



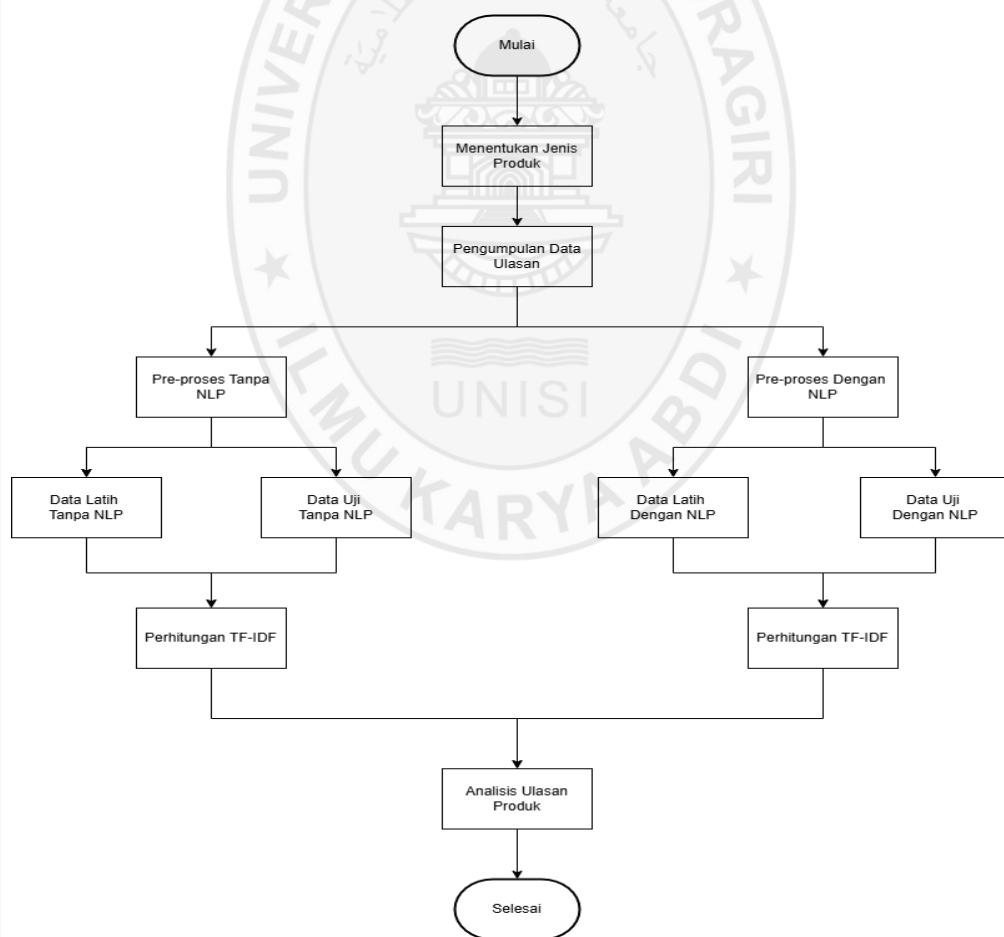
BAB III

METODE PENELITIAN

Pada penelitian ini menerapkan metode kuantitatif dengan pendekatan eksperimen komputasional untuk mendapatkan pemahaman yang lebih komprehensif dan mendalam tentang penelitian *Natural Language Processing (NLP)*.

3.1 Kerangka Penelitian

Tahapan penelitian mengenai implementasi natural language processing memiliki tahapan-tahapan yang dilakukan dengan tahapan penelitian sebagai berikut:



Gambar 3.1 Kerangka Penelitian [16]

- Hak Cipta Dilangi Undang-Undang**
1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang. Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.
 2. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia.
 3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.



3.1.1 Menentukan Jenis Produk

Penentuan jenis produk ini adalah untuk memberikan nilai tambah yang konkret bagi pengguna, baik itu perusahaan, peneliti, atau konsumen. produk yang dihasilkan diharapkan dapat membantu dalam pengambilan keputusan, pemahaman pasar, atau peningkatan kualitas produk. singkatnya, penentuan produk dalam penelitian ini adalah proses mengidentifikasi hasil akhir yang berguna dan relevan, yang dapat memberikan manfaat nyata bagi pengguna.

3.1.2 Pegumpulan Data Ulasan

Pengumpulan data ulasan adalah kita melihat secara mendalam bagaimana data ulasan produk dikumpulkan untuk kemudian dianalisis menggunakan teknik *Natural Language Processing*. proses pengumpulan data ini sangat krusial karena kualitas dan kuantitas data yang terkumpul akan sangat mempengaruhi akurasi hasil analisis sentimen, serta menentukan reliabilitas dan validitas temuan penelitian secara keseluruhan.

3.1.3 Pre-Proses Tanpa NLP

Pre-Proses dalam konteks umum, tanpa melibatkan *NLP (Natural Language Processing)*, merujuk pada serangkaian langkah yang dilakukan untuk mempersiapkan data mentah agar siap untuk dianalisis atau diproses lebih lanjut. Tujuan utama dari pre-proses adalah untuk membersihkan data, mengubah format data, menghilangkan inkonsistensi, dan memastikan data tersebut konsisten serta sesuai dengan kebutuhan analisis yang akan dilakukan dalam tahapan pemodelan, evaluasi, dan visualisasi hasil secara menyeluruh.

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang. Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.
2. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia.
3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.



3.1.4 Pre-Proses Dengan NLP

Pre-Proses dengan *NLP* adalah serangkaian langkah yang dilakukan untuk membersihkan dan mengubah data teks mentah menjadi format yang dapat dipahami dan diproses oleh komputer, khususnya dalam konteks *Natural Language Processing (NLP)*. jika pre-proses tanpa *NLP* lebih fokus pada pembersihan data secara umum, maka pre-proses dengan *NLP* memiliki tujuan yang lebih spesifik, yaitu mempersiapkan teks untuk tugas-tugas *NLP* seperti analisis sentimen, pengenalan entitas, atau penerjemahan mesin.

3.1.5 Data Latih Tanpa NLP

Data latih tanpa *NLP* mengacu pada kumpulan data yang digunakan untuk melatih model machine learning sebelum diterapkan teknik *Natural Language Processing (NLP)*. data ini belum melalui proses pembersihan, transformasi, atau pengolahan khusus yang biasanya dilakukan dalam *NLP*. dengan kata lain, data ini masih dalam bentuk mentah atau mendekati bentuk aslinya seperti yang didapatkan dari sumber data.

3.1.6 Data Uji Tanpa NLP

Data uji tanpa *NLP* berarti kita melakukan evaluasi terhadap suatu model atau sistem yang dirancang untuk memproses bahasa alami, namun tanpa melibatkan teknik-teknik khusus dari *Natural Language Processing*. dalam konteks ini, kita ingin mengetahui seberapa baik model atau sistem tersebut dapat bekerja pada data teks mentah, tanpa melalui tahap pembersihan, transformasi, atau pengolahan, sehingga terlihat perbandingan nyata antara data mentah dan data terolah.

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.
2. Penggunaan tanpa izin untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.
3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.



3.1.7 Data Latih Dengan NLP

Data latih dengan *NLP* adalah kumpulan data teks yang telah diproses secara khusus menggunakan teknik-teknik *Natural Language Processing (NLP)* agar siap digunakan untuk melatih model *machine learning*. proses pengolahan ini bertujuan untuk mengubah teks mentah menjadi representasi numerik yang dapat dipahami oleh komputer. dengan kata lain, data ini sudah "bersih" dan siap untuk digunakan dalam algoritma pembelajaran mesin.

3.1.8 Data Uji Dengan NLP

Data uji dengan *NLP* artinya kita mengevaluasi kinerja suatu model atau sistem *Natural Language Processing* pada data teks yang telah diproses secara khusus menggunakan teknik-teknik *NLP*. Setelah model dilatih dengan data pelatihan yang telah diproses, kita akan mengujinya pada data pengujian yang terpisah untuk melihat seberapa baik model tersebut dapat melakukan tugas yang telah ditentukan, serta memastikan konsistensi, akurasi, dan generalisasi hasil analisis pada berbagai jenis data.

3.1.9 Perhitungan TF-IDF

TF-IDF (Term Frequency-Inverse Document Frequency) adalah suatu metode statistik yang digunakan untuk mengukur seberapa penting sebuah kata dalam suatu dokumen tertentu di dalam sebuah kumpulan dokumen korpus. dengan kata lain, *TF-IDF* memberikan bobot atau skor kepada setiap kata dalam sebuah dokumen, di mana kata-kata yang lebih penting akan memiliki skor yang lebih tinggi dibandingkan kata-kata umum yang sering muncul dalam banyak dokumen tanpa memberikan informasi.

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang. Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.
2. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia.
3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.



Metode ini sangat berguna dalam berbagai aplikasi pengolahan bahasa alami, seperti analisis sentimen, pencarian informasi, sistem rekomendasi, dan klasifikasi teks, karena dapat menyoroti kata-kata kunci yang benar-benar memiliki kontribusi signifikan terhadap pemahaman isi dokumen.

3.1.10 Analisis Ulasan Produk

Analisis Ulasan Produk adalah suatu proses yang melibatkan pengumpulan, evaluasi, dan interpretasi informasi dari berbagai sumber yang relevan dengan topik tertentu. dalam analisis tinjauan, kita akan mengumpulkan data dari berbagai penelitian, artikel, laporan, atau sumber lainnya yang berkaitan dengan topik yang ingin kita teliti. kemudian, data-data tersebut akan kita evaluasi secara kritis untuk mengidentifikasi pola, tren, kesenjangan, dan kontribusi baru dalam bidang tersebut. tujuan utama dari analisis tinjauan adalah untuk mendapatkan pemahaman yang komprehensif dan mendalam tentang suatu topik, serta untuk mengidentifikasi arah penelitian di masa depan.

Hak Cipta Dilengkapi dengan Undang-Undang

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang. Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.
2. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia.
3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.