

BAB III

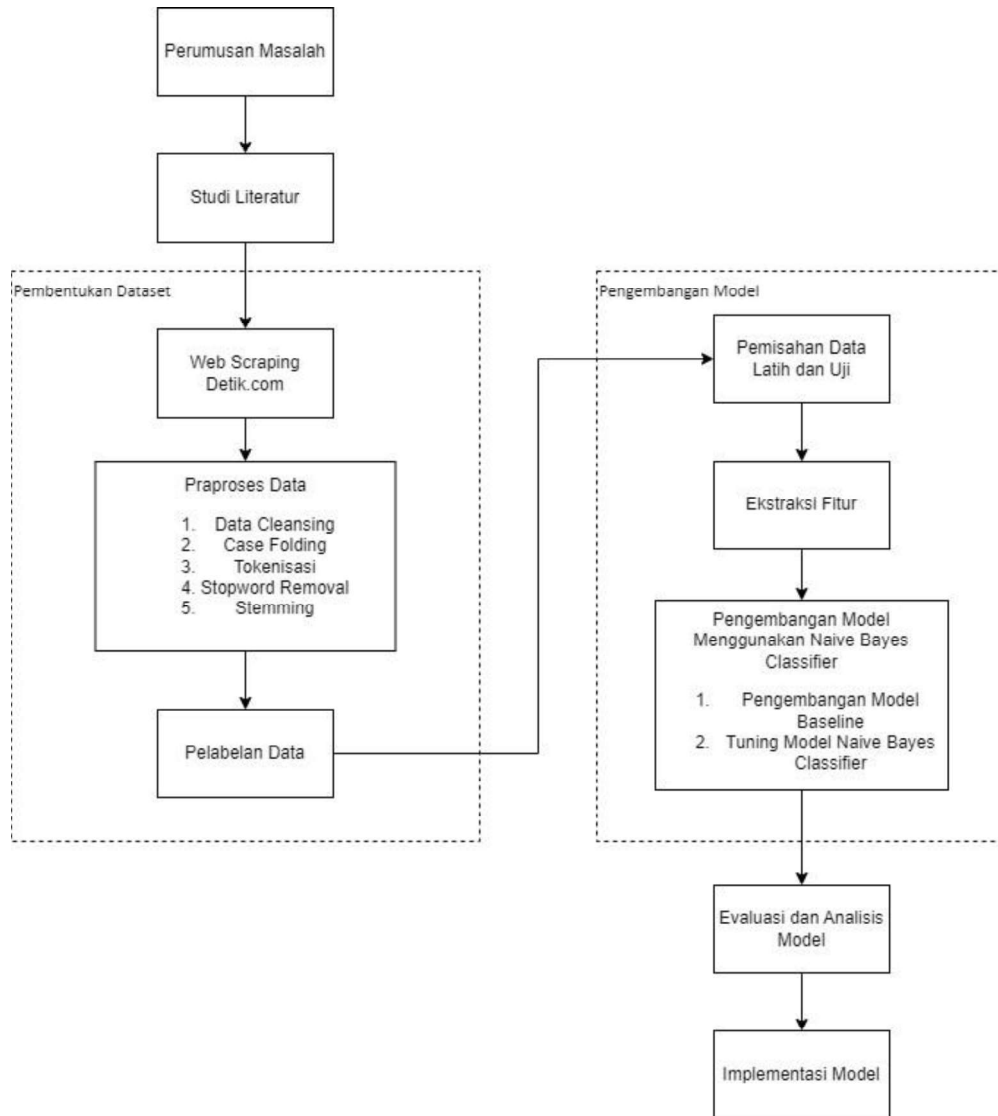
METODOLOGI PENELITIAN

3.1 Jenis Penelitian

Berdasarkan dari tujuan penelitian, maka jenis penelitian yang digunakan pada penelitian ini adalah penelitian kuantitatif. Penelitian kuantitatif adalah penelitian yang menghasilkan penemuan, yang dilakukan menggunakan prosedur statistik atau cara lain secara kuantitatif (pengukuran) (Tersiana, 2018). Pendekatan kuantitatif memusatkan perhatian pada gejala yang mempunyai karakteristik atau yang disebut variabel. Penelitian ini termasuk dalam kategori penelitian kuantitatif karena melibatkan pengumpulan data berita dalam jumlah besar yang diproses menjadi data numerik untuk analisis sentimen. Metode klasifikasi Naive Bayes, yang digunakan dalam penelitian ini, adalah teknik statistik yang memodelkan data dan mengklasifikasikannya ke dalam kategori seperti positif, netral, atau negatif. Penelitian ini juga mengevaluasi kinerja model menggunakan metrik kuantitatif seperti akurasi, presisi, *recall*, dan *F1 score*, memberikan hasil numerik yang menilai seberapa baik model dalam klasifikasi sentimen. Selain itu, penelitian ini melibatkan eksperimen terukur dan proses evaluasi yang memungkinkan reproduksi dan validasi hasil secara objektif. Semua elemen ini menunjukkan bahwa pendekatan kuantitatif diterapkan dalam analisis dan pengukuran performa model analisis sentimen berita.

3.2 Desain Penelitian

Desain penelitian adalah kerangka kerja yang digunakan untuk melakukan penelitian. Pada bagian ini penulis akan memaparkan kerangka kerja dari mulai penelitian sampai selesai. Desain penelitian diperlihatkan pada Gambar 3.1.



Gambar 3. 1 Desain Penelitian

Pada Gambar 3.1 dijelaskan alur penelitian yang akan dilakukan, penjelasan alur penelitiannya adalah sebagai berikut:

1. Perumusan Masalah

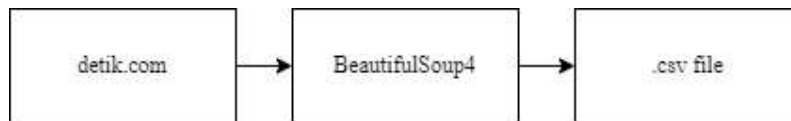
Pada tahapan ini bertujuan untuk mendefinisikan secara jelas dan spesifik permasalahan yang akan diteliti. Pada tahap ini, peneliti mengidentifikasi pertanyaan penelitian yang ingin dijawab dan menentukan tujuan, metode, serta ruang lingkup penelitian.

2. Studi Literatur

Studi literatur dilakukan oleh peneliti untuk melakukan kajian terhadap penelitian-penelitian terdahulu yang relevan dengan topik yang diteliti. Tujuan dari studi literatur adalah untuk memahami perkembangan terkini dalam bidang tersebut serta mengidentifikasi celah yang dapat diisi oleh penelitian ini. Literatur yang dikaji berkenaan dengan topik *text mining*, *natural language processing*, analisis sentimen, situs detik dan algoritma Naive Bayes.

3. Pembentukan *Dataset*

Dalam pembentukan *dataset* langkah pertama adalah mengumpulkan data berita di detik.com menggunakan metode *web scraping framework* yang tersedia di *library* Python bernama BeautifulSoup4. BeautifulSoup4 adalah sebuah *library* Python yang dirancang untuk *web scraping* dengan menggunakan teknik HTML *parsing* (Lawson, 2015). HTML *parsing* digunakan untuk mengambil konten yang terdapat di halaman HTML. Data dikumpulkan dengan memasukkan *starting* url dan HTML *tag* yang kontennya ingin dikumpulkan. Berikut adalah alur proses dari web scraping.



Gambar 3. 2 Proses Web Scraping

Langkah selanjutnya adalah melakukan praproses data yang prosesnya meliputi data cleansing, *case folding*, tokenisasi, dan *stemming*. Hasil praproses tersebut akan digunakan untuk pelabelan data menggunakan *library* NLTK dari bahasa pemrograman python.

4. Pengembangan Model Naive Bayes *Classifier*

Pada tahap ini, peneliti mengembangkan model Naive Bayes dengan melakukan proses pemisahan data tes dan uji, ekstraksi fitur untuk mengubah data menjadi vektor, pengembangan model Naive Bayes dasar, dan *tuning* model Naive Bayes yang telah dikembangkan.

5. Implementasi Model Naive Bayes *Classifier*.

Setelah model dikembangkan, tahap ini fokus pada penerapan model untuk melakukan klasifikasi terhadap data uji. Peneliti menerapkan model yang telah dikembangkan dan menganalisis hasil klasifikasi yang diperoleh.

6. Evaluasi dan Analisis Hasil

Tahap ini bertujuan untuk mengevaluasi kinerja model dengan menggunakan metrik evaluasi seperti akurasi, *precision*, *recall*, F1-score, dan *confusion matrix*. Analisis hasil dilakukan untuk menilai efektivitas model dan memberikan wawasan terkait hasil yang diperoleh dari implementasi model.

7. Penarikan Kesimpulan

Pada tahap akhir, peneliti menarik kesimpulan dari temuan penelitian dan hasil evaluasi. Ini mencakup ringkasan temuan utama, implikasi dari hasil penelitian, serta rekomendasi untuk penelitian lebih lanjut atau aplikasi praktis dari hasil penelitian yang telah dilakukan.

3.3 Evaluasi Model

Evaluasi model adalah proses pengujian evaluasi model untuk mengetahui apakah model yang telah dibuat sudah akurat atau belum. *Confusion matrix* merupakan salah satu evaluasi model yang umum dipakai saat mengevaluasi model dari *machine learning* khususnya klasifikasi Naive Bayes. *Confusion matrix* adalah tabel matrix 2x2 yang merepresentasikan hasil klasifikasi biner pada suatu *dataset* (Andika et al., 2019).

<i>Class</i>	<i>Predicted</i>	
	<i>Positive</i>	<i>Negative</i>
<i>Actual Positive</i>	<i>True Positive</i> (TP)	<i>False Negative</i> (FN)
<i>Actual Negative</i>	<i>False Positive</i> (FP)	<i>True Negative</i> (TN)

Tabel 3. 1 *Confusion Matrix*

1. *True Positive* = Kelas data yang aktualnya positif, model juga memprediksi positif.
2. *False Negative* = Kelas data yang aktualnya positif, tapi model memprediksi negatif

3. *False Positive* = Kelas data yang aktualnya negatif, tapi model memprediksi positif.
4. *True Negative* = Kelas data yang aktualnya negatif, model juga memprediksi negatif

Dari hasil *confusion matrix* diatas maka pengukuran model Akurasi (Acc), Presisi (P), *Recall* (R), dan F-score (F1) bisa diukur.

1. Akurasi (*Accuracy*):

$$\text{Akurasi} = (TP + TN) / (TP + TN + FP + FN)$$

2. Presisi (*Precision*):

$$\text{Presisi} = TP / (TP + FP)$$

3. *Recall* (*Sensitivity* atau *True Positive Rate*):

$$\text{Recall} = TP / (TP + FN)$$

4. F1-Score:

$$\text{F1-Score} = 2 * (\text{Presisi} * \text{Recall}) / (\text{Presisi} + \text{Recall})$$

3.4 Lingkungan Komputasi

Penelitian ini akan membutuhkan beberapa alat untuk membantu jalannya penelitian. Alat-alat ini berupa perangkat keras dan perangkat lunak. Berikut alat-alat tersebut:

3.4.1 Perangkat Keras

1. Spesifikasi Komputer
 - i. Prosesor: Intel® Core™ i5-8300H
 - ii. RAM: 16GB
 - iii. Sistem Operasi: Windows 10 Version 20H2

3.4.2 Perangkat Lunak

1. Google Colaboratory
2. Python 3.8.2
3. Google Chrome