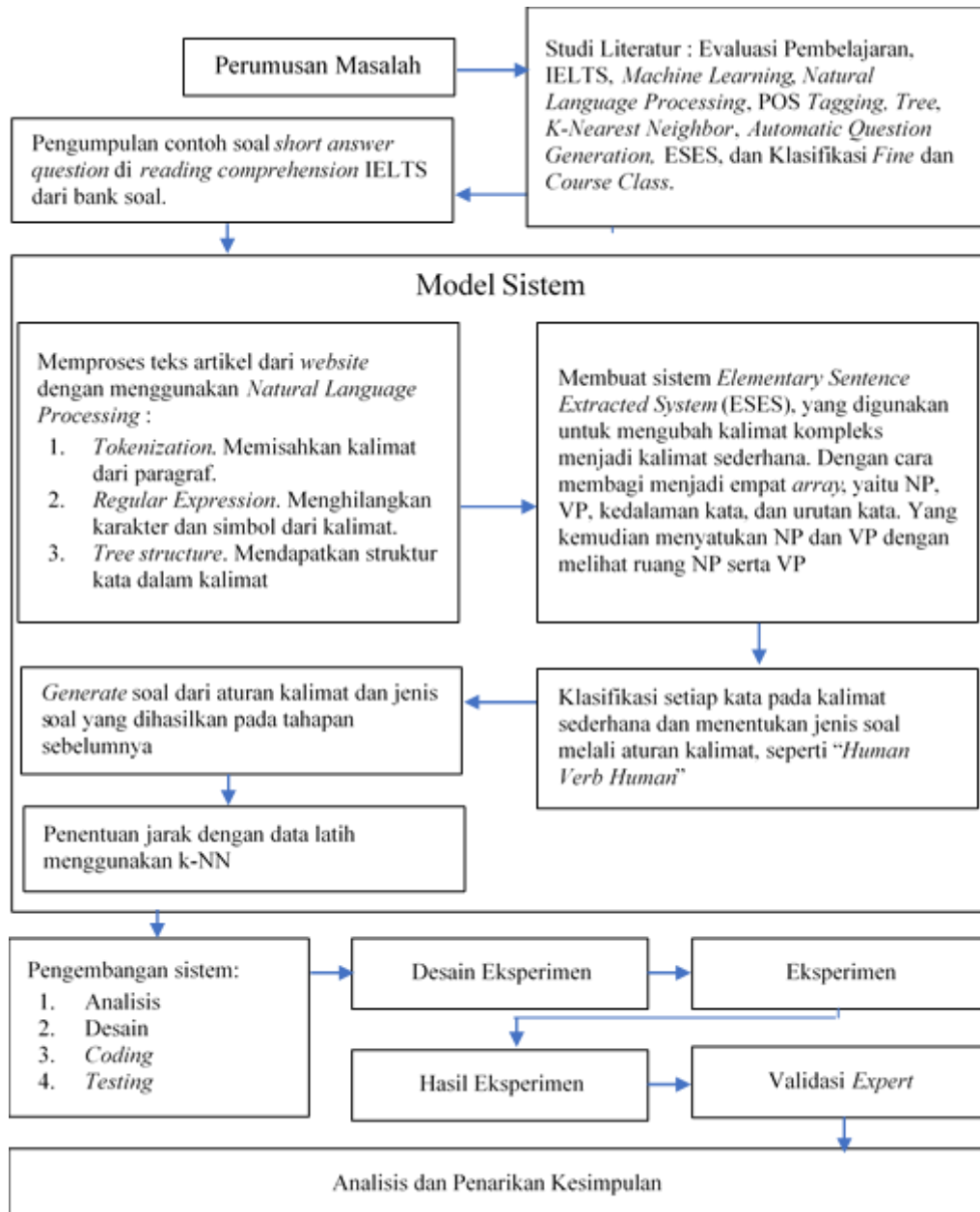


BAB III

METODOLOGI PENELITIAN

3.1 Desain Penelitian

Desain penelitian merupakan sebuah kerangka atau gambaran dari proses kerja yang digunakan untuk melakukan penelitian. Pada Gambar 3.1 menunjukkan desain dari penelitian yang akan dilakukan.



Gambar 3.1 Desain Penelitian

3.1.1. Perumusan Masalah

Tahap ini merupakan tahapan pertama yang akan dilakukan dalam penelitian ini. Pada tahap ini merupakan tahap memahami latar belakang yang nantinya akan merumuskan masalah dan akan dicari solusi dari masalah yang disebutkan.

Tahapan ini yang nantinya akan sangat berguna bagi peneliti, yang akan menjadi sebagai panduan dalam melakukan penelitiannya. Tentu hal tersebut didasari pada pencarian referensi yang tepat dan pemahaman yang baik. Yang akan menghasilkan tujuan serta manfaat, dan menjawab semua rumusan masalah.

3.1.2. Studi Literatur

Tahap ini merupakan tahapan yang dimana dilakukannya mempelajari serta memahami dari teori-teori yang sudah ada untuk menyelesaikan penelitian dan menjawab rumusan masalah. Teori-teori yang akan mendukung untuk berjalannya penelitian ini yaitu evaluasi pembelajaran, IELTS, *machine learning*, *natural language processing*, *POS tagging*, algoritma *k-nearest neighbor*, dan penelitian terkait *automatic question generation*.

Tahapan ini peneliti mempelajari teori-teori yang akan mendukung dalam berjalannya penelitian, dengan cara membaca dan memahami dari beberapa literatur seperti buku, jurnal, *paper*, dan juga artikel yang memiliki topik serupa dengan teori-teori yang telah disebutkan sebelumnya.

3.1.3. Perancangan Model Sistem *Automatic Question Generation*

Tahap ini merupakan tahapan dimana peneliti membuat model dari rancangan sistem *automatic question generation*. Model tersebut dibangun berdasarkan beberapa tahapan yang penting, yaitu:

1. Pengumpulan data dari situs media asing dan buku IELTS

Pada tahapan ini merupakan tahapan pengumpulan data training dan data testing. Kemudian data tersebut akan diolah oleh sistem untuk mendapatkan informasi mengenai pola dari soal tersebut atau lainnya. *Data training* diperoleh dari buku-buku latihan IELTS. *Data testing* diperoleh dari

situs media asing yang terpercaya mulai dari berita serta tata bahasanya ataupun dari situs IELTS itu sendiri.

2. *Natural Language Processing*

Natural Language Processing (NLP) atau bisa disebut juga pengolahan bahasa alami merupakan sebuah kecerdasan buatan yang berfokus pada pemrosesan teks, memberikan komputer sebuah kemampuan untuk memahami bahasa manusia. Tidak sedikit teknik NLP yang bisa digunakan untuk memproses teks.

Pemrosesan teks ini dilakukan dengan tujuan untuk menghilangkan karakter dan simbol yang ada dalam kalimat. Karena dalam melakukan konversi menjadi *tree* tidak dapat mengkonversi karakter dan simbol. Selain itu, juga digunakan untuk memisahkan paragraf menjadi beberapa kalimat dengan syarat-syarat yang sudah ditentukan.

Teknik yang digunakan yaitu, *regular expression*, *tokenization*, dan *tree structure*. *Regular expression* atau bisa disebut dengan *regex* digunakan untuk membersihkan kalimat dari karakter dan juga simbol, agar bisa masuk ke tahapan selanjutnya. *Tokenization* digunakan untuk ekstraksi kalimat-kalimat dari paragraf dengan syarat disebut sebagai kalimat yaitu diawali dengan huruf kapital dan diakhiri dengan titik, syarat ini diambil dari penelitian Ali (2010). Yang terakhir Konversi dari kalimat menjadi *tree structure*, yang berguna untuk proses selanjutnya yaitu ekstraksi kalimat sederhana dari kalimat kompleks, oleh karena itu *tree structure* ini berguna untuk mendapatkan aturan setiap kata dalam kalimat.

3. *K-Nearest Neighbor*

K-Nearest Neighbor (k-NN) merupakan salah satu dari algoritma *supervised learning* yang termasuk kedalam tipe klasifikasi yang terdapat pada *machine learning*. Algoritma ini digunakan untuk menghitung jarak antara data latih dengan kandidat soal. Semakin kecil jaraknya maka semakin mirip kandidat soal dengan soal IELTS, atau semakin laya kandidat soal menjadi soal.

3.1.4. Pengembangan Sistem

Pada tahapan ini penulis akan melakukan pengembangan sistem *automatic question generation* sesuai dengan alur model. Model pengembangan perangkat lunak yang penulis gunakan dalam penelitian ini yaitu model *waterfall* atau bisa disebut juga model air terjun, karena alur pengembangan perangkat lunak ini sesuai dengan air terjun.

Tahapan untuk model *waterfall* ini antara lain *Requirements Definition, System and Software Design, Implementation and Unit Testing, Integration and System Testing*, dan *Operation and Maintenance*. Pada model ini jika ada yang perlu diperbaiki pada tahapan sebelumnya maka dapat kembali ke tahapan sebelumnya. Atau dapat dikatakan bahwa untuk Langkah pertama yaitu menentukan kebutuhan dari sistem, fitur apa saja yang harus dimiliki oleh sistem. Tahapan kedua membuat desain dari sistem, atau membangun alur model sistem. Tahapan ketiga yaitu memulai *coding* dan *testing* sistem dengan memperhatikan kebutuhan sistem serta alur model yang telah ditentukan sebelumnya. Tahapan keempat melakukan *testing* terhadap sistem yang telah dibuat. Tahapan kelima yaitu memperbaiki kesalahan-kesalahan yang muncul saat tahapan sebelumnya atau *testing* dilakukan.

3.1.5. Desain Eksperimen

Tahapan desain eksperimen ini merupakan tahapan dimana merancang eksperimen yang akan dilakukan dengan cara menentukan skenario eksperimen. Penelitian ini penulis menentukan tiga skenario, yaitu yang pertama skenario ekstraksi kalimat sederhana dari kalimat kompleks, yang kedua adalah skenario klasifikasi kalimat kandidat soal, dan yang ketiga adalah skenario evaluasi kualitas soal.

3.1.6. Eksperimen

Pada tahapan ini eksperimen akan dilakukan sesuai dengan skenario yang telah ditentukan sebelumnya. Dalam melakukan eksperimen ini melibatkan data kandidat soal yang menjadi data uji dan data latih yang merupakan kumpulan soal IELTS yang diambil dari *ebook* dan juga *website*. Eksperimen ini akan menguji sistem yang sudah dikembangkan oleh penulis.

3.1.7. Analisis dan Validasi *Expert*

Hasil dari sistem yang dibuat yaitu kandidat soal untuk jenis *short answer question* bagian *reading comprehension* IELTS, selanjutnya adalah melakukan validasi soal oleh *expert*, hal ini dilakukan untuk dapat menentukan kelayakan dan kesesuaian soal yang dihasilkan oleh sistem. Setelah itu melakukan analisis oleh penulis dari hasil validasi yang dilakukan oleh *expert*. Validasi ini dilakukan dengan parameter yang ditentukan pada tahapan sebelumnya.

3.2 Alat dan Bahan Penelitian

Pada subbab ini akan menjelaskan alat serta bahan yang menunjang dalam melakukan penelitian *automatic question generation* ini.

3.2.1. Alat Penelitian

Melakukan penelitian tentu membutuhkan alat-alat yang dapat menunjang penelitian agar dapat selesai. Alat-alat penelitian yang dibutuhkan ini dibagi menjadi dua aspek, yaitu yang pertama dari perangkat keras dan yang kedua dari perangkat lunak. Lebih jelasnya yaitu sebagai berikut:

a. Perangkat Keras (*Hardware*) yaitu sebuah laptop dengan spesifikasi :

- 1) *Processor* AMD A9-9420 RADEON R5
- 2) *Random Access Memory* (RAM) 12 GB
- 3) *Operating system* 64-bit

b. Perangkat Lunak (*Software*) yaitu diantaranya:

- 1) Python 3.68

Perangkat lunak ini digunakan untuk menjalankan kode program di komputer dengan bahasa pemrograman python.

- 2) *Web Browser* Google Chrome

Perangkat lunak ini digunakan untuk menjalankan google colabs, karena google colabs dijalankan secara online via *website*.

- 3) Sublime text editor 3

Perangkat lunak ini digunakan untuk melakukan edit program ataupun sebagai pengetikan program sederhana yang menjadi proses dari system.

- 4) *Command Prompt*

Perangkat lunak ini digunakan untuk melakukan instalasi dari *library* StanfordCoreNLP.

5) Google Colabs

Perangkat lunak ini digunakan untuk mengembangkan sistem *automatic generate question*.

6) StanfordCoreNLP

Perangkat lunak ini digunakan untuk melakukan konversi dari kalimat menjadi *tree structure*.

7) Flask

Perangkat lunak ini merupakan *framework* dari python yang berupa *website*, digunakan menjadi *user interface* pada sistem ini.

3.2.2. Bahan Penelitian

Bahan penelitian yang digunakan dalam penelitian ini adalah artikel yang berasal dari *website* berita berbahasa inggris, dengan penulisan *grammar* yang terpercaya seperti CNN, BBC, The Jakarta Post, ataupun New York Times. Selain itu, pada penelitian ini mengumpulkan soal *short answer question* pada bagian *reading comprehension* IELTS juga diperlukan sebagai data latihan dari ebook.

3.3 Metode Penelitian

Pada penelitian ini menggunakan metode yang dibagi menjadi dua bagian, yaitu antara lain metode pengumpulan data dan metode pengembangan perangkat lunak. Penjelasan lebih lengkapnya sebagai berikut.

3.3.1. Metode Pengumpulan Data

Dalam metode pengumpulan data yang digunakan, ada dua metode yang dilakukan oleh penulis, yaitu sebagai berikut:

1. Mengumpulkan data soal IELTS

Dalam penelitian ini membutuhkan dataset atau data latihan yang berguna untuk membandingkan antara data latihan dengan data uji, agar dapat menjaga kualitas soal yang dihasilkan, kualitas yang dimaksud yaitu kualitas dari jenis soal *short answer question* pada bagian *reading comprehension* IELTS. Data latihan ini dikumpulkan dari dua sumber, yaitu *ebook* dan *website* resmi, tentunya

jenis soal yang dikumpulkan adalah *short answer question* pada bagian *reading comprehension* IELTS.

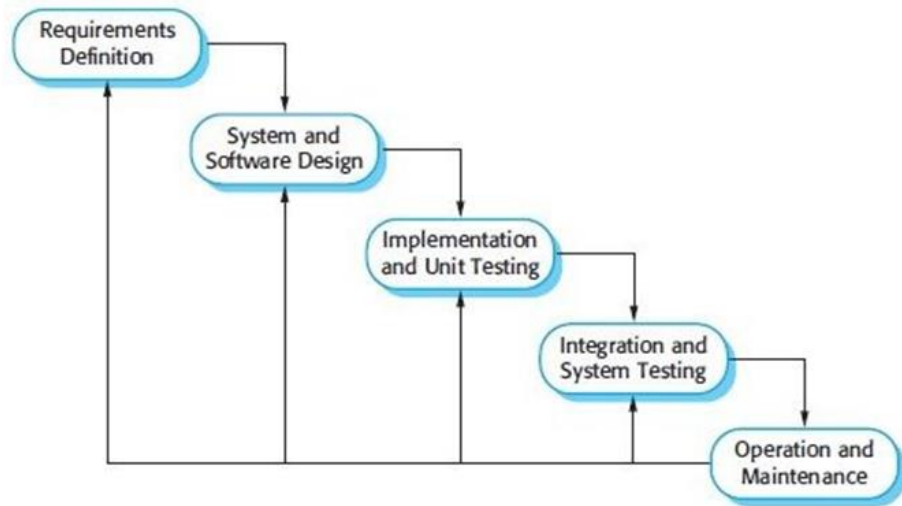
2. Mengunduh artikel berita

Pengunduhan artikel berita ini digunakan untuk melakukan eksperimen agar mendapatkan kandidat soal yang menjadi data uji. Artikel berita dikumpulkan dengan secara otomatis menggunakan *library newspaper3k* dengan cara *scrapping* isi artikel. *Library* ini disediakan oleh *python*, hasil dari *scrapping* ini berupa isi artikel dari *url website* yang dimasukkan, kemudian artikel akan disimpan dalam bentuk *file*.

3.3.2. Metode Pengembangan Perangkat Lunak

Ada banyak metode pengembangan perangkat lunak, akan tetapi metode pengembangan perangkat lunak yang dipakai dalam penelitian ini adalah metode *waterfall* atau bisa disebut juga metode air terjun. Karena dalam metode ini dalam pengembangannya seperti air terjun. Untuk tahapan dan diagram bagaimana metode ini dilakukan dapat dilihat pada Gambar 3.2.

Dapat diketahui pada Gambar 3.2, bahwa terdapat lima tahapan yang harus dilalui pada pengembangan perangkat lunak dengan metode jenis ini, yaitu diantaranya *Requirements Definition*, *System and Software Design*, *Implementation and Unit Testing*, *Integration and System Testing*, dan *Operation and Maintenance*. Pada model ini jika ada yang perlu diperbaiki pada tahapan sebelumnya maka dapat kembali ke tahapan sebelumnya. Atau dapat dikatakan bahwa untuk Langkah pertama yaitu menentukan kebutuhan dari sistem, fitur apa saja yang harus dimiliki oleh sistem. Tahapan kedua membuat desain dari sistem, atau membangun alur model sistem. Tahapan ketiga yaitu memulai *coding* dan *testing* sistem dengan memperhatikan kebutuhan sistem serta alur model yang telah ditentukan sebelumnya. Tahapan keempat melakukan *testing* terhadap sistem yang telah dibuat. Tahapan kelima yaitu memperbaiki kesalahan-kesalahan yang muncul saat tahapan sebelumnya atau *testing* dilakukan.



Gambar 3. 2 Model *Waterfall* (Sommerville, 2011).