

## **BAB III**

### **METODE PENELITIAN**

#### **3.1 Alat dan Bahan Penelitian**

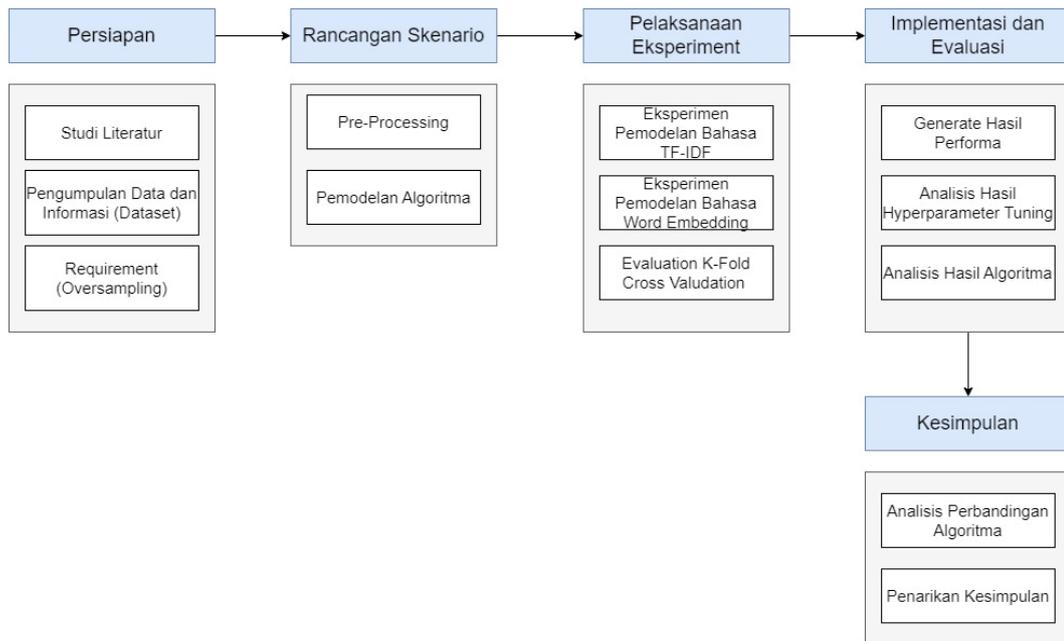
Pada Penelitian ini peneliti menggunakan alat dan bahan sebagai berikut :

1. Perangkat Keras (*Hardware*)
  - a. Laptop Asus Core I7
  - b. Hardisk 1 TB
  - c. Wifi Biznet 100 Mbps
2. Perangkat Lunak (*Software*)
  - a. Node JS
  - b. Google Colab
  - c. OS Windows 11
  - d. Ms. Word 2019
  - e. Ms. Excel 2019
  - f. Ms. Power Point 2019

#### **3.2 Tahapan Penelitian**

Ini termasuk algoritma, rute, model, dan desain yang terkait dengan desain sistem.

Gambar 2 di bawah ini menunjukkan alur investigasi data Twitter ke dalam sistem analisis sentimen menggunakan presentasi online kata kunci Undang-Undang Ibu Kota Nusantara.



Gambar 3. 1 Tahapan Penelitian

1. Pada fase persiapan ini, peneliti membaca buku, jurnal, esai, dan laporan penelitian tentang topik penelitian mereka. Setelah itu, kumpulkan data dan informasi berikut: menggunakan kata kunci Undang-Undang Ibu Kota Nusantara untuk memperoleh data dari Twitter. Selain itu, pendefinisian kebutuhan dilakukan dengan mengidentifikasi data yang diperlukan, menampilkan langkah-langkah yang sedang berjalan, menganalisis sistem yang sedang berjalan, dan menghasilkan hasil evaluasi sistem.
2. Rancangan Skenario Pada tahap ini dilakukan perancangan untuk mengolah data yang diperoleh dari proses sebelumnya. Ada dua fase dalam fase ini. Artinya, oversampling atau pengambilan sampel dataset dari penelitian sebelumnya sebagai referensi dan melakukan proses preprocessing untuk mengelola data yang digunakan dan diproses dalam format atau sistem yang ideal.

3. Pelaksanaan Experiment Fase ini mencakup proses desain sistem dan perangkat lunak seperti inferensi *Naive Bayesian* dan pemodelan mesin vektor pendukung, serta pemodelan proses dengan alur terperinci dari awal hingga akhir dan juga pemodelan Bahasa *Tf-Idf* dan juga *Word Embedding*.
4. Implementasi dan Evaluasi Pada tahap ini, hasil penelitian diperoleh dengan membandingkan kedua algoritma dan mengukur kinerja akurasi, akurasi, memori, dan waktu pelatihan dalam bentuk angka, tabel, dan grafik.
5. Kesimpulan Pada fase ini, proses pendokumentasian dan penerbitan disertasi mencakup hasil penelitian yang dilakukan.

### **3.3 Metode yang digunakan**

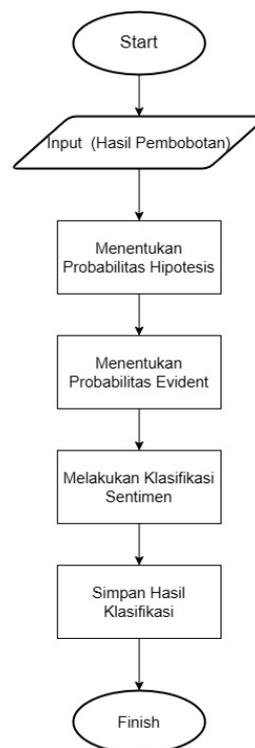
Peneliti melakukan review literatur tentang klasifikasi *Bayesian* dan *Support Vector Machine*. Kedua metode ini biasa digunakan untuk pembelajaran mesin untuk memprediksi atau memprediksi sesuatu dalam penelitian ini. Saat melakukan analisis sentimen data Twitter dengan kata kunci UU Ibu Kota Nusantara, data yang diuji dalam penelitian ini adalah akurasi, presisi, recall, dan waktu pelatihan dari perbandingan algoritma *Naive Bayes* dengan *Support Vector Machine*.

#### **3.3.1 Pemodelan Algoritma**

Pada fase ini dilakukan setiap model algoritma, *support vector machine* dan metode pemodelan *Naive Bayes*. Rincian lebih lanjut tentang bagaimana model diberikan dalam subbagian berikutnya.

## 1. *Naïve Bayes*

Pada tahap ini, pemodelan model algoritmik dilakukan dengan menambahkan model pada konfigurasi yang ditunjukkan pada Gambar 3.3 atau *arsitektur Naive Bayes*. Alur yang dibuat :

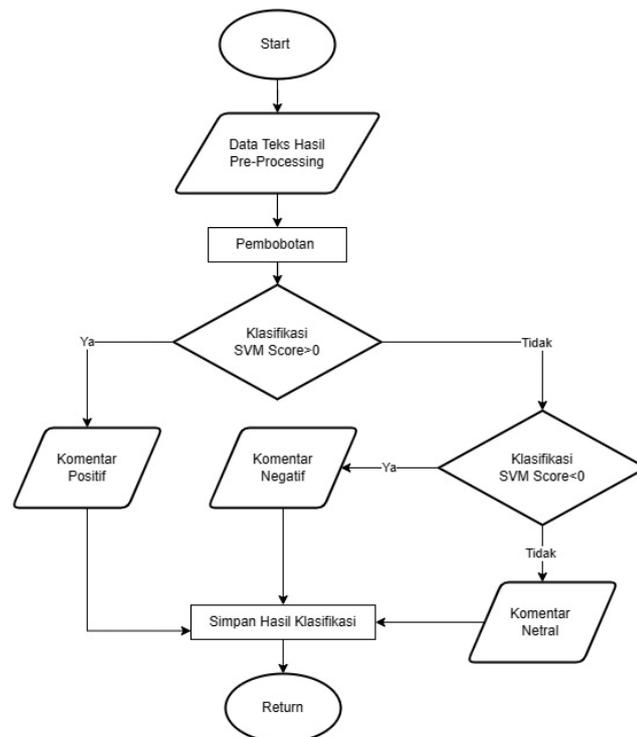


Gambar 3. 2 Model Arsitektur Arsitektur Naive Bayes

Pada Gambar 3 proses analisis dilakukan dengan menggunakan *Naïve Bayes*, dokumen berupa matriks word-document, dan setelah mendapatkan pembobotan TF-IDF, proses selanjutnya adalah analisis *Naïve Bayes*. H. Setiap dokumen ditandai sebagai positif, negatif, dan netral. Proses Naive Bayes kemudian dijalankan dan bobot antara data uji dan semua data latih dalam dokumen artikel dihitung menurut metode probabilistik.

## 2. Support Vector Machine

Pada fase ini dilakukan pemodelan model algoritmik dengan membuat model atau mendukung arsitektur mesin vektor dengan konfigurasi seperti pada Gambar 4 :



Gambar 3. 3 Model Arsitektur SVM

Pada Gambar 4 proses analisis dilakukan dengan menggunakan SVM dan dimulai dengan mengubah teks menjadi data vektor. Vektor penelitian ini memiliki dua komponen yaitu dimensi (word ID) dan bobot. Bobot ini adalah nilai TFIDF yang menggunakan model ruang vektor untuk memberikan bobot setiap kata dalam dokumen berdasarkan ID (dimensi) dan kepentingannya dalam dokumen (posisi dokumen di dalamnya). Ukuran). SVM mencoba menemukan baris

yang paling baik membagi tiga kelas dan mengkategorikan dokumen pengujian berdasarkan sisi baris mana yang muncul.

### **3.4 Evaluasi Metode**

#### **3.4.1 Metode Pengumpulan Data**

Selain itu, penelitian ini menggunakan tahap eksperimen sebagai metode pengumpulan data, namun selain itu peneliti juga mengumpulkan data dengan menggunakan Node JS yang di install pada google colab lalu dilanjutkan dengan memasukan API twitter menggunakan Bahasa pemrograman python. Namun pada tahap ini peneliti melakukan Teknik pengambilan data secara bertahap karena twitter membatasi dalam pengambilan data. Tahap eksperimen dilakukan dengan membandingkan model algoritma *Naïve Bayes* dengan *Support Vector Machine* sebagai cara untuk membandingkan. Selanjutnya hasil perbandingan kedua model algoritma tersebut dievaluasi berdasarkan akurasi, akurasi, recall, dan waktu pelatihan, serta hasil perbandingan tersebut. Proses pengumpulan data dimulai dengan menggunakan Twitter API untuk mendapatkan parameter yang sedang diselidiki dan dilanjutkan dengan proses pengumpulan data untuk objek yang diselidiki.

Dalam hal ini peneliti mendapatkan data Twitter berdasarkan kata kunci Undang-Undang Ibu Kota Nusantara yang disebut dengan dataset. Dataset yang digunakan adalah 1300 Dataset Twitter. Dataset yang diperoleh akan digunakan sebagai model data untuk membandingkan dua model algoritma yang sedang diselidiki. Berikut sampel tweet positif, negatif dan netral pada tabel 3.

Tabel 3. 1 Sampel Data Tweet

Id	Tweet	Kategori		
		Positif	Negatif	Netral
135585343	Para konglomerat Indonesia turut serta dalam pembangunan Ibu Kota Nusantara (IKN). Otorita IKN mengungkap 15 investor bakal melakukan groundbreaking proyek baru di IKN pada Januari hingga Februari 2024 mendatang.	✓		
159000034	Selain itu Prabowo-Gibran juga berkomitmen terhadap pembangunan infrastruktur serta melanjutkan pembangunan Ibu kota Nusantara (IKN)	✓		
886000045	Dah peras APBN lalu sedot anggaran daerah pula. Kementerian Dalam Negeri meminta Pemerintah Provinsi DKI Jakarta menyediakan alokasi anggaran dalam APBD 2024 untuk kepindahan ibu kota negara dari Jakarta ke Ibu Kota Nusantara (IKN) di Kalimantan Timur		✓	
68930552	WARTEG DILARANG DI IKN IBU KOTA NUSANTARA ini LUCU KEDENGARAN NYA		✓	
255866913	Simak cerita penginapan milik Jatmiko atau yang biasa disapa Kapon. Berada sekitar 10 kilometer dari Titik Nol Ibu Kota Nusantara atau IKN.			✓

	Tepatnya di Desa Bukit Raya Kecamatan Sepaku Penajam Paser Utara Kalimantan Timur.			
298641815	Konsep Struktur dalam Otorita Ibu Kota Nusantara Berbeda dengan Pemerintah Daerah pada Umumnya			✓

### 3.4.2 Metode Analisis Data

Untuk penelitian ini, analisis data dilakukan dengan membandingkan hasil eksperimen dari awal hingga akhir. Percobaan dimulai dengan membersihkan atau mensanitasi data. Langkah selanjutnya adalah mengubah informasi. Pada tahap ini, pengelompokan dilakukan untuk memecah menjadi beberapa kelompok atau kelompok data. Setelah data dikelompokkan, langkah selanjutnya adalah pemodelan data. Proses ini mengubah data yang dikelompokkan menjadi nilai komposisi terpisah atau nilai komposisi terpisah (CSV) sebagai format data yang dimasukkan ke dalam database. Setelah proses pemodelan data atau data dalam database input selesai, langkah selanjutnya adalah membandingkan data yang diperoleh dengan dua model algoritma: Naive Bayes dan Support Vector Machine. Pembagian sepuluh subset dan validasi cross-K-FOLD digunakan untuk membandingkan kedua model algoritme ini. Tujuan skema ini adalah untuk mengukur konsistensi performa terhadap jumlah data yang besar., dikarenakan jumlah data tersebut mempengaruhi proses training dan pembobotan pada setiap algoritma. Metode ini secara acak membagi subset dari K. Setiap subset  $n = 1$  adalah data latih dan  $n = 2, 3, \dots, n$  adalah data uji. Kemudian dilakukan iterasi K menggunakan data uji  $n+1$ , dan kondisi data uji tidak sama dengan kondisi data

latih. Dengan mengukur akurasi, memori, dan nilai kinerja waktu pelatihan, ditarik kesimpulan dari hasil evaluasi dua model algoritme yang dipilih. Pengukuran tersebut digunakan sebagai acuan atau pedoman dalam mengidentifikasi hasil dan menganalisis emosi dalam