

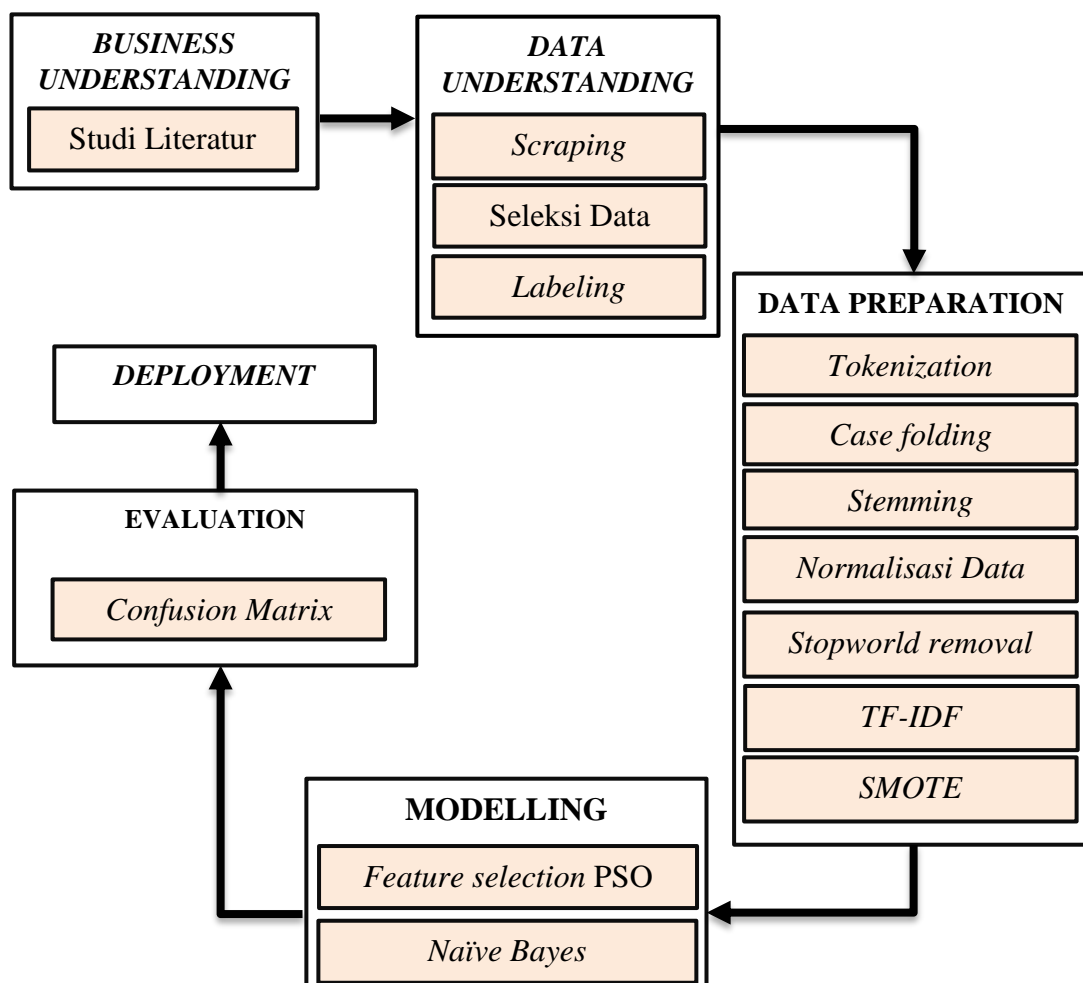
## BAB III

### METODOLOGI

Metodologi penelitian menjelaskan tahapan-tahapan yang digunakan dalam penelitian “*Sentiment Analysis Terhadap Pembelajaran Jarak Jauh Pada Masa Pandemi Covid-19 dengan Metode PSO (Particle Swarm Optimization) dan Naïve Bayes*”

#### 1.1 Tahapan Penelitian

Langkah langkah pada penelitian divisualisasikan pada gambar berikut,



Gambar 3.1 Tahapan Penelitian dengan CRISP-DM

Gambar 3.1 merupakan langkah-langkah dalam penelitian pengenalan pola yang dilakukan. Langkah-langkah tersebut diuraikan sebagai berikut :

### **1.1.1 Business Understanding**

Tahapan ini diperlukan pemahaman mengenai objek penelitian dengan menggali informasi dari media sosial yaitu data *tweet*, yang bersumber dari twitter mengenai *sentiment analysis* tentang pembelajaran jarak jauh. Postingan di twitter bisa mengetahui cuitan masyarakat, klasifikasi dilakukan untuk melihat *tweet* yang bernilai positif dan negatif. Berdasarkan ini, pendekatan model klasifikasi *tweet* yang akan digunakan algoritma *Naïve Bayes* berbasis *Particle Swarm Optimization*.

### **1.1.2 Data Understanding**

Proses pengambilan data mentah twitter dilakukan dengan teknik *scraping* sesuai dengan kata kunci yang dibutuhkan. Data yang digunakan adalah tweet bahasa Indonesia, data yang diambil dari periode 16 Januari 2022 sampai 18 maret 2022 mengenai pembelajaran jarak jauh. Data yang didapat dilakukan seleksi sesuai kebutuhan untuk dilakukan pembersihan dan diberi label sentimen yaitu *sentiment positif* dan *sentiment negatif* secara manual.

### **1.1.3 Data Preparation**

Proses mempersiapkan data untuk melakukan langkah-langkah yang disebut dengan *text preprocesing*. Proses yang dilakukan diantaranya:

1. *Tokenize*, proses pemotongan sebuah string kalimat menjadi bagian- bagian atau sekumpulan karakter akan dipecah menjadi satuan kata.

2. *Case folding*, proses penyamaan case dalam sebuah dokumen atau mengubah setiap huruf dalam post menjadi huruf kecil secara keseluruhan atau lowercase
3. *Stemming*, merupakan proses mengubah kata yang berimbuhan menjadi kata dasar.
4. Normalisasi Data, mengubah kata tidak baku, singkatan, bahasa asing menjadi bentuk kata baku.
5. *Stopword Removal*, memproses agar kata-kata yang tidak penting atau tidak bermakna dihapus untuk analisis sentimen. Contoh kata-kata tersebut adalah atau, yang, dengan, di, ke, dan tetapi.
6. TF-IDF, merupakan pembobotan tiap-tiap kata agar dapat menaikkan kemampuan analisis sentimen pada proses text mining.
7. *SMOTE*, memproses data setelah tahapan preprocessing, maka selanjutnya dengan proses metode *SMOTE* sebagai penyeimbang kelas data dengan mereplikasi kelas minoritas sebanyak kelas mayoritas.

#### **1.1.4 Modelling**

Proses ini secara langsung melibatkan teknik data mining yaitu dengan cara melakukan pemilihan teknik data mining dan menerapkan algoritma yang akan digunakan. *Tool* yang digunakan adalah RapidMiner. Hasil dalam pengujian model yang dilakukan adalah mengklasifikasi sentimen positif dan sentimen negatif menggunakan algoritma *Naïve Bayes* dan *Particle Swarm Optimization* (PSO) sebagai *feature selection* untuk meningkatkan bobot dari atribut terhadap seluruh atribut yang digunakan, sehingga akurasi yang dihasilkan lebih optimal.

### **1.1.5 Evaluation**

Proses ini untuk melihat hasil dari model yang telah dibuat menggunakan *Confusion Matrix* untuk memberikan informasi perbandingan hasil yang telah dilakukan oleh model dengan hasil klasifikasi sebenarnya dengan melihat nilai akurasi, *presisi* dan *recall*.

### **1.1.6 Deployment**

*Deployment* dilakukan untuk menampilkan hasil penelitian agar lebih mudah dipahami dan memaparkan hasil penelitian berdasarkan dokumen berlabel positif dan negatif, sehingga menjadi informasi yang berguna.

