

BAB III

METODE PENELITIAN

3.1. Objek Penelitian

Pada penelitian ini terdapat objek penelitian yaitu data yang diperoleh dari *twitter (X)* yang berupa *tweet* yang berjumlah 487 data. Sehingga data tersebut dapat diolah menggunakan *Python* serta dapat menerapkan algoritma *Support Vector Machine* untuk dilakukan klasifikasi sentimen analisis.

3.2. Peralatan Penelitian

Penelitian ini dilakukan dengan menggunakan *equipment* yang terdiri dari perangkat keras dan perangkat lunak.

1. Perangkat keras yang digunakan:
 - a. Laptop ASUS-A470UA.
 - b. Processor Intel(R) Core i3-7020U CPU @ 2.30GHz.
 - c. Intel(R) HD Graphics 620.
 - d. RAM 4.00 GB.
 - e. Harddisk 1 TB.
2. Perangkat lunak yang digunakan:
 - a. 64-bit *Operating System* Windows 10 Home Single Language.
 - b. Browser.
 - c. Microsoft Office Word 2016.
 - d. Microsoft Office Excel 2016.
 - e. Microsoft Office Visio 2016.
 - f. Microsoft Visual Code Studio.
 - g. Google Colaboratory.
 - h. Spreadsheet.

3.3. Lokasi dan Waktu Penelitian

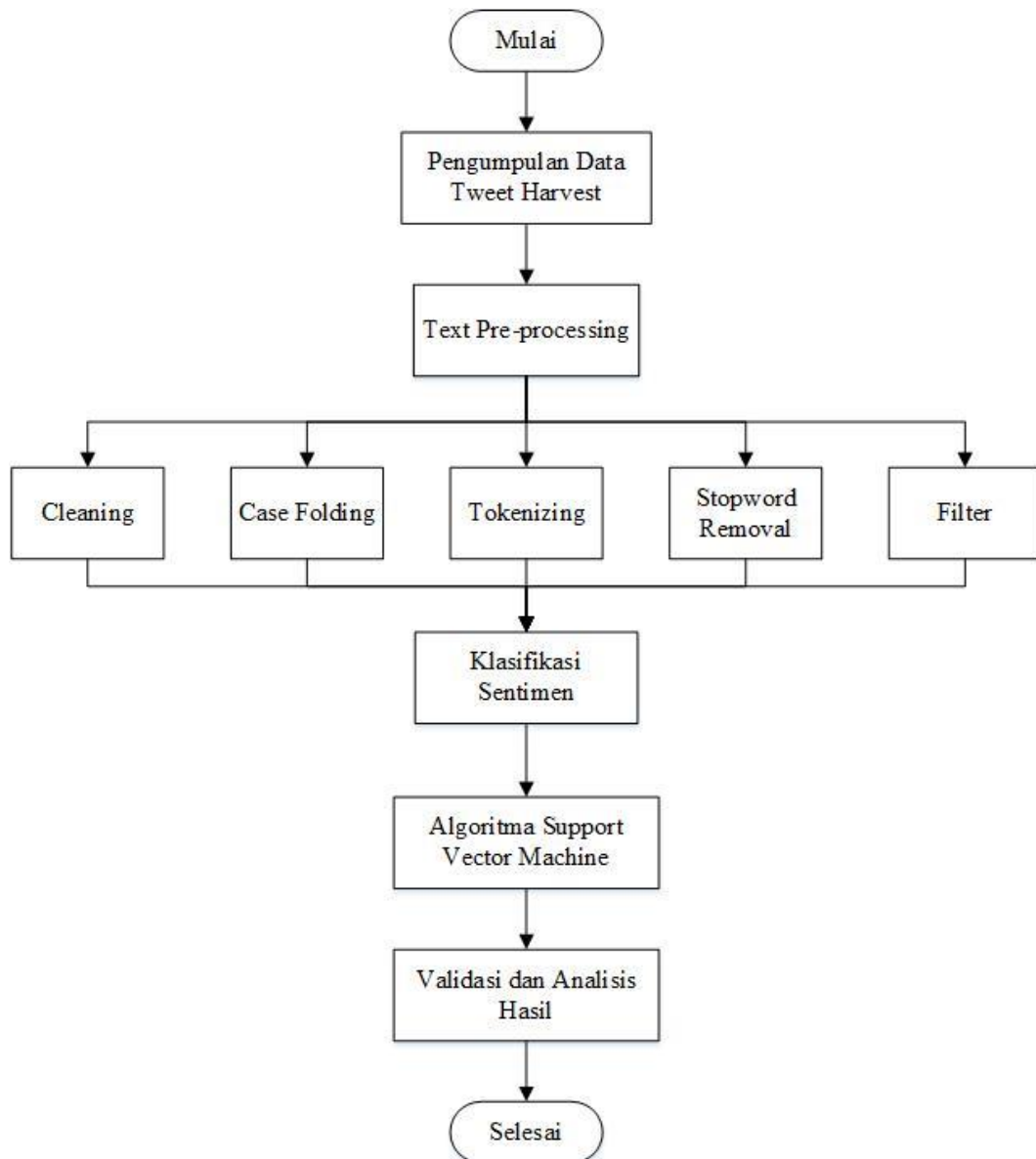
Lokasi penelitian ini akan dilakukan di rumah dan di kost, karena data yang diperoleh berasal dari media sosial *twitter (X)*. Kemudian data diolah menggunakan aplikasi *RapidMiner* dan *Python*. Adapun waktu penelitian terdapat pada tabel di bawah ini:

Tabel 3. 1 Waktu Penelitian.

NO	KEGIATAN	BULAN											
		NOVEMBER				DESEMBER				JANUARI			
		2023				2023				2024			
		1	2	3	4	1	2	3	4	1	2	3	4
1	Analisis Kebutuhan												
2	Pengumpulan Data												
3	Analisis Data												
4	Seleksi Data												
5	Penyusunan Tugas Akhir												

3.4. Prosedur Penelitian

Pada penelitian ini terdapat langkah-langkah dan tahapan yang berupa kerangka pemikiran dan penelitian dasar yang akan dilakukan. Secara umum, sistem yang akan dibuat adalah sistem yang dapat menghasilkan sentimen negatif ataupun positif terhadap kondisi rumput *Jakarta International Stadium* pada saat berlangsungnya Piala Dunia U-17 di Indonesia. Berikut adalah *flowchart* alur penelitian.



Gambar 3. 1 Flowchart Alur Prosedur Penelitian.

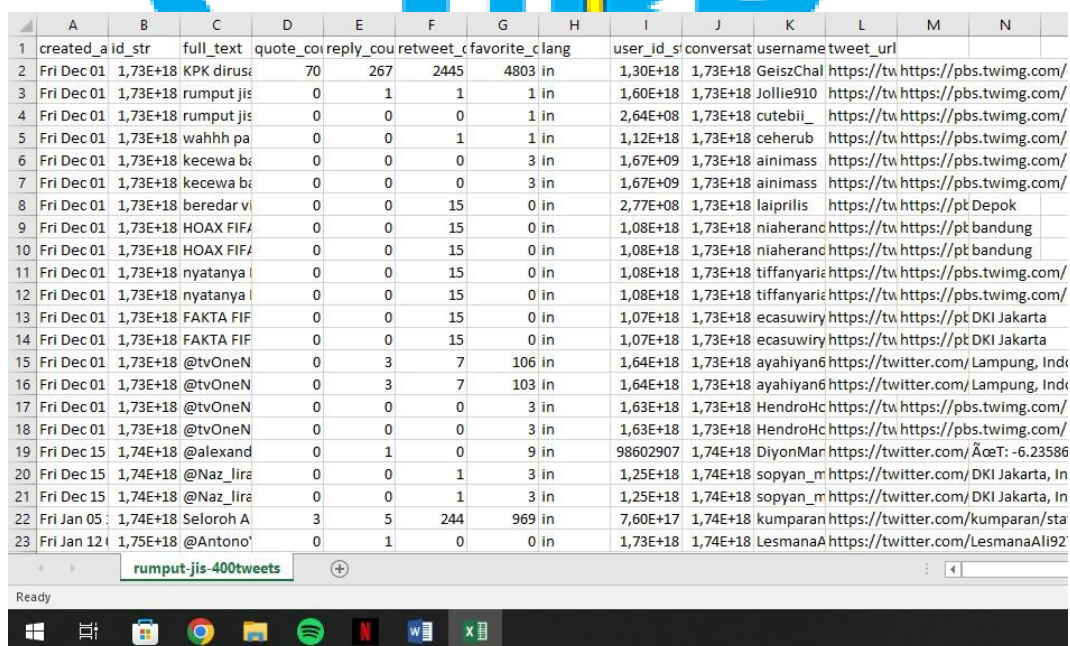
1. Pengumpulan Data.

Tahap pengumpulan data bertujuan untuk mendapatkan data yang berkaitan dengan penelitian dan informasi penelitian. Data yang akan digunakan dan diolah berupa *tweet* yang diperoleh dari *twitter (X)* dengan kata kunci “rumput jis” menggunakan *tweet harvest* pada tanggal 11 November 2023 sampai 12

Januari 2024 berjumlah 487 *tweet* dengan format .csv. Data yang telah diperoleh dapat diakses menggunakan aplikasi *Microsoft Excel*, *Python*.

2. Preprocessing Data.

Tahap *preprocessing* data adalah proses pembersihan data dari fitur yang tidak diperlukan untuk pengolahan *data mining*. Langkah ini dilakukan untuk menghilangkan data yang tidak memiliki nilai atau *missing value*, membuat *dataset* sesuai kebutuhan dan memudahkan pada saat pemrosesan data pada sistem. Tahap *text preprocessing* pada penelitian ini antara lain: *cleaning*, *case folding*, *tokenizing*, *stopword removal* dan *filter*. Pada data yang belum melalui proses *preprocessing* terdapat 12 jumlah atribut pada masing-masing data. Untuk melakukan analisis sentimen hanya dibutuhkan dua atribut yaitu isi *full text* dan klasifikasi. Pada atribut klasifikasi dilabeli dua kelas yaitu label kelas positif dan negatif. Berikut adalah contoh data yang belum melalui proses *preprocessing data*.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	created_at	id_str	full_text	quote_count	reply_count	retweet_count	favorite_count	lang	user_id_str	conversation_id_str	username	tweet_url		
2	Fri Dec 01	1,73E+18	KPK dirusi	70	267	2445	4803	in	1,30E+18	1,73E+18	GeiszChal	https://twitter.com/GeiszChal/status/1730000000000000000		
3	Fri Dec 01	1,73E+18	rumpu jjs	0	1	1	1	in	1,60E+18	1,73E+18	Jollie910	https://twitter.com/Jollie910/status/1730000000000000000		
4	Fri Dec 01	1,73E+18	rumpu jjs	0	0	0	1	in	2,64E+08	1,73E+18	cutebii	https://twitter.com/cutebii/status/1730000000000000000		
5	Fri Dec 01	1,73E+18	wahhh pa	0	0	1	1	in	1,12E+18	1,73E+18	ceherub	https://twitter.com/ceherub/status/1730000000000000000		
6	Fri Dec 01	1,73E+18	kecewa b	0	0	0	3	in	1,67E+09	1,73E+18	ainimass	https://twitter.com/ainimass/status/1730000000000000000		
7	Fri Dec 01	1,73E+18	kecewa b	0	0	0	3	in	1,67E+09	1,73E+18	ainimass	https://twitter.com/ainimass/status/1730000000000000000		
8	Fri Dec 01	1,73E+18	beredar vi	0	0	15	0	in	2,77E+08	1,73E+18	laiprili	https://twitter.com/laiprili/status/1730000000000000000		
9	Fri Dec 01	1,73E+18	HOAX FIF	0	0	15	0	in	1,08E+18	1,73E+18	niaherand	https://twitter.com/niaherand/status/1730000000000000000		
10	Fri Dec 01	1,73E+18	HOAX FIF	0	0	15	0	in	1,08E+18	1,73E+18	niaherand	https://twitter.com/niaherand/status/1730000000000000000		
11	Fri Dec 01	1,73E+18	nyatanya i	0	0	15	0	in	1,08E+18	1,73E+18	tiffanyari	https://twitter.com/tiffanyari/status/1730000000000000000		
12	Fri Dec 01	1,73E+18	nyatanya i	0	0	15	0	in	1,08E+18	1,73E+18	tiffanyari	https://twitter.com/tiffanyari/status/1730000000000000000		
13	Fri Dec 01	1,73E+18	FAKTA FIF	0	0	15	0	in	1,07E+18	1,73E+18	ecasuwi	https://twitter.com/ecasuwi/status/1730000000000000000		
14	Fri Dec 01	1,73E+18	FAKTA FIF	0	0	15	0	in	1,07E+18	1,73E+18	ecasuwi	https://twitter.com/ecasuwi/status/1730000000000000000		
15	Fri Dec 01	1,73E+18	@tvOneN	0	3	7	106	in	1,64E+18	1,73E+18	ayahiyani6	https://twitter.com/ayahiyani6/status/1730000000000000000		
16	Fri Dec 01	1,73E+18	@tvOneN	0	3	7	103	in	1,64E+18	1,73E+18	ayahiyani6	https://twitter.com/ayahiyani6/status/1730000000000000000		
17	Fri Dec 01	1,73E+18	@tvOneN	0	0	0	3	in	1,63E+18	1,73E+18	HendroHc	https://twitter.com/HendroHc/status/1730000000000000000		
18	Fri Dec 01	1,73E+18	@tvOneN	0	0	0	3	in	1,63E+18	1,73E+18	HendroHc	https://twitter.com/HendroHc/status/1730000000000000000		
19	Fri Dec 15	1,74E+18	@alexand	0	1	0	9	in	98602907	1,74E+18	DiyonMar	https://twitter.com/DiyonMar/status/1740000000000000000		
20	Fri Dec 15	1,74E+18	@Naz_lira	0	0	1	3	in	1,25E+18	1,74E+18	sopyan_m	https://twitter.com/sopyan_m/status/1740000000000000000		
21	Fri Dec 15	1,74E+18	@Naz_lira	0	0	1	3	in	1,25E+18	1,74E+18	sopyan_m	https://twitter.com/sopyan_m/status/1740000000000000000		
22	Fri Jan 05	1,74E+18	Seluruh A	3	5	244	969	in	7,60E+17	1,74E+18	kumpanan	https://twitter.com/kumpanan/status/1740000000000000000		
23	Fri Jan 12	1,75E+18	@Antono	0	1	0	0	in	1,73E+18	1,74E+18	LesmanaA	https://twitter.com/LesmanaA/status/1750000000000000000		

Gambar 3. 2 Data Sebelum Melalui Proses Preprocessing.

3. Klasifikasi Sentimen.

Pada tahap ini, setelah dataset melalui proses *text preprocessing* maka diperoleh dataset bersih yang kemudian diberi label kelas positif dan negatif.

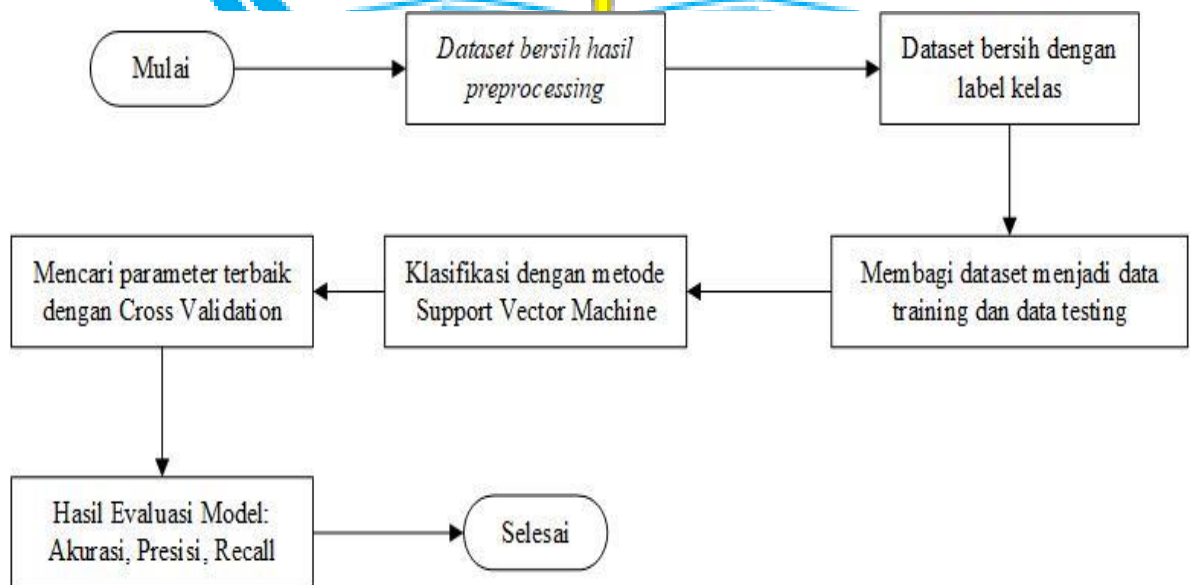
Setelah itu data dilabeli secara manual 80% sebagai data *training* dan 20% sebagai data *tesing*.

4. Implementasi Algoritma *Support Vector Machine*.

Pada proses ini dilakukan implementasi dari algoritma SVM menggunakan *Python*. Kelebihan menggunakan metode ini adalah kemampuannya untuk membuat model klasifikasi yang baik dengan formulasi yang jelas dan hanya sedikit parameter yang diatur. Berikut alur metode *Support Vector Machine* (Ailiyya, 2020):

- Melakukan transformasi data ke dalam ruang.
- Menghitung margin antar kelas.
- Mengetahui parameter kernel.
- Menemukan *hyperplane* terbaik.

Berikut adalah *flowchart* alur implementasi dari algoritma *support vector machine*:



Gambar 3. 3 *Flowchart* Algoritma *Support Vector Machine*.

Pada proses ini *dataset* telah melalui proses *preprocessing* dan *labeling* kelas positif atau negatif sebagai data *training*. Kemudian data *testing* akan melalui proses klasifikasi menggunakan algoritma *support vector machine*

dengan mencari parameter terbaik berdasarkan pelabelan data *training* yang telah dibuat. *Output* yang dihasilkan adalah hasil klasifikasi sentimen positif atau negatif yang akan digunakan untuk proses validasi hasil.

5. Validasi dan Analisis Hasil.

Pada tahap validasi dilakukan perhitungan untuk memperoleh nilai akurasi, presisi dan *recall* yang diambil dari hasil persamaan **Error! Reference source not found.**, **Error! Reference source not found.** dan **Error! Reference source not found.** pada algoritma *support vector machine*. Hasil dari analisis sentimen diklasifikasikan menjadi dua jenis yaitu positif dan negatif. Hasil yang diperoleh dapat dikatakan baik atau layak apabila tingkat akurasi, presisi dan *recall* yang dihasilkan tinggi.

