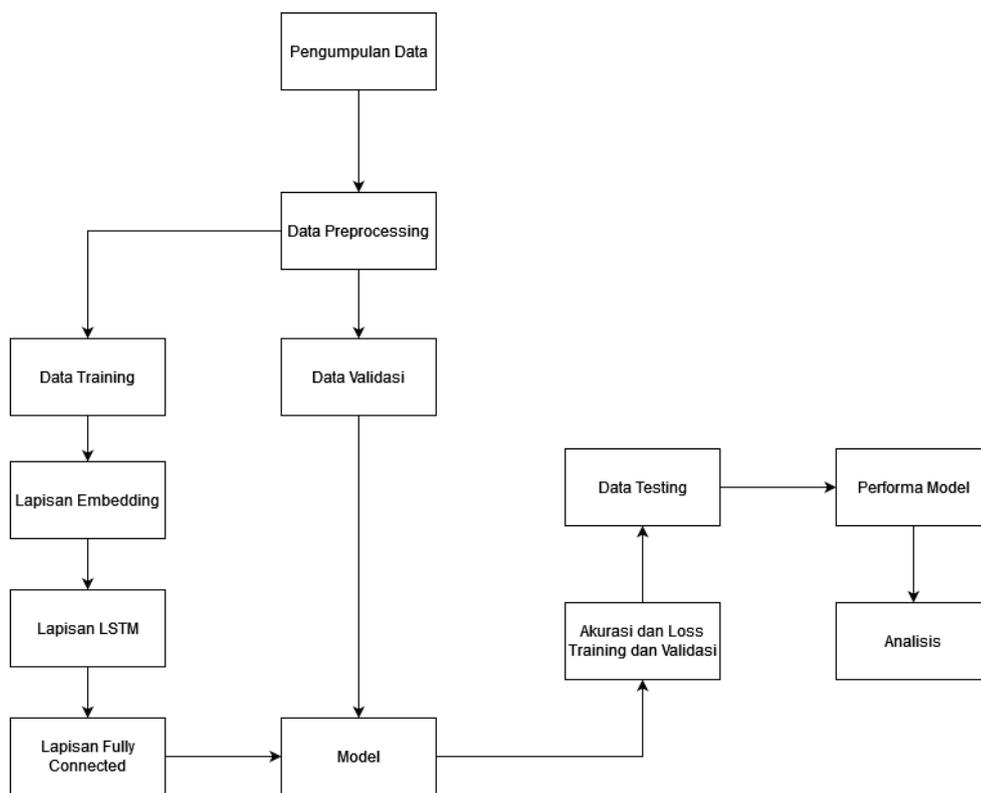


BAB III

METODOLOGI PENELITIAN

3.1 Metode Penelitian

Metode penelitian yang dilakukan secara garis besar diperlihatkan pada gambar 3.1 dan dijelaskan secara dalam sub bab 3.1.



Gambar 3.1 Alur Penelitian

3.1.1 Pengumpulan Data

Dalam penelitian ini, data yang digunakan adalah hasil pengisian kuesioner kepuasan mahasiswa Universitas Siliwangi pada bulan Januari tahun 2020. Setiap data akan diberikan label sesuai dengan sentimen yang terdapat pada data

tersebut. Label bernilai 0 untuk sentimen negatif dan 1 untuk sentimen positif. Label diberikan dengan cara manual, yang disesuaikan dengan kandungan pada setiap data.

3.1.2 *Data Preprocessing*

Data yang telah terkumpul tidak dapat langsung diproses oleh model *machine learning*, tetapi harus dilakukan pengolahan terlebih dahulu. Perlu dilakukan pemrosesan sebelumnya untuk pengurangan dimensi pada datanya sehingga meningkatkan tingkat akurasi pada klasifikasinya.

Pada penelitian ini, tahapan *data preprocessing* yang dilalui adalah sebagai berikut:

1. *Data Cleaning*, dilakukan untuk menghapus tanda baca, *hyperlink*, tanda baca.
2. *Stopword Removal*, dilakukan untuk menghapus kata-kata yang terdapat pada *stoplist*. *Stoplist* yang digunakan pada penelitian ini diambil dari *Python NLTK*.
3. *Case Folding*, dilakukan untuk mengubah setiap karakter dalam kalimat menjadi huruf kecil.
4. *Stemming*, dilakukan untuk mengubah kata ke bentuk dasarnya.
5. *Tokenizing*, dilakukan untuk membagi kalimat menjadi beberapa bagian, bisa dibagi per kata ataupun bagian lain yang dibutuhkan oleh model *Machine Learning*.

3.1.3 Membuat dan Melatih Model

Penelitian ini menggunakan algoritma *Long-Short Term Memory* (LSTM). Algoritma LSTM dipilih karena merupakan salah satu pengembangan dari algoritma RNN, yang mana algoritma RNN mampu menangani data yang bersifat sekuensial dengan baik (Rozi et al., 2020), sedangkan algoritma LSTM dapat menangani data yang bersifat sekuensial lebih baik dari RNN karena kemampuannya dalam menangani data masukan sebelumnya lebih banyak (Patel et al., 2021).

Pada penelitian ini, model menggunakan *embedding layer* sebagai *input layer* yang kemudian data akan diproses menggunakan *word embedding*.

3.1.4 Analisis

Performa dari setiap model kemudian akan dilakukan evaluasi dengan menggunakan metode *confusion matrix*, dengan menghitung jumlah *True Positive*, *True Negative*, *Fale Positive*, dan *False Negative*.

3.2 Alat dan Bahan Penelitian

3.2.1 Alat Penelitian

Dalam penelitian ini, digunakan perangkat keras komputer dengan spesifikasi sebagai berikut:

1. Processor Intel Core-i3 2348M
2. Memori 4 GB
3. *Harddisk* berkapasitas 500 GB

Adapun perangkat lunak yang digunakan adalah sebagai berikut:

1. Ubuntu Linux
2. Jupyter Notebook
3. Python 3.

3.2.2 Bahan Penelitian

Penelitian ini menggunakan data yang diambil dari hasil pengisian kuesioner Mahasiswa Universitas Siliwangi di Universitas Siliwangi pada bulan Januari tahun 2020. Bahasa yang digunakan adalah bahasa Indonesia dengan sebagian kuesioner terdapat bahasa yang tidak baku atau terdapat kesalahan penulisan kata.