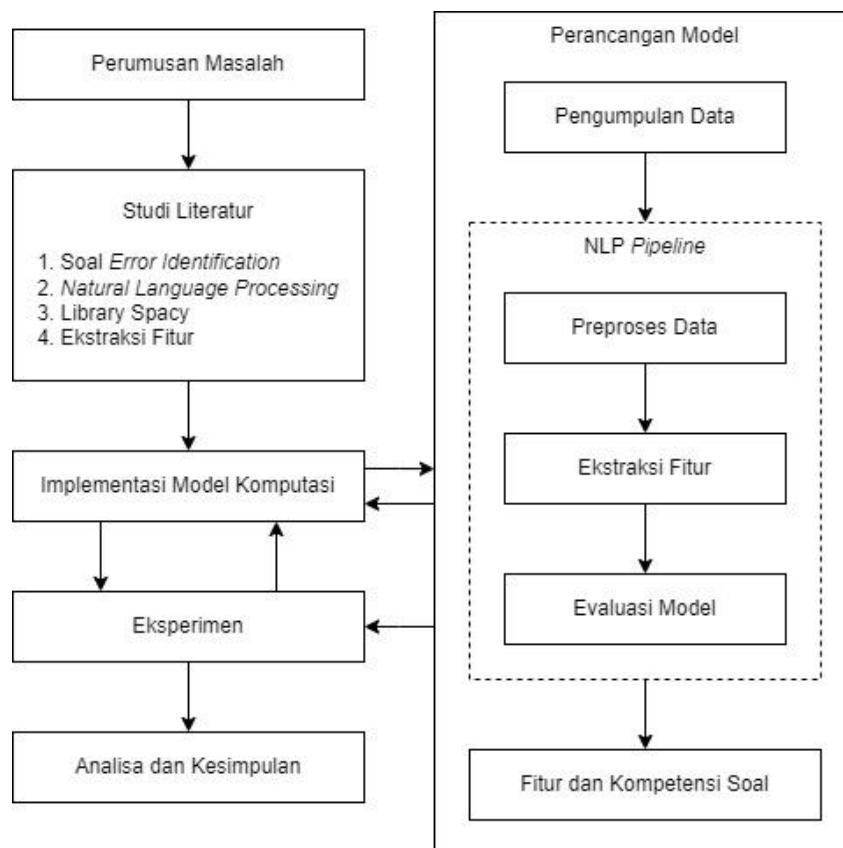


BAB III METODE PENELITIAN

3.1 Desain Penelitian

Desain penelitian merupakan kerangka kerja yang digunakan untuk melakukan penelitian. Kerangka kerja penelitian dijelaskan dalam bab ini, dimulai dari awal hingga akhir penelitian. Gambar 3.1 menggambarkan desain penelitian secara keseluruhan.



Gambar 3.1 Diagram Desain Penelitian.

Terdapat 6 tahap utama dalam desain penelitian yang dilakukan, yang akan dijelaskan sebagai berikut:

1. Perumusan Masalah

Pada tahap awal penelitian, rumusan masalah ditetapkan sebagai fokus penelitian. Kemudian, metode algoritma yang tepat dipilih untuk menyelesaikan permasalahan yang diidentifikasi. Tahap selanjutnya adalah pemilihan model

penelitian yang sesuai untuk membantu proses penyelesaian masalah dengan metode yang telah dipilih. Ketiga tahap ini merupakan langkah krusial dalam membangun fondasi penelitian yang kokoh dan terarah.

2. Studi Literatur

Studi literatur berkaitan dengan proses belajar dan literasi topik yang telah disetujui sebelumnya. Penulis melakukan studi literatur mengenai soal *Error Identification*, *Natural Language Processing*, *library Spacy*, dan *feature extraction* sebagai tempat berjalannya penelitian yang dilakukan oleh penulis. Studi literatur berasal dari beberapa sumber seperti buku, jurnal, juga internet, dan bacaan lainnya yang dapat dijadikan bahan literatur.

3. Pengumpulan Data

Data yang digunakan adalah data soal-soal Written Expression TOEFL ITP khususnya pada bagian Error Identification. Data tersebut diperoleh dari buku-buku latihan TOEFL ITP dan diinput secara manual menjadi data dengan format csv. Data tersebut berjumlah 50 soal yang terdiri dari teks soal, opsi jawaban, dan kunci jawaban.

4. Perancangan Model Komputasi

Tahap pertama pada perancangan model komputasi adalah dengan melakukan data preprocessing pada data soal yang telah diperoleh sebelumnya. Lalu penentuan fitur-fitur yang akan diekstrak dari soal yang merujuk pada sumber buku. Selanjutnya diperlukan juga proses data transformation untuk mendapatkan atribut-atribut yang diperlukan pada tahap ekstraksi fitur. Setelah data dipreproses dan ditransformasi, *library Spacy* dan *Natural Language Processing* digunakan pada data untuk mengekstraksi fitur-fitur yang telah ditentukan sebelumnya.

5. Eksperimen

Pada tahap ini, penulis membuat skenario eksperimen yang dapat digunakan untuk menguji keakuratan dari model komputasi yang telah dibuat dan melakukan

uji coba pada model komputasi yang telah dibangun dilakukan dengan memvalidasi hasil ekstraksi fitur oleh ahli.

6. Hasil dan Analisis Eksperimen

Setelah skenario dilakukan, penulis akan mengevaluasi hasilnya. Dari hasil ini, penulis akan menilai apakah model komputasi yang dipakai berjalan sesuai atau tidak. Kemudian, penulis akan menyimpulkan langkah-langkah sebelumnya secara berurutan.

3.2 Metode Penelitian

Penelitian dapat dibagi menjadi dua bagian, yaitu metode pengumpulan data dan metode pengembangan perangkat lunak.

3.2.1 Metode Pengumpulan Data

Penulis berupaya mendapatkan data yang valid dan relevan untuk mendukung penelitian. Berikut adalah metode pengumpulan data yang digunakan pada penelitian ini:

3.2.2.1 Studi Literatur

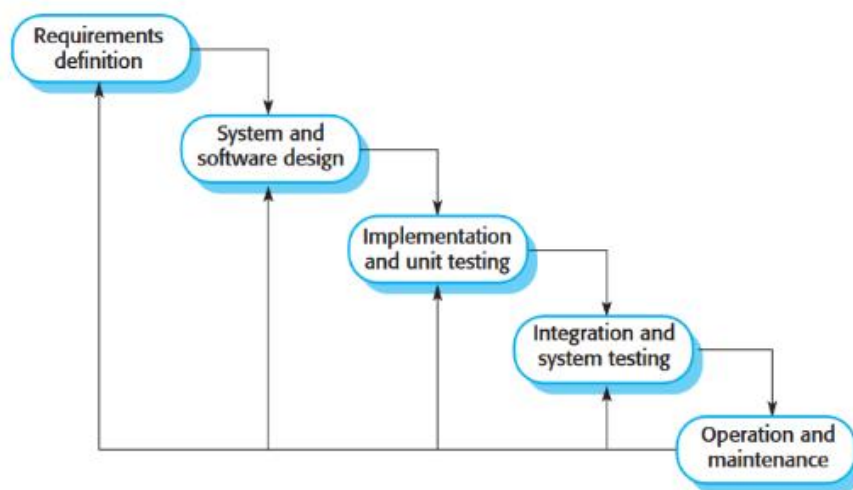
Studi literatur dilakukan untuk mendapatkan ilmu dan teori yang bisa digunakan penulis dalam penelitian ini, yaitu tentang Soal *Error Identification* TOEFL ITP, *Natural Language Processing*, *Feature Extraction*, dan *library Spacy*.

3.2.2.2 Mendapatkan Data Soal

Metode untuk mendapatkan data soal-soal TOEFL ITP tipe Error Identification yang valid yaitu didapat dari contoh soal pada sumber buku-buku latihan soal TOEFL ITP.

3.2.2 Metode Pengembangan Perangkat Lunak

Terdapat berbagai jenis metode pengembangan perangkat lunak. Salah satu metode yang digunakan oleh penulis adalah metode *waterfall* atau air terjun. Model ini juga dikenal sebagai model *linear sequential*. Tahapan dalam model *waterfall* dapat dilihat pada Gambar 3.2 berikut ini:



Gambar 3.2 Model Waterfall (Pressman, 2010)

Pada Gambar 3.2 ditunjukkan bahwa model waterfall terdiri dari lima tahapan. Tahapan-tahapan ini dapat membantu dalam proses penyelesaian perangkat lunak. Penjelasan mengenai masing-masing tahapan dari Gambar 3.2 diberikan oleh Pressman pada tahun 2010 sebagai berikut:

1. Requirement

Pada tahap ini, komunikasi sangat penting bagi perancang sistem untuk memahami perangkat lunak yang diinginkan oleh pengguna serta batasannya. Informasi dapat dikumpulkan melalui wawancara, diskusi, atau pertanyaan langsung. Data yang diperoleh kemudian dianalisis untuk menentukan kebutuhan pengguna.

2. Design

Di tahap ini, perancang mengembangkan desain sistem yang berguna untuk menentukan kebutuhan perangkat keras dan sistem, serta membantu mendefinisikan arsitektur keseluruhan sistem.

3. Implementation

Pada tahap ini, sistem awalnya dikembangkan dalam program-program kecil yang disebut unit. Unit-unit ini kemudian diintegrasikan pada fase berikutnya.

Muhammad Melvyn Maunti, 2024

EKSTRAKSI FITUR KOMPETENSI SOAL ERROR IDENTIFICATION TOEFL MENGGUNAKAN NATURAL LANGUAGE PROCESSING

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

Setiap unit dikembangkan dan diuji untuk fitur-fitur tertentu dalam proses yang disebut pengujian unit.

4. Verification

Pada tahap ini, sistem divalidasi dan diuji untuk memastikan bahwa persyaratan sistem terpenuhi secara keseluruhan atau sebagian. Pengujian dikategorikan menjadi pengujian unit (dijalankan pada modul kode tertentu), pengujian sistem (untuk mengevaluasi respons sistem ketika semua modul terintegrasi), dan pengujian penerimaan (dilakukan dengan atau atas nama pelanggan untuk memastikan semua kebutuhan pelanggan terpenuhi).

5. Maintenance

Ini adalah tahap akhir dari metode waterfall. Perangkat lunak yang telah selesai dioperasikan dan dipelihara. Pemeliharaan mencakup perbaikan bug yang tidak terdeteksi pada langkah-langkah sebelumnya.

3.3 Alat dan Bahan Penelitian

Sub bab ini menjelaskan apa saja komponen hardware dan software yang digunakan saat melakukan penelitian.

3.3.1 Alat Penelitian

1. Perangkat Keras berupa sebuah Laptop dengan spesifikasi:
 - Processor Intel(R) Core(TM) i3-4170U CPU @ 3.70GHz
 - RAM 10 GB
 - Nvidia GeForce GT 630
 - HDD 500 GB
2. Perangkat Lunak yang diperlukan adalah sebagai berikut:
 - Sistem Operasi Windows 10 64 bit
 - Visual Studio Code
 - Python 3.10

3.3.2 Bahan Penelitian

Bahan yang diperlukan untuk melakukan penelitian adalah data soal *Error Identification* TOEFL yang dikumpulkan dalam format csv. Data soal diperoleh dari buku latihan TOEFL. Data berisi teks soal, opsi A, opsi B, opsi C, opsi D dan kunci jawaban.