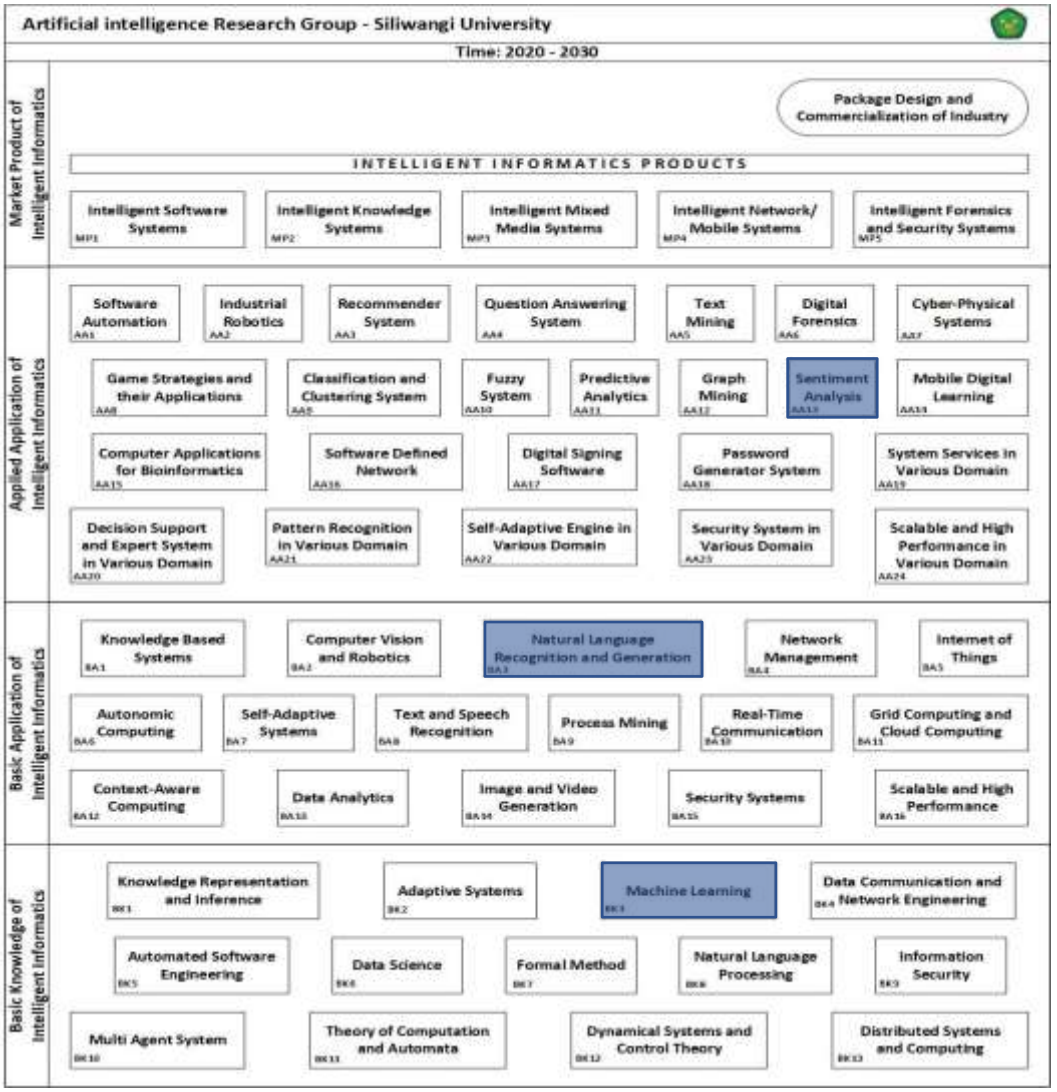
BAB III
METODOLOGI PENELITIAN

3.1 Peta Jalan Penelitian

Berikut merupakan topik penelitian sejalan dengan Peta Jalan Kelompok Keahlian Informatika dan Sistem Inteligen (ISI) seperti yang digambarkan di pada Gambar 3.1.



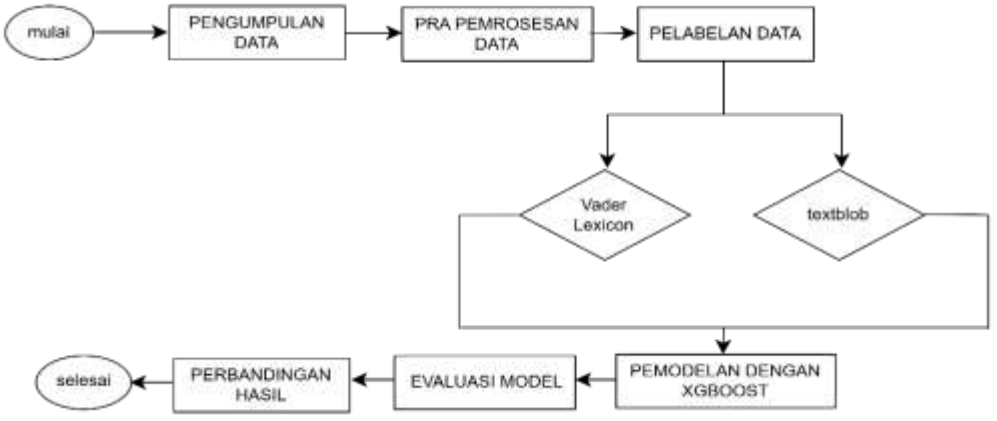
Gambar 3.1 Peta Jalan Penelitian

Gambar 3.1 Peta jalan penelitian menunjukkan keterkaitan antara

machine learning, *natural language processing* (NLP), dan sentiment analysis menjadi sangat penting. *Machine learning*, khususnya algoritma *XGBoost*, digunakan untuk membangun model prediktif yang dapat belajar dari pola data. Sementara itu, NLP berperan dalam memahami dan memproses ulasan teks dari aplikasi TikTok. Teknik pelabelan menggunakan *VADER Lexicon* dan *TextBlob* keduanya merupakan metode NLP dilakukan untuk mengekstraksi dan mengklasifikasikan sentimen dari teks ulasan. VADER *Lexicon* lebih cocok untuk data yang bersifat sosial media, sementara *TextBlob* menggunakan pendekatan berbasis *lexicon* untuk menilai polaritas sentimen. Dengan kombinasi ini, penelitian membandingkan efektivitas pelabelan otomatis untuk menghasilkan prediksi yang lebih akurat menggunakan *XGBoost* dalam analisis sentimen.

3.2 Tahapan Penelitian

Tahapan penelitian ini berdasarkan pada penelitian sebelumnya yaitu (Dava Maulana dkk., 2023), (Fathoni Naufal Fernanda dkk., 2024) yang telah dikembangkan menjadi seperti pada Gambar 3.2.



Gambar 3.2 Tahapan Penelitian

Berdasarkan tahapan penelitian pada Gambar 3.2 dijabarkan lebih rinci pada penjelasan dibawah ini:

3.2.1 Pengumpulan Data

Mengidentifikasi sumber data yang relevan. Dalam penelitian ini, sumber data yang dipilih adalah ulasan pengguna aplikasi TikTok yang tersedia di *Google Playstore* dari Maret 2023 sampai Oktober 2024. Alasan pemilihan platform ini adalah karena *Google Playstore* merupakan platform utama di mana pengguna aplikasi Android memberikan ulasan terkait pengalaman, keluhan, maupun kepuasan mereka terhadap aplikasi, yang relevan dengan fokus penelitian. Setelah memilih *Google Playstore* sebagai sumber data, langkah selanjutnya adalah memastikan bahwa ulasan yang diambil dapat mewakili berbagai jenis pengalaman pengguna. Proses ini mencakup mengidentifikasi aspek-aspek utama yang dibahas dalam ulasan tersebut, seperti performa aplikasi, fitur dan antarmuka pengguna (Sukirman dkk., 2024).

Setelah sumber data diidentifikasi, data ulasan diekstraksi menggunakan teknik web scraping, yaitu metode otomatis yang digunakan untuk mengumpulkan informasi dari situs web. Web scraping dilakukan dengan menggunakan program atau skrip yang mengekstraksi data ulasan secara sistematis dari halaman G*oogle Playstore*.

3.2.2 Pra-Pemrosesan Data

Data yang telah dikumpulkan sering kali mengandung noise seperti tanda baca, angka, dan karakter khusus yang tidak relevan. Oleh karena itu, langkah pra-pemrosesan dilakukan untuk membersihkan data. Proses ini meliputi:

a. *Case Folding*

Proses mengubah semua huruf dalam teks menjadi huruf kecil (*lowercase*) untuk menyamakan formatnya. Proses ini sangat penting untuk memastikan konsistensi data, terutama dalam analisis teks atau pencarian informasi, sehingga kata dengan huruf besar dan kecil dianggap sama, seperti "Data" dan "data" (Samrin & Nur Akbar, 2023).

b. *Filtering*

Proses membersihkan teks dengan menghapus elemen-elemen yang tidak relevan, seperti emoji, URL, mention (@username), hashtag (#tag), dan link. Selain itu, filtering juga mencakup penghapusan karakter byte, karakter selain huruf, tanda baca, serta menghilangkan spasi berlebih. Langkah ini bertujuan untuk menyederhanakan data sehingga lebih mudah diproses dalam analisis atau model teks (Azwarni & Shah, 2024).

c. Penghapusan *Stop Words*

Menghapus kata-kata umum yang tidak memberikan nilai signifikan pada analisis sentimen, seperti “dan”, “atau”, “tetapi”. Tujuannya agar analisis sentimen lebih akurat dan efisien engan menghapusnya, model dapat lebih fokus pada kata-kata yang benar-benar mempengaruhi sentimen suatu teks, sehingga hasil klasifikasi sentimen menjadi lebih relevan dan tepat (Thakur dkk., 2024).

d. *Stemming*

proses mengubah kata berimbuhan menjadi bentuk dasar atau akar katanya

dengan menghilangkan imbuhan seperti awalan, akhiran, atau sisipan. Teknik ini digunakan untuk menyederhanakan teks, sehingga kata-kata dengan makna serupa dapat diidentifikasi sebagai entitas yang sama. Misalnya, kata "makan", "memakan", dan "dimakan" akan direduksi menjadi "makan" (*K. M. Azharul Hasan, dkk*, 2021).

e. Normalisasi

Proses standarisasi teks dengan mengganti kata atau frasa tidak baku menjadi bentuk baku. Contohnya, mengganti "gk" menjadi "tidak" atau "tdk". Normalisasi bertujuan untuk meningkatkan konsistensi data. Agar datanya lebih bersih dan mudah diolah dalam analisis sentimen. Dengan normalisasi, variasi penulisan yang berbeda tetapi memiliki makna yang sama dapat disatukan, sehingga model dapat mengenali pola dengan lebih baik dan menghasilkan prediksi yang lebih akurat (Kurnia dkk., 2024).

f. Tokenisasi

Proses memecah teks menjadi unit-unit kecil. Token bisa berupa kata, frasa, atau kalimat. Tokenisasi digunakan untuk memisahkan teks menjadi komponen yang lebih kecil sehingga lebih mudah dianalisis atau diproses oleh algoritma (*K. M. Azharul Hasan, dkk*, 2021).

3.2.3 Pelabelan Data

Pelabelan data merupakan proses penting dalam pembelajaran mesin yang melibatkan pemberian label atau anotasi pada data mentah agar mesin dapat memahami dan belajar dari data tersebut. Data mentah bisa berupa teks, gambar, audio, atau video yang belum memiliki informasi tambahan yang dapat digunakan

oleh model pembelajaran mesin (Bose, 2020).

Proses pelabelan data biasanya dimulai dengan manusia yang menilai dan menambahkan label pada data. Misalnya, dalam pengolahan bahasa alami (NLP), teks dapat diberi label untuk menunjukkan sentimen (positif, negatif, netral) atau kategori topik tertentu. Dalam tahapan pemrograman, gambar dapat diberi label untuk mengidentifikasi objek yang ada di dalamnya, seperti mobil, pohon, atau hewan (Harris Pratama & Hayaty, 2023b). Pelabelan data sangat penting karena model pembelajaran mesin membutuhkan data berlabel untuk *training*. Data berlabel ini digunakan sebagai "*ground truth*" atau kebenaran dasar yang menjadi acuan bagi model untuk mempelajari pola dan membuat prediksi yang akurat. Keakuratan model sangat bergantung pada kualitas dan keakuratan label yang diberikan pada data pelatihan.

3.2.4 *VADER Lexicon*

VADER (*Valence Aware Dictionary and sEntiment Reasoner*) merupakan alat analisis sentimen berbasis leksikon yang dirancang khusus untuk media sosial. V*ADER* memberikan skor sentimen berdasarkan kata-kata yang ada dalam teks (Azwarni & Shah, 2024). Proses pelabelan setiap ulasan aplikasi TikTok dianalisis menggunakan *VADER* untuk mendapatkan skor sentimen. Skor ini berkisar antara -1 (sangat negatif) hingga +1 (sangat positif). Berdasarkan skor ini, ulasan dikategorikan sebagai positif, negatif, atau netral. Misalnya, skor di atas 0.05 dianggap positif, di bawah -0.05 dianggap negatif, dan antara -0.05 hingga 0.05 dianggap netral (Afrianto Singgalen, 2024).

3.2.5 Textblob

TextBlob merupakan pustaka Python untuk pemrosesan teks yang menyediakan alat sederhana untuk analisis sentimen. *TextBlob* menggunakan pendekatan berbasis leksikon dan aturan untuk menentukan sentimen teks. Proses Pelabelan Setiap ulasan aplikasi *TikTok* dianalisis menggunakan *TextBlob* untuk mendapatkan skor sentiment (Singh Shekhawat, 2019). *TextBlob* memberikan dua skor *polarity* (kecenderungan) dan *subjectivity* (subjektivitas). Skor *polarity* berkisar antara -1 (sangat negatif) hingga +1 (sangat positif). Berdasarkan skor *polarity*, ulasan dikategorikan sebagai positif, negatif, atau netral. Misalnya, skor di atas 0 dianggap positif, di bawah 0 dianggap negatif, dan 0 dianggap netral (Putri Meliani dkk., 2022b).

3.2.6 Pemodelan Dengan XGBoost

Pemodelan dengan *XGBoost* merupakan proses pembangunan model klasifikasi berbasis *gradient boosting* yang menggabungkan banyak *decision tree* untuk menghasilkan prediksi yang akurat. Setelah proses pelabelan sentimen menggunakan *TextBlob* dan *VADER Lexicon*, tahap selanjutnya adalah membangun model klasifikasi sentimen dengan algoritma XGBoost (Asri dkk., 2022).

Label sentimen yang semula berbentuk teks, yaitu netral, positif, dan negatif, dikonversi ke dalam bentuk numerik agar dapat diproses oleh algoritma pembelajaran mesin, dengan pemetaan netral sebagai 0, positif sebagai 1, dan negatif sebagai 2. Dataset kemudian dibagi menjadi data latih dan data uji dengan perbandingan 80:20. Setelah pembagian data, dilakukan tahap feature engineering menggunakan metode TF-IDF (*Term Frequency–Inverse Document Frequency*) untuk mengubah data teks

ulasan menjadi representasi numerik dalam bentuk vektor (Sukirman dkk., 2024). Proses TF-IDF dilakukan dengan mempelajari distribusi kata pada data latih untuk menghitung nilai *term frequency* dan *inverse document frequency*, kemudian data uji ditransformasikan menggunakan bobot yang sama agar berada pada ruang fitur yang identik. Tahapan ini bertujuan untuk mencegah kebocoran data dan memastikan proses pelatihan serta pengujian model berlangsung secara objektif. Matriks fitur hasil TF-IDF selanjutnya digunakan sebagai masukan dalam proses pelatihan model *XGBoost* (Natras dkk., 2022).

Pelatihan model dengan *XGBoost* melibatkan penggunaan data pelatihan untuk membangun model yang dapat memprediksi hasil dengan baik. *XGBoost* menggunakan teknik boosting, di mana model dibangun secara bertahap dengan menambahkan pohon keputusan baru yang memperbaiki kesalahan dari pohon sebelumnya . Proses ini berlanjut hingga jumlah pohon yang ditentukan tercapai atau hingga kesalahan tidak lagi berkurang secara signifikan. Selama pelatihan, model belajar dari kesalahan sebelumnya dan mencoba meminimalkan fungsi kerugian yang telah ditentukan (Thakur dkk., 2024).

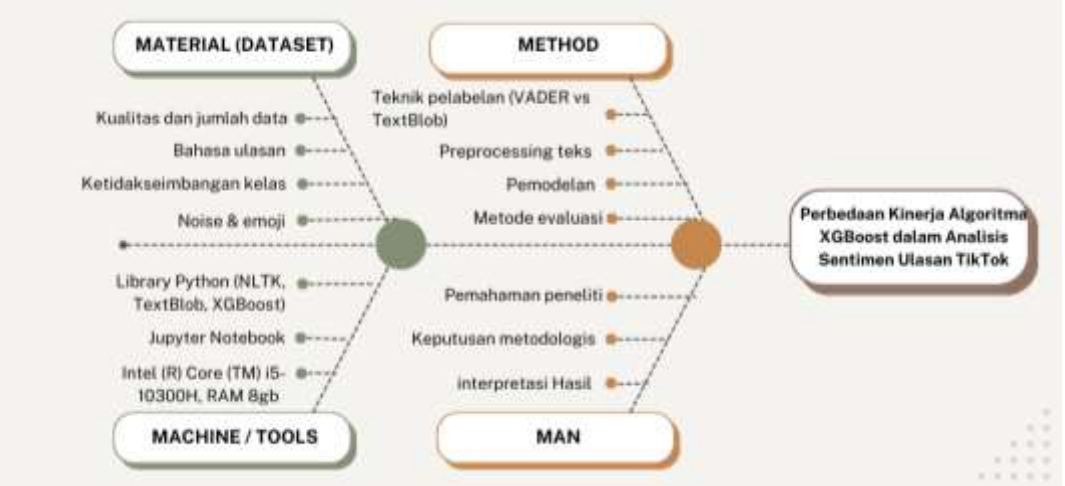
3.3 Evaluasi Model

Evaluasi model merupakan proses penting dalam pengembangan dan penerapan model pembelajaran mesin atau statistik. Tujuannya untuk menilai kinerja model berdasarkan data yang digunakan serta membandingkan hasil yang diperoleh dengan harapan atau standar tertentu. Metode evaluasi dapat bervariasi tergantung pada jenis model dan tujuan yang ingin dicapai, tetapi umumnya melibatkan penggunaan metrik-metrik tertentu untuk mengukur akurasi, presisi,

recall, F1-score, dan lainnya (Setiawan dkk., 2024).

3.4 Perbandingan Hasil

Membandingkan hasil yang diperoleh dengan model lain atau dengan standar yang telah ditetapkan. Perbandingan ini membantu untuk menentukan apakah model yang dikembangkan sudah cukup baik atau apakah perlu dilakukan penyesuaian lebih lanjut. Misalnya, jika model yang digunakan menghasilkan akurasi yang lebih rendah dibandingkan dengan model lain yang sudah ada, maka mungkin perlu dilakukan revisi atau perbaikan pada model tersebut.



Gambar 3. 3 Diagram Fishbone

Diagram Fishbone yang dibuat pada Gambar 3.3 merujuk pada (Holifahtus Sakdiyah dkk., 2022), menggambarkan apa saja yang membangun pada penelitian ini. Kepala ikan yaitu sebagai masalah utama pada penelitian, Kemudian pada bagian tulang penyebab utama yang memiliki keterangan sebagai berikut:

a. *Man*

Faktor yang berkaitan dengan kompetensi peneliti, termasuk pemahaman terhadap konsep analisis sentimen, ketelitian dalam proses *preprocessing* data,

ketepatan dalam menentukan parameter pada algoritma *XGBoost*, serta penetapan nilai ambang (*threshold*) sentimen yang digunakan.

b. *Method*

Mencakup penerapan teknik pelabelan berbasis *lexicon* seperti *VADER lexicon* dan *TextBlob*, proses pembagian dataset menjadi data pelatihan dan data pengujian, serta metode yang digunakan untuk mengevaluasi kinerja model.

c. *Machine / Tools*

Meliputi penggunaan algoritma *XGBoost* sebagai model klasifikasi, pemanfaatan pustaka pendukung seperti NLTK, *TextBlob*, dan *Scikit-learn*, serta perangkat keras yang digunakan selama proses pengolahan dan pemodelan data.

d. *Material*

Mencakup karakteristik dan kualitas data ulasan, perbedaan bahasa yang digunakan (Bahasa Indonesia maupun Bahasa Inggris), ketidakseimbangan distribusi kelas sentimen, serta keberadaan noise pada data seperti penggunaan emoji dan bahasa informal.