

BAB III

METODOLOGI PENELITIAN

3.1. Bahan Penelitian

Pada Penelitian ini, terdapat sebuah dataset yang berisi data kalimat pengujian dan pelatihan untuk membangun model *Text Classification* untuk menemukan sentimen kalimat bahasa Indonesia. Struktur dataset ini terdiri dari kalimat yang sudah diberi label positif, negatif, dan netral. Penulis mengambil dataset ini dari berbagai sumber. Seperti ulasan aplikasi, *movie reviews*, komentar-komentar netizen di Facebook, Youtube, Instagram, dan Twitter. Dataset berjumlah 2550 kalimat yang terdiri dari 2100 untuk pelatihan dan 450 untuk pengujian.

3.2. Alat Penelitian

Adapun alat penelitian yang mendukung penelitian ini terdiri dari dua komponen, yaitu perangkat lunak dan perangkat keras. Berikut spesifikasi perangkat lunak dan perangkat keras yang digunakan.

3.3.1. Perangkat Lunak

a. Nama : Ubuntu 16.04 LTS Desktop

Sumber : Canonical Ltd

Sebagai sistem operasi yang digunakan untuk meneliti, mengembangkan dan menjalankan pelatihan dan pengujian sentimen kalimat bahasa Indonesia

b. Nama : Python IDLE Version 2.7.14

Sumber : Python

Sebagai IDE (Integrated Development Environment) untuk membantu mengembangkan code dalam melakukan pengujian dan pelatihan dataset.

c. Nama : Library Spacy 2.x

Sumber : Spacy.io

Sebagai library dalam membuat model

3.3.2. Perangkat Keras

Perangkat keras yang digunakan untuk mendukung penelitian ini adalah sebagai berikut.

a. PC (Personal Computer) dengan spesifikasi:

CPU : Intel Core i7-370K CPU @3.5 GHz

RAM : 16384 MB

HDD : 2TB

OS : Ubuntu Linux 16.04 (Xenial Xerus) 64bit

b. Laptop Sony Vaio dengan spesifikasi

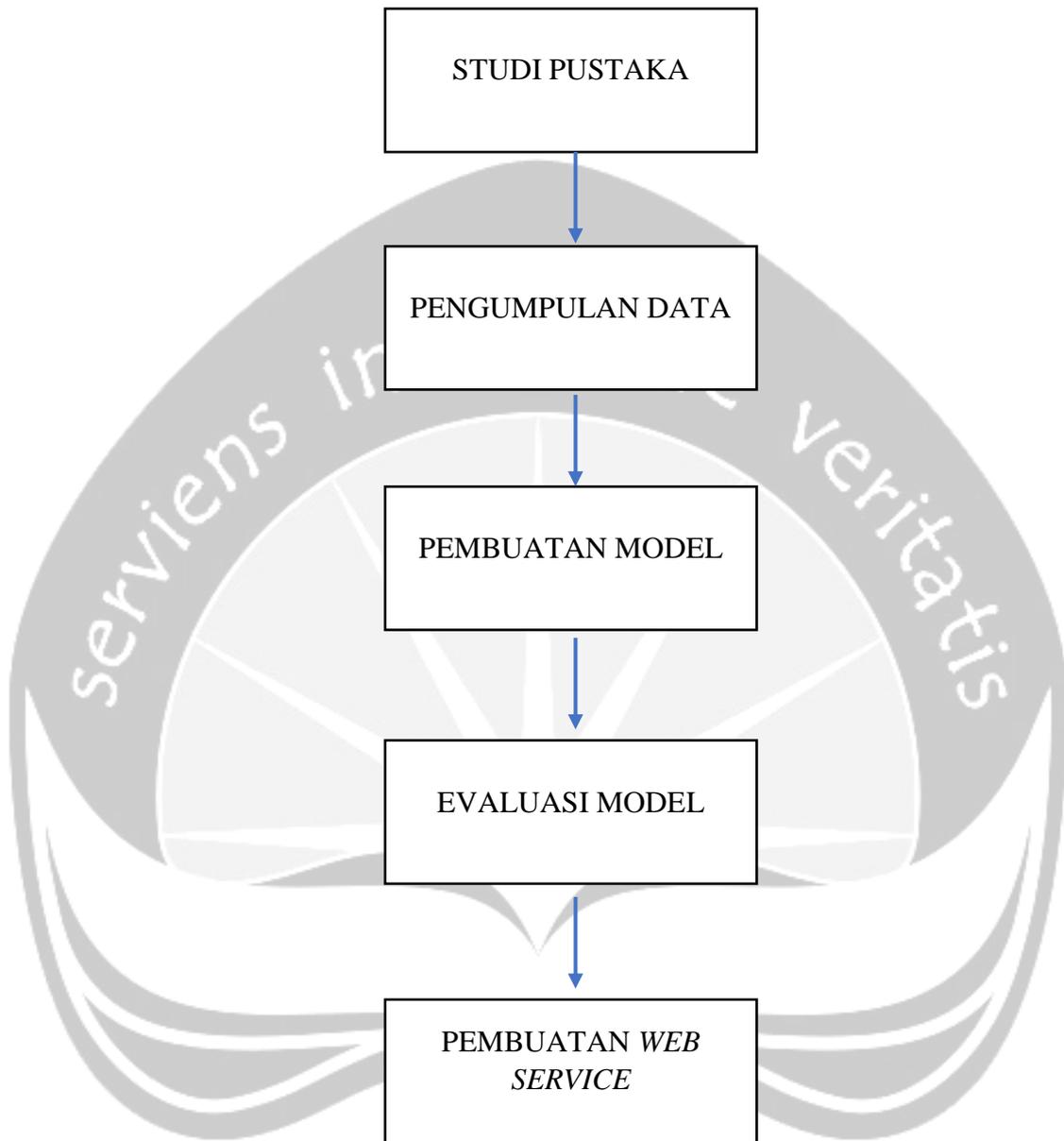
CPU : Processor Intel i3

RAM : 8GB

OS : Windows 10 64bit

3.3. Langkah-langkah Penelitian

Penelitian ini akan merancang sebuah model *Text Classification* dengan menggunakan *library Spacy*. Adapun beberapa tahap yang dilakukan dalam penelitian ini, yaitu studi pustaka, pengumpulan data, pembuatan model, evaluasi model, dan pembuatan *web service*, Berikut bagian yang menunjukkan tahap-tahap penelitian.



Gambar 3.2 Langkah- Langkah Penelitian

3.3.1. Studi Pustaka

Studi Pustaka merupakan langkah pertama dalam penelitian ini, Studi pustaka berguna untuk mendapatkan informasi dan teori-teori yang dibutuhkan dalam melakukan penelitian. Studi Pustaka ini dilakukan dengan cara mempelajari dan membaca jurnal-jurnal dan referensi-referensi buku lainnya yang berkaitan dengan konsep dasar *Machine Learning* dan implementasi *Machine Learning* untuk membuat sebuah model yang kemudian bisa digunakan dalam penelitian ini.

3.3.2. Pengumpulan Bahan

Tahap pengumpulan bahan bertujuan untuk mengumpulkan dan mencari sumber korpus kalimat Bahasa Indonesia dari berbagai sumber yang bisa digunakan untuk membuat model *Text Classification*.

3.3.3. Pembuatan Model

Tahap pembuatan model dilakukan setelah mengumpulkan bahan yang terdiri dari korpus dan lain-lain. Tahap pembuatan model dilakukan dengan 2 tahap yaitu :

3.3.3.1. Training Model Vektor

Tahap pembuatan model vektor dilakukan dengan melatih korpus kalimat Bahasa Indonesia menjadi model vektor agar dapat menyimpan arti dari tiap kata yang terdapat di kalimat pelatihan.

3.3.3.2. Training Model Text Classification

Tahap pembuatan model *Text Classification* dilakukan dengan memberi label kalimat yang didapat menjadi positif negatif, dan netral. Pada tahap ini juga dilakukan analisis pengujian hasil pelatihan. Jika nilai *losses* masih terlalu tinggi maka akan dilakukan proses penambahan dataset untuk mendapatkan nilai *losses* yang lebih kecil.

3.3.4. Evaluasi Model

Tahap evaluasi model dilakukan setelah melakukan pelatihan terhadap vektor dan dataset *Text Classification*. Di tahap ini akan dilakukan penyatuan kedua data tersebut untuk mendapatkan nilai evaluasi. Nilai evaluasi didapat dari hasil pengujian terhadap 50 kalimat tes.

3.3.5. Pembuatan Web Services

Pembuatan *web services* dilakukan dengan mengubah format model ke format Json yang dapat memudahkan untuk mengakses dari berbagai *platform*.

