammadiyah Kisaran menggunakan Algoritma *Naïve Bayes* harus lebih mudah dipahami dan dimengerti.

3. Economy Analysis

Economy Analysis sistem yang dirancang dapat sangat membantu calon siswa mendapatkan gambaran perihal jurusan yang sesuai dengan adanya sistem penentuan jurusan berbasis web.

4. Control Analysis

Control Analysis dilakukan untuk mendeteksi kesalahan atau kecurangan yang mungkin terjadi dengan penyalahgunaan sistem dan juga dapat menjamin keamanan data mengenai penentuan jurusan berbasis web. Dengan adanya *control analysis*, maka gangguan yang terjadi dapat diperbaiki dengan cepat.

5. Efficiency Analysis

Efficiency Analysis dilakukan untuk mengetahui tingkat efisiensi sistem penentuan jurusan berbasis web sehingga dapat ditingkatkan agar kualitasnya menjadi semakin baik.

6. Service Analysis

Dengan adanya *service analysis* calon siswa dapat mendapatkan informasi perihal jurusan yang sesuai dimana saja selagi mereka terhubung dengan jaringan internet.

3.3. Analisis Data

Data yang digunakan merupakan data siswa yang didapat melalui proses pengisian kuisisioner yang sebelumnya peneliti lakukan ditempat penelitian. Sebanyak 321 data diperoleh dari kegiatan kuisisioner tersebut. Sampel data awal yang telah diperoleh dapat dilihat pada tabel 1.

Tabel 1. Data Awal

No	Nama	Jenis Kelamin	Asal	Minat	PT	SOT	MM	BING	IPA	KPTS
1	Abeliazra Agusta	L	SMP	SAV	TB	CM	70	81	79	TAV
2	Afriandi	L	SMP	SAV	B	TM	79	80	81	TAV
3	Agus Firmansyah Nasution	L	SMP	KD	B	CM	81	81	88	TAV
...
320	Al-Anshar Fauzan Rangkuti	L	MTs	O	B	CM	82	91	82	TBSM
321	Arifin	L	SMP	O	B	CM	88	76	79	TBSM

Namun, peneliti memilih beberapa data untuk dijadikan data training yang mana data training tersebut nantinya akan diolah menggunakan

perhitungan matematika sehingga menghasilkan output yang diinginkan. Data *training* pada masalah di penelitian ini dapat dilihat pada tabel 2.

Tabel 2. Data Training

No	Asal	Minat	PT	SOT	MM	BING	IPA	JRSN
1	SMP	SAV	TB	CM	70	81	79	TAV
2	SMP	SAV	B	TM	79	80	81	TAV
3	SMP	KD	B	CM	81	81	88	TKJ
4	SMP	JR	TB	M	83	78	88	TKJ
5	MTs	JR	B	M	92	82	87	TKJ
6	SMP	JR	TB	M	82	97	82	TKJ
7	SMP	OTO	CB	CM	95	91	89	TBSM
8	MTs	OTO	B	M	77	95	83	TBSM

Setelah ditentukannya data *training* dan dilakukannya berbagai proses perhitungan, maka peneliti menghasilkan data output atau yang disebut dengan data *testing* pada sistem penentuan jurusan

siswa berbasis web. Data *testing* pada masalah di penelitian ini dapat dilihat pada tabel 3.

Tabel 3. Data Testing

No	Asal	Minat	PT	SOT	MM	BING	IPA	KPTS	JRSN
1	SMP	SAV	TB	CM	70	81	79	TAV	TAV
2	SMP	SAV	B	TM	79	80	81	TAV	TAV
3	SMP	KD	B	CM	81	81	88	TAV	TKJ
4	SMP	JR	TB	M	83	78	88	TKJ	TKJ
5	MTs	JR	B	M	92	82	87	TKJ	TKJ
6	SMP	JR	TB	M	82	97	82	TKJ	TKJ
7	SMP	OTO	CB	CM	95	91	89	TBSM	TBSM
8	MTs	OTO	B	M	77	95	83	TBSM	TBSM

3.4. Transformasi Data Nilai

Transformasi data yaitu mengubah data ke dalam bentuk yang lebih sesuai untuk proses *data mining* yang dibutuhkan. Data akan diubah ke dalam format yang dapat memudahkan proses untuk menentukan jurusan siswa berbasis web. Data transformasi pada penelitian ini dapat dilihat pada tabel 4.

Tabel 4. Transformasi Data

Nilai	Transformasi
80 - 100	A

70 - 79	B
60 - 69	C
45 - 59	D
0 - 44	E

3.5. Penerapan Metode Klasifikasi Algoritma *Naïve Bayes*

Terdapat beberapa tahapan dalam proses perhitungan algoritma *Naïve Bayes* dan dari perhitungan probabilitas yang sudah dilakukan maka dapat di buatkan seperti pada tabel 5.

Tabel 5. Probabilitas *Data Sheet*

Jenis Kelamin	Jumlah Kejadian "dipilih"			Probabilitas			
	TAV	TKJ	TBSM	TAV	TKJ	TBSM	
Perempuan	2	81	1	0,0952	0,5912	0,0061	
Laki-Laki	19	56	162	0,9048	0,4088	0,9939	
Jumlah	21	137	163	1,00	1,00	1,00	
Asal Sekolah	Jumlah Kejadian "dipilih"			Probabilitas			
	TAV	TKJ	TBSM	TAV	TKJ	TBSM	
	SMP	15	110	117	0,7143	0,8029	0,7178
	MTS	6	27	46	0,2857	0,1971	0,2822
Jumlah	21	137	163	1,00	1,00	1,00	
Pengaruh Teman	Jumlah Kejadian "dipilih"			Keputusan			
	TAV	TKJ	TBSM	TAV	TKJ	TBSM	
Berpengaruh	9	75	80	0,4286	0,5474	0,4908	

Cukup Berpengaruh	4	28	47	0,1905	0,2044	0,2883
Tidak Berpengaruh	8	34	36	0,3810	0,2482	0,2209
Jumlah	21	137	163	1,00	1,00	1,00
Jumlah Kejadian "dipilih"						
Saran Orang Tua	TAV	TKJ	TBSM	TAV	TKJ	TBSM
Menyarankan	6	45	72	0,2857	0,3285	0,4417
Cukup Menyarankan	8	51	57	0,3810	0,3723	0,3497
Tidak Menyarankan	7	41	34	0,3333	0,2993	0,2086
Jumlah	21	137	163	1,00	1,00	1,00
Jumlah Kejadian "dipilih"						
Minat	TAV	TKJ	TBSM	TAV	TKJ	TBSM
Design Grafis	1	17	2	0,0476	0,1241	0,0123
Jaringan	2	72	4	0,0952	0,5255	0,0245
Komunikasi Digital	6	15	2	0,2857	0,1095	0,0123
Otomotif	1	2	131	0,0476	0,0146	0,8037
Programming	1	28	5	0,0476	0,2044	0,0307
Sistem Audio Vidio	7	2	3	0,3333	0,0146	0,0184
Listrik dan Elektronika	3	1	16	0,1429	0,0073	0,0982
Jumlah	21	137	163	1,00	1,00	1,00
Jumlah Kejadian "dipilih"						
Nilai	TAV	TKJ	TBSM	TAV	TKJ	TBSM
A	17,9	117,9	130,9	0,8483	0,8600	0,8026
B	2,9	18,9	31,9	0,1374	0,1379	0,1956
C	0,1	0,1	0,1	0,0047	0,0007	0,0006
D	0,1	0,1	0,1	0,0047	0,0007	0,0006
E	0,1	0,1	0,1	0,0047	0,0007	0,0006
Jumlah	21	137	163	1,00	1,00	1,00
Jumlah Kejadian "dipilih"						
Nilai	TAV	TKJ	TBSM	TAV	TKJ	TBSM
A	17,9	108,9	130,9	0,8483	0,7943	0,8026
B	2,9	27,9	31,9	0,1374	0,2035	0,1956
C	0,1	0,1	0,1	0,0047	0,0007	0,0006
D	0,1	0,1	0,1	0,0047	0,0007	0,0006
E	0,1	0,1	0,1	0,0047	0,0007	0,0006
Jumlah	21	137	163	1,00	1,00	1,00
Jumlah Kejadian "dipilih"						
Nilai	TAV	TKJ	TBSM	TAV	TKJ	TBSM
A	15,9	115,9	136,9	0,7536	0,8454	0,8394
B	4,9	20,9	25,9	0,2322	0,1524	0,1588
C	0,1	0,1	0,1	0,0047	0,0007	0,0006
D	0,1	0,1	0,1	0,0047	0,0007	0,0006
E	0,1	0,1	0,1	0,0047	0,0007	0,0006
Jumlah	21	137	163	1,00	1,00	1,00

Berdasarkan perhitungan prediksi data *sheet* dan data *testing* menggunakan Algoritma *Naïve Bayes* tingkat akurasi yang didapat menggunakan Model *Confusion Matrix* sebesar 89%. Hasil dari *Confusion Matrix* dapat dilihat pada tabel 6.

Tabel 6. *Confusion Matrix*

Aktual	Prediksi		
	TAV	TKJ	TBSM
TAV	2	1	0
TKJ	1	3	0
TBSM	0	0	3

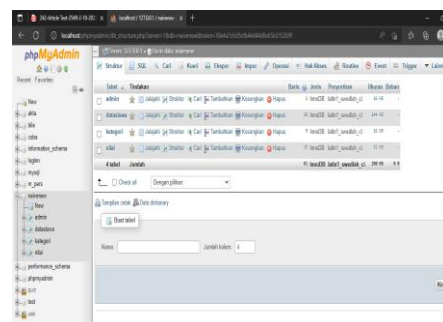
Maka nilai akurasi confusion matrix berdasarkan tabel 6 diperoleh dengan persamaan sebagai berikut :

$$a = \frac{t}{n} \times 100\% \quad (2)$$

Dimana a adalah nilai akurasi yang akan dihasilkan, t adalah banyak prediksi yang benar dan n adalah total prediksi keseluruhan.

3.6. Perancangan Sistem dan Database

Sebelum menggunakan aplikasi sistem informasi penentuan jurusan berbasis web maka harus dilakukan perancangan database terlebih dahulu, yang sudah tersedia dan terpasang pada komputer dengan benar. Perancangan instalasi diawali dengan menginstalasi Xampp untuk database berbasis MySQL. Tampilan database dapat dilihat pada Gambar 3.



Gambar 3. Database.

3.7. Pembangunan Sistem

Pembangunan sistem yaitu tahap tahap yang dilakukan dalam menyusun program aplikasi yang mana aplikasi tersebut nantinya dapat digunakan dengan baik sebagai suatu cara yang digunakan untuk menemukan solusi dari sebuah permasalahan yang terjadi. Sistem ini dibangun dengan menggunakan bahasa pemrograman PHP dan database MySQL.[15].

3.8. Implementasi Sistem

Implementasi sistem yaitu langkah-langkah atau prosedur-prosedur yang dilakukan dalam menyelesaikan desain sistem yang telah disetujui, untuk menginstal, menguji dan memulai sistem baru atau sistem yang diperbaiki. Sebelum sistem ini dijalankan pada *server* internet, sistem ini akan dijalankan pada *server* komputer lokal terlebih dahulu.

a. Tampilan Halaman Utama

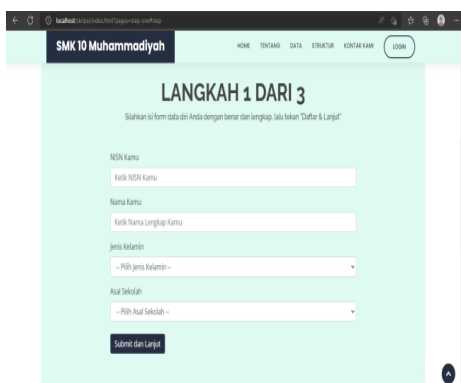
Halaman Utama merupakan Halaman yang pertama kali dijumpai saat user atau admin mengakses web, terdapat beberapa menu yang dapat dipilih pada halaman utama, tampilan pada halaman utama dapat dilihat pada gambar 4.



Gambar 4. Halaman Utama

b. Tampilan Halaman Prediksi

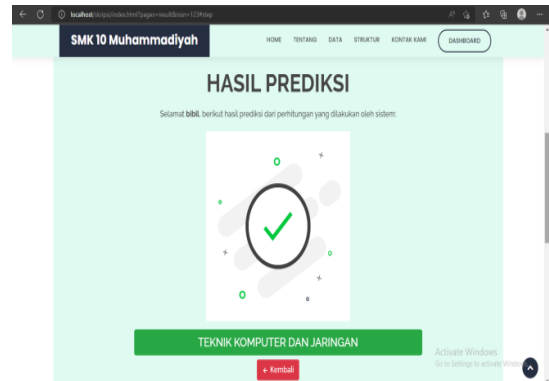
Halaman prediksi merupakan Halaman yang dapat dipilih oleh user untuk melakukan proses prediksi jurusan, menampilkan beberapa form pertanyaan yg harus diisi untuk mendapatkan hasil prediksi. Adapun halaman prediksi aplikasi dapat dilihat pada gambar 5.



Gambar 5. Halaman Prediksi

c. Tampilan Halaman Hasil Prediksi

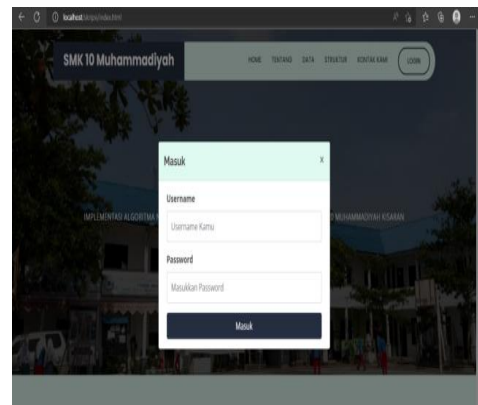
Halaman hasil prediksi merupakan halaman yang menampilkan hasil dari prediksi atau penentuan jurusan berdasarkan data yang telah diinputkan oleh *user* atau siswa. Adapun tampilan halaman hasil prediksi dapat dilihat pada gambar 6.



Gambar 6. Halaman Dashboard

d. Tampilan Form Login Admin

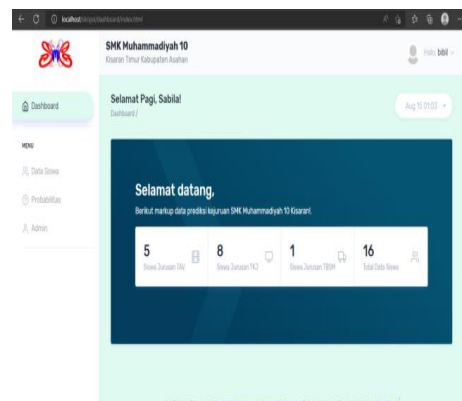
Form Login berfungsi untuk verifikasi data *admin* yang sudah terdaftar pada database untuk masuk ke dalam sistem. Adapun tampilan *form login* dapat dilihat pada gambar 7.



Gambar 7. Form Login

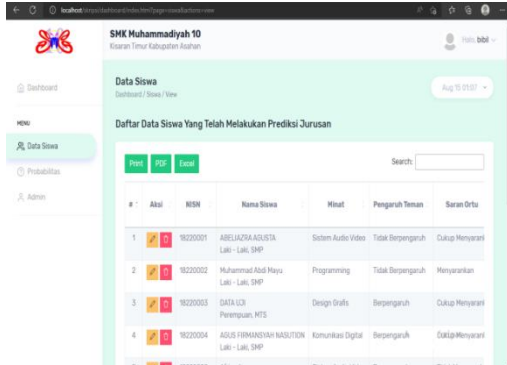
e. Tampilan Halaman Dashboard

Halaman *Dashboard* merupakan Halaman yang akan dijumpai setelah admin berhasil melakukan proses *login* dan halaman ini hanya dapat diakses oleh admin saja. Adapun halaman prediksi aplikasi dapat dilihat pada gambar 8.



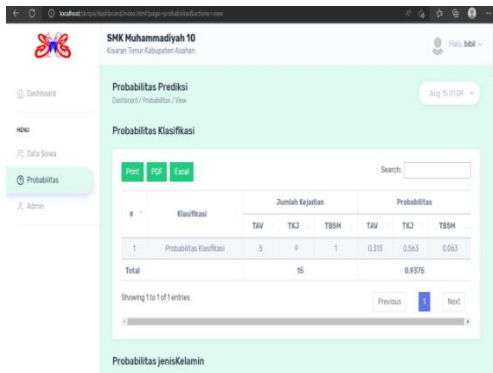
Gambar 8. Halaman Dashboard

f. Tampilan Menu Data Siswa
 Menu ini akan menampilkan seluruh daftar siswa yang telah melakukan proses prediksi pada aplikasi ini. Adapun menu Data Siswa pada aplikasi ini dapat dilihat pada gambar 9.



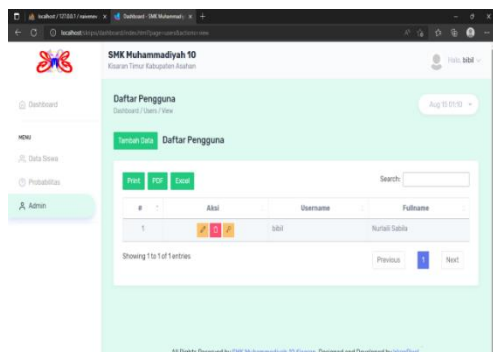
Gambar 9 Tampilan Menu Data Siswa

g. Tampilan Menu Probabilitas
 Menu Probabilitas merupakan menu yang akan menampilkan probabilitas dari data siswa yang telah melakukan proses prediksi pada aplikasi ini. Adapun menu Probabilitas pada aplikasi ini dapat dilihat pada gambar 10.



Gambar 10 Tampilan Menu Probabilitas

h. Tampilan Menu Admin
 Halaman ini akan menampilkan seluruh data admin yang sudah tersimpan di database, pada menu ini admin juga bisa menambah, mengedit, serta menghapus data admin yang sudah terdaftar. Adapun menu Admin pada aplikasi ini dapat dilihat pada gambar 11.



Gambar 11 Tampilan Menu Admin

4. KESIMPULAN

Berdasarkan hasil penelitian, data mining menggunakan metode naive bayes yaitu, Sistem yang telah dirancang mampu memprediksi jurusan siswa SMK 10 Muhammadiyah Kisaran sehingga mereka dapat mengetahui jurusan yang sesuai berdasarkan informasi yang diberikan dengan tingkat akurasi sebesar 89%.

Berdasarkan pengujian sistem dengan *black box testing*, sistem yang dirancang dapat berfungsi dengan baik dan benar sesuai dengan fungsi-fungsi yang ada dalam program dan dapat diimplementasikan untuk penentuan jurusan berbasis web di SMK 10 Muhammadiyah Kisaran.

DAFTAR PUSTAKA

- [1] A. I. Rizmayanti, N. Hidayati, F. S. Nugraha, And W. Gata, "Penerapan Data Mining Untuk Memprediksi Kompetensi Siswa Menggunakan Metode Decision Tree (Studi Kasus Smk Multicomp Depok)," *Swabumi*, Vol. 9, No. 1, Pp. 9–18, 2021, Doi: 10.31294/Swabumi.V9i1.8363.
- [2] M. Kusmira And R. E. Indrajit, "Penerapan Algoritma Klasifikasi Data Mining Decision Tree Untuk Menentukan Penjurusan Siswa Sma 6 Tasikmalaya," *Simp. Nas. Ilmu Pengetah. Dan Teknol.*, Pp. 49–53, 2017.
- [3] M. Rahmayu And R. K. Serli, "Sistem Pendukung Keputusan Pemilihan Jurusan Pada Smk Putra Nusantara Jakarta Menggunakan Metode Analytical Hierarchy Process (Ahp)," *Simetris J. Tek. Mesin, Elektro Dan Ilmu Komput.*, Vol. 9, No. 1, Pp. 551–564, 2018, [Online]. Available: <https://jurnal.umk.ac.id/index.php/Simet/Article/View/2022>.
- [4] A. W. Kusuma, D. Mahdiana, M. I. Komputer, F. Teknologi, I. Universitas, And B. Luhur, "Development Of Data Warehouse To Predicate The Regarding Of Umroh Congregations Using The Manearest Neighbour Algorithm (Case Study Pt . Bahana Sukses Sejahtera) Pengembangan Data Warehouse Untuk Memprediksi Pengunduran Diri Jemaah Umroh Menggunakan (Studi Kasus Pt . Bahana Sukses Sejahtera)," Vol. 3, No. 4, Pp. 1007–1012, 2022.
- [5] M. K. Hartono, "Prediction Of Baby Birth Rate Using Naïve Bayes Classification Algorithm In Randau Village," Vol. 3, No. 4, Pp. 863–869, 2022.
- [6] A. N. Yuliarina, "Comparison Of Prediction Analysis Of Gofood Service Users Using The Knn & Naive Bayes Algorithm With Rapidminer Software Perbandingan Analisis Prediksi Kepuasan Pengguna Layanan

- Gofood Menggunakan Algoritma Knn & Naive Bayes Dengan Software Rapidminer,” Vol. 3, No. 4, Pp. 847–856, 2022.
- [7] I. Yunanto And S. Yulianto, “Twitter Sentiment Analysis Pedulilindungi Application Using Naïve Bayes And Support Vector Machine Analisis Sentimen Twitter Aplikasi Pedulilindungi,” Vol. 3, No. 4, Pp. 807–814, 2022.
- [8] Idris, “Implementasi Data Mining Dengan Algoritma Naive Bayes Untuk Memprediksi Angka Kelahiran,” *J. Pelita Inform.*, Vol. 7, No. 3, Pp. 421–428, 2019, [Online]. Available: <https://ejournal.stmik-budidarma.ac.id/index.php/pelita/article/view/1154>.
- [9] R. Setiawan, M. R. Nashrulloh, R. Ramadhani, And A. Sutedi, “Enterprise Architecture System In Private Vocational School Using Togaf Adm (Case Study Of Smk Al-Hikmah) Arsitektur Enterprise Sistem Pada Sekolah Menengah Kejuruan Swasta Menggunakan Togaf Adm (Studi Kasus Smk Al-Hikmah),” Vol. 3, No. 1, Pp. 183–191, 2022.
- [10] D. Pertiwi, A. A. A. Arifin, S. S. Utama, And M. A. Sembiring, “Pengaruh Implementasi Aplikasi Penentu Program Studi Berbasis Android Untuk Calon Mahasiswa Stmik Royal,” *J. Sci. Soc. Res.*, Vol. 4, No. 3, P. 299, 2021, Doi: 10.54314/Jssr.V4i3.659.
- [11] R. Kurniasari And A. Fatmawati, “Penerapan Algoritma C4.5 Untuk Penjurusan Siswa Sekolah Menengah Atas,” *Komputa J. Ilm. Komput. Dan Inform.*, Vol. 8, No. 1, Pp. 19–27, 2019, Doi: 10.34010/Komputa.V8i1.3045.
- [12] S. Widaningsih, “Perbandingan Metode Data Mining Untuk Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma C4,5, Naïve Bayes, Knn Dan Svm,” *J. Tekno Insentif*, Vol. 13, No. 1, Pp. 16–25, 2019, Doi: 10.36787/Jti.V13i1.78.
- [13] D. P. Utomo And B. Purba, “Penerapan Datamining Pada Data Gempa Bumi Terhadap Potensi Tsunami Di Indonesia,” *Pros. Semin. Nas. Ris. Inf. Sci.*, Vol. 1, No. September, P. 846, 2019, Doi: 10.30645/Senaris.V1i0.91.
- [14] H. D. Wijaya And S. Dwiasnati, “Implementasi Data Mining Dengan Algoritma Naïve Bayes Pada Penjualan Obat,” *J. Inform.*, Vol. 7, No. 1, Pp. 1–7, 2020, Doi: 10.31311/Ji.V7i1.6203.
- [15] I. K. Siregar And M. Ihsan, “Application Of The Certainty Factor Method For Diagnose Palm Oil Disease Web-Based Penerapan Metode Certainty Factor Untuk Mendiagnosa,” Vol. 3, No. 3, Pp. 581–590, 2022.