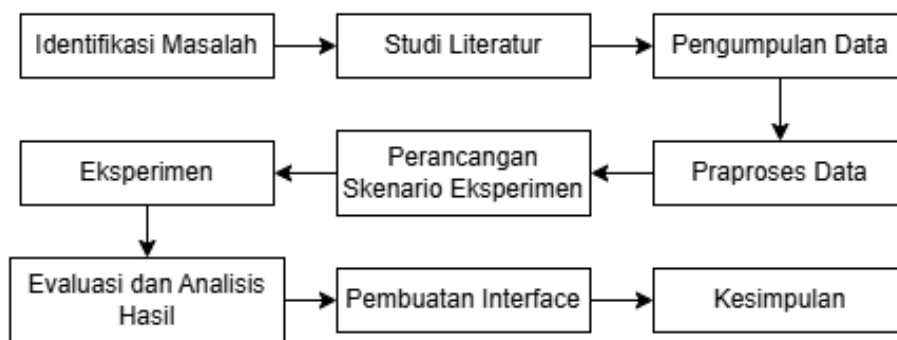


## BAB III METODOLOGI PENELITIAN

### 3.1 Desain Penelitian

Desain penelitian merupakan kerangka tahapan atau gambaran alur kerja yang akan dilaksanakan dalam pelaksanaan penelitian. Desain penelitian dibuat dengan tujuan untuk menyediakan struktur dan panduan sistematis sehingga memudahkan peneliti dalam pelaksanaan penelitian. Alur penelitian yang digunakan diilustrasikan dalam Gambar 3.1.



**Gambar 3.1** Diagram Alur Penelitian

Tahapan-tahapan di atas dijelaskan secara lebih rinci sebagai berikut.

#### 1. Identifikasi Masalah

Identifikasi masalah merupakan tahap awal yang menjadi fondasi dalam membangun arah penelitian. Proses ini diawali dengan mengamati fenomena atau kesenjangan antara kondisi ideal dan realita, lalu mengumpulkan serta menyaring isu-isu yang muncul untuk dipilih sebagai fokus utama. Selain itu, tahap ini juga membantu peneliti menentukan batasan penelitian, memprioritaskan masalah berdasarkan urgensi, serta merancang pertanyaan dan tujuan yang jelas. Masalah yang ingin dipecahkan dalam penelitian ini adalah bagaimana menerapkan metode *fine-tuning* model berbasis Transformer untuk analisis sentimen dan metode berbasis Transformer untuk *task* pemodelan topik pada pembahasan merek teknologi di media sosial X, sekaligus mengevaluasi kualitas dan hasil penerapannya untuk dilakukan analisis.

## 2. Studi Literatur

Studi literatur merupakan proses sistematis untuk mengumpulkan, menganalisis, dan memadukan sumber-sumber referensi yang relevan dengan topik penelitian. Tahap ini diawali dengan menetapkan fokus penelitian, lalu dilanjutkan dengan pencarian sumber literatur seperti buku, jurnal, artikel ilmiah, dan internet. Literatur yang dikumpulkan akan digunakan sebagai referensi untuk mendukung penyelesaian penelitian. Literatur yang dikumpulkan dan dipelajari mencakup topik mengenai *electronic word-of-mouth*, media sosial X, NLP, analisis sentimen, pemodelan topik, Transformer, BERT (mencakup proses *pre-training & fine-tuning*, model IndoBERT, dan BERTopic), metode evaluasi, *dataset*, dan penelitian-penelitian terdahulu.

## 3. Pengumpulan Data

Pengumpulan data adalah tahap mengumpulkan bahan yang akan digunakan dalam penelitian. Dalam penelitian ini, data dikumpulkan dari platform media sosial X dengan teknik *scraping*. Data yang dikumpulkan adalah *post* publik di media sosial X yang mengandung kata kunci "xiaomi" atau "xiaomiindonesia" serta memenuhi sejumlah kriteria tambahan, yaitu berbahasa Indonesia, tidak mengandung tautan pada isi teks *post*, dan hanya *post* yang mengandung konten teks. Kriteria tambahan tersebut digunakan sebagai upaya awal dalam mendapatkan data yang relevan dengan fokus penelitian. Seluruh proses ini dilakukan dengan tetap memperhatikan ketentuan privasi dari platform X dan prinsip etika penelitian, termasuk menjaga kerahasiaan identitas pengguna. Data pendukung berupa *dataset* yang relevan untuk pelatihan model klasifikasi sentimen juga dikumpulkan.

## 4. Praproses Data

Praproses data adalah tahap yang bertujuan untuk memeriksa, membersihkan, menstandarisasi, dan mempersiapkan data mentah agar siap digunakan untuk tahap eksperimen dengan model. Pada penelitian ini, tahap praproses yang dilakukan yaitu penanganan data duplikat, pemilihan data, pelabelan data, serta tahapan praproses teks seperti *case folding* (merubah huruf besar menjadi huruf kecil), *punctuation removal* (menghapus tanda baca dan karakter spesial lainnya),

*stopwords removal* (menghapus kata hubung yang dianggap tidak bermakna), dan *slang normalization* (pengubahan kata dalam bentuk tidak baku menjadi bentuk baku/formal).

## 5. Perancangan Skenario Eksperimen

Perancangan skenario eksperimen adalah tahapan yang meliputi pemilihan variabel data yang akan digunakan, penetapan metode dan teknik analisis, hingga penyusunan alur tahapan penelitian. Tahapan ini juga melakukan penentuan metrik-metrik evaluasi yang akan menjadi standar pengukuran keberhasilan dari eksperimen.

## 6. Eksperimen

### a. Eksperimen Pembangunan Model Klasifikasi Sentimen

Model *pre-trained* IndoBERT akan dilakukan *fine-tuning* menggunakan data latih dari data yang telah berlabel. *Hyperparameter* seperti *learning rate*, *batch size*, dan *epoch* akan dieksplorasi melalui eksperimen untuk mendapatkan konfigurasi terbaik. Model yang telah dilakukan *fine-tuning* kemudian dievaluasi pada data uji untuk mengukur performa model.

### b. Eksperimen Pemodelan Topik dengan BERTopic

Pada tahap eksperimen pemodelan topik, model BERTopic akan diterapkan pada data *post* yang telah melalui praproses. Tahapan ini diawali dengan mengubah setiap *post* menjadi representasi vektor numerik (*embeddings*). Representasi vektor ini kemudian diproses dengan algoritma reduksi dimensi, UMAP, untuk mengurangi kompleksitasnya. Selanjutnya, algoritma *clustering* HDBSCAN digunakan untuk mengelompokkan *post* yang dengan jarak vektor berdekatan sebagai satu *cluster*, di mana setiap *cluster* yang terbentuk dianggap sebagai satu topik. Terakhir, untuk mengekstrak makna dari setiap topik, kata-kata kunci yang paling representatif dihasilkan melalui mekanisme *c-TF-IDF*. Sebagai bagian dari eksperimen, eksplorasi konfigurasi model juga akan dilakukan dengan memvariasikan parameter untuk menentukan pengaturan

yang mampu menghasilkan topik-topik dengan kualitas yang dianggap interpretatif dan informatif.

## 7. Evaluasi dan Analisis Hasil

Evaluasi dilakukan untuk mengukur performa model dalam klasifikasi sentimen dan kualitas hasil pemodelan topik. Metrik evaluasi yang digunakan untuk model klasifikasi sentimen adalah *accuracy*, *precision*, *recall*, *F1-score*, dan kurva ROC. Untuk *task* pemodelan topik, metrik kuantitatif yang digunakan adalah *topic coherence score* ( $C_v$ ). Untuk *task* pemodelan topik, dilakukan juga evaluasi dengan pendekatan kualitatif dengan *human judgement* terhadap kata-kata kunci representatif tiap topik yang dihasilkan untuk memastikan relevansi dan kebermaknaannya.

Dilakukan juga analisis terhadap hasil eksperimen yang berfokus pada interpretasi dan pembahasan temuan dari eksperimen yang dilakukan. Analisis diawali dengan pembahasan mengenai performa *fine-tuning* model IndoBERT untuk klasifikasi sentimen dan dilanjut dengan analisis terhadap hasil pemodelan topik menggunakan BERTopic. Selanjutnya, hasil dari kedua eksperimen tersebut digabungkan untuk melihat sebaran sentimen di tiap topik.

## 8. Pembuatan *Interface*

Pada tahap ini berfokus pada pembuatan sebuah *interface* sederhana untuk mendemonstrasikan penggunaan model yang telah dibangun. Hasil analisis yang akan dilakukan yakni berupa visualisasi sebaran sentimen dan kelompok topik-topik dari *input* data yang diberikan.

## 9. Penarikan Kesimpulan

Tahap akhir ini berfokus pada merumuskan kesimpulan dari hasil analisis yang telah dilakukan. Kesimpulan mencakup temuan utama mengenai pola sentimen publik terhadap merek Xiaomi di media sosial X, topik-topik yang paling sering dibahas, serta evaluasi terhadap performa model *fine-tuning* IndoBERT dan BERTopic dalam konteks penelitian ini.

## 3.2 Alat dan Bahan Penelitian

### 3.2.1 Alat Penelitian

Dalam penelitian ini, peneliti menggunakan berbagai alat bantu yang mencakup perangkat keras dan perangkat lunak. Perangkat keras yang digunakan adalah sebuah laptop dengan spesifikasi sebagai berikut:

#### 1. Perangkat Keras

- Processor Intel Core i5-1035G1
- GPU NVIDIA GeForce MX330
- RAM 12 GB
- SSD 256 GB + HDD 1 TB

#### 2. Perangkat Lunak

- Sistem Operasi Microsoft Windows 11 Pro 64 bit
- Microsoft Excel
- Google Chrome Web Browser
- Google Colab (GPU Tesla T4U 16 GB)
- Visual Studio Code
- Bahasa Pemrograman Python

### 3.2.2 Bahan Penelitian

Bahan penelitian yang digunakan dalam penelitian ini yakni kumpulan *post* publik di media sosial X yang mengandung kata kunci “xiaomi” atau “xiaomiindonesia”, dalam rentang waktu unggahan antara bulan Mei 2023 hingga Mei 2025. *Post* yang digunakan merupakan *post* yang memenuhi kriteria tertentu ketika proses pengumpulan, seperti berbahasa Indonesia, tidak mengandung tautan (*link*), hanya terdiri dari konten teks, serta bukan merupakan *reply* atau *post* balasan.