# BAB III

# METODOLOGI PENELITIAN

## 3.1 Gambaran Umum Objek Penelitian

Objek penenelitian yang akan digunakan pada penelitian ini adalah data pada sosial media *Twitter* dalam bentuk *Text* yang telah dilakukan *Text Mining* dengan menggunakan aplikasi *Rapidminer* pada komentar dan postingan mengenai mata uang kripto khususnya *Bitcoin, Ethereum, Binance Coin, DOGE, dan Ripple.*

## 3.2 Variabel Penelitian

### 3.2.1 Variabel Independen

Variabel Independen adalah variabel yang menjadi penyebab terjadinya atau terpengaruhnya suatu variabel terikat atau variable idependen [34]. Berikut adalah yang merupakan variabel independen atau bebas pada penelitian ini adalah:

1. Postingan mengenai mata uang kripto atau *cryptocurrency* di media sosial *Twitter*.
2. Komentar dan tanggapan mengenai mata uang kripto atau *cryptocurrency* di sosial media *Twitter*.

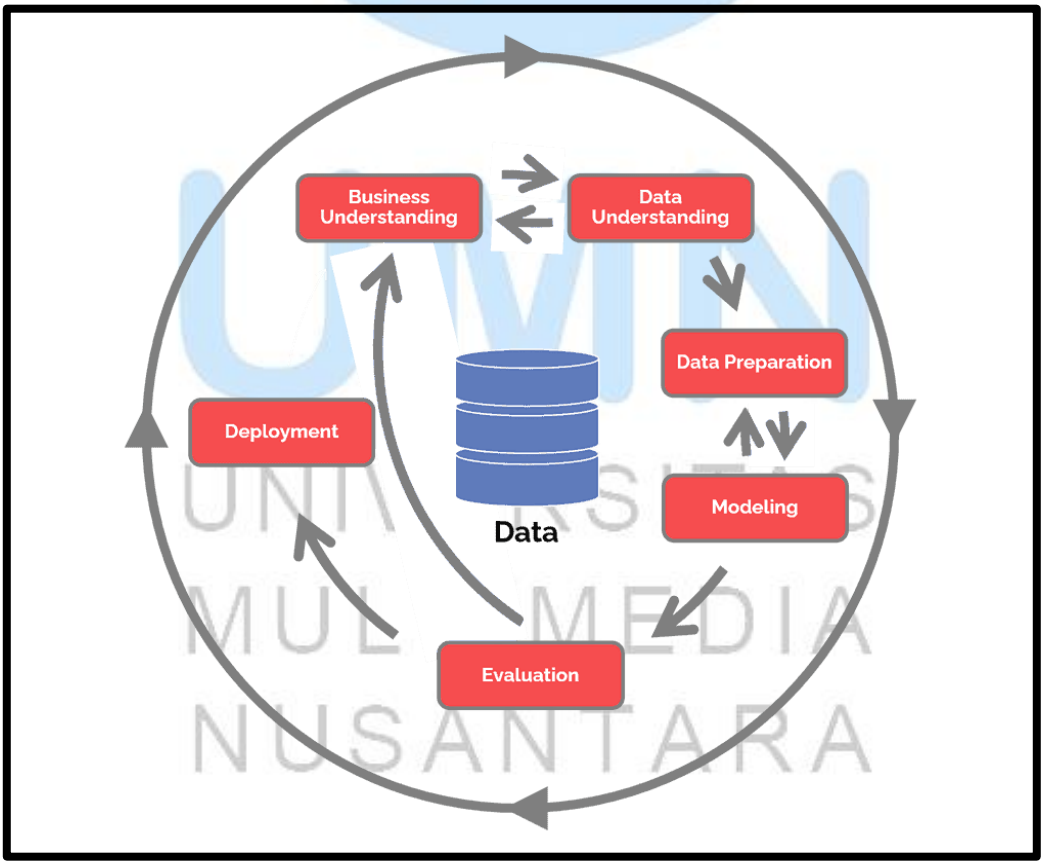
### 3.2.2 Variabel Dependen

Variabel dependen adalah variabel terikat atau variabel dependen yang dipengaruhi karena adanya variabel bebas atau variabel independent [34]. Variabel dependen atau variabel terikat pada penelitian ini adalah klasifikasi sentimen yang terdiri positif dan negatif terhadap mata uang *kripto Bitcoin, Ethereum, Binance Coin, DOGE,* dan *Ripple* pada komentar dan positngan di sosial media *Twitter*

## 3.3 Metode Pengumpulan *Data*

Metode pengumpulan *data* pada penelitian ini menggunakan *Rapid Miner* untuk melakukan *text mining* dan akan mengambil postingan dan komentar di sosial media *Twitter* yang terdiri dari koin *Bitcoin, Ethereum, Binance Coin, DOGE,* dan *Ripple.* Data yang akan ditargetkan untuk dikumpulkan adalah 7500 *data. Data* dikumpulkan dari tanggal 1 Agustus 2021 sampai dengan tanggal 22 Agustus 2021 dan dilakukan *Text Mining* setiap minggunya. Data ditargetkan sebanyak 7500 dikarenakan pada penelitian sebelumnya yang digunakan sebagai referensi menggunakan data sebanyak 3000, penelitian ini menggunakan data sebanyak 7500 agar hasil yang didapat lebih akurat [10].

## 3.4 Metode Penyelesaian Masalah




**Gambar 3.1 Gambar Metode Penyelesaian Masalah CRISP – DM** [35]

Gambar 3.1 diatas adalah gambar metode penyelesaian masalah dengan kerangka CRISP – DM. Metode yang digunakan dalam melakukan penelitian adalah *CRISP - DM* yang dibagi menjadi 6 tahapan yaitu:

### 3.4.1 *Business Understanding Phase*

Tujuan dari penelitian ini adalah mencari urutan mata uang kripto berdasarkan klasifikasi sentimen positif yang terdiri dari Bitcoin, Ethereum, Binance Coin, DOGE, dan Ripple. Penelitian ini juga akan melakukan perbandingan algoritma klasifikasi untuk mencari performa terbaik dari algoritma klasifikasi yang teridiri dari algoritma Support Vector Machine dan Naïve Bayes.

### 3.4.2 *Data Understanding Phase*

Pada tahapan ini akan dilakukan Text Mining pada sosial media Twitter dengan menggunakan aplikasi yang bernama Rapidminer. Data yang akan dikumpulkan melalui sosial media Twitter adalah data yang mengenai mata uang kripto Bitcoin, Ethereum, Binance Coin, DOGE, dan Ripple. Proses Text Mining ini dilakukan sebanyak 1 kali dalam seminggu yang dimulai dari tanggal 1 Agustus 2021 sampai dengan 22 Agustus 2021 dikarenakan target data yang akan dikumpulkan sudah tercapai. Data ditargetkan sebanyak 7500 dikarenakan pada penelitian sebelumnya yang digunakan sebagai referensi menggunakan data sebanyak 3000, penelitian ini menggunakan data sebanyak 7500 agar hasil yang didapat lebih akurat [10].

### 3.4.3 *Data preparation Phase*

Pada tahapan penelitian ini, data yang sebelumnya telah dikumpulkan akan dioleh dan dibersihkan sehingga data tersebut layak untuk digunakan. Tahapan ini terdiri dari 3 proses yaitu:

*1) Data Union*

Pada proses ini data – data yang sebelumnya sudah dikumpulkan akan digabungkan untuk setiap koin.

*2) Data Cleansing*

Data yang sudah digabungkan akan dibersihkan dengan menggunakan beberapa langkah – langkah yaitu:

*I. Remove URL*

Menghapus seluruh data yang berisikan link atau website pada data *Twitter*

*II. Remove Non Alfabet*

Menghapus seluruh lambang pada data *text Twitter.*

*III. Remove RT*

Menghapus seluruh data *Twitter* yang berisikanan kata *RT.*

*IV. Remove Duplicate*

Melakukan pembuangan pada data yang sama dan duplikat.

*3) Data Labelling*

Memberikan label keterangan positif, negatif, dan netral untuk nantinya dilakukan uji klasifikasi menggunakan algoritma *Support Vector Machine* dan *Naïve Bayes*. Berikut adalah tahapan dari *Data Labeling:*

*I. Extract Sentiment*

Menggunakan algoritma *VADER* untuk memberikan *score* atau nilai pada suatu data Tweets yang sebelumnya sudah melewati tahapan *Data Union* dan *Data Cleansing.*

*II. Labelling Data*

Memberikan label pada data yang sebelumnya sudah diberikan *score* atau nilai. Label positif diberikan pada data dengan *score* positif, label negatif pada data *score* negatif, dan label netral pada data dengan score 0.

III. Penghapusan Klasifikasi

Menghapuskan data dengan label netral dikarenakan pada penelitian ini hanya ada 2 jenis klaisifkasi yaitu positif dan negatif.

### 3.4.4 *Modeling Phase*

Pada tahapan ini akan dipilih model algoritma klasifikasi yaitu algoritma *Support Vector Machine* dan *Naïve Bayes* dalam melakukan analisa sentimen. *Confusion Matrix* digunakan untuk menghitung akurasi dan ketepatan algoritma.

### 3.4.5 *Evaluation Phase*

Pada tahapan ini akan mendapatkan hasil dari perhitungan dalam menggunakan *Confusion Matrix* yang terdiri dari *accuracy, precision, dan recall*. Hasil dari setiap *Confusion Matrix* dari algoritma klasifikasi *Support Vector Machine* dan *Naïve Bayes* akan dibandingkan dengan mata uang kripto seperti *Bitcoin, Ethereum, Binance Coin, DOGE, dan Ripple.*

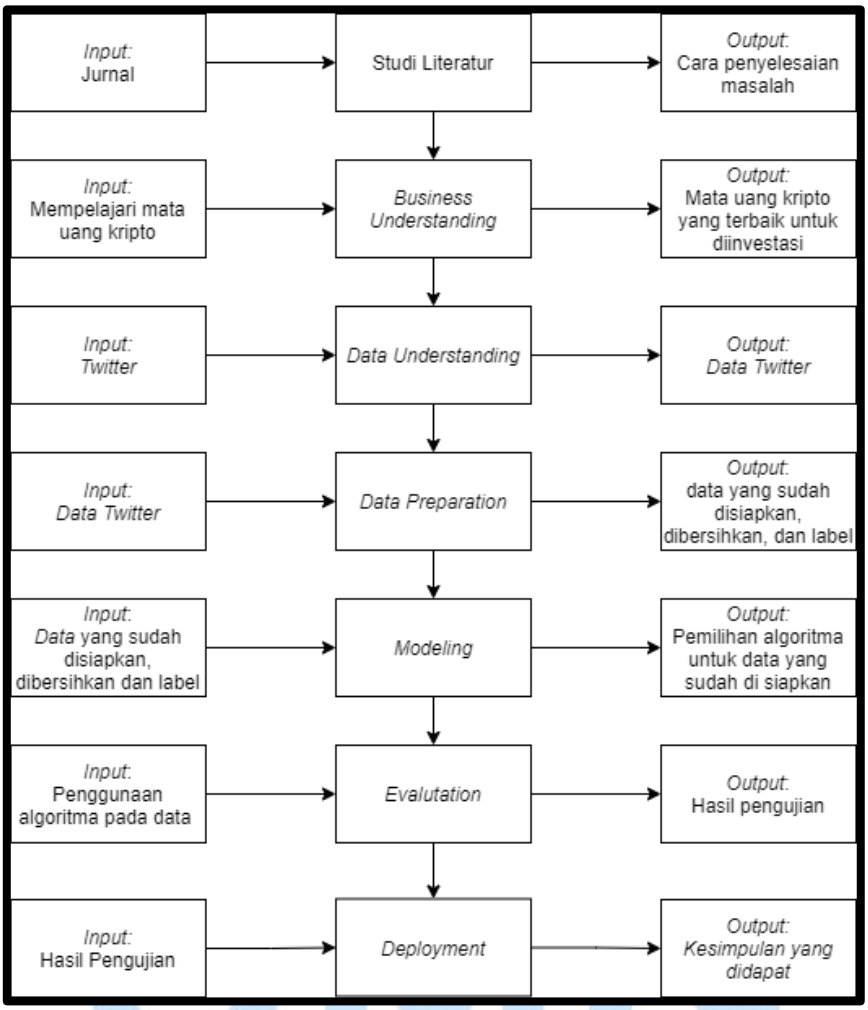
### 3.4.6 *Deployment Phase*

Pada tahapan ini akan ditampilkan suatu visualisasi dalam bentuk *Pie Chart* dan *Bar Chart* dari seluruh data yang didapat dan hasil klasifikasi dan perbandingan performa dari algoritma *Support Vector Machine* dan *Naïve Bayes*.

## 3.5 Kerangka Pikir

Dalam melakukan penelitian ini, penulis menggunakan kerangka piker sebagai berikut.

UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA




**Gambar 3.2 Kerangka Pikir**

Gambar 3.2 diatas adalah kerangka pikir yang terinspirasi dari penelitian terdahulu yang sebelumnya telah dipelajari oleh penulis. Berikut adalah penjelasan mengenai langkah – langkah penelitian pada kerangka pikir penelitian ini:

1) Studi Literatur

   ***Input:*** Mengumpulkan jurnal – jurnal penelitian terdahulu

   ***Process:*** Mempelajari penemuan pada jurnal penelitian terdahulu

   ***Output:*** Memperoleh metode penyelesaian masalah dalam analisa sentimen

2) *Business Understanding*

***Input:*** Informasi mengenai mata uang kripto khususnya dengan kapitulasi pasar terbesar

***Process:*** *Business Understanding* memahami inti dari suatu kegiatan melalui sudut pandang dalam melakukan suatu bisnis.

***Output:*** Mencari mata uang kripto yang terbaik untuk dilakukan investasi dengan menggunakan analisa sentimen

3) *Data Understanding*

***Input:*** Data yang tersedia di sosial media *Twitter* mengenai *Bitcoin, Ethereum, Binance Coin, DOGE,* dan *Ripple*

***Process:*** *Date Preparation* melakukan *Text Mining* mengenai *Bitcoin, Ethereum, Binance Coin, DOGE,* dan *Ripple* pada sosial media *Twitter*

***Output:*** *Data Twitter* yang sudah berhasil dikumpulkan dengan menggunakan *Rapidminer* dalam melakukan *Text Mining.*

4) *Data Preparation*

***Input:*** *Data Twitter* yang telah dikumpulkan dengan *Rapidminer* dalam melakukan *Text Mining*

***Process:*** Melakukan 3 tahapan dalam *Data Preparation* yaitu: *Data Union, Data Cleansing, dan Data Labelling.*

***Output:*** Data yang sudah melewati proses *Data Union, Data Cleasing,* dan *Data Labelling*

5) *Modeling*

***Input:*** Data yang sudah melewati proses *Data Union, Data Cleasing,* dan *Data Labelling*

***Process:*** Modelling yaitu memilih model algoritma klasifikasi *Support Vector Machine* dan *Naïve Bayes*

***Output:*** Menggunakan algoritma klasifikasi *Support Vector Machine* dan *Naïve Bayes* dalam melakukan analisa sentimen.

6) *Evaluation*

***Input:*** Menggunakan algoritma klasifikasi *Support Vector Machine* dan *Naïve Bayes* dalam melakukan analisa sentimen.

***Process:*** Evaluation yaitu melakukan perhitungan analisa sentimen dengan menggunakan algoritma klasifikasi *Support Vector Machine* dan *Naïve Bayes*

***Output:*** Hasil yang telah didapatkan berupa Confusion Matrix yang terdiri dari accuracy, precision dan recall pada algoritma Support Vector Machine dan Naïve Bayes.

7) *Deployment*

***Input:*** Hasil yang telah didapatkan berupa Confusion Matrix yang terdiri dari accuracy, precision dan recall pada algoritma Support Vector Machine dan Naïve Bayes.

***Process:*** Deployment yaitu melakukan visualisasi data terhadap penelitian yang telah dilakukan dalam bentuk *Pie Chart* dan *Bar Chart*

***Output:*** Visualisasi dari kesimpulan yang telah didapatkan dari penelitian yang telah dilakukan dalam bentuk *Pie Chart* dan *Bar Chart*

UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA