BAB II

TINJAUAN PUSTAKA

2.1 Teori Dasar

2.1.1 *Platform Streaming*

Streaming adalah teknologi yang memungkinkan pengguna yang membutuhkan *player* untuk pengiriman data, *video*, atau *audio* dalam format terkompresi secara *realtime* melalui jaringan internet yang ditampilkan. Aplikasi *streaming* dicirikan oleh aliran data digital media secara simultan dan real-time dari server ke pengguna, serta distribusi *multimedia, audio*, dan *video* secara *real-time* atau sesuai permintaan melalui jaringan. Server menyediakan data yang diperlukan pada interval yang telah ditentukan, pengguna tidak perlu menunggu hingga seluruh data diunduh (Anjani et al., 2023). *Streaming* juga merupakan teknologi yang memungkinkan *file audio* atau *video* dijalankan langsung melalui jaringan lokal atau internet. Sistem membaca data dari *buffer* saat mengunduh *file*, sehinnga memungkinkan proses *streaming* berlanjut ke mesin klien (Nurindahsari & Zen, 2021).

Transfer data *audio* maupun *video* yang konstan dari *server* kepada klien dikenal sebagai *streaming*. *Streaming* mirip dengan menonton *TV, file video* di*transfer* oleh *TV byte* demi *byte*. Ini tidak sama dengan mengunduh *video* dan menontonnya sampai selesai mengunduh. Dengan streaming berjalan secara *real-time*, konten dapat diakses secara langsung tanpa harus mengunduh seluruh file, hal ini menjadikannya lebih efisien daripada proses pengunduhan. Sejumlah besar

RAM terbuang ketika komputer menyimpan semua data ke *hard drive* selama proses pengunduhan. Komputer tidak benar-benar menyalin dan menyimpan data ke memori saat streaming. Biasanya, streaming menggunakan protokol seperti *MPEG-DASH* atau *HLS* (Ananda & Nama, 2024).

2.1.2 Analisis Sentimen

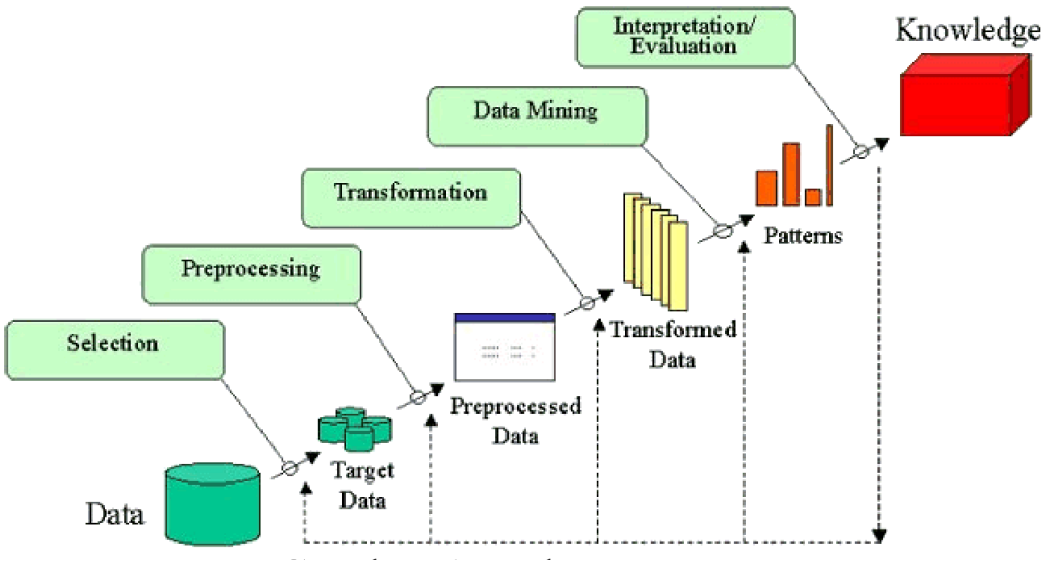
Studi dengan melibatkan penelitian serta analisis argumen, sentimen, evaluasi, perilaku, bahkan perasaan mengenai suatu substansi bentuk tulisan merupakan definisi analisis sentimen. Jasa dan layanan, individu, organisasi, masalah, subjek, atau produk dapat dianggap sebagai entitas. Proses mengidentifikasi informasi berupa sentimen positif, negatif, atau netral yang berasal dari data dikenal sebagai analisis sentimen. Pengguna internet memberikan analisis sentimen pada *platform* media sosial untuk mengekspresikan pendapat atau penilaian individu (Raisa Priskilla, 2024).

Pendekatan berbasis perhitungan yang dipakai dalam mendeteksi dan menginterpretasikan opini, emosi, tingkah seseorang pada objek atau peristiwa tertentu merupakan pengertian analisis sentimen (Arsadhana et al., 2025). Teknik untuk pengetahuan, perolehan, pemrosesan data guna memperoleh suatu sentimen dalam ulasan disebut dengan analisis sentimen (Supriadi, 2021).

2.1.3 *Knowledge Discovery In Database*

Knowledge Discovery in Databases (KDD) mengacu pada tahapan dan prosedur untuk mengekstrak data dari *database* yang ada (Alghifari & Juardi, 2021). *data selection, preprocessing*, *transformation, data mining* dan *evaluation* adalah beberapa langkah yang biasanya disertakan pada *Knowledge Discovery in*

Database (KDD) dalam analisis *datamining* (Takdirillah, 2020). Berikut merupakan tahapan *Knowledge Discovery in Database* (*KDD*):



Gambar 2.1 *Knowledge Discovery In Database*

1. *Data selection*

   Sebelum menuju langkah ekstraksi data *KDD*, serangkaian data operasional harus dilakukan. Berdasarkan data operasional, hasil dari proses pemilihan data ditampilkan pada satu halaman untuk digunakan dalam tahap *data mining*.

2. *Preprocessing*

   Data yang menjadi subjek *KDD* memerlukan teknik pembersihan sebelum beralih ke tahap berikutnya. Tahapan ini mencakup pembersihan data dari duplikasi, pemeriksaan terhadap ketidaksesuaian, serta koreksi terhadap kesalahan entri dan penulisan.

3. *Transformation*

   Transformasi data bertujuan untuk menyesuaikan format dan struktur data dengan langkah *datamining*. Jenis data atau model yang diekstraksi dari kumpulan data sangat memengaruhi proses kreatif transformasi.

4. *Data mining*

   Menemukan pola atau hasil yang menarik merupakan tujuan dari penerapan teknik pada langkah ini. Memilih algoritma yang tepat sangat dipengaruhi oleh teknik *datamining*, metodologi, dan algoritma serta keragaman, maksud, serta tahapan *KDD* secara menyeluruh.

5. *Interpretation* dan *Evaluation*

   temuan dapat lebih mudah dipahami oleh pihak yang berkepentingan berdasarkan model data yang diperoleh dari tahapan *data mining*. Tujuan bagian ini adalah untuk menentukan apakah informasi atau kebijakan yang diperoleh bertentangan dengan informasi yang sudah ada sebelumnya (Pebdika et al., 2023).

2.1.4 *Data mining*

Praktik mengekstraksi informasi serta data dalam jumlah masif yang tidak terindentifikasi sebelumnya. *Data mining* merupakan istilah yang merujuk pada proses untuk menemukan model yang sebelumnya tidak teridentifikasi dari data yang sudah diperoleh (Felicia Watratan et al., 2020). *Data mining* yaitu praktik penggunaan metode statistik, matematika, *artificial intelligence* (*ai*), *machine learning* guna mengambil dan menemukan pemahaman terkait serta informasi

berguna dari basis data besar (Laila Sari et al., 2022). Proses *data mining*, yang mencakup pencarian dan evaluasi sejumlah besar data untuk mengungkap informasi tersembunyi yang baru, sah, dan bermanfaat, mengarah pada penemuan pola atau rumus dalam data. (Pebdika et al., 2023).

2.1.5 Klasifikasi

Dalam data mining, klasifikasi digunakan sebagai metode untuk mengelompokkan data berdasarkan prediksi kelas yang belum ditentukan sebelumnya. Salah satu cara untuk menunjukkan bahwa objek data termasuk dalam salah satu jenis yang telah dideskripsikan sebelumnya adalah dengan klasifikasi (Alghifari & Juardi, 2021). Klasifikasi memegang peran penting dalam keseluruhan proses *data mining*. Tahapan mengkategorikan atau memberi label data baru berdasar kualitas tertentu dikenal sebagai klasifikasi. Memeriksa variabel yang berasal dari data yang sudah ada sebelumnya merupakan metode pendekatan klasifikasi (Pebdika et al., 2023). Klasifikasi adalah teknik pengelompokan data yang menggunakan algoritma klasifikasi untuk menganalisis data pelatihan. *k-Nearest Neighbor, Decision Tree, Naive Bayes*, dan *Support Vector Machine* adalah beberapa teknik pada klasifikasi (Khasanah et al., 2022).

2.1.6 Algoritma *Naïve Bayes*

Merupakan metode analisis dan statistik pengklasifikasi yang dapat memperkirakan kemungkinan bergabung dalam suatu kelas adalah *Naïve Bayes Classifier (NBC).* Menurut teori *Bayesian* yang menjadi dasar *NBC,* nilai atribut tidak berhubungan satu sama lain. *NBC* memiliki keunggulan karena mudah dipahami dan sangat akurat. *Naive Bayes* adalah metode peramalan probabilistik

yang mudah dipahami yang menggunakan estimasi independensi yang kuat sesuai dengan penerapan teorema *Bayes* (hukum *Bayes*). Manfaat pendekatan ini yaitu hanya membutuhkan data pelatihan seadanya guna menjamin parameter yang tepat di seluruh fase klasifikasi (Pebdika et al., 2023). Berikut merupakan persamaan dari teorema Bayes:

$$P(H|X) = \frac{\mathrm{P(X|H).P(H)}}{P(X)}$$

Rumus 2.1 *Teorema Bayes*

Penjelasan:

P(H|X) = Peluang posterior kelas C (target) diberikan fitur X (prediktor). Ini merupakan probabilitas yang ingin dihitung.

P(X|H) = Kemungkinan (likelihood), yaitu probabilitas fitur X berdasarkan kelas C.

P(H) = Probabilitas posterior kelas C, yaitu peluang kelas C terjadi tanpa memperhitungkan fitur.

P(X) = Konstanta bukti (*evidence*) atau normalisasi (*normalizing),* yaitu kemungkinan karakteristik X akan terjadi. Karena tujuan dari proses klasifikasi adalah untuk mengidentifikasi kelas dengan probabilitas posterior tertinggi, nilai ini sering diabaikan karena sama untuk semua kelas yang dibandingkan (Felicia Watratan et al., 2020)

Teorema Bayes dijelaskan dengan pengetahuan bahwasanya proses klasifikasi membutuhkan berbagai indikator dalam mengidentifikasi kelas mana

yang sesuai untuk sampel yang sedang dipelajari. Oleh karena itu, teorema Bayes sebelumnya di uraikan sebagai:

$$P(X \mid C) = P(x1 \mid C) \cdot P(x2 \mid C) \cdot \ldots \cdot P(xn \mid C)$$

Rumus 2.2 Penjabaran Teorema *Bayes* 1

di mana X=(x1,x2,…,xn) adalah vektor fitur.

$$P(C \mid x1, x2, \ldots, xn) \propto P(C) \cdot P(x1 \mid C) \cdot P(x2 \mid C) \cdot \ldots \cdot P(xn \mid C)$$

Rumus 2.3 Penjabaran Teorema *Bayes* 2

Menemukan kelas C yang memaksimalkan probabilitas posterior P(C|X) merupakan tujuan dari klasifikasi. Ketika kelas direpresentasikan oleh variabel C, dan kualitas petunjuk yang diperlukan untuk menyelesaikan klasifikasi direpresentasikan oleh variabel X1... Xn. Rumus ini menjelaskan bahwa kemungkinan suatu sampel tergolong dalam kelas C (posterior) bergantung pada probabilitas awal kelas C (prior) sebelum data sampel dianalisis, dikalikan dengan kemungkinan munculnya karakteristik sampel di kelas C, Kemudian dibagi dengan probabilitas kemunculan unsur sampel tersebut secara keseluruhan (*evidence*). Oleh karena itu, theorema ini juga bisa dinyatakan sebagai berikut (Putri & Wijayanto, 2022).

$$Posterior = \frac{\text{Prior} \times \text{likelihood}}{evidence}$$

Rumus 2.3 Menghitung *Posterior*

2.1.7 *Google Colab*

Google Colab is another name for Google's free Jupyter Cloud computing platform, which is used to write Python code and instructional materials in a web browser. It is also frequently used to teach machine learning. This eliminates the need for specialized hardware and makes it simple to train machine learning models that demand a lot of processing power. Google Colab allows you to share experiments online. Users may now create and share documents using narrative text, graphics, code, and mathematical formulae using Google Colab (Wulandari et al., 2024).

Kode Python dapat ditulis dan dijalankan di peramban menggunakan *Google Colab* merupakan *platform* berbasis *cloud*. Seperti analisis data, pendidikan, dan pembelajaran mesin banyak memanfaatkan *google colab* secara ekstensif (Ramdan Adi Surya & Hayati, 2024). Kode *Python* dapat diketik, dijalankan, dan dibagikan secara gratis menggunakan *Google Colab,* yang merupakan *platform* berbasis *cloud.* Layanan ini memfasilitasi pemrosesan data dan penelitian matematika, pembelajaran mesin, dan statistik dengan *GPU* dan *TPU*. Sebagian besar library *software* yang dibutuhkan telah disiapkan oleh *Google Colab*. Fitur perangkat keras yang ditawarkan oleh *Google Colab* meliputi *CPU, GPU, TPU, RAM,* media penyimpanan yang terintegrasi dengan *Google Drive*, dan dukungan tambahan *Matplotlib* dalam visualisasi grafik. Tersedia juga model *Python* 2.x dan 3.x serta *TensorFlow* 1.x dan 2.x (Nursinggah et al., 2024)

2.1.8 Python

One of the most widely used languages for data analytics, machine learning, and the Internet of Things (IoT) is Python, a multi-platform interpretive programming language that prioritizes code readability. The Python text data processing module offers a standardized API for evaluating natural language processing (NLP) tasks like sentiment analysis, noun extraction, and part of speech tagging. Python is an easy-to-learn programming language, which is why the researchers chose it. Additionally, Python offers a multitude of libraries that facilitate statistical analysis, machine learning, data analysis, and approximation (Afandi et al., 2022). *Python is a computer programming language, along with other languages like C, C++, Java, PHP, and others. Python is a computer language that differs from other programming languages in its lexicon, dialect, and set of rules* (Handayani, 2020)

Salah satu alat yang disarankan untuk hal tersebut adalah *Python. Python* telah menjadi pengganti yang ampuh untuk pekerjaan analisis data dalam beberapa tahun terakhir karena dukungan pustaka yang semakin meningkat, terutama untuk panda. *Python* adalah pilihan yang baik sebagai bahasa tunggal untuk membuat aplikasi yang berpusat pada data sentris karena kemampuan pemrogramannya yang bersifat umum. Perusahaan besar dan para pengembang (*developer*) memanfaatkan *Python* yang merupakan salah satu bahasa pemrograman, untuk membuat berbagai aplikasi *desktop*, daring, dan seluler. Pada tahun 1990, *van Rossum* mendirikan *Python* di Belanda, yang diberi nama berdasarkan program televisi kegemarannya, *Guido Monty Python's Flying Circus*. Setelah diciptakan oleh *Van Rossum* sebagai

hobi, *Python* memperoleh popularitas menjadi bahasa pemrograman baik di industri maupun akademis karena sederhana, kemudahan penggunaannya, keringkasan, sintaksis, dan pustaka yang luas (Qisthiano et al., 2021).

2.2 Penelitian Terdahulu

Algoritma Naive Bayes digunakan dalam penelitian sebelumnya untuk memperkirakan penyebaran COVID-19 di Indonesia. Klasifikasi permasalahan, tinjuan pustaka, pengambilan data, dan pendekatan metode adalah beberapa teknik penelitian yang digunakan. Temuan studi ini menyatakan Algoritma *Naive Bayes* dapat mengklasifikasikan data *COVID-19* tiap provinsi dengan tingkat akurasi 48.4848%. Meskipun demikian, penelitian ini menyarankan perlunya pengujian lebih lanjut dengan metode lain untuk membandingkan akurasi prediksi (Felicia Watratan et al., 2020).

Penelitian selanjutnya dengan tujuan untuk untuk mengukur sejauh mana mahasiswa merasa puas dengan pembelajaran *online* dan membantu memberikan saran untuk akademi membuat peraturan yang akan meningkatkan standar pembelajaran daring. Pada riset ini, pendekatan Naïve Bayes digunakan untuk menganalisis data kuesioner yang diberikan kepada 110 siswa. Menurut temuan penelitian, tingkat kepuasan siswa terhadap pendidikan daring dapat diprediksi dengan akurasi yang sangat tinggi (Damanik et al., 2021).

Riset berikutnya, *presentase* kelulusan mahasiswa Sekolah Tinggi Pariwisata (STP) Mataram diprediksi dengan *Naive Bayes* dan *K-NN*. Dua pendekatan tersebut dianggap sebagai metode pengklasifikasi dengan tingkat akurasi yang tinggi, sehingga dilakukan perbandingan. Hasil penelitian ini digunakan untuk membuat

aplikasi yang memprediksi kelulusan mahasiswa dalam kurun waktu yang ditetapkan atau melebihi batas waktu. Hasil riset menunjukkan bahwa akurasi pendekatan *K-NN* mencapai 96,18%, sedangkan akurasi pendekatan *Naive Bayes* dengan 91,94%. Berdasarkan hasil penelitian ini disimpulkan bahwasanya pendekatan *K-NN* cocok untuk memprediksikan kelulusan mahasiswa (Ketut Sriwinarti & luh Putu Juniarti, 2021).

Penelitian ini bertujuan untuk menggunakan penerapan *Naïve Bayes, KNN*, *Decision Tree* guna menguji bagaimana sentiment mahasiswa terhadap pembelajaran daring. Data Twitter yang dikumpulkan menggunakan proses *crawling*. Mencapai akurasi 61,92%, presisi sebesar 73,63%, dan *recall* sebanyak 11,42%, temuan studi menunjukkan bahwa pendekatan *Decision Tree* memiliki akurasi tertinggi dari metode pendekatan lainnya (Wiratama Putra & Triayudi, 2022).

Studi relevan lainnya mengklasifikasikan pengguna layanan Surat Keterangan Tidak Mampu (SKTM) di Rukun Tetangga 002, Kelurahan Meruya Selatan, memakai pendekatan *naïve Bayes*. Observasi, tinjauan pustaka, dan kuesioner merupakan beberapa teknik yang digunakan untuk mengumpulkan data dalam penelitian ini. menerapkan *naïve bayes* dalam mengklasifikasikan penerima bantuan Surat Keterangan Tidak Mampu (SKTM) Di Rukun Warga 002, Kelurahan Meruya Selatan. Kuesioner, studi pustaka, dan observasi adalah tahapan yang diterapkan dalam mengumpulkan data pada studi ini. Pengolahan data dan penilaian keakuratan dataset dilakukan dengan menggunakan *data mining* dan aplikasi *Rapidminer*. Hasil studi menunjukan, *Naive Bayes* memiliki akurasi sistem

sebanyak 62,86%, *recall* sebesar 78,57%, dan precision sebesar 52,38% dalam memprediksi penerima bantuan SKTM (Riyanah et al., 2021).

Penelitian terdahulu lainnya, membandingkan metode analisis sentiment, yaitu, *KNN, Decision Tree*, dan *Naïve Bayes* pada data *Twitter* yang berkaitan dengan layanan BPJS. Setelah menghilangkan data duplikat, 1000 tweet yang digunakan difilter menjadi 903 data. Berdasarkan hasil penelitian, pendekatan *Decision Tree* memiliki akurasi tertinggi (96,13%), diikuti oleh *KNN* (95,58%) dan *Naïve Bayes* (89,14%). *RapidMiner* digunakan dalam penelitian ini untuk mengumpulkan, memproses, dan menganalisis data (Puspita & Widodo, 2021).

Studi relevan lainnya, bertujuan menganalisis kelancaran pembayaran pinjaman dengan menerapkan data mining terhadap data anggota Koperasi Kredit Sejahtera. Karena kemampuannya dalam mengkategorikan data dan memprediksikan kemungkinan keanggotaan suatu kelas, dipilih algoritma *Naïve Bayes* sebagai metode penelitian ini. Dalam penelitian ini, menggunakan enam fitur aplikasi *WEKA* dan *RapidMiner* untuk analisis 1064 data pelatihan dan 300 data pengujian. Temuan penelitian menunjukkan bahwa data anggota Koperasi Kredit Sejahtera yang berkaitan dengan gagal bayar pinjaman dapat diklasifikasikan menggunakan algoritma *Naïve Bayes*, dengan tingkat akurasi pengujian sebesar 70,33% (Borman & Wati, 2020).

Studi selanjutnya, meneliti kinerja algoritma *Naïve Bayes Classifier* terhadap beragam tipe dataset yang tidak seimbang (*unbalanced dataset*). Dataset yang digunakan berisi data yang tidak seimbang dan diperoleh dari *Kaggle*. Dengan menggunakan akurasi, presisi, *recall*, dan *f-measure*, kinerja *Naïve Bayes Classifier*

dievaluasi. Temuan, menunjukkan pendekatan *Naïve Bayes Classifier* menghasilkan angka kinerja tidak menentu terhadap *unbalanced dataset* (Apriliyani & Salim, 2022).

Literatur yang sudah ada ini memiliki tujuan untuk meneliti ulasan pengguna *Netflix* pada *Google Play Store*. Untuk mengkategorikan sentimen dari 893 data ulasan, peneliti menggunakan pendekatan *Naïve Bayes*. Perangkat lunak *RapidMiner* digunakan untuk memproses data setelah dikumpulkan melalui *web scraping* menggunakan *Python*. hasil menyatakan bahwasanya akurasi *Naïve Bayes* sebesar 93,39% dalam mengidentifikasi sentimen ulasan (Bagas Pranata et al., 2024).

Studi terdahulu melakukan penelitian sentimen pengguna terhadap aplikasi *Vidio* dengan penerapan *naïve bayes*. Pada bulan Februari 2024, 1500 ulasan dikumpulkan melalui *Google Play Store* sebagai data yang diterapkan di riset ini. Terdapat 1475 komentar dengan 165 ulasan positif dan 1310 ulasan negatif setelah melakukan pembersihan data (*data cleaning*). Data diperoleh dari *Web scraping*, pelabelan data, prapemroroses, dan pembobotan *TF-IDF*, diikuti dengan analisis menggunakan *cross validation* dan *Naïve Bayes* adalah prosedur riset yang dipergunakan. Menurut temuan, akurasi algoritma *Naïve Bayes* adalah 80,28%, presisi 24,18%, dan *recall* 35,76% (Siregar et al., 2024).

Selanjutnya pada penelitian ini, analisis sikap konsumen terhadap UMKM yang berjualan di platform Tokopedia menjadi fokus utama penelitian ini. Kepuasan konsumen diukur menggunakan *Net Promoter Score* (*NPS*) dengan pendekatan algoritma *naïve bayes*. Berdasarkan data, sentimen penguna yang positif mencapai

75,69%, netral 11,08%, dan negatif sebesar 13,23%. Dengan presisi*, recall,* dan *F1 Score* sebanyak 80%, penerapan *Multinomial Naive Bayes* akurasinya 80%. Selanjutnya, skor *NPS* yang dihitung 62,46%, dapat disimpulkan bahwasanya sebagian besar konsumen memiliki tingkat kepuasan yang tinggi terhadap UMKM di Tokopedia (Arsadhana et al., 2025).

Pada penelitian selanjutnya, membandingkan dua teknik, *Support Vector Machine* (*SVM*), *Naive Bayes*, untuk menganalisis ulasan teks *Netflix*, guna menentukan antara dua metode dapat berkinerja lebih unggul dalam hal akurasi. Data, yang mencakup hingga 1000 ulasan, diambil dari *Google Play Store* dan dianalisis dengan *Python*. Data evaluasi aplikasi *Netflix* yang diperoleh kemudian dibagi antara 70% *data training* serta *data testing* 30%. Akurasinya sebesar 82% dicapai dengan menerapkan algoritma *Naive Bayes*, sedangkan 85% diperoleh dari *Support Vector Machine* (*SVM*). Hal ini menjadi bukti Support Vector Machine (SVM) jauh unggul dari *Naive Bayes*. (Khoirunnisaa et al., 2024).

Selanjutnya tujuan dari studi ini adalah menilai keakurasian pendekatan *naïve bayes* setelah optimasi dengan *genetic algorithm* dan *bagging*. *Naive Bayes* merupakan pendekatan yang diterapkan pada jenis data tidak seimbang. Dalam hal klasifikasi, *Naive Bayes* memiliki kinerja yang baik, meskipun demikian, optimasi ini perlu supaya memliki hasil klasifikasi yang unggul. Menurut temuan, kinerja *naïve bayes* dapat meningkat hingga 4,57% menggunakan penggabungan *genetic algorithm* dan *bagging* (Nugroho & Religia, 2021)

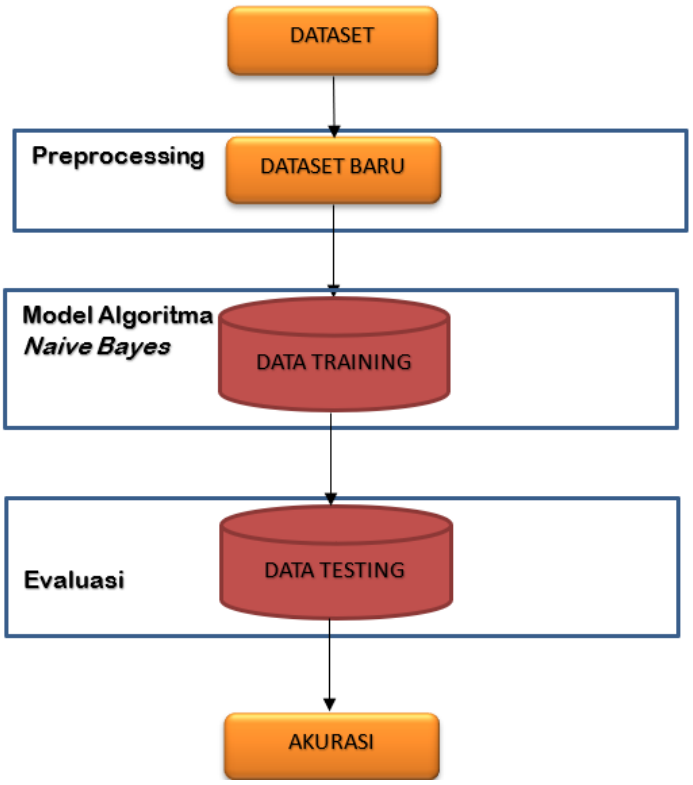
Penelitian relevan lainnya dengan pendekatan *Naive Bayes* untuk menelaah ulasan pengguna *TikTok* dan *Instagram* di *Google Play Store*. Menurut survei ini,

87,5% evaluasi pengguna tergolong memiliki sentimen negatif, membuktikan kekecewaan pengguna terhadap *platform* tersebut. Meskipun beberapa evaluasi seharusnya dikategorikan sebagai netral atau positif, model tersebut cenderung mendukung prediksi negatif (Zainuddin Siregar et al., 2025).

Penelitian selanjutnya untuk untuk membandingkan usia produktif ibu hamil dan memastikan persyaratan minimum dan maksimum bagi seorang wanita untuk hamil serta apakah bayi yang dilahirkan memenuhi persyaratan. Penerapan pengolahan datanya dengan metode *C4.5* dan *Naive Bayes*. Data kelahiran dari tahun 2017–2021, temuan akurasi selama lima tahun mempunyai tingkat akurasi 92,84% (Afandi et al., 2022)

2.3 Kerangka Pemikiran

Kerangka penelitian ini dikembangkan dengan menganalisis hubungan antara variabel kunci dan melakukan analisis menyeluruh terhadap literatur yang ada dalam upaya untuk memahami secara komprehensif fenomena yang diteliti. Berikut merupakan kerangka konseptual hasil dari proses ini:



Gambar 2.2 Kerangka Pemikiran

2.4 Hipotesis

Hipotesis studi yang dikembangkan berlandaskan uraian latar belakang permasalahan serta studi literatur yakni sebagai:

1. Diduga bahwa pendekatan *Naïve Bayes* dapat secara otomatis mengkategorikan pemikiran penonton *film* di *Netflix* dan *Disney+* menjadi sentimen, positif, negatif, dan netral.
2. Diduga pendekatan *Naïve Bayes* mengklasifikasikan sentimen komentar pengguna pada *platform streaming video* dengan akurasi, presisi, *recall*, dan skor-F1 yang cukup baik.