# BAB III

# METODOLOGI PENELITIAN

## Data Penelitian

*Dataset* atau data yang digunakan dalam penelitian ini bersumber dari Twitter berupa data teks kicauan (*tweet)* yang diperoleh dari tanggal 1 Desember 2020 hingga 31 Januari 2021. Data tersebut diperoleh menggunakan pustaka Tweepy melalui proses *crawling*. *Dataset* yang diperoleh dikumpulkan berdasarkan beberapa parameter kata kunci yang terkait dengan sistem pembelajaran daring antara lain: X, Y dan Z. Berikut beberapa contoh data *tweet* hasil dari proses *crawling* yang dapat dilihat pada Tabel X berikut:

**Tabel X Sampel data *tweet***

| *Tweet ID* | *Username* | *Tweet* | Waktu *tweet* |
| --- | --- | --- | --- |
| 1335989354792103936 | LRomdani | Tetap memakai masker meski dirumah sendiri  Tetap semangat belajar dari rumah dimasa pandemi  #DiktiMengajarDariRumah  #DiktiDutaEdukasiPerubahanPrilaku https://t.co/c1WMa5SVSj | 2020-12-07 16:47:38 |
| 1336520460255724032 | kelaskitadotcom | Gunakan hak suara kamu dengan bijak, ya! Selamat memilih! #kelaskita #carabarubelajarseru #belajardirumah #elearning #belajaronline #dirumahaja #quotes https://t.co/1anyTiETlA | 2020-12-09 03:58:03 |
| 1336742494122340096 | fandimas16 | @collegemenfess 1. Jenuh banget di rumah 2. Gw dri dulu suka ama suasana kelas, dan suasana itu mendukung gw untuk belajar dan memahami suatu materi | 2020-12-09 18:40:20 |
| 1338003730587812096 | kumparan | Tanpa smartphone di masa pandemi, bisa berarti putus sekolah, karena kini dilakukan belajar online atau pembelajaran jarak jauh. https://t.co/rVW6xOgrfI | 2020-12-13 06:12:02 |

## Penerapan Metode

Dalam membangun aplikasi analisis sentimen yang dilakukan pada penelitian ini, terdapat beberapa tahapan yang dilakukan. Tahapan tersebut merepresentasikan setiap proses dan rancangan dalam penelitian, dari awal hingga akhir aplikasi berjalan. Tahapan yang dilakukan dapat dilihat pada Gambar X berikut:



**Gambar X Tahapan metode**

Pada Gambar X, pengumpulan data dilakukan melalui proses *crawling* untuk mendapatkan *dataset* berupakicauan (*tweet*). Selanjutnya, *tweet* yang telah diperoleh dalam bentuk *excel*, kemudian dimasukkan kedalam basis data (*database*) untuk dilakukan proses *preprocessing*, pada proses *preprocessing* dilakukan penyaringan, pembuangan dan perbaikan kata. Hasil dari proses *preprocessing* menghasilkan kalimat yang lebih terstruktur (*clean text*) yang kemudian digunakan pada tahap selanjutnya. *Clean text* yang diperoleh dari proses *preprocessing* akan diproses dalam tahap *labeling* untuk menentukan kelas (*label*) berupa sentimen positif atau negatif, kemudian *tweet* yang telah ber*label* akan dibagi menjadi dua (2) buah bagian antara lain: data uji dan data latih. Data latih merupakan data yang berfungsi sebagai pembangun pengetahuan untuk proses klasifikasi, proses pembangunan pengetahuan tersebut dilakukan melalui proses *modeling* dan menghasilkan sebuah model latih menggunakan data latih yang tersedia. Sementara data uji merupakan data yang disiapkan untuk menguji tingkat keakuratan model latih yang dihasilkan oleh proses *modeling*, proses pengujian tingkat akurasi tersebut dilakukan melalui proses pengujian dan evaluasi. Setelah dilakukan proses pengujiandan evaluasi model latih menggunakan data uji yang tersedia, hasil pengujian tersebut dipaparkan dalam bentuk persentase dan grafik.

### Pengumpulan Data

Pada tahapan pengumpulan data dilakukan melalui proses *crawling*. Proses tersebut meliputi: mendapatkan API *key* Twitter melalui akun *developer* Twitter (<https://developer.twitter.com/>). API *key* Twitter yang diperoleh antara lain: *Customer* API *key*, *Customer* API *Secret key*, *Access Token*, dan *Access Token Secret*. Proses selanjutnya adalah penambangan data yang bersumber pada media sosial Twitter menggunakan pustaka Tweepydengan akses dari API *key* yang telah didapatkan. Data *tweet* yang berhasil di kumpulkan akan disimpan ke dalam sebuah file *excel (.xlsx)*, yang kemudian disimpan ke dalam basis data (*database*) MySQL. Ilustrasi tahap pengumpulan data dapat dilihat pada Gambar X berikut:



**Gambar X Tahap pengumpulan data**

### *Preprocessing*

Pada tahapan *preprocessing* dilakukan penyaringan, pembuangan dan perbaikan kata melalui beberapa proses. Hal tersebut dimaksudkan untuk menghasilkan data kicauan (*tweet*) yang lebih terstruktur atau disebut dengan *clean text.* Proses yang dilakukan dalam tahap *preprocessing* antara lain*: casefolding, cleansing, mengubah slang word, menghapus stop word, dan stemming.*

#### *Casefolding*

Pada Gambar X proses *casefolding* dilakukan penyetaraan teks menjadi huruf kecil secara keseluruhan, misalnya: ‘Kalian’ akan diubah menjadi ‘kalian’, ‘PASTI BISA’ akan diubah menjadi ‘pasti bisa’, dan seterusnya.



**Gambar X Proses *casefolding***

#### *Cleansing*

Pada proses *cleansing* dilakukan penyaringan dan pembuangan teks yang untuk proses analisis. Proses *cleansing* terdiri atas beberapa tahapan antara lain: menghapus URL, *mention* (*@mention*), *hastag* (*#hastag*), angka (0-9), tanda baca, dan spasi berlebih.

##### Menghapus URL

Pada Gambar X proses penghapusan URL akan menghapus semua teks yang diawali dengan ‘http’, karena dianggap kurang memiliki makna namun sering disisipkan dalam sebuah kicauan (*tweet*).

****

**Gambar X Proses menghapus URL**

##### Meghapus *mention* (*@mention*)

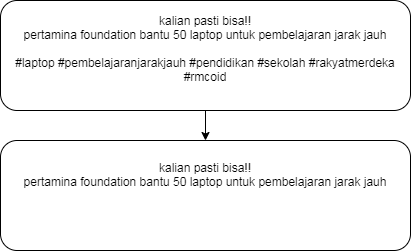
Pada Gambar X proses penghapusan *mention* (*@mention*) akan menghapus semua teks yang diawali dengan ‘@’.



**Gambar X Proses menghapus *mention***

##### Menghapus *hastag* (*#hastag*)

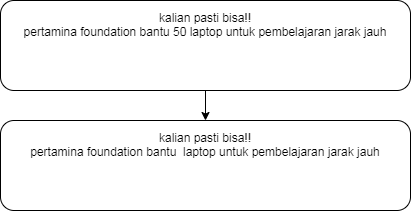
Pada Gambar X proses penghapusan tagar atau *hastag* (#hastag) akan menghapus semua teks yang diawali dengan ‘#’.



**Gambar X Proses menghapus *hastag***

##### Menghapus angka

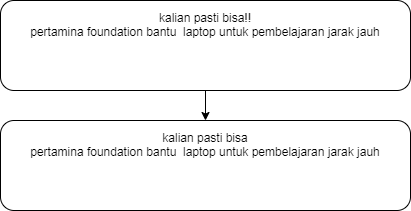
Pada Gambar X proses penghapusan angka akan menghapus angka (0-9) pada teks, misalnya: ‘bantu 50 laptop’ menjadi ‘bantu laptop’.



**Gambar X Proses menghapus angka**

##### Menghapus tanda baca

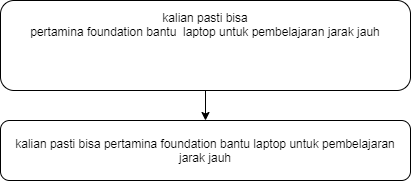
Pada Gambar X proses penghapusan tanda baca akan menghapus tanda baca pada teks, misalnya: ‘bisa!!’ menjadi ‘bisa’.



**Gambar X Proses menghapus tanda baca**

##### Menghapus spasi berlebih

Pada Gambar X proses penghapusan spasi atau (*whitespace)* berlebih akan menghapus baris dan *whitespace* yang lebih dari satu diantara kata, misalnya: ‘bantu laptop’ menjadi ‘bantu laptop’.



**Gambar X Proses menghapus spasi berlebih**

#### Merubah *slang word*

Pada Gambar X proses merubah *slang word* akan merubah setiap kata gaul, kata singkatan atau kata tidak baku ke bentuk bakunya, misalnya: ‘utk’ menjadi ‘untuk’, ‘yng’ menjadi ‘yang’ dan seterusnya. Proses pengubahan tersebut melibatkan kamus *slang word* yang terdapat dalam basis data (*database*).



**Gambar X Proses merubah *slang word***

#### Menghapus *stop word*

Pada Gambar X proses menghapus *stop word* akan menghapus setiap kata yang kurang memiliki makna namun sering dijumpai dalam sebuah teks, misalnya kata: ‘untuk’ ‘yang’, dan seterusnya. Proses penghapusan tersebut melibatkan kamus *stop word* yang terdapat dalam basis data (*database*).



**Gambar X Proses menghapus *stop word***

#### *Stemming*

### *Labeling*

### Pemisahan Data

#### Data Uji

#### Data Latih

### *Modeling*

### Pengujian& Evaluasi

### Visualisasi Hasil

## Rancangan Pengujian