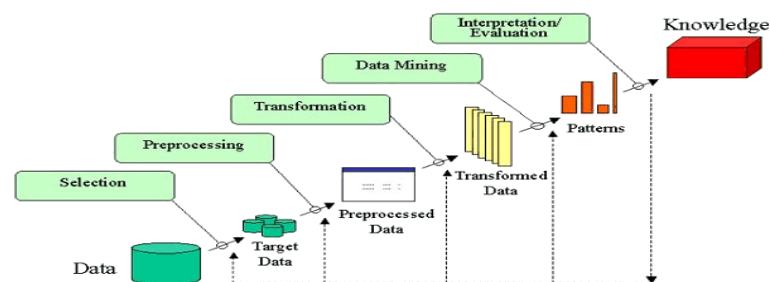


BAB II

LANDASAN TEORI

2.1 *Knowledge Discovery in Database (KDD)*

Knowledge Discovery in Databases (KDD) adalah suatu metode yang digunakan untuk mengekstrak pengetahuan dari database yang ada. Dalam database, terdapat tabel-tabel yang saling berhubungan dan berelasi. Pengetahuan yang dihasilkan dari proses ini dapat dimanfaatkan sebagai basis pengetahuan untuk mendukung pengambilan keputusan. Meskipun istilah KDD dan data mining sering digunakan secara bergantian untuk menggambarkan proses penggalian informasi tersembunyi dari basis data besar, sebenarnya keduanya memiliki konsep yang berbeda namun saling terkait. Salah satu tahapan dalam keseluruhan proses KDD adalah data mining itu sendiri. [1]. Terdapat beberapa tahapan di dalam proses KDD diantaranya yaitu selection, preprocessing, transformation, data mining, dan evaluation yang terdapat pada gambar 2.1



Gambar 2.1 Tahapan Proses KDD

Sumber : Srirahayu, Pribadie [2]

Berikut ini penjelasan setiap tahapan dalam proses KDD :

1. *Selection*

Sebelum tahap penggalian informasi dalam Knowledge Discovery in Database (KDD) dimulai, penting untuk melakukan pemilihan atau seleksi data dari sekumpulan data operasional. Data yang telah diseleksi ini akan digunakan dalam proses data mining dan disimpan dalam berkas terpisah dari basis data operasional.

2. *Pre-Processing*

Sebelum melaksanakan proses data mining, langkah awal yang penting adalah melakukan pembersihan data yang akan menjadi fokus dalam Knowledge Discovery in Database (KDD). Proses pembersihan ini mencakup beberapa tahapan, antara lain menghapus duplikasi data, memeriksa konsistensi data, serta memperbaiki kesalahan yang terdapat dalam data, seperti kesalahan pengetikan. Selain itu, proses pengayaan juga dilakukan, yang bertujuan untuk memperkaya data yang sudah ada dengan informasi lain yang relevan dan dibutuhkan dalam rangka Knowledge Discovery in Database (KDD), termasuk data atau informasi eksternal yang juga penting.

3. *Transformation*

Transformasi adalah proses mengubah data menjadi format yang sesuai untuk keperluan data mining. Tujuan dari tahap ini adalah untuk menyesuaikan data yang telah dikumpulkan, sehingga siap untuk diolah menggunakan algoritma dan perangkat lunak yang digunakan dalam pengolahan data.

4. *Data Mining*

Proses data mining adalah langkah yang melibatkan penggunaan metode atau algoritma tertentu untuk mengidentifikasi pola atau informasi menarik dari data yang telah dipilih. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan yang ingin dicapai serta keseluruhan proses KDD (Knowledge Discovery in Databases). Dengan melalui data mining, kita dapat memproses data untuk menghasilkan prediksi dan klasifikasi, membangun model analitis, serta menciptakan alat visualisasi yang efektif.

5. *Evaluation*

Evaluasi adalah proses yang mengubah pola-pola yang dihasilkan melalui data mining. Pola-pola ini harus disajikan dalam format yang jelas dan mudah dipahami oleh para pemangku kepentingan. Tahap ini merupakan bagian dari proses KDD dan melibatkan pemeriksaan untuk memastikan apakah pola yang dihasilkan sesuai dengan fakta atau asumsi yang telah ada sebelumnya. [3].

2.2 **Data Mining**

Data mining adalah kumpulan penemuan data besar yang berguna seperti matematika, metode statistik, dan kecerdasan buatan yang dapat memberikan pola dan wawasan berguna dari kumpulan data yang kompleks. Dalam penelitian ilmiah, data mining dikenal juga dengan istilah *Knowledge Discovery in Database* (KDD). Data mining muncul pada pertengahan tahun 1990an sebagai teknologi inovatif yang dapat menganalisis dan mengekstrak informasi berharga yang sebelumnya tidak diketahui dari kumpulan data yang sangat besar. Berbagai penelitian menyoroti bahwa teknik data mining dapat membantu pemilik data

menganalisis dan menemukan hubungan tak terduga antara data yang ada, sehingga dapat membantu dalam pengambilan keputusan [3].

Terdapat beragam definisi mengenai data mining yang diambil dari berbagai sumber, antara lain:

1. Data mining adalah proses pencarian dan analisis data dalam jumlah besar untuk menemukan pola-pola dan aturan yang signifikan.
2. Data mining merupakan teknik yang digunakan untuk menganalisis sekumpulan data besar, dengan tujuan menemukan hubungan yang tak terduga namun berguna bagi pemilik data.
3. Data mining adalah langkah untuk mengidentifikasi pola dan hubungan yang terdapat dalam suatu data.
4. Data mining adalah perangkat lunak yang dirancang untuk mengekstrak pola-pola tersembunyi dari database besar dan menghasilkan aturan-aturan yang dapat digunakan untuk memprediksi perilaku di masa depan.
5. Data mining adalah proses, baik otomatis maupun semi otomatis, yang bertujuan menemukan informasi baru dan berpotensi dari sekumpulan data.

Dari berbagai definisi tersebut, dapat disimpulkan bahwa data mining memiliki keterkaitan yang erat dengan penemuan informasi atau pengetahuan baru yang bersifat potensial dan tidak terduga dalam suatu database. [4].

2.3 *Naïve Bayes*

Metode *Naïve Bayes* merupakan teknik yang digunakan untuk menghitung nilai probabilitas dengan cara menjumlahkan frekuensi serta kombinasi nilai dari sebuah dataset. Dalam pendekatan ini, *Naïve Bayes* mengasumsikan bahwa atribut-atribut yang ada saling bebas secara kondisional, asalkan diberi nilai output tertentu. Salah satu keunggulan penggunaan *Naïve Bayes* adalah bahwa metode ini hanya memerlukan dua jenis data, yaitu data pelatihan (training set) dan data pengujian (testing set), untuk melakukan pengujian terhadap data yang ingin dianalisis. [5]. *Naïve bayes* adalah suatu teknik klasifikasi yang berakar pada Teorema Bayes. Pendekatan ini digunakan untuk memproyeksikan data di waktu mendatang berdasarkan informasi yang sudah ada atau yang telah terkumpul sebelumnya. Salah satu karakteristik utama dari metode *Naive bayes* ini adalah keyakinan yang kuat dalam independensi antara kondisi-kondisi yang berbeda [6]. Ciri *Utama Naïve bayes* adalah asumsi yang sangat kuat akan kebebasan dari masing-masing atribut. Algoritma ini mudah untuk digunakan dan sangat berguna dalam kondisi dimana input tinggi terhadap kelas tertentu sebagai positif atau negatif. Sehingga baik ada atau tidaknya suatu ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri kelas lain.

2.3.1 Persamaan Metode *Naïve Bayes*

Berikut adalah Persamaan *Teorema Bayes*:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Penjelasan dari persamaan rumus yaitu:

- $P(A|B)$ adalah probabilitas kejadian A terjadi, dengan syarat B telah terjadi (*posterior*).
- $P(B|A)$ adalah probabilitas kejadian B terjadi, dengan syarat A telah terjadi (*likelihood*).
- $P(A)$ adalah probabilitas kejadian A terjadi sebelum ada bukti B (*prior*).
- $P(B)$ adalah probabilitas kejadian B terjadi (*evidence*)

2.4 *Support Vector Machine (SVM)*

Support Vector Machine (SVM) adalah sebuah teknik yang digunakan untuk melakukan prediksi, baik dalam konteks klasifikasi maupun regresi. Prinsip dasar SVM berasal dari classifier linier, yang berarti bahwa ia dirancang untuk menangani kasus-kasus klasifikasi yang bisa dipisahkan secara linier. Namun, SVM telah mengalami pengembangan sehingga dapat digunakan untuk mengatasi masalah non-linier dengan memperkenalkan konsep kernel dalam ruang berdimensi tinggi. Dalam ruang dimensi tinggi ini, SVM akan mencari *hyperplane* yang mampu memaksimalkan jarak (margin) antara berbagai kelas data.

Metode SVM memiliki beberapa keunggulan, antara lain tingkat akurasi yang tinggi dan kemampuan yang baik saat dihadapkan pada data dengan dimensi yang tinggi. Dalam prosesnya, pengklasifikasi SVM hanya menggunakan sebagian kecil dari titik-titik pelatihan, sehingga membutuhkan memori yang relatif rendah. Namun, terdapat juga beberapa kelemahan dari metode ini. SVM

kurang cocok digunakan pada kumpulan data yang besar, dan pengklasifikasi ini seringkali kurang efektif dalam menangani klasifikasi dengan banyak kelas sentimen. SVM sendiri lebih baik digunakan hanya untuk pengklasifikasian kelas sentimen positif dan negatif [7].

2.4.1 Persamaan Metode *Support Vector Machine* (SVM)

Berikut adalah model persamaan *Support Vector Machine* (SVM) :

$$F(X) = W.X + B$$

W = parameter *hyperplane* yang dicari (garis yang tegak lurus antara garis *hyperplane* dan titik *support vector machine*)

X = titik data masukan *Support Vector Machine*

B = parameter *hyperplane* yang dicari (nilai bias)

F = fungsi *hyperplane*[8].

2.5 *Confusion Matrix*

Matriks kebingungan, atau confusion matrix, adalah sebuah tabel yang menggambarkan jumlah klasifikasi data uji yang berhasil dan yang gagal. Metode ini digunakan untuk mengevaluasi kinerja sistem klasifikasi. Pada dasarnya, matriks kebingungan menyajikan informasi yang membandingkan hasil klasifikasi yang dihasilkan oleh sistem dengan hasil klasifikasi yang diharapkan. Dalam pengukuran kinerja menggunakan matriks kebingungan, terdapat empat istilah

yang menjadi representasi hasil dari proses klasifikasi, yaitu *True Positive (TP)*, *True Negative (TN)*, *False Positive (FP)*, dan *False Negative (FN)*. Kondisi untuk melakukan pengujian klasifikasi menggunakan metode *Confusion Matrix*. Berikut ini pada tabel 2.1 adalah table dari *confusion matrix*:

Tabel 2.1 Tabel *Confusion Matrix*

		True Class	
Predicted Class		Class 1	Class 2
	Class 1	True Positive	False Negative
	Class 2	False Positive	True Negative

Keterangan:

TP (True Positive) : jumlah data positif yang berhasil diidentifikasi dengan akurat oleh sistem.

TN (True Negative): merujuk pada jumlah data negatif yang juga terklasifikasi dengan benar.

FN (False Negative): jumlah data positif yang salah terklasifikasi sebagai negatif oleh sistem.

FP (False Positive) : mencakup jumlah data negatif yang keliru diidentifikasi sebagai positif oleh sistem.

Nilai-Nilai yang dihasilkan dari kinerja klasifikasi sebuah algoritma yaitu:

1. *Accuracy*

untuk menghitung jumlah nilai rasio yang diprediksi benar dari keseluruhan data. Rumus :

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

2. *Precision*

Precision untuk menghitung jumlah nilai rasio yang diprediksi benar bernilai positif lalu dibandingkan dari keseluruhan hasil yang benar bernilai positif. Rumus :

$$Precision = \frac{TP}{TP + FP}$$

3. *Recall*

Recall untuk menghitung jumlah nilai rasio yang diprediksi benar bernilai positif lalu dibandingkan dari keseluruhan data yang benar bernilai positif. Rumus :

$$Recall = \frac{TP}{TP + FN}$$

4. *F1-Score*

F1-Score merupakan kombinasi antara *Precision* dan *Recall* yang berguna untuk mengukur kinerja model pengklasifikasian [9].

$$F1-Score = 2 \times \frac{(Recall \times Precision)}{(Recall + Precision)}$$

2.6 K-Fold Cross Validation

Cross Validation adalah teknik yang memverifikasi keakuratan model yang dibangun pada kumpulan data tertentu. Data yang digunakan dalam proses pembangunan model disebut data pelatihan, dan data yang digunakan untuk memvalidasi model disebut data pengujian. Pendekatan *K-Fold Cross Validation* membagi data menjadi k bagian kumpulan data dengan ukuran yang sama. Kemudian, serangkaian eksperimen dijalankan sebanyak k kali, dengan setiap eksperimen menggunakan data partisi ke-k sebagai data pengujian dan partisi sisanya sebagai data pelatihan. Pada percobaan pertama, subset S1 diperlakukan sebagai data uji dan subset lainnya diperlakukan sebagai data pelatihan[3].

2.7 Klasifikasi

Klasifikasi merupakan proses evaluasi dan penempatan objek data ke dalam kategori tertentu dari sekian banyak kategori yang tersedia. Proses ini membangun model berdasarkan data pelatihan yang ada, lalu menggunakan model tersebut untuk mengklasifikasikan data baru. Secara sederhana, klasifikasi dapat dipahami sebagai tugas yang menggunakan metode pembelajaran untuk memetakan setiap sekumpulan atribut (fitur) ke dalam kumpulan label kelas yang telah ditentukan. Idealnya, sebuah sistem klasifikasi diharapkan dapat mengklasifikasikan seluruh kumpulan data dengan akurasi tinggi. Namun, tidak dapat dipungkiri bahwa mencapai kinerja yang 100% benar adalah hal yang sulit, sehingga penting bagi sistem klasifikasi untuk secara berkala mengukur efektivitasnya sendiri.[10].

Klasifikasi merupakan teknik data mining yang menggunakan data latih berlabel kelas sebagai dasar pembelajaran memprediksi kelas suatu objek yang tidak diketahui label kelasnya. Ada 2 proses dalam klasifikasi yaitu proses training dan testing. Untuk melakukan pembangunan model digunakan data training dan melakukan tes menggunakan data testing. Model klasifikasi dapat membantu mengidentifikasi pola perilaku dan preferensi peserta. Algoritma klasifikasi berupaya membentuk serangkaian aturan, yang digunakan sebagai panduan untuk memprediksi kelas dari data prediksi [3].

2.8 Kesehatan Mental

Kesehatan mental adalah kondisi psikologis yang mencerminkan kemampuan seseorang untuk menyesuaikan diri dan menyelesaikan masalah, baik yang berasal dari dalam dirinya sendiri maupun yang muncul dari lingkungan sekitar. Kesehatan mental berkaitan dengan cara berpikir, merasakan, dan bertindak yang efektif dan efisien dalam menghadapi tantangan dan stres kehidupan. Menurut Organisasi Kesehatan Dunia (WHO), kesehatan mental adalah kondisi kesejahteraan yang disadari oleh individu, di mana terdapat kemampuan untuk mengelola stres kehidupan yang wajar, bekerja dengan produktif, dan berkontribusi dalam komunitas. Kesehatan mental yang baik menunjukkan bahwa fungsi-fungsi mental seseorang berjalan dengan optimal. [11]

Kesehatan mental merupakan kesehatan yang merujuk seluruh aspek perkembangan seseorang, baik fisik maupun psikis. Kesehatan mental juga meliputi upaya yang dilakukan untuk mengatasi stres, ketidakmampuan seseorang

dalam menyesuaikan diri, dan sikap seseorang dalam menjalani hubungan dengan orang lain, serta berkaitan erat dengan pengambilan keputusan. Dalam buku *Mental Hygiene* kesehatan mental berkaitan dengan beberapa hal.

1. Cara seseorang memikirkan, merasakan, dan menjalani keseharian dalam kehidupan.
2. Cara seseorang memandang diri sendiri dan orang lain
3. Cara seseorang mengevaluasi berbagai alternative solusi dan cara seseorang mengambil keputusan terhadap keadaan yang sedang terjadi.

Menurut Daradjat, kesehatan mental adalah keharmonisan dalam kehidupan yang terwujud melalui keseimbangan fungsi jiwa, kemampuan dalam menghadapi masalah, serta kemampuan untuk merasakan kebahagiaan dan menghargai diri sendiri secara positif. Ia menekankan bahwa kesehatan mental menggambarkan kondisi di mana individu bebas dari gejala gangguan jiwa (neurose) dan penyakit mental lainnya. [12].

Kesehatan mental yang baik membuat seseorang lebih tahan terhadap stresor, yaitu penyebab terjadinya stres. Individu dengan mental yang sehat mampu mengendalikan diri dari tekanan yang muncul baik dari dalam dirinya sendiri maupun dari lingkungan sekitarnya. Noto Soedirdjo menyatakan bahwa ciri-ciri orang yang memiliki kesehatan mental yang baik adalah kemampuan mereka untuk bertahan menghadapi tekanan yang berasal dari lingkungan. Di sisi lain, Clausen menjelaskan bahwa ketahanan seseorang terhadap stresor bervariasi, dipengaruhi oleh faktor genetik, proses belajar, budaya di sekitarnya,

serta intensitas stresor yang dialami, yang dapat berbeda antara satu orang dengan yang lainnya. [13].

2.8.1 *Depresi Anxiety Stress Scales*

Depresi Anxiety Stress Scales atau disebut DASS-21 merupakan bentuk singkat dari DASS-42, yaitu sebuah alat ukur yang digunakan untuk mengukur kondisi emosi dalam bentuk negative dari depresi, kecemasan dan stress seseorang. Skala alat ukur ini dapat digunakan baik dalam pengaturan klinis seperti rumah sakit atau non klinis seperti penelitian atau populasi. Menurut riset psikologi, masing-masing dari tiga skala DASS-21 memiliki 7 pernyataan yang telah dipilih untuk merepresentasikan skala versi panjangnya. Selain itu versi 21 item lebih layak digunakan dalam praktik yang sibuk dan juga kurang membebani responden. DASS-21 merupakan bentuk pendek dari DASS-42 yang sudah diterjemahkan ke dalam 45 bahasa berbeda dan salah satu instrument pengukuran yang banyak digunakan. DASS-21 sudah teruji validitas dan realibilitasnya di berbagai negara [3].

2.9 Mahasiswa Akhir

Mahasiswa tingkat akhir seringkali mengalami tekanan yang besar, terutama dalam menyelesaikan skripsi, yang merupakan karya tulis ilmiah hasil penelitian untuk memenuhi syarat kelulusan dari program studi di universitas. Salah satu permasalahan yang umum dihadapi mahasiswa adalah kesulitan dalam menemukan ide atau topik yang tepat serta mencari referensi yang relevan. Selain

itu, terdapat berbagai kendala lain, seperti kesulitan dalam menentukan sampel dan alat ukur, terbatasnya waktu untuk melakukan penelitian, revisi yang harus dilakukan berulang kali, serta kurangnya waktu bimbingan dari dosen pembimbing. Lamanya waktu respons dari dosen juga seringkali menjadi tantangan tersendiri. Permasalahan-permasalahan ini perlu diatasi dengan baik agar mahasiswa dapat terus maju dan menyelesaikan tugas akademis mereka. Mahasiswa yang mampu bangkit menghadapi tantangan dalam menyusun skripsi menunjukkan ketahanan yang kuat dalam mengatasi masalah. Pada umumnya, hambatan-hambatan ini dari awal hingga akhir proses penyusunan skripsi merupakan tuntutan yang dapat menyebabkan stres bagi mahasiswa. [9].

2.10 RapidMiner



Gambar 2.2 RapidMiner

Sumber : Afrahul, Dohardo,etc. 2024 [14].

RapidMiner merupakan perangkat lunak yang dibuat oleh Dr. Markus Hofman dari *Institute of Teknologi Blanchardstown* dan Ralf Klinkenberg (rapid-i.com). RapidMiner dirancang mudah untuk digunakan, bahkan tanpa pengalaman pemrograman [3]. RapidMiner merupakan platform perangkat lunak yang sangat kuat untuk bidang ilmu data dan pembelajaran mesin. Dengan menyediakan berbagai alat untuk persiapan data, pemodelan, evaluasi, dan implementasi,

RapidMiner dirancang agar mudah digunakan. Hal ini memungkinkan pengguna untuk dengan mudah membangun dan menguji berbagai model, bahkan bagi mereka yang tidak memiliki pengalaman dalam pemrograman. Salah satu fitur menarik dari RapidMiner adalah antarmuka drag-and-drop yang memudahkan pengguna dalam merancang alur kerja untuk memproses dan menganalisis data. Platform ini mendukung berbagai sumber data, termasuk file data datar, basis data, dan platform big data seperti Hadoop dan Spark. Selain itu, perangkat lunak ini dilengkapi dengan berbagai operator yang telah dibangun sebelumnya, yang berfungsi sebagai blok bangunan dalam alur kerja. Operator ini mencakup semua tahapan penting dalam proses data mining, mulai dari pembersihan data hingga pemilihan fitur dan pemodelan. [15].

2.11 Microsoft Excel



Gambar 2.3 Microsoft Excel

Sumber : Rizaldi, Dewi Anggraeni etc.2024 [16].

Microsoft office excel adalah *Software* lembar kerja *Sphreadsheet* yang dibuat dan didistribusikan oleh Microsoft Corporation yang dapat dijalankan pada Microsoft Windows dan Mac OS. sejak diluncurkan pertama kali pada tahun 1985, Excel telah menjadi salah satu alat yang paling banyak digunakan untuk

mengelola dan menganalisis data dalam berbagai bidang, termasuk keuangan, ilmu pengetahuan, pendidikan, dan bisnis. Excel memungkinkan penggunaannya untuk membuat tabel, menghitung angka, menganalisis data, serta membuat grafik dan diagram. Software ini memiliki banyak fitur yang membuat Microsoft Excel menjadi salah satu program komputer yang populer digunakan di seluruh dunia. Kelebihan Microsoft Excel diantaranya yaitu, *userfriendly*, tersedia fitur pembuatan grafik data, adanya fitur pemrograman VBA, dan kemampuan untuk menyimpan data dalam jumlah besar.

2.12 Metode Penelitian

Metode penelitian yang digunakan dalam penelitian ini meliputi :

2.12.1 Metode Pengumpulan Data

Data adalah instrument penting untuk melakukan suatu penelitian. Data pada suatu penelitian dapat berupa angka, teks, gambar, atau hal lain yang berhubungan dengan tujuan penelitian. Kualitas suatu data bergantung pada metode pengumpulannya, sehingga menentukan suatu metode yang tepat sangat berpengaruh untuk memastikan bahwa data yang dikumpulkan akurat dan relevan untuk penelitian yang dilakukan.

Terdapat dua jenis data pada penelitian, yaitu data primer dan data sekunder. Data primer adalah data yang didapatkan secara langsung dari objek penelitian atau merupakan informasi yang pertama kali dikumpulkan, sedangkan

data sekunder adalah data yang tidak diperoleh secara langsung atau merupakan data yang telah dikumpulkan oleh pihak lain sebelumnya.

2.12.2 Studi Pustaka

Studi pustaka adalah pendekatan yang mengumpulkan data dengan mengkombinasikan berbagai literature Data itu dapat diperoleh melalui berbagai sumber, seperti skripsi, buku dan jurnal penelitian yang sudah ada.

2.12.3 Observasi

Observasi adalah salah satu teknik yang digunakan untuk mengumpulkan data. Ini merupakan metode yang dilakukan secara sistematis dan selektif, di mana peneliti secara sadar memperhatikan dan mendengarkan interaksi atau fenomena yang terjadi.

2.12.4 Kuesioner

Kuesioner merupakan salah satu instrument pengumpulan data primer dengan metode survei yang fleksibel dan mudah digunakan untuk memperoleh opini responden. Kuesioner dapat didistribusikan kepada responden secara langsung oleh peneliti dapat dibuat dengan Google Formulir yang nantinya dapat disebar melalui sosial media seperti WhatsApp. Kuesioner dapat digunakan untuk memperoleh data pribadi responden misalnya sikap, opini, harapan dan keinginan responden.

2.13 Kerangka Kerja Penelitian

Kerangka kerja penelitian yang akan dilakukan disajikan pada tabel dibawah ini yaitu sebagai berikut :

Tabel 2.2 Kerangka Kerja Penelitian

No	Kegiatan	Nov	Des	Jan	Feb	Mar
1	Penentuan Topik Penelitian	✓				
2	Pendefenisian Masalah	✓	✓			
3	Menganalisa Masalah	✓	✓			
4	Menentukan Tujuan		✓			
5	Pengumpulan Data			✓	✓	
6	Cleaning Data			✓	✓	
7	Tranfsormasi Data			✓	✓	
8	Merancang Algoritma				✓	
9	Pengujian Algoritma				✓	
10	Evaluasi Akhir				✓	
11	Pengajuan Seminar Proposal				✓	
12	Seminar Proposal					✓

Catatan :

- Durasi setiap kegiatan dapat disesuaikan dengan kebutuhan spesifikasi penelitian.
- Waktu pelaksanaan bergantung pada kalender akademik dan kesiapan

2.14 Peneliti Terdahulu

Tabel 2.3 Penelitian Terdahulu

Judul	Nama Peneliti	Tahun	Hasil
Penerapan Metode <i>Naive Bayes</i> Untuk Memprediksi Tingkat Kesehatan Mental Siswa Menjelang Akhir Masa Sekolah	Nirwana, Tioria Pasaribu etc.	2024	Penelitian ini menggunakan metode <i>Naïve Bayes</i> untuk memprediksi tingkat kesehatan mental siswa menjelang akhir masa sekolah. Sistem yang dibangun mampu memprediksi secara efektif tanpa perlu konsultasi langsung dengan psikolog atau dokter, serta membantu memberikan solusi yang relevan berdasarkan tingkat kesehatan mental yang teridentifikasi.
Analisis Minat Masyarakat Menggunakan Media Sosial Menggunakan Algoritma C4.5 dan Metode <i>Naïve Bayes</i>	Panjaitan, Harahap etc.	2024	Penelitian ini menganalisis minat masyarakat menggunakan media sosial dengan menggunakan algoritma C4.5 dan metode <i>Naïve Bayes</i> . Hasilnya menunjukkan bahwa kedua metode mampu memberikan klasifikasi yang akurat dan efisien, menghasilkan wawasan mendalam mengenai pola minat masyarakat pada media sosial.
Perbandingan Algoritma <i>Naive</i>	Putra, Khairani etc.	2023	Algoritma <i>Support Vector Machine</i> mencapai

<i>Bayes Classifier</i> Dan <i>Support Vector Machine</i> Untuk Klasifikasi Data Kesehatan Mental Mahasiswa			akurasi 94,37% dalam klasifikasi kesehatan mental mahasiswa, sedangkan <i>Naive Bayes</i> mencapai akurasi 86,87%. <i>Support Vector Marchine</i> lebih optimal dibandingkan <i>Naive Bayes</i> .
Implementasi Data Mining untuk Memprediksi Kesehatan Mental Mahasiswa menggunakan Algoritma <i>Naïve Bayes</i>	Juwita, Rizky etc.	2023	Penelitian ini menggunakan algoritma <i>Naive Bayes</i> untuk memprediksi kesehatan mental mahasiswa dengan data dari dataset <i>Student Mental Health</i> (Kaggle). Hasil pengujian <i>Confusion Matrix</i> menunjukkan akurasi 93,33%, dengan <i>precision</i> 93,33%, dan <i>recall</i> 100%. Penelitian ini memberikan prediksi bahwa 28 mahasiswa mengalami gangguan kesehatan mental tanpa pengobatan.