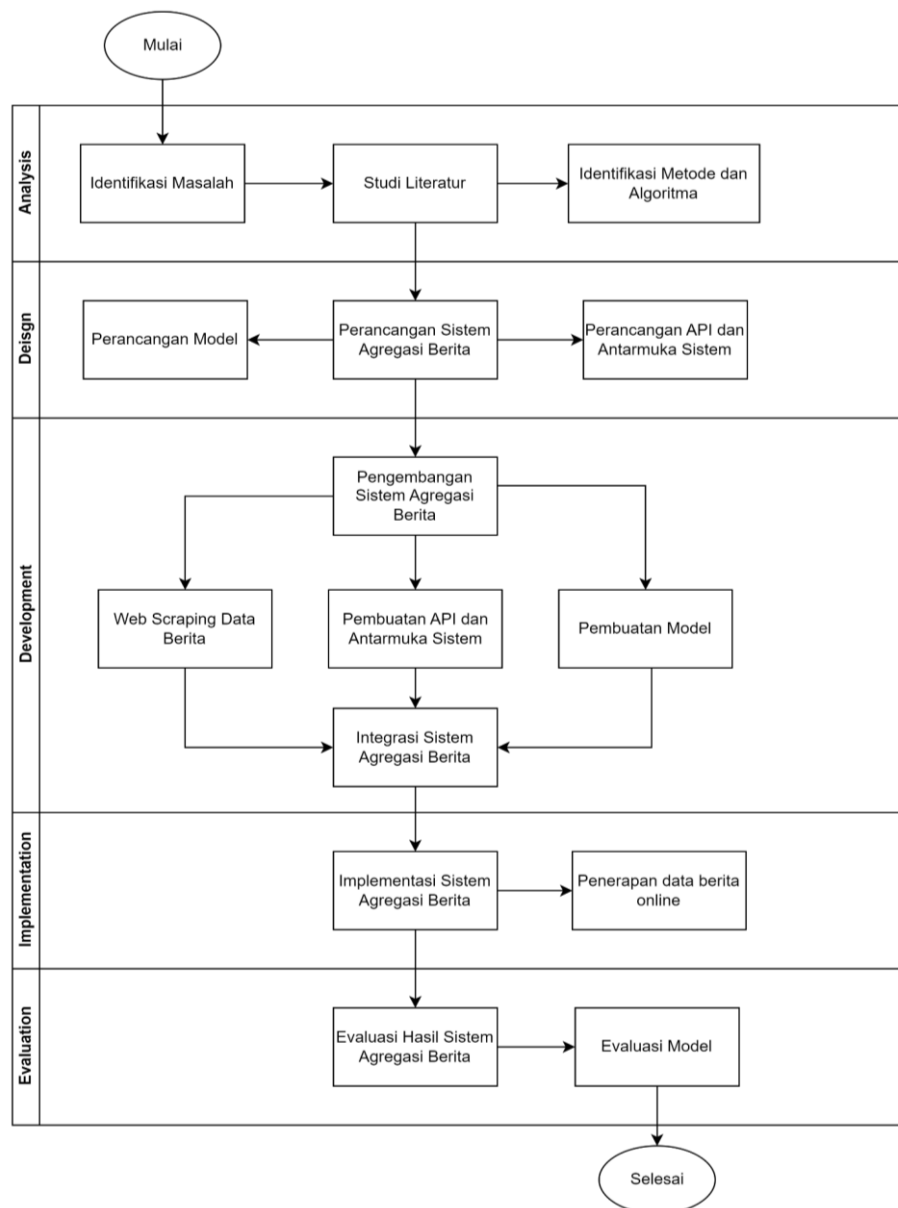


BAB III

METODE PENELITIAN

3.1 Desain Penelitian

Desain penelitian adalah suatu kerangka kerja yang digunakan oleh peneliti untuk menggambarkan langkah-langkah yang dilakukan selama proses penelitian. Penelitian ini mengikuti model ADDIE (*Analysis, Design, Development, Implementation, Evaluation*) ditunjukkan dalam Gambar 3.1.



Gambar 3.1 Desain Penelitian

Dengan memperhatikan tahapan di atas, maka langkah-langkah yang dilakukan dalam penelitian ini meliputi:

1. *Analysis*

Penelitian dimulai dengan tahap analisis. Pada tahap ini akan dilakukan identifikasi masalah, studi literatur, serta identifikasi metode dan algoritma yang akan digunakan dalam penelitian.

a. Identifikasi Masalah

Pada tahap ini peneliti mencoba mengidentifikasi dan merumuskan masalah. Dari permasalahan yang ada, selanjutnya dapat diidentifikasi manfaat serta tujuan yang dapat dicapai dengan dilakukannya penelitian ini. Masalah yang diangkat pada penelitian ini yaitu terkait klasifikasi, *clustering*, dan peringkasan teks secara otomatis.

b. Studi Literatur

Studi literatur merupakan proses belajar untuk memperoleh informasi yang berkaitan dengan topik penelitian. Pada penelitian ini, studi literatur yang dilakukan mengenai pengembangan perangkat lunak, klasifikasi teks, *clustering* teks, dan peringkasan teks. Dilakukan juga pemahaman pada dataset yang akan digunakan serta penelitian yang pernah dilakukan sebelumnya.

c. Identifikasi Metode dan Algoritma

Pada tahap ini, peneliti mengidentifikasi metode serta algoritma yang akan digunakan untuk klasifikasi, *clustering*, dan peringkasan berita secara otomatis.

2. *Design*

Setelah menganalisis permasalahan dan mendapat ide untuk membantu menyelesaikan hal tersebut, penelitian dilanjutkan dengan tahapan desain. Pada tahap ini dilakukan perancangan antarmuka dan gambaran umum sistem agregasi berita.

a. Perancangan Sistem Agregasi Berita

Tahapan perancangan sistem agregasi berita dilakukan untuk mendapat gambaran dari sistem yang akan dibuat.

b. Perancangan API (*Application Programming Interface*) dan Antarmuka Sistem

Pada tahap ini, peneliti merancang struktur API serta tampilan sistem sebagai antarmuka pengguna. Dengan antarmuka ini, pengguna dapat melihat dan mengakses artikel berita yang terdapat pada sistem agregasi berita.

c. Perancangan Model

Pembuatan model terbagi menjadi tiga bagian utama, yaitu model klasifikasi, clustering, dan peringkasan.

3. *Development*

Perangkat lunak yang telah dirancang untuk melakukan klasifikasi, *clustering*, dan peringkasan berita secara otomatis, kemudian dikembangkan pada tahap ini.

a. Pengembangan Sistem Agregasi Berita

Pada penelitian ini, model pengembangan perangkat lunak yang digunakan adalah *waterfall*, yang terdiri dari spesifikasi kebutuhan, desain dan implementasi, integrasi dan testing, serta operasi dan pemeliharaan.

b. *Web Scraping* Data Berita

Proses pengambilan data artikel berita dilakukan melalui scraping terhadap situs berita daring berbahasa Indonesia.

c. Pembuatan Model

Pembuatan model terbagi menjadi tiga, yakni pelatihan model klasifikasi teks, perancangan model *clustering*, dan *fine tuning* model peringkasan teks.

d. Pembuatan API dan Antarmuka Sistem

Pada tahap ini, peneliti membuat API dan antarmuka sistem yang akan digunakan sebagai sistem agregasi berita.

e. Integrasi Sistem Agregasi Berita

Setelah pengembangan antarmuka dan backend sistem agregasi berita serta ketiga model selesai dilakukan, tahapan selanjutnya adalah

mengintegrasikan bagian-bagian tersebut menjadi satu perangkat lunak agregasi berita. Proses pengolahan data yang dilakukan dalam perangkat lunak ini secara berurutan yaitu *fetching* data berita, klasifikasi kategori berita, *clustering* topik berita berdasarkan kategori dan rentang waktu, dan terakhir peringkasan abstraktif untuk setiap topik berita.

4. *Implementation*

Pada tahap ini, dilakukan penerapan antarmuka sistem dan model didalamnya terhadap data berita dari berbagai sumber. Sistem harus mampu mengklasifikasi, meng-*cluster*, serta meringkas berita yang didapat dari berbagai sumber berita *online*.

5. *Evaluation*

Tahap terakhir yaitu evaluasi dilakukan untuk menilai efektivitas sistem agregasi berita yang telah dikembangkan dan diimplementasikan pada tahapan sebelumnya. Evaluasi yang dilakukan termasuk fungsional perangkat lunak serta kualitas hasil model yang dibuat. Metode evaluasi untuk model klasifikasi yang digunakan yaitu akurasi dan *F1-Score*, sementara untuk *clustering* menggunakan CH-index (*Calinski-Harabasz Index*), dan untuk peringkasan menggunakan ROUGE (*Recall-Oriented Understudy for Gisting Evaluation*). Evaluasi CH-index akan dilakukan terhadap hasil *clustering* pada setiap kategori berita yang sebelumnya telah diklasifikasi dan dikelompokkan berdasarkan rentang waktu, lalu dihitung rata-rata dan median nya.

3.2 Alat Penelitian

Dalam penelitian ini beberapa perangkat keras dan perangkat lunak yang digunakan untuk menunjang kebutuhan selama penelitian adalah sebagai berikut:

1. Perangkat keras (hardware) yaitu sebuah laptop dengan spesifikasi:
 - a. Prosesor: Intel(R) Core(TM) i5-7300HQ CPU @ 2.50GHz
 - b. RAM: 20 GB
 - c. SSD: 256 GB
 - d. Monitor: resolusi 1920 x 1080

2. Perangkat lunak (*software*) sebagai berikut:

- a. Sistem Operasi Windows 10
- b. Google Chrome Web Browser
- c. Kaggle
- d. Visual Studio Code

3.3 Bahan Penelitian

Penelitian ini menggunakan bahan berupa dataset berita bernama IndoSum. Dataset ini terdiri dari sekitar 20 ribu artikel berita, dengan setiap artikel memiliki judul, kategori, sumber berita, link berita asli, serta ringkasan abstraktif yang dibuat secara manual. Artikel berita yang terdapat pada dataset IndoSum berasal dari sumber media berita Indonesia, seperti CNN Indonesia dan Kumparan. Pada dataset ini sudah terdapat kategori untuk setiap artikel beritanya sehingga memungkinkan untuk dilakukan klasifikasi tanpa harus melalui proses *labeling*. Selain itu, terdapat juga ringkasan yang telah dibuat secara manual yang dapat digunakan dalam pengukuran metrik evaluasi ROUGE.