

**Prediksi Kelulusan Mata Kuliah Mahasiswa Teknologi Informasi  
Menggunakan Algoritma K-Nearest Neighbor**

*The Prediction for Graduation for Information Technology Student's Course  
Using The K-Nearest Neighbor Algorithm*

**Nazaruddin Ahmad<sup>1\*</sup>, Saifan Hafizh<sup>2</sup>, Rana Sulthanah<sup>3</sup>**

Program Studi Teknologi Informasi, Universitas Islam Negeri Ar-Raniry Banda Aceh, Indonesia

\*E-mail: [nazar.ahmad@ar-raniry.ac.id](mailto:nazar.ahmad@ar-raniry.ac.id)

**Abstrak**

Penelitian ini dilakukan untuk mengembangkan sebuah model prediksi dengan menggunakan metode K-Nearest Neighbor (KNN) yang bertujuan melakukan prediksi kelulusan mata kuliah mahasiswa Program Studi Teknologi Informasi. Permasalahannya adalah ketidakpastian dalam melakukan peramalan kelulusan mahasiswa berdasarkan data historis dan atribut tertentu. Yang menjadi perhatian pada penelitian ini adalah pentingnya memahami faktor yang dapat memberikan pengaruh kelulusan mahasiswa dalam mata kuliah Sistem Manajemen Basis Data, sehingga dapat memberikan prediksi yang akurat dan membantu dalam meningkatkan tingkat kelulusan mahasiswa pada mata kuliah Sistem Manajemen Basis Data. Tujuan penelitian ini adalah membangun model prediktif dengan algoritma KNN dan menerapkan model algoritma KNN ke dalam bahasa pemrograman PHP. Dari penelitian ini diharapkan dapat memberikan wawasan yang berharga bagi institusi membantu meningkatkan proses pengajaran dan pembelajaran, dan menambah pemahaman dalam penerapan konsep data mining pada studi kasus tertentu. Prediksi yang dilakukan adalah untuk mengetahui seorang mahasiswa itu akan masuk ke kelas LULUS atau TIDAK LULUS pada mata kuliah Sistem Manajemen Basis Data berdasarkan data latih dan data uji. Dilakukan perhitungan untuk mendapatkan nilai jarak terdekat antar data latih dengan data uji. Hasil nya adalah didapatkan nilai akurasi sebesar 90% untuk proses prediksi kelulusan mata kuliah dengan menggunakan  $k=5$  dengan 40 data latih dan 20 data uji.

**Kata kunci:** data mining, algoritma K-Nearest Neighbor, php, prediksi, database.

**Abstract**

This research aims to develop a predictive model using the K-Nearest Neighbor (KNN) method to forecast the course completion of students in the Information Technology program. The issue at hand is the uncertainty in predicting student success based on historical data and specific attributes. This study focuses on the importance of understanding the factors that influence student success in the Database Management Systems course to provide accurate predictions and help improve student pass rates in this course. The objective of this research is to build a predictive model using the KNN algorithm and to implement this model using the PHP programming language. The study aims to offer valuable insights for educational institutions to enhance teaching and learning processes and to expand understanding of data mining concepts in specific case studies. The prediction aims to determine whether a student will pass or fail the Database Management Systems course based on predetermined training and testing data. By calculating the nearest distance between the training data and the test data. The results showed an accuracy rate of 90% for predicting course completion using  $k=5$ , with a dataset consisting of 40 training data points and 20 testing data points.

**Keywords:** data mining, K-Nearest Neighbor algorithm, php, prediktion, database.

Naskah diterima 03 Mar. 2024; direvisi 14 Jun. 2024; dipublikasikan 01 Okt. 2024.

JAMIKA is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.



**I. PENDAHULUAN**

Mahasiswa merupakan satu diantara asset penting yang harus dikelola dengan baik pada sebuah institusi pendidikan. Setiap perguruan tinggi dituntut untuk dapat menyelenggarakan pendidikan yang berkualitas. Kelulusan pada setiap mata kuliah yang diambil oleh mahasiswa pada setiap semester menjadi hal yang perlu diperhatikan dikarenakan dapat mempengaruhi lamanya masa studi yang ditempuh. Ketidakstabilan persentase terhadap kemampuan mahasiswa untuk menyelesaikan mata kuliah sangat berpengaruh terhadap elemen-elemen penilaian lainnya. Untuk itu, perlu adanya pemantauan dan evaluasi kelulusan mata kuliah yang diambil oleh mahasiswa.

Nilai kelulusan suatu mata kuliah menjadi patokan bagi mahasiswa bahwa dia mampu menguasai mata kuliah tersebut dan mahasiswa tersebut dapat memastikan bahwa dia tidak akan mengulang mata kuliah tersebut. Nilai kelulusan mata kuliah yang diberikan oleh dosen juga menjadi indikator bahwa pola mengajar

yang diberikan sudah sesuai dengan standar pengajaran yang ditentukan. Nilai mata kuliah yang dikumpulkan oleh mahasiswa tidak saja hanya satu komponen saja tapi terdiri dari komponen nilai quiz, nilai tugas mandiri, nilai tugas terstruktur, nilai ujian tengah semester (UTS), nilai ujian praktik, dan nilai ujian akhir semester (UAS). Dosen perlu melakukan prediksi terhadap nilai-nilai yang sudah didapatkan oleh mahasiswa sehingga dosen akan dapat mengelompokkan mahasiswa yang lulus mata kuliah dan mahasiswa yang tidak lulus mata kuliah yang diajarkan.

Proses prediksi kelulusan mata kuliah yang dilakukan oleh dosen sangatlah penting karena akan menjadi dasar pertimbangan bagi dosen tersebut untuk melakukan tindak lanjut terhadap nilai yang didapatkan oleh mahasiswa yang berpotensi tidak lulus mata kuliah yang diambil. Dosen harus mendapatkan capaian pembelajaran yang tuntas dengan acuan bahwa setiap mahasiswa yang mengambil mata kuliah yang diampu oleh dosen tersebut mampu menguasainya, mendapatkan nilai yang maksimal, dan masuk ke dalam kategori lulus. Diperlukan sebuah cara untuk mengetahui pola data nilai yang ada kemudian diklasifikasi ke dalam kelas-kelas yang sudah didefinisikan sebelumnya.

Terkait dengan cara untuk mengetahui pola dan melakukan klasifikasi maka penulis mencoba untuk menerapkan algoritma *K-Nearest Neighbor* yang menjadi salah satu metode di dalam data mining untuk mendapatkan informasi terkait dengan lulus atau tidak lulus seorang mahasiswa pada suatu mata kuliah yang diambil berdasarkan komponen-komponen nilai-nilai yang sudah didapatkan oleh mahasiswa [1]. Sehingga dengan menerapkan algoritma *K-Nearest Neighbor* akan didapatkan informasi mahasiswa yang lulus dan tidak lulus mata kuliah yang diajarkan oleh dosen, dan dosen dapat melakukan tindak lanjut terhadap mahasiswa yang tidak lulus mata kuliah tersebut [2].

Data mining adalah sebuah metode yang dimungkinkan untuk dapat mengakses data dalam jumlah yang besar pada suatu waktu yang dapat menampilkan informasi dari data tersebut menjadi informasi terbaru dengan menggunakan cara mencari pola dari data tersebut [3]. Data mining juga dapat diartikan dengan metode yang digunakan untuk menggali informasi dari sekumpulan data yang ada atau juga melakukan proses penambahan data dari jumlah data yang besar untuk mendapatkan informasi baru dengan algoritma tertentu [4], mampu untuk mengelompokkan data tertentu berdasarkan karakteristik dari data [5]. Data mining dapat menggunakan konsep ilmu statistik, konsep matematika, penerapan konsep kecerdasan buatan, dan juga pembelajaran mesin (*machine learning*) yang digunakan untuk mengekstrak informasi [6], [7].

Salah satu konsep algoritma yang umum digunakan di dalam data mining adalah algoritma *K-Nearest Neighbor* yang sering disebut dengan Algoritman KNN [8]. Metode *K-Nearest Neighbor* (KNN) adalah salah satu model algoritma klasifikasi di dalam data mining untuk melakukan prediksi terhadap data-data yang digunakan untuk melihat kemiripan, kesesuaian, dan juga kedekatan jarak antar data tersebut [9], [10]. Tujuan dari penentuan jarak terdekat antara data latih dengan data uji adalah untuk menentukan klasifikasi atau kelas dari data yang diuji sehingga akan didapatkan informasi terbaru dari data yang digunakan [11].

Algoritma KNN adalah algoritma yang digunakan untuk melakukan prediksi berdasarkan kedekatan jarak antar data yang ada yang menggunakan konsep pembelajaran data [12]. Menurut [13] terdapat 2 proses di dalam model klasifikasi (1) adanya proses *learning*, yaitu membangun model dengan menggunakan data latih (*data training*), (2) adalah proses *testing*, yaitu menggunakan model yang telah didapatkan dari proses *training*. Algoritma KNN ini umum digunakan karena mampu melakukan prediksi berdasarkan data uji kemudian diproses menggunakan data latih yang menghasilkan nilai akurasi yang tinggi dalam klasifikasi data [14].

Algoritma KNN ini sudah digunakan dalam berbagai penelitian yang dilakukan seperti penerapan algoritma KNN pada klasifikasi data penerima zakat di kota pekanbaru yang mendapatkan nilai akurasi 89% [15]. Penelitian yang dilakukan oleh [16] menggunakan algoritma KNN untuk melakukan deteksi daun pada tanaman menggunakan perangkat mobile yang dianggap mampu mendeteksi dan melakukan klasifikasi jenis daun tanaman. Penelitian yang lain juga dilakukan oleh Qurotul A'yuniyah dan Muhammad Reza pada tahun 2023 dengan menggunakan algoritma KNN untuk melakukan klasifikasi jurusan siswa SMA dengan mendapatkan nilai akurasi sebesar 93,52% dengan nilai  $k=3$  [17]. Algoritma KNN ini juga dapat digunakan pada penelitian bidang kesehatan seperti penelitian yang dilakukan oleh Andi Maulida Argina pada tahun 2020 untuk melakukan klasifikasi penderita penyakit diabetes sehingga mendapatkan informasi diabetes lebih awal dengan nilai akurasi sebesar 39% dengan nilai  $k=3$  [18]. Penelitian yang dilakukan ini adalah untuk melihat hasil prediksi LULUS atau TIDAK LULUS pada mata kuliah Sistem Manajemen Basis Data dimana data latih nya dimasukkan ke dalam database, algoritma KNN dan proses perhitungannya dimasukkan ke dalam coding bahasa pemrograman PHP sehingga memberikan kemudahan kepada pengguna untuk melakukan pengujian data.

Metode algoritma *K-Nearest Neighbor* dipilih sebagai metode prediksi dalam penelitian ini karena metode ini dianggap mudah untuk dipahami dan dapat menjadi langkah awal dalam mempelajari model-model

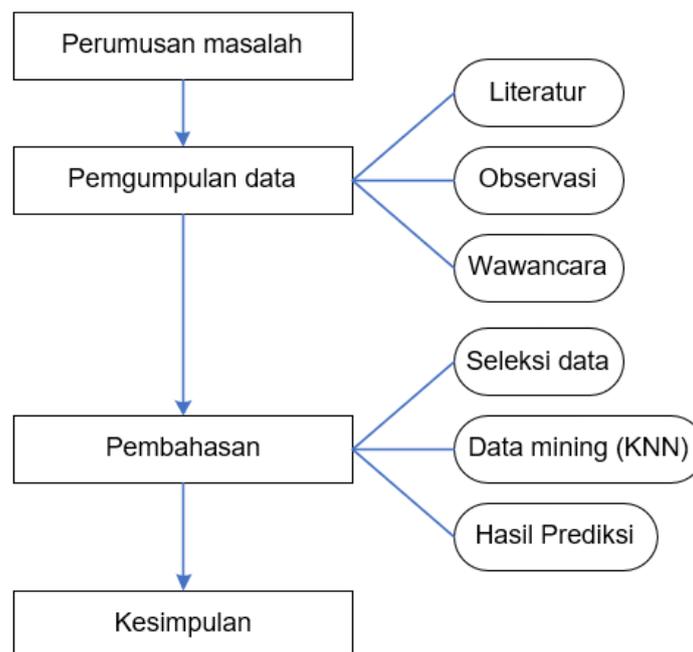
algoritma prediksi lainnya. Tentunya dalam mempelajari sesuatu harus dari hal yang mudah dipahami terlebih dahulu sehingga akan memudahkan untuk mempelajari model-model prediksi lainnya. Hasil perhitungan manual yang dilakukan juga tidak sulit sehingga hasil yang didapatkan dapat dipahami secara baik. Adapun tujuan dari penelitian ini adalah untuk mengetahui keterangan LULUS atau TIDAK LULUS bagi mahasiswa yang mengambil mata kuliah Sistem Manajemen Basis Data dengan mengimplementasikan algoritma *K-Nearest Neighbor*. Antar muka yang digunakan adalah sebuah program sederhana menggunakan bahasa pemrograman PHP dengan data nilai mahasiswa disimpan pada tabel di dalam *database*. Data uji dimasukkan dengan menggunakan form input untuk melihat proses klasifikasi yang dilakukan dengan data latih yang sudah tersimpan di dalam *database*.

## II. METODE PENELITIAN

Penelitian dilakukan terhadap mahasiswa program studi Teknologi Informasi Fakultas Sains dan Teknologi UIN Ar-Raniry Banda Aceh yang mengambil mata kuliah Sistem Manajemen Basis Data. Data yang digunakan adalah hasil penilaian perkuliahan selama 1 semester.

Penelitian ini menggunakan model studi kasus. Dibangun sebuah program sederhana menggunakan bahasa PHP untuk melakukan proses perhitungan menggunakan algoritma KNN. Database menggunakan XAMPP MySQL untuk menyimpan data nilai mahasiswa. Menggunakan aplikasi *visual studio code* untuk mengetikkan perintah program dalam bahasa PHP.

Tahapan penelitian yang dilakukan adalah sebagai berikut:



Gambar 1. Alur tahapan penelitian

**Perumusan masalah** : melakukan pengkajian masalah terkait dengan adanya ketidakpastian di dalam peramalan terhadap ketidakpastian kelulusan mahasiswa pada mata kuliah Sistem Manajemen Basis Data.

**Penggumpulan data** : data dikumpulkan dari pendataan nilai mahasiswa pada semester genap tahun akademik 2022/2023 sebanyak 80 data mahasiswa. Dilakukan tahapan studi literatur untuk mendapatkan dasar teori tentang data mining menggunakan metode algoritma KNN. Tahapan observasi dilakukan untuk mengamati secara langsung objek yang terlibat di dalam penelitian ini. Data yang dikumpulkan adalah data nilai mahasiswa pada mata kuliah Sistem Manajemen Basis Data pada tahun ajaran semester genap tahun akademik 2022/2024. Dan melakukan tahapan wawancara untuk mendapatkan informasi langsung dari mahasiswa terkait mata kuliah Sistem Manajemen Basis Data terkait dengan perolehan nilai.

**Pembahasan** : melakukan pemisahan data mata kuliah dengan melakukan seleksi data nilai mahasiswa. Kemudian memberikan pemahaman tentang data mining menggunakan metode KNN dan melakukan perhitungan dengan konsep metode KNN sehingga akan mendapatkan hasil prediksi.

**Kesimpulan** : memberikan hasil dari peroleh proses yang dilakukan untuk memberikan penjelasan terkait penggunaan metode algoritma KNN.

### Analisis Sistem

Proses analisis sistem dilakukan untuk melihat prediksi kelulusan mahasiswa pada mata kuliah Sistem Manajemen Basis Data dengan menggunakan algoritma KNN dengan menggunakan persamaan dan langkah-langkah sebagai berikut [19]:

1. Menentukan parameter K sebagai banyaknya jumlah tetangga terdekat dengan objek baru.
2. Menghitung jarak antar objek/data baru terhadap semua objek/data yang telah di training menggunakan rumus:

$$euc = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

3. Urutkan hasil perhitungan tersebut.
4. Tentukan tetangga terdekat berdasarkan jarak minimum ke K.
5. Tentukan kategori dari tetangga terdekat dengan objek/data.
6. Gunakan kategori mayoritas sebagai klasifikasi objek/data baru.

### Euclidean Distance

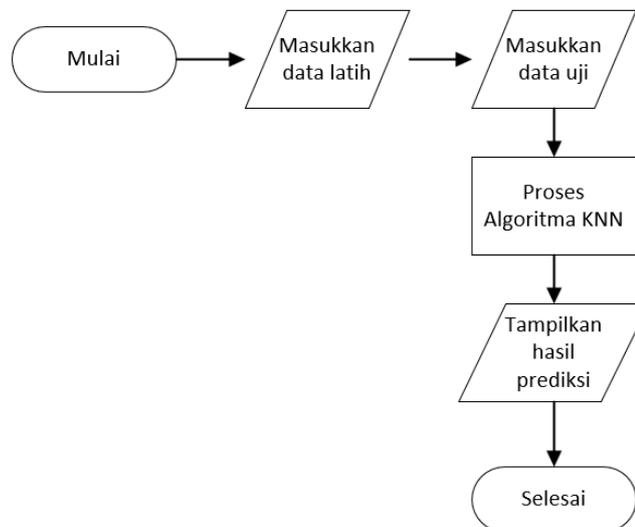
*Euclidean distance* merupakan suatu metode perhitungan yang digunakan untuk mengukur jarak (*distance*) [20]. Jarak dihitung adalah jarak antara data *testing* terhadap data *training* yang digunakan sehingga akan didapatkan nilai dengan jarak terdekat terhadap data-data yang ada. Persamaan yang digunakan adalah sebagai berikut:

$$euc = d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2)$$

dimana:  $x_i$  = data latih (data *trainig*)  
 $y_i$  = data uji (data *testing*)  
 $i$  = variabel data  
 $euc = d(x,y)$  = jarak  
 $n$  = jumlah data

### Alur Prediksi Kelulusan

Diagram alur prediksi kelulusan dapat dilihat seperti gambar 1 berikut ini:



Gambar 2. Diagram alur prediksi lulus mata kuliah

Diagram alur pada gambar 2 di atas memperlihatkan proses pengerjaan algoritma secara umum untuk menyelesaikan perhitungan menggunakan metode KNN. Dimulai dengan memasukkan data mahasiswa yang sudah mendapatkan keterangan LULUS dan TIDAK LULUS yang akan digunakan sebagai data latih (*training*). Kemudian memasukkan data yang akan diuji untuk dilakukan proses perhitungan dengan metode KNN yang ingin di prediksi LULUS atau TIDAK LULUS. Kemudian dilakukan perhitungan dengan algoritma KNN dan selanjutnya didapatkan hasil prediksi.

#### Data

Data yang digunakan ada 2, yaitu:

1. Data latih (*training*), yaitu sampel data nilai mahasiswa yang telah dinyatakan LULUS dan TIDAK LULUS pada mata kuliah Sistem Manajemen Basis Data. Atribut data yang digunakan adalah Quiz, TKT, TKM, Praktik, UTS, dan UAS.
2. Data uji (*testing*), yaitu data yang digunakan sebagai data yang diuji dimana data tersebut belum diketahui keterangannya LULUS atau TIDAK LULUS yang memiliki atribut yang sama dengan data latih (*training*).

#### Pengumpulan Data

Data yang digunakan untuk penelitian ini diperoleh dari data nilai mata kuliah Sistem Manajemen Basis Data dengan total jumlah mahasiswa yang dijadikan sebagai data adalah 60 orang mahasiswa, dimana 40 orang data mahasiswa yang digunakan sebagai data latih, dan 20 orang data mahasiswa sebagai data uji. Data ini didapatkan dari pendataan nilai mahasiswa yang dilakukan pada semester genap 2022/2023 untuk mata kuliah Sistem Manajemen Basis Data. Data yang digunakan terdiri dari nilai quiz, nilai tugas mandiri, nilai tugas terstruktur, nilai ujian tengah semester (UTS), nilai ujian praktik, dan nilai ujian akhir semester (UAS). Pada tabel 1 berikut ini merupakan data latih (*training*) yang digunakan, yaitu:

TABEL 1  
DATA LATIH (TRAINING)

	Data Mahasiswa	Quiz	TKM	TKT	Praktik	UTS	UAS	TOTAL	KETERANGAN
1	NABIL ANANDA	30	80	75	79	41	65	61	TIDAK LULUS
2	ORIZA SEPTIAN	70	0	0	0	20	65	34	TIDAK LULUS
3	INTAN WAHYUNI	70	100	95	96	90	75	85	LULUS
4	HAIFA DAYANA	70	100	95	96	20	65	68	LULUS
5	ARIEF FATHIN ABRAR	70	95	95	91	35	90	78	LULUS
6	FIYES FALEFI	70	95	95	88	30	65	68	LULUS
7	RIYASHA	50	100	95	93	50	65	71	LULUS
8	ZABRINA DEWI	40	80	95	85	25	65	62	TIDAK LULUS
9	SELIA PUTRI	30	95	95	88	25	65	63	TIDAK LULUS
10	ZURIATI FITRI	60	100	95	93	10	65	64	TIDAK LULUS
11	ARMIS	60	100	95	89	15	65	65	TIDAK LULUS
12	NOVI NURFARIZA	40	95	95	92	30	65	66	LULUS
13	M. AZUWAN RAFSANJANI	60	100	95	92	25	90	76	LULUS
14	DIVA ERLANGGA PRAMUDYA ZUHDI	55	80	70	65	45	65	62	TIDAK LULUS
15	AIDIL GHUFRAN	60	95	95	91	25	65	66	LULUS
16	RANA SULTHANAH	70	100	95	97	95	95	93	LULUS
17	SRI RAHMI DAHLIA WULANDARI	40	95	95	89	70	65	73	LULUS
18	SYAHRU RIZKI	70	85	95	85	25	65	66	LULUS
19	ARDY RAMADHANA	70	90	95	86	35	65	68	LULUS
20	HADI ALBAR	70	95	95	90	35	65	69	LULUS
21	DEWI AYU SAPUTRI	50	80	80	84	25	65	61	TIDAK LULUS
22	MUHAMMAD ARIQ RUNA	80	100	95	86	45	65	72	LULUS
23	MUHAMMAD IKHSAN RAJABNA	60	95	95	89	16	65	64	TIDAK LULUS
24	FURQAN	50	100	95	91	10	65	63	TIDAK LULUS
25	ADHE MERA ULFA	40	100	95	91	15	65	63	TIDAK LULUS
26	MUHAMMAD FIKRI NAUFALDI	40	80	80	86	56	75	70	LULUS
27	MUHAMMAD FERDIAN	60	100	95	87	25	65	66	LULUS
28	LILIS MAULINA	60	100	95	87	20	65	65	LULUS
29	FAYYADH ZHAFRAN	60	95	95	91	10	85	70	LULUS
30	TEUKU GHIFARI TS	50	100	95	92	35	85	75	LULUS

Data Mahasiswa	Quiz	TKM	TKT	Praktik	UTS	UAS	TOTAL	KETERANGAN
31 MUAMMAR AL-FARUQI	80	100	70	90	25	65	66	LULUS
32 YUDHI JAUHARI	60	100	95	93	45	90	80	LULUS
33 KISMULLAH	40	95	95	79	0	65	58	TIDAK LULUS
34 RAUDHATUNNISA	60	95	95	88	25	75	69	LULUS
35 IDHAM ANGGA SETIAWAN	60	100	95	90	35	65	69	LULUS
36 YUDI ARAMI	40	100	95	88	10	65	61	TIDAK LULUS
37 ARJA DANIALA	45	75	85	70	45	65	63	TIDAK LULUS
38 SHVVIEL AZKA	80	100	80	81	0	65	61	TIDAK LULUS
39 IRFANI ARIJA	60	100	95	89	56	65	73	LULUS
40 MUHAMMAD FADHLAN	45	95	95	79	32	60	63	TIDAK LULUS

TABEL 2  
DATA UJI (TESTING)

Nama Mahasiswa	Quiz	TKM	TKT	Praktik	UTS	UAS	TOTAL	KETERANGAN
1 SISNA ELLYANA	45	95	95	93	90	85	85	?
2 ALBIRAN NISA	40	95	95	85	90	85	84	?
3 HAURA SALSABILA	30	100	95	98	90	55	74	?
4 CUT RAUDHATUL ILMI	45	95	95	96	90	55	75	?
5 M. SYAUKI	25	100	95	85	40	55	62	?
6 MAULINATUL HAKIKI	45	100	95	99	100	55	78	?
7 PUTRI MAULAYA	45	97	95	98	90	55	76	?
8 DAFFAUL AISAR	50	100	95	98	50	55	68	?
9 M. HAYKAL. MZ	25	100	95	98	40	55	64	?
10 AMIRUL 'ADLY	40	97	95	96	40	55	65	?
11 M. ABDUL AZIZ	45	80	95	78	30	55	59	?
12 FAIZ RAMADHANA SURYA	40	85	95	95	40	85	74	?
13 HARIST ISLAMI RIDHA	40	95	95	97	40	55	65	?
14 T. AUDIANSYAH	30	95	95	94	60	85	78	?
15 MUHAMMAD RIAND	60	100	95	98	70	100	89	?
16 MUHAMMAD SAYYIDAL HAMDY INYU	50	95	95	95	50	55	68	?
17 MUHAMMAD NIZAM	25	95	95	90	24	70	64	?
18 M. MARTUNIS	25	97	95	93	56	55	66	?
19 ERA SYAFINA	35	95	95	92	90	100	89	?
20 AFIF MOELANA	25	90	95	87	60	100	81	?

### Tahapan Algoritma KNN

Tahapan yang dilakukan dalam melakukan perhitungan algoritma KNN yang dilakukan secara manual adalah sebagai berikut:

- Tentukan nilai  $k$  yang akan digunakan.
- Melakukan penghitungan *Euclidean distance* masing-masing objek terhadap sampel data yang ada menggunakan persamaan (2).
- Menghitung jarak data latih ke data uji.
- Melakukan perbandingan jarak data latih ke data uji dari jarak yang terkecil.
- Menentukan label kelas didasarkan pada label mayoritas pada  $k$ .

### Menghitung Nilai Euclidean Distance

Ada dua objek yang akan diukur tingkat kedekatannya menggunakan *Euclidean distance*. Untuk objek data latih kita namakan variabel  $X$  dan untuk objek data uji kita namakan variabel  $Y$ . Untuk objek data latih dengan variabel  $X$  adalah sebagai berikut:

Data Mahasiswa	Quiz	TKM	TKT	Praktik	UTS	UAS	TOTAL	KETERANGAN
	X1	X2	X3	X4	X5	X6	X7	
Data Latih-3	70	100	95	96	90	75	85	LULUS

Yang akan dilakukan perhitungan adalah variable  $X_1, X_2, X_3, X_4, X_5, X_6$ , dan  $X_7$ . Adapun untuk objek data uji dengan variabel  $Y$  adalah sebagai berikut:

Data Mahasiswa	Quiz	TKM	TKT	Praktik	UTS	UAS	TOTAL	KETERANGAN
	Y1	Y2	Y3	Y4	Y5	Y6	Y7	
Data Uji-1	45	95	95	93	90	85	85	?

Yang akan dilakukan perhitungan adalah Y1, Y2, Y3, Y4, Y5, Y6, dan Y7.

Kemudian dilakukan perhitungan *Euclidean distance* untuk variabel X dan variabel Y menggunakan persamaan 2 sebagai berikut:

$$(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + (x_4 - y_4)^2 + (x_5 - y_5)^2 + (x_6 - y_6)^2 + (x_7 - y_7)^2}$$

$$(x, y) = \sqrt{(70 - 45)^2 + (100 - 95)^2 + (95 - 95)^2 + (96 - 93)^2 + (90 - 90)^2 + (75 - 85)^2 + (85 - 85)^2}$$

$$(x, y) = \sqrt{(25)^2 + (5)^2 + (0)^2 + (3)^2 + (0)^2 + (-10)^2 + (0)^2}$$

$$(x, y) = \sqrt{625 + 25 + 0 + 9 + 0 + 100 + 0}$$

$$(x, y) = \sqrt{759}$$

$$(x, y) = 27,55$$

Sehingga, dari perhitungan di atas dapat ditentukan bahwa tingkat kedekatan X dan Y adalah 27,55. Model perhitungan seperti di atas nanti yang akan dilakukan untuk data latih dan data uji yang sudah ditentukan.

### Menghitung Confusion Matrix

*Confusion Matrix* adalah metode evaluasi yang disajikan dalam bentuk tabel yang digunakan untuk menghitung tingkat akurasi dari hasil prediksi dengan membandingkan jumlah data yang benar dan data yang salah dari hasil prediksi data [21]. *Confusion matrix* juga digunakan untuk menghitung kinerja terkait dengan kebenaran di dalam suatu proses klasifikasi [22]. Di dalam penelitian ini menggunakan model *confusion matrix* 2x2. Seperti terlihat pada tabel 3 di bawah ini:

TABEL 3  
CONFUSION MATRIX 2x2

		Prediksi	
		LULUS	TIDAK LULUS
Aktual	LULUS	TP	FP
	TIDAK LULUS	FN	TN

Model persamaan yang digunakan untuk menghitung nilai akurasi dari proses pengujian data adalah:

$$Akurasi = \frac{\text{jumlah data yang diprediksi benar}}{\text{jumlah data yang diuji}} \times 100\% \quad (3)$$

## III. HASIL DAN PEMBAHASAN

### Implementasi Menggunakan Sistem

Sistem yang dibuat dengan menggunakan bahasa pemrograman php bertujuan untuk melakukan pengujian penerapan konsep data mining pada bahasa pemrograman php. Jumlah data yang digunakan sebanyak 60 data mahasiswa, dimana data latih berjumlah 40 data dengan nilai mahasiswa yang beragam dengan keterangan Lulus dan Tidak Lulus, dan 20 data sebagai data uji.

Tahapan dalam menjalankan sistem ini adalah dengan menjalankan aplikasi pada XAMPP sebagai control panel untuk menjalankan mysql dan apache. Kemudian membuka browser untuk mengaktifkan sistem melalui localhost.

### Halaman Data Latih

Pada halaman data latih sudah terdapat data nilai mahasiswa yang diinputkan ke dalam tabel data latih pada *database* yang ada. Data latih yang digunakan adalah sebanyak 40 data mahasiswa yang sudah memiliki nilai dan sudah memiliki keterangan LULUS atau TIDAK LULUS. Adapun data latih ditunjukkan pada gambar 3.

**DATA LATIH NILAI MAHASISWA**

Input Data Uji

No.	Nim	Nama Mahasiswa	Nilai Quiz	Nilai TKM	Nilai TKT	Nilai Praktek	Nilai UTS	Nilai UAS	Nilai Total	Kesimpulan
1	200705009	NABIL ANANDA	30	80	75	79	41	65	61	TIDAK LULUS
2	200705087	ORIZA SEPTIAN	70	0	0	0	20	65	34	TIDAK LULUS
3	210705001	INTAN WAHYUNI	70	100	95	96	90	75	85	LULUS
4	210705002	HAIFA DAYANA	70	100	95	96	20	65	68	LULUS
5	210705003	ARIEF FATHIN ABRAR	70	95	95	91	35	90	78	LULUS
6	210705007	FIYES FALEFI	70	95	95	88	30	65	68	LULUS
7	210705009	RIYASHA	50	100	95	93	50	65	71	LULUS
8	210705011	ZABRINA DEWI	40	80	95	85	25	65	62	TIDAK LULUS
9	210705012	SELIA PUTRI	30	95	95	88	25	65	63	TIDAK LULUS
10	210705014	ZURIATI FITRI	60	100	95	93	10	65	64	TIDAK LULUS

Gambar 3. Data Latih Nilai Mahasiswa

### Form Input Data Uji

Untuk memasukkan data uji maka di desain form untuk melakukan input data uji. Nanti, setelah selesai memasukkan data uji selanjutnya adalah menekan tombol Submit. Tampilan form data uji adalah seperti gambar 4 berikut ini:

**INPUT DATA UJI** ×

Nama Mahasiswa

Nilai Quiz

Nilai TKM

Nilai TKT

Nilai Praktek

Nilai UTS

Nilai UAS

Nilai Total

K

Gambar 4. Form Input Data Uji

### Hasil Pengujian

Tampilan data uji yang telah di input terlihat seperti gambar 5 berikut ini:

DATA UJI NILAI MAHASISWA								
Nama Mahasiswa	Nilai Quiz	Nilai TKM	Nilai TKT	Nilai Praktek	Nilai UTS	Nilai UAS	Nilai Total	k
SISNA ELLYANA	45	95	95	93	90	85	85	3

Gambar 5. Tampilan Data Uji

Setelah dilakukan perhitungan menggunakan metode KNN untuk mendapatkan nilai *Euclidean distance* dengan menentukan nilai  $k=3$ , maka didapatkan nilai jarak antara data latih dengan data uji. Hasil perhitungannya terlihat seperti gambar 6 berikut ini:

DATA HASIL HITUNG KNN									
Nama	Nilai_quiz	Nilai_tkm	Nilai_tkt	Nilai_praktek	Nilai_uts	Nilai_uas	Total	Kesimpulan	Jarak
NABIL ANANDA	30	80	75	79	41	65	61	TIDAK LULUS	60.42
ORIZA SEPTIAN	70	0	0	0	20	65	34	TIDAK LULUS	154.84
INTAN WAHYUNI	70	100	95	96	90	75	85	LULUS	27.39
HAIFA DAYANA	70	100	95	96	20	65	68	LULUS	77.14
ARIEF FATHIN ABRAR	70	95	95	91	35	90	78	LULUS	60.62
FIYES FALEFI	70	95	95	88	30	65	68	LULUS	68.01
RIYASHA	50	100	95	93	50	65	71	LULUS	45.28
ZABRINA DEWI	40	80	95	85	25	65	62	TIDAK LULUS	69.82
SELIA PUTRI	30	95	95	88	25	65	63	TIDAK LULUS	69.64
ZURIATI FITRI	60	100	95	93	10	65	64	TIDAK LULUS	83.96
ARMIS	60	100	95	89	15	65	65	TIDAK LULUS	79.21

Gambar 6. Data latih dengan jarak *Euclidean*

Setelah mendapatkan jarak *Euclidean* untuk setiap data latih, selanjutnya akan dilakukan proses pengurutan jarak dimulai dari jarak yang terkecil, artinya adalah dari hasil perhitungan KNN tersebut, akan mendapatkan data latih yang memiliki jarak terdekat dengan nilai data uji. Berikut ini adalah 11 data dari 40 data yang telah dilakukan proses pengurutan dari urutan yang terkecil sampai dengan urutan yang terbesar. Tampilannya dapat dilihat seperti gambar 7.

Kemudian dilakukan proses pengurutan dari urutan yang terkecil ke urutan yang terbesar, selanjutnya adalah mengambil 3 data berdasarkan tiga urutan jarak yang terdekat dari data latih dengan data uji. Pada proses pengurutan ini, dapat dikatakan sebagai proses untuk penentuan kelas terhadap data uji, sehingga 3 data yang memiliki jarak *Euclidean distance* terdekat dengan data uji akan menghasilkan prediksi LULUS untuk data uji tersebut. Hasil jarak terdekat ditunjukkan pada gambar 8.

PENGURUTAN EUCLIDEAN DISTANCE										
Nama	Nilai_quiz	Nilai_tkm	Nilai_tkt	Nilai_praktek	Nilai_uts	Nilai_uas	Total	Kesimpulan	Jarak	Urutan
INTAN WAHYUNI	70	100	95	96	90	75	85	LULUS	27.39	1
RANA SULTHANAH	70	100	95	97	95	95	93	LULUS	27.84	2
SRI RAHMI DAHLIA WULANDARI	40	95	95	89	70	65	73	LULUS	28.72	3
MUHAMMAD FIKRI NAUFALDI	40	80	80	86	56	75	70	LULUS	41.61	4
IRFANI ARIJA	60	100	95	89	56	65	73	LULUS	42.5	5
RIYASHA	50	100	95	93	50	65	71	LULUS	45.28	6
YUDHI JAUHARI	60	100	95	93	45	90	80	LULUS	47.96	7
ARJA DANIALA	45	75	85	70	45	65	62	TIDAK LULUS	54.08	8
TEUKU GHIFARI TS	50	100	95	92	35	85	75	LULUS	55.45	9
DIVA ERLANGGA PRAMUDYA ZUHDI	55	80	70	65	45	65	62	TIDAK LULUS	58.09	10
NABIL ANANDA	30	80	75	79	41	65	61	TIDAK LULUS	60.42	11

Gambar 7. Pengurutan jarak yang terkecil sampai terbesar

HASIL JARAK TERDEKAT										
Nama	Nilai_quiz	Nilai_tkm	Nilai_tkt	Nilai_praktek	Nilai_uts	Nilai_uas	Total	Kesimpulan	Jarak	Urutan
INTAN WAHYUNI	70	100	95	96	90	75	85	LULUS	27.39	1
RANA SULTHANAH	70	100	95	97	95	95	93	LULUS	27.84	2
SRI RAHMI DAHLIA WULANDARI	40	95	95	89	70	65	73	LULUS	28.72	3

Gambar 8. Urutan jarak *Euclidean distance* terdekat

Untuk hasil perhitungan data uji atas nama SISNA ELLYANA termasuk ke dalam kelas LULUS karena mayoritas tetangganya merupakan kelas LULUS. Terlihat seperti gambar 9 berikut ini:

HASIL PREDIKSI								
Nama	Nilai Quiz	Nilai TKM	Nilai TKT	Nilai Praktek	Nilai UTS	Nilai UAS	Nilai Total	Hasil Prediksi
SISNA ELLYANA	45	95	95	93	90	85	85	LULUS

Gambar 9. Hasil prediksi data uji

Selanjutnya dilakukan perhitungan untuk kelas TIDAK LULUS dengan data uji seperti gambar 10 berikut ini:

DATA UJI NILAI MAHASISWA								
Nama Mahasiswa	Nilai Quiz	Nilai TKM	Nilai TKT	Nilai Praktek	Nilai UTS	Nilai UAS	Nilai Total	k
M. ABDUL AZIZ	45	80	95	78	30	55	59	3

Gambar 10. Data uji atas nama mahasiswa M. Abdul Aziz dengan nilai  $k=3$

Untuk data hasil hitung KNN sama seperti gambar 6, selanjutnya akan ditampilkan hasil pengurutan jarak *Euclidean distance* mulai dari urutan yang terkecil sampai ke urutan yang terbesar. Hasilnya terlihat seperti gambar 11 berikut ini:

PENGURUTAN EUCLIDEAN DISTANCE										
Nama	Nilai_quiz	Nilai_tkm	Nilai_tkt	Nilai_praktek	Nilai_uts	Nilai_uas	Total	Kesimpulan	Jarak	Urutan
ZABRINA DEWI	40	80	95	85	25	65	62	TIDAK LULUS	12.25	1
MUHAMMAD FADHLAN	45	95	95	79	32	60	63	TIDAK LULUS	15.94	2
NOVI NURFARIZA	40	95	95	92	30	65	66	LULUS	18.71	3
DEWI AYU SAPUTRI	50	80	80	84	25	65	61	TIDAK LULUS	19.36	4
ARJA DANIALA	45	75	85	70	45	65	62	TIDAK LULUS	21.21	5
AIDIL GHUFRAN	60	95	95	91	25	65	66	LULUS	23.98	6
SELIA PUTRI	30	95	95	88	25	65	63	TIDAK LULUS	23.98	7
MUHAMMAD IKHSAN RAJABNA	60	95	95	89	16	65	64	TIDAK LULUS	27.31	8
ADHE MERA ULFA	40	100	95	91	15	65	63	TIDAK LULUS	27.39	9
IDHAM ANGGA SETIAWAN	60	100	95	90	35	65	69	LULUS	27.39	10
MUHAMMAD FERDIAN	60	100	95	87	25	65	66	LULUS	27.39	11

Gambar 11. Pengurutan jarak yang terkecil sampai terbesar (11 data latih dari 40 data latih)

Selanjutnya menentukan 3 jarak terdekat dengan data uji dari 40 data latih yang telah dilakukan proses perhitungan dengan metode KNN. Hasilnya dapat dilihat seperti gambar 12 berikut ini:

HASIL JARAK TERDEKAT										
Nama	Nilai_quiz	Nilai_tkm	Nilai_tkt	Nilai_praktek	Nilai_uts	Nilai_uas	Total	Kesimpulan	Jarak	Urutan
ZABRINA DEWI	40	80	95	85	25	65	62	TIDAK LULUS	12.25	1
MUHAMMAD FADHLAN	45	95	95	79	32	60	63	TIDAK LULUS	15.94	2
NOVI NURFARIZA	40	95	95	92	30	65	66	LULUS	18.71	3

Gambar 12. Urutan jarak *Euclidean distance* terdekat

Kemudian didapatkan hasil prediksi untuk data uji dengan data atas nama mahasiswa M. ABDUL AZIZ dengan hasil prediksinya adalah TIDAK LULUS, dikarenakan mayoritas tetangganya adalah kelas TIDAK LULUS. Hasil seperti gambar 13 berikut ini:

HASIL PREDIKSI								
Nama	Nilai Quiz	Nilai TKM	Nilai TKT	Nilai Praktek	Nilai UTS	Nilai UAS	Nilai Total	Hasil Prediksi
M. ABDUL AZIZ	45	80	95	78	30	55	59	TIDAK LULUS

Gambar 13. Hasil prediksi data uji atas nama M. ABDUL AZIZ

Berikut akan ditampilkan tabel hasil penentuan hasil prediksi data uji dengan data latih yang dilakukan secara manual (menggunakan excel) dengan proses perhitungan KNN dengan menggunakan sistem.

TABEL 4  
KLASIFIKASI HASIL DATA UJI (*DATA TESTING*) DENGAN  $K = 3$

No	Nama Mahasiswa	Quiz	TKT	TKM	Praktik	UTS	UAS	KET	Hasil Prediksi	
									Excel	Sistem
1	SISNA ELLYANA	45	95	95	93	90	85	LULUS	LULUS	SESUAI
2	ALBIRAN NISA	40	95	95	85	90	85	LULUS	LULUS	SESUAI
3	HAURA SALSABILA	30	100	95	98	90	55	LULUS	LULUS	SESUAI
4	CUT RAUDHATUL ILMU	45	95	95	96	90	55	LULUS	LULUS	SESUAI

	Nama Mahasiswa	Quiz	TKT	TKM	Praktik	UTS	UAS	KET		SESUAI / TIDAK SESUAI
								Excel	Sistem	
5	M. SYAUKI	25	100	95	85	40	55	TIDAK LULUS	LULUS	TIDAK SESUAI
6	MAULINATUL HAKIKI	45	100	95	99	100	55	LULUS	LULUS	SESUAI
7	PUTRI MAULAYA	45	97	95	98	90	55	LULUS	LULUS	SESUAI
8	DAFFAUL AISAR	50	100	95	98	50	55	LULUS	LULUS	SESUAI
9	M. HAYKAL. MZ	25	100	95	98	40	55	TIDAK LULUS	LULUS	TIDAK SESUAI
10	AMIRUL 'ADLY	40	97	95	96	40	55	TIDAK LULUS	LULUS	TIDAK SESUAI
11	M. ABDUL AZIZ	45	80	95	78	30	55	TIDAK LULUS	TIDAK LULUS	SESUAI
12	FAIZ RAMADHANA SURYA	40	85	95	95	40	85	LULUS	LULUS	SESUAI
13	HARIST ISLAMI RIDHA	40	95	95	97	40	55	TIDAK LULUS	LULUS	TIDAK SESUAI
14	T. AUDIANSYAH	30	95	95	94	60	85	LULUS	LULUS	SESUAI
15	MUHAMMAD RIAND	60	100	95	98	70	100	LULUS	LULUS	SESUAI
16	MUHAMMAD SAYYIDAL HAMDY INYU	50	95	95	95	50	55	LULUS	LULUS	SESUAI
17	MUHAMMAD NIZAM	25	95	95	90	24	70	TIDAK LULUS	TIDAK LULUS	SESUAI
18	M. MARTUNIS	25	97	95	93	56	55	LULUS	LULUS	SESUAI
19	ERA SYAFINA	35	95	95	92	90	100	LULUS	LULUS	SESUAI
20	AFIF MOELANA	25	90	95	87	60	100	LULUS	LULUS	SESUAI

Kemudian akan dihitung nilai akurasinya menggunakan model *confusion matrix* nya sebagai berikut:

		Prediksi	
		LULUS	TIDAK LULUS
Aktual	LULUS	14	4
	TIDAK LULUS	0	2

$$Akurasi = \frac{TP + TN}{TP + FP + FN + TN} \times 100\%$$

$$Akurasi = \frac{14 + 2}{14 + 4 + 0 + 2} \times 100\%$$

$$Akurasi = 0,8 \times 100\%$$

$$Akurasi = 80\%$$

Dengan menggunakan  $k = 3$ , didapatkan hasil prediksi kelulusan mahasiswa untuk mata kuliah sistem manajemen basis data dengan akurasi sebesar 80%. Dari tabel 4 di atas masih didapatkan 4 data yang masih tidak sesuai. Selanjutnya disusun tabel 5 untuk melihat data uji dengan menggunakan  $k=5$  adalah sebagai berikut:

TABEL 5  
KLASIFIKASI HASIL DATA UJI (*DATA TESTING*) DENGAN  $K = 5$

	Nama Mahasiswa	Quiz	TKT	TKM	Praktik	UTS	UAS	KET		SESUAI / TIDAK SESUAI
								Excel	Sistem	
1	SISNA ELLYANA	45	95	95	93	90	85	LULUS	LULUS	SESUAI
2	ALBIRAN NISA	40	95	95	85	90	85	LULUS	LULUS	SESUAI
3	HAURA SALSABILA	30	100	95	98	90	55	LULUS	LULUS	SESUAI
4	CUT RAUDHATUL ILMI	45	95	95	96	90	55	LULUS	LULUS	SESUAI
5	M. SYAUKI	25	100	95	85	40	55	TIDAK LULUS	TIDAK LULUS	SESUAI
6	MAULINATUL HAKIKI	45	100	95	99	100	55	LULUS	LULUS	SESUAI
7	PUTRI MAULAYA	45	97	95	98	90	55	LULUS	LULUS	SESUAI

	Nama Mahasiswa	Quiz	TKT	TKM	Praktik	UTS	UAS	KET		SESUAI/ TIDAK SESUAI
								Excel	Sistem	
8	DAFFAUL AISAR	50	100	95	98	50	55	LULUS	LULUS	SESUAI
9	M. HAYKAL. MZ	25	100	95	98	40	55	TIDAK LULUS	TIDAK LULUS	SESUAI
10	AMIRUL 'ADLY	40	97	95	96	40	55	TIDAK LULUS	LULUS	TIDAK SESUAI
11	M. ABDUL AZIZ	45	80	95	78	30	55	TIDAK LULUS	TIDAK LULUS	SESUAI
12	FAIZ RAMADHANA SURYA	40	85	95	95	40	85	LULUS	LULUS	SESUAI
13	HARIST ISLAMI RIDHA	40	95	95	97	40	55	TIDAK LULUS	LULUS	TIDAK SESUAI
14	T. AUDIANSYAH	30	95	95	94	60	85	LULUS	LULUS	SESUAI
15	MUHAMMAD RIAND	60	100	95	98	70	100	LULUS	LULUS	SESUAI
16	MUHAMMAD SAYYIDAL HAMDI INYU	50	95	95	95	50	55	LULUS	LULUS	SESUAI
17	MUHAMMAD NIZAM	25	95	95	90	24	70	TIDAK LULUS	TIDAK LULUS	SESUAI
18	M. MARTUNIS	25	97	95	93	56	55	LULUS	LULUS	SESUAI
19	ERA SYAFINA	35	95	95	92	90	100	LULUS	LULUS	SESUAI
20	AFIF MOELANA	25	90	95	87	60	100	LULUS	LULUS	SESUAI

Akan dihitung nilai akurasinya menggunakan model *confusion matrix* nya sebagai berikut:

		Prediksi	
		LULUS	TIDAK LULUS
Aktual	LULUS	14	2
	TIDAK LULUS	0	4

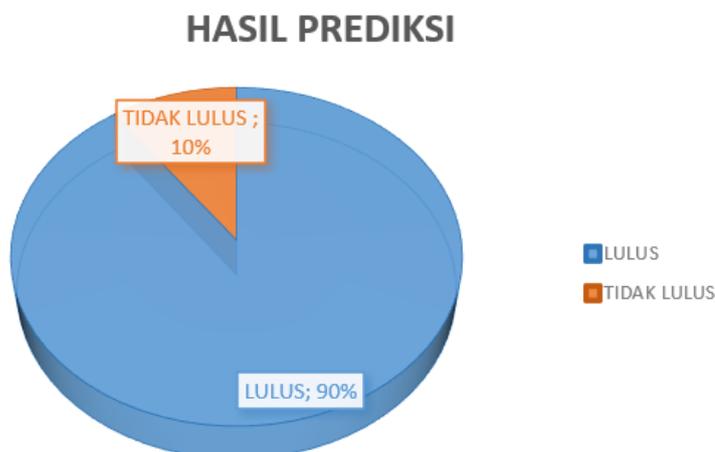
$$Akurasi = \frac{TP + TN}{TP + FP + FN + TN} \times 100\%$$

$$Akurasi = \frac{14 + 4}{14 + 2 + 0 + 4} \times 100\%$$

$$Akurasi = 0,9 \times 100\%$$

$$Akurasi = 90\%$$

Dengan menggunakan  $k = 5$ , didapatkan hasil prediksi kelulusan mahasiswa untuk mata kuliah sistem manajemen basis data dengan akurasi sebesar 90%.



Gambar 14. Hasil prediksi

Hasil prediksi yang dilakukan menggunakan metode KNN dari 20 data testing didapatkan hasil prediksinya adalah 18 SESUAI dan 2 TIDAK SESUAI, artinya adalah dengan menggunakan algoritma KNN hasil prediksi yang dilakukan secara manual dan dengan menggunakan Bahasa pemrograman PHP memiliki hasil yang sama, yaitu 18 mahasiswa lulus dan 2 mahasiswa tidak lulus. Dengan menggunakan jumlah data latihan sebanyak 40 data dan data training adalah 20 data testing maka algoritma KNN mempunyai akurasi 90%.

#### IV. KESIMPULAN

Dari hasil perancangan sebuah sistem sederhana berbasis pemrograman php, dapat diterapkan konsep data mining untuk perhitungan model KNN secara sederhana. Bahasa pemrograman web dapat dijadikan sebagai pilihan untuk menerapkan konsep data mining dengan membangun sistem yang lebih kompleks lagi. Proses pengujian akan lebih cepat dikarenakan data latihan (*data training*) telah dimasukkan ke dalam *database* sehingga memudahkan proses perhitungan. Dengan menggunakan jumlah data uji (*data testing*) sebanyak 20 data uji didapatkan akurasi prediksi menggunakan sistem sebesar 80% dengan menggunakan  $k=3$ , dan nilai akurasi sebesar 90% dengan menggunakan  $k=5$ . Dengan menggunakan nilai  $k$  yang bervariasi, akan menghasilkan perbedaan nilai akurasi yang mendekati nilai prediksi dengan nilai aktual. Semakin tinggi nilai  $k$  yang digunakan maka akan mendapatkan hasil prediksi yang lebih baik karena memiliki jumlah data yang besar. Maka untuk mendapatkan hasil prediksi dengan akurasi yang tinggi dan optimal sebaiknya menggunakan data dalam jumlah yang besar.

#### DAFTAR PUSTAKA

- [1] R. A. Manullang and F. A. Sianturi, "Penerapan Algoritma K-Nearest Neighbour Untuk Memprediksi Kelulusan Mahasiswa," *JIKOMSI (Jurnal Ilmu Komput. dan Sist. Informasi)*, vol. 4, no. 2, pp. 42–50, 2021.
- [2] D. Prasetyawan and R. Gatra, "Algoritma KNN Untuk Memprediksi Prestasi Mahasiswa Berdasarkan Latar Belakang Pendidikan.pdf," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 7, no. 1, pp. 56–67, 2022.
- [3] M. Mustika *et al.*, *Data Mining dan Aplikasinya*. Penerbit Widina Bhakti Persada, 2021.
- [4] A. D. A. Putra and S. Juanita, "Analisis Sentimen pada Ulasan pengguna Aplikasi Bibit Dan Bareksa dengan Algoritma KNN," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 8, no. 2, pp. 636–646, 2021, doi: 10.35957/jatisi.v8i2.962.
- [5] P. Subekti, T. D. Andini, and M. Islamiyah, "Sistem Penentuan Konsentrasi Jurusan Bagi Mahasiswa Informatika Menggunakan Metode K-Means Di Institut Asia Malang," *J. Manaj. Inform.*, vol. 12, no. 1, pp. 25–39, 2022, doi: 10.34010/jamika.v12i1.6452.
- [6] S. B. Rahardjo, W. W. A. Sulistyohati, and U. U., "Penerapan Data Mining Untuk Menganalisa Pola Pembelian Sayuran Hidroponik Menggunakan Metode Algoritma Apriori," *J. Pract. Comput. Sci.*, vol. 1, no. 2, pp. 38–49, 2022, doi: 10.37366/jpcs.v1i2.939.
- [7] R. Harun, K. C. Pelangi, and Y. Lasena, "Penerapan Data Mining Untuk Menentukan Potensi Hujan Harian Dengan Menggunakan Algoritma K Nearest Neighbor (KNN)," *MISI (Jurnal Manaj. Inform. Sist. Inf.)*, vol. 3, no. 1, pp. 8–15, 2020.
- [8] N. Wati, "Prediksi Kelulusan Mahasiswa Menggunakan K-Nearest Neighbor Berbasis Particle Swarm Optimization," *JTIJurnal Teknol. Inf. Indones.*, vol. 6, no. 2, pp. 118–127, 2021.
- [9] F. Marisa *et al.*, "Pengukuran Tingkat Kematangan Kopi Arabika Menggunakan Algoritma K-Nearest Neighbour," *JIMP J. Inform. Merdeka Pasuruan*, vol. 6, no. 3, pp. 4–8, 2021.
- [10] M. N. Maskuri, K. Sukerti, and R. M. Herdian Bhakti, "Penerapan Algoritma K-Nearest Neighbor (KNN) untuk Memprediksi Penyakit Stroke Stroke Disease Predict Using KNN Algorithm," *J. Ilm. Intech Inf. Technol. J. UMUS*, vol. 4, no. 1, pp. 130–140, 2022.
- [11] T. K. Janubiya, S. Andryana, and I. D. Sholihati, "E-Recruitment Menggunakan Metode Simple Additive Weighting dan Algoritma K-Nearest Neighbor," *J. Sains Komput. Inform.*, vol. 6, pp. 161–171, 2022.
- [12] S. Sukamto, Y. Adriyani, and R. Aulia, "Prediksi Kelompok UKT Mahasiswa Menggunakan Algoritma K-Nearest Neighbor," *JUITA J. Inform.*, vol. 8, no. 1, p. 121, 2020, doi: 10.30595/juita.v8i1.6267.
- [13] S. R. Raysyah, Veri Arinal, and Dadang Iskandar Mulyana, "Klasifikasi Tingkat Kematangan Buah Kopi Berdasarkan Deteksi Warna Menggunakan Metode Knn Dan Pca," *JSiI (Jurnal Sist. Informasi)*, vol. 8, no. 2, pp. 88–95, 2021, doi: 10.30656/jsii.v8i2.3638.
- [14] N. A. Widiastuti, M. Azhar, and H. Mulyo, "Implementasi Algoritma K-Nearest Neighbor untuk

- Klasifikasi Jurusan pada Peserta Didik Baru,” *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 14, no. 2, pp. 195–208, 2023, doi: 10.24176/simet.v14i2.10092.
- [15] A. Hernandez, S. Kurnia Gusti, F. Syafria, L. Handayani, and S. Ramadhani, “Klasifikasi Data Penerimaan Zakat dengan Algoritma K-Nearest Neighbor,” *Media Online*, vol. 4, no. 3, pp. 1632–1640, 2023, doi: 10.30865/klik.v4i3.1528.
- [16] A. Muzakir, “Perangkat Lunak Mobile Untuk Mendeteksi Daun Pada Tanaman Menggunakan Algoritma K-Nearest Neighbor (K-NN),” *J. Inf. Technol. Ampera*, vol. 2, no. 2, pp. 117–126, 2021, doi: 10.51519/journalita.volume2.issue2.year2021.page117-126.
- [17] Q. A’yuniyah and M. Reza, “Penerapan Algoritma K-Nearest Neighbor Untuk Klasifikasi Jurusan Siswa Di Sma Negeri 15 Pekanbaru,” *Indones. J. Inform. Res. Softw. Eng.*, vol. 3, no. 1, pp. 39–45, 2023, doi: 10.57152/ijirse.v3i1.484.
- [18] A. M. Argina, “Penerapan Metode Klasifikasi K-Nearest Neighbor pada Dataset Penderita Penyakit Diabetes,” *Indones. J. Data Sci.*, vol. 1, no. 2, pp. 29–33, 2020, doi: 10.33096/ijodas.v1i2.11.
- [19] S. R. Cholil, T. Handayani, R. Prathivi, and T. Ardianita, “Implementasi Algoritma Klasifikasi K-Nearest Neighbor (KNN) Untuk Klasifikasi Seleksi Penerima Beasiswa,” *IJCIT (Indonesian J. Comput. Inf. Technol.)*, vol. 6, no. 2, pp. 118–127, 2021, doi: 10.31294/ijcit.v6i2.10438.
- [20] M. Jannah and N. Humaira, “Implementasi Metode Euclidean Distance Untuk Ekstraksi Fitur Jarak Pada Citra Skeleton,” *J. Ilm. Inform. Komput.*, vol. 24, no. 2, pp. 134–139, 2019, doi: 10.35760/ik.2019.v24i2.2368.
- [21] R. M. Sagala, “Prediksi Kelulusan Mahasiswa Menggunakan Data Mining Algoritma K-Means,” *TeIka*, vol. 11, no. 2, pp. 131–142, 2021, doi: 10.36342/teika.v11i2.2610.
- [22] R. Musfikar, H. Apriadinata, and B. Yusuf, “Aplikasi Prediksi Prestasi pada Siswa Menggunakan Algoritma C4.5,” *J. Manaj. Inform.*, vol. 13, no. 2, pp. 148–162, 2023, doi: 10.34010/jamika.v13i2.10649.