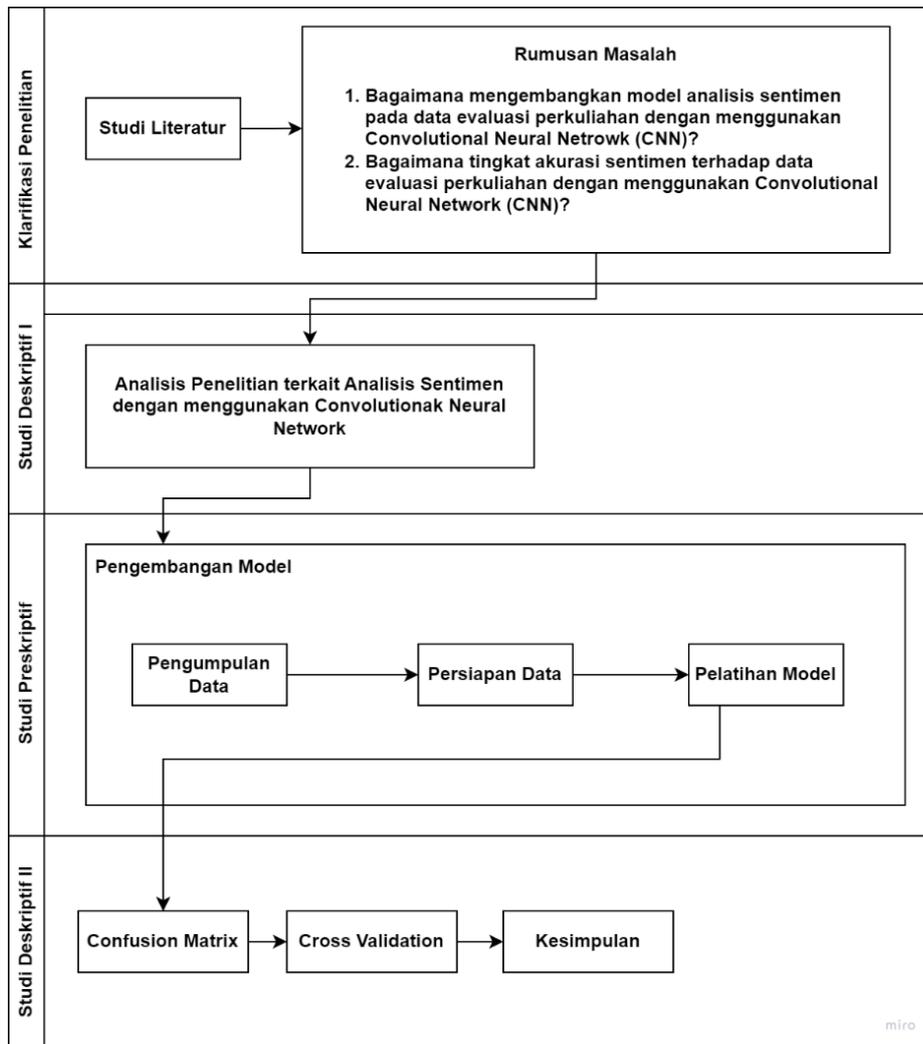


BAB III METODOLOGI PENELITIAN

3.1 Desain Penelitian

Desain penelitian menjelaskan sebuah kerangka penelitian yang akan dilakukan. Metode penelitian yang digunakan pada penelitian ini adalah *DRM* (*Design Research Methodology*). DRM adalah sebuah metode dan pedoman pendukung yang akan digunakan sebagai kerangka kerja selama melaksanakan sebuah penelitian desain (Blessing dan Chakrabakti, 2009). Terdapat empat tahapan utama dalam DRM yaitu klarifikasi penelitian, studi deskriptif i, studi preskriptif, dan studi deskriptif ii. Berikut merupakan desain penelitian pada penelitian ini:



Gambar 3. 1 Diagram Desain Penelitian

3.1.1 Klarifikasi Penelitian

Tahap pertama yang dilakukan pada penelitian adalah klarifikasi penelitian. Tahap ini dilakukan untuk mengidentifikasi tujuan dari penelitian. Pada penelitian ini peneliti melakukan studi literatur untuk dapat mengidentifikasi permasalahan dan tujuan dari penelitian. Studi literatur dilakukan dengan dengan membaca sumber-sumber bacaan dari artikel, jurnal, skripsi, buku, dan bacaan di internet. Topik literatur yang menjadi sumber adalah yang berhubungan dengan analisis sentimen, klasifikasi teks dengan algoritma *Convolutional Neural Network*, dan evaluasi perkuliahan. Luaran pada tahap ini adalah topik permasalahan yaitu analisis sentimen untuk mengevaluasi umpan balik perkuliahan dengan menggunakan *Convolutional Neural Network* (CNN). Topik permasalahan ini dijelaskan dalam rumusan masalah pada Bab I. Hasil studi literatur penelitian ini di dibahas pada Bab II.

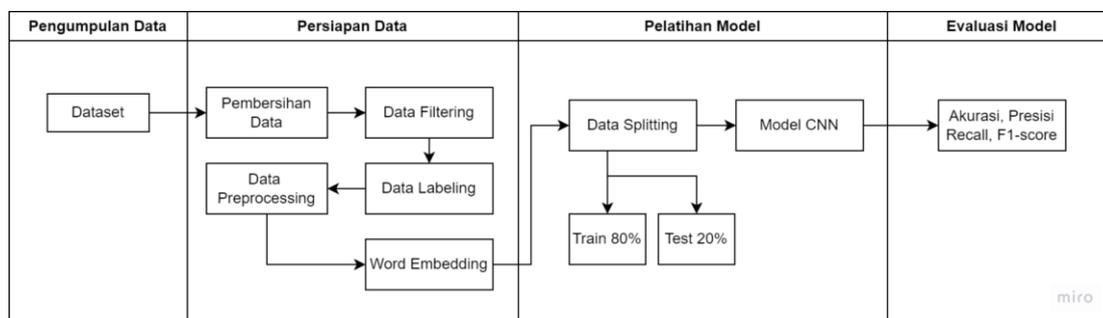
3.1.2 Studi Deskriptif I

Tahap selanjutnya ialah studi deskriptif I, pada tahap ini dilakukan pemahaman mendalam terkait kondisi yang berkaitan dengan penelitian. Pada penelitian ini, peneliti melakukan pemahaman mendalam mengenai analisis sentimen dengan algoritma CNN. Pemahaman mendalam ini dilakukan dengan menganalisis penelitian terdahulu yang memiliki topik permasalahan yang sama. Penelitian-penelitian terdahulu itu dianalisis dari implementasi algoritma CNN, dataset yang digunakan, metrik yang digunakan, hingga hasil yang didapatkan dari penelitian tersebut. Hasil dan pembahasan analisis ini dapat dilihat pada Bab II bagian penelitian terkait. Hasil analisis ini berupa batasan masalah dan hal-hal yang menjadi pembeda penelitian ini dengan penelitian sebelumnya.

3.1.3 Studi Preskriptif

Tahap selanjutnya ialah studi preskriptif, tahap ini adalah tahapan pengembangan model yang merupakan solusi dari pemahaman yang didapatkan pada tahap sebelumnya. Pada penelitian ini tahap studi preskriptif dimulai dengan pengumpulan data. Pengumpulan data dilakukan melalui kuesioner terkait PBM atau perkuliahan di UPI, kuesioner ini terdapat pada sistem akademik mahasiswa. Pengambilan data dilakukan melalui Direktorat Sistem dan Teknologi Informasi.

Data kemudian masuk ke tahap persiapan agar data dapat digunakan sesuai dengan kebutuhan. Kemudian dilakukan pelatihan data untuk mendapatkan model solusi. Berikut merupakan alur dari pengembangan model:



Gambar 3. 2 Diagram Alur Pengembangan Model

Pengumpulan data dilakukan melalui kuesioner terkait PBM atau perkuliahan di UPI, kuesioner ini terdapat pada sistem akademik mahasiswa. Pengambilan data dilakukan melalui Direktorat Sistem dan Teknologi Informasi melalui surat izin penelitian. Kemudian tahapan kedua adalah melakukan persiapan terhadap data yang akan digunakan. Pada persiapan data dilakukan pembersihan data agar terhindar dari *noise* atau data yang tidak diperlukan. Kemudian dilakukan penyaringan atau *filter* data untuk mengurangi konteks data agar lebih mudah selama penelitian. Kemudian dilakukan *data filtering* pada kolom mata kuliah.

Proses selanjutnya adalah *data preprocessing*, pada proses ini dilakukan beberapa hal seperti menghapus data duplikat, menghapus angka dan karakter selain huruf, *case folding*, *stopwords removal*, *tokenization*, dan *stemming*. Berikut merupakan deskripsi dari tahapan di dalam *data preprocessing*:

Tabel 3. 1

Deskripsi Tahapan pada *Data Preprocessing*

No	Tahapan	Deskripsi
1	Menghapus data duplikat	Melakukan penghapusan terhadap data yang terduplikasi.
2	<i>Remove Punctuation</i>	Melakukan penghapusan terhadap angka, tanda baca, simbol, dan karakter selain huruf.
3	<i>Case folding</i>	Melakukan perubahan terhadap huruf menjadi huruf kecil.

4	<i>Tokenization</i>	Melakukan tokenisasi atau pembagian menjadi token-token dalam sebuah kalimat. Pembagian token ini berdasarkan spasi.
5	<i>Stopwords removal</i>	Melakukan proses penghapusan pada kata yang dianggap kurang penting misalnya kata sambung, kata depan, dan sebagainya.
6	<i>Stemming</i>	Melakukan perubahan terhadap kata-kata menjadi ke bentuk kata dasarnya.

Proses selanjutnya adalah *labeling data*, proses ini dilakukan untuk memberi label sentimen pada setiap data, sentimen ini akan menjadi target variabel dalam proses *modeling*. Proses *labeling data* dilakukan menggunakan bantuan leksikon SentiStrength_id digunakan pada penelitian ini karena memiliki kemampuan yang cukup baik dalam memberikan sentimen pada teks (Fauziah dkk., 2021). Pada penelitian ini juga menggunakan bantuan Vander (*valance aware dictionary and sentiment reasoner*) dalam menentukan polaritas suatu teks. Vander akan melakukan perhitungan polaritas suatu teks berdasarkan leksikon SentiStrength_id. Nilai polaritas tersebut dilakukan normalisasi kemudian dihasilkan suatu nilai jika nilai tersebut lebih besar dari nol maka teks berlabel positif, jika kurang dari nol maka berlabel negatif, dan jika nilai sama dengan nol maka teks berlabel netral.

Setelah itu dilakukan *word embedding* yaitu proses mengubah data teks menjadi bentuk vektor. Pada penelitian ini dilakukan perbandingan antara dua pendekatan *word embedding*. Pelatihan model dilakukan untuk mengembangkan model dengan *layer embedding* dan model dengan *word2vec*. *Layer embedding* merupakan proses representasi kata menjadi bentuk vektor dengan melakukan pelatihan pada dataset yang akan digunakan. Sedangkan *word2vec* merupakan proses representasi teks ke dalam bentuk vektor dengan menggunakan model yang sudah ada, jadi nilai vektor yang didapat merupakan hasil pelatihan yang telah dilakukan sebelumnya menggunakan dataset wikipedia. Setelah dilakukan proses representasi data teks menjadi bentuk vektor data tersebut akan dilakukan proses pelatihan dengan algoritma CNN.

Sebelum pelatihan dilakukan pembagian data dengan skema data latih sebanyak 80% dan data uji sebanyak 20%. Skema pembagian data ini merujuk pada penelitian sebelumnya yang melakukan pembagian yang sama dengan hasil model yang cukup baik. Proses pelatihan data menggunakan CNN dilakukan dengan data yang telah berbentuk numerik kemudian dilakukan proses konvolusi untuk mendeteksi fitur penting, kemudian lapisan pooling untuk mengurangi dimensi dan mempertahankan fitur penting. Setelah itu masuk ke lapisan *fully connected* untuk menggabungkan informasi dan menghasilkan prediksi. Selanjutnya fungsi loss bekerja untuk menghitung selisih antara prediksi dan label asli, yang kemudian digunakan dalam *backpropagation* untuk memperbaiki bobot model melalui *optimizer*. Terdapat beberapa parameter yang digunakan dalam melakukan proses pelatihan pada algoritma CNN, berikut Tabel 3.2 menjelaskan terkait parameter yang terdapat pada algoritma CNN.

Tabel 3. 2

Deskripsi Parameter

Parameter	Deskripsi
Dimensi embedding	Jumlah dimensi dari data input yang akan dilakukan pelatihan.
Filter	Jumlah filter/kernel yang digunakan.
Ukuran kernel	Ukuran jendela filter yang melintasi input.
Unit dense	Jumlah neuron yang saling terhubung untuk dilakukan proses klasifikasi.
Nilai dropout	Nilai regulasi untuk menghilangkan neuron selama pelatihan.
Learning rate	Parameter untuk melakukan kontrol terhadap seberapa besar perubahan yang dibuat terhadap

bobot jaringan selama
pelatihan.

3.1.4 Studi Deskriptif II

Tahap terakhir adalah studi deskriptif II, tahap ini merupakan tahap untuk mengidentifikasi apakah solusi sesuai dengan apa yang diharapkan atau tidak. Proses identifikasi ini dilakukan dengan mengevaluasi model yang telah dikembangkan. Evaluasi dilakukan dengan membandingkan hasil akurasi dan *loss* dari model yang telah dikembangkan. Selanjutnya dilakukan evaluasi dengan *confusion matrix* untuk mengetahui metrik akurasi, presisi, recall, dan f1-score yang dihasilkan model. Selain evaluasi dilakukan juga validasi untuk menilai kinerja model, validasi pada penelitian ini dilakukan menggunakan *k-fold cross validation* dengan *fold* sebanyak 10. Dari hasil validasi ini bisa didapatkan nilai standar deviasi yang menunjukkan perbedaan tiap akurasi yang dihasilkan oleh model.

3.2 Populasi dan Sampel

Pada penelitian ini populasi dapat berupa data umpan balik dari kuesioner evaluasi Proses Belajar Mengajar (PBM) di Universitas Pendidikan Indonesia. Namun, karena jumlah data umpan balik dari kuesioner evaluasi PBM berjumlah banyak maka akan diambil sampel untuk merepresentasikan data tersebut. Data umpan balik pada mata kuliah umum Bahasa Indonesia, Pendidikan Agama Islam, dan Pendidikan Kewarganegaraan akan menjadi sampel pada penelitian ini. Pemilihan ketiga mata kuliah tersebut karena mata kuliah tersebut merupakan mata kuliah umum sehingga data yang didapatkan akan banyak. Selain itu data yang akan digunakan merupakan data dengan rentang waktu 2020-2021.

3.3 Instrumen Penelitian

Instrumen penelitian yang digunakan pada penelitian ini adalah kuesioner evaluasi PBM. Kuesioner ini dibagikan melalui portal sistem akademik mahasiswa UPI (SIAKKu). Proses pengumpulan data dilakukan setiap akhir semester kepada mahasiswa. Pada kuesioner terdapat dua bagian, yaitu bagian pilihan ganda dan bagian saran yang berisi umpan balik selama perkuliahan. Umpan balik tersebut berisi opini mahasiswa selama perkuliahan yang dapat

menjadi salah satu komponen dalam evaluasi perkuliahan. Pengambilan data dilakukan melalui Direktorat Sistem dan Teknologi Informasi (STI) Universitas Pendidikan Indonesia. Adapun perangkat yang digunakan pada penelitian ini adalah sebagai berikut:

Tabel 3. 3
Perangkat yang Digunakan pada Penelitian

Perangkat	Keterangan
Perangkat keras	Laptop intel core i3
	RAM 8GB
	SSD 512
Perangkat lunak	Google Colaboratory
	Google Drive
	Google Chrome
	Python 3.10.0