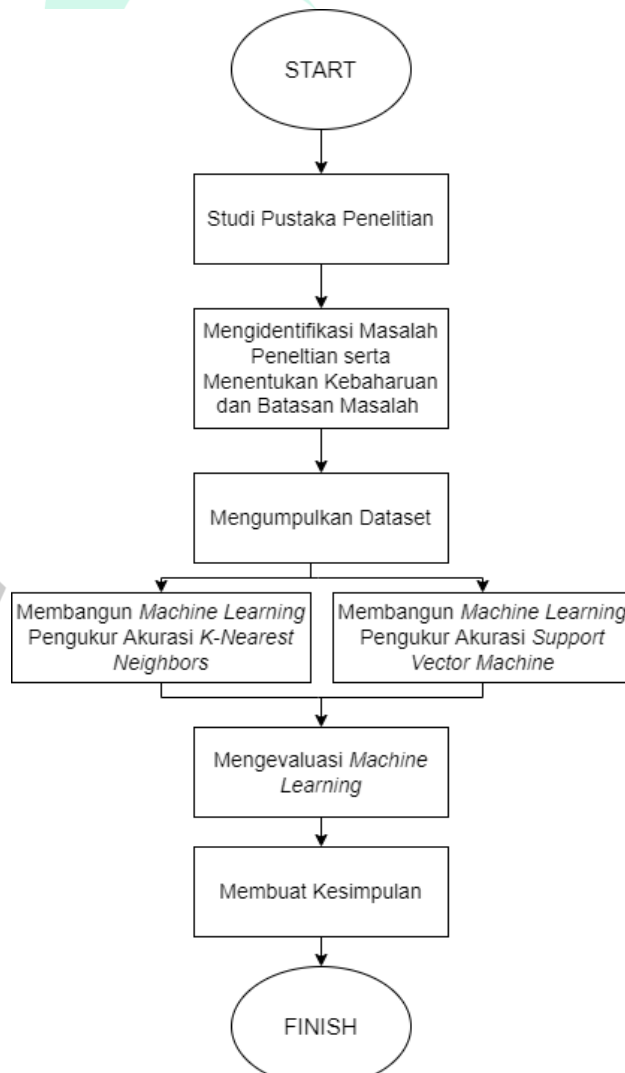


BAB IV PERENCANAAN

Skripsi ini dilakukan dengan menggunakan langkah – langkah dan rancangan penelitian. Setiap langkan dan rancangan yang dibuat merupakan pedoman untuk melakukan penelitian secara berurutan sesuai tujuan skripsi ini.

4.1 Langkah – Langkah Penelitian

Penelitian dapat berjalan dengan baik dengan adanya langkah – langkah yang dirancang secara sistematis dan berurutan. Langkah – langkah tersebut nantinya digunakan sebagai gambaran untuk melakukan penelitian. Berikut langkah – langkah penelitian yang digambarkan dalam bentuk diagram *flowchart*.



Gambar 4. 1 Flowchart Langkah - Langkah Penelitian

Berdasarkan diagram *flowchart* diatas, langkah – langkah penelitian dapat diuraikan sebagai berikut.

- (1) Studi pustakan dilakukan dengan mencari topik yang akan diteliti. Topik dibuat dengan melakukan riset pada pencapaian terdahulu terkait topik yang sudah terbit sebelumnya. Riset dilakukan dengan mencari jurnal penelitian – penelitian yang sudah pernah dilakukan oleh peneliti sebelumnya. Berdasarkan riset tersebut, peneliti dapat mempersempit topik penelitian dengan menentukan detail seperti penentuan algoritma yang akan digunakan penelitian.
- (2) Mengidentifikasi masalah penelitian dengan merumuskan masalah yang akan diteliti. Rumusan masalah tersebut merupakan penelitian yang akan dilaksanakan. Selanjutnya dengan menentukan kebaharuan berdasarkan penelitian yang sudah dilakukan sebelumnya terkait dengan topik yang diangkat pada penelitian ini, serta menentukan batasan penelitian yang akan dilakukan.
- (3) Mengumpulkan data penelitian berupa *dataset* saham PT Bank Mandiri Tbk sebagai data histori yang akan digunakan dalam penelitian. Data tersebut diperoleh dari situs *Yahoo Finance* sesuai dengan batasan penelitian.
- (4) Membangun rancangan *machine learning* untuk menjalankan algoritma KNN dan SVM yang digunakan sebagai pengukur tingkat akurasi dalam melakukan penelitian. Awal dari rancangan ini dengan membuat sistem komputasi dalam bentuk *flowchart* sebagai gambaran alur dari *machine learning* dengan menggunakan algoritma yang sudah ditentukan. Lalu rancangan *machine learning* diprogram dengan menggunakan bahasa Python.
- (5) Membahas hasil penelitian dengan mengevaluasi tingkat akurasi yang diperoleh pada algoritma yang digunakan. Evaluasi tersebut juga dilakukan untuk melihat langkah – langkah yang sudah dilakukan berjalan sesuai dengan topik penelitian. Selanjutnya hasil dari kedua algoritma tersebut akan dilakukan perbandingan tingkat akurasi yang memiliki nilai yang lebih akurat.

(6) Membuat kesimpulan berdasarkan hasil penelitian yang sudah dilakukan dengan mengetahui tingkat akurasi mana yang lebih akurat dari setiap algoritma yang digunakan dalam memprediksi saham PT Bank Mandiri Tbk dengan parameter RMSE sebagai perbandingan antara dua algoritma. Peneliti juga memberikan saran kepada pembaca terkait dengan penelitian yang sudah dilakukan agar nantinya penelitian ini dapat dilanjutkan dan dijadikan sebagai sumber referensi.

4.2 Konsep Data Dan *Machine Learning*

4.2.1 Data

Data yang digunakan dalam penelitian ini diambil dari situs Yahoo Finance dimana data tersebut merupakan histori data harga saham PT Bank Mandiri dalam bentuk data harian. Data harian tersebut menjelaskan harga saham yang terjadi pada hari – hari sebelumnya. Sebelumnya, dari data yang diperoleh dalam penelitian ini terdapat beberapa fitur yang menjelaskan kategori harga saham seperti *Date*, *Open*, *High*, *Low*, *Close*, *Adj Close*, dan *Volume* dimana masing – masing fitur memiliki penjelasan sebagai berikut.

- a. *Date*, merupakan tanggal dari suatu harga saham yang sudah terjadi.
- b. *Open*, merupakan harga saham pembukaan yang terjadi di suatu hari.
- c. *High*, merupakan harga saham dengan nilai tertinggi yang terjadi di suatu hari.
- d. *Low*, merupakan harga saham dengan nilai terendah yang terjadi di suatu hari
- e. *Close*, merupakan harga penutupan dari suatu hari.
- f. *Adj Close*, merupakan harga saham penutupan yang telah disesuaikan oleh perusahaan
- g. *Volume*, merupakan jumlah saham yang diperdagangkan pada suatu hari dalam jumlah lembar.

Dari beberapa fitur tersebut, penelitian ini menggunakan satu sampai empat indikator fitur *Close*, *Open*, *High*, dan *Low* untuk memprediksi harga saham PT Bank Mandiri yang terjadi pada hari berikutnya. Selanjutnya data tersebut dilakukan perubahan dengan menggunakan data *time series* karena data tersebut merupakan data harian yang diperoleh dari hari sebelumnya untuk dilakukan

prediksi data harga saham untuk hari – hari berikutnya. Data *time series* dibuat dengan menambahkan fitur – fitur yang menjelaskan terkait data 5 hari sebelum hari yang akan diprediksi sudah terjadi. Fitur yang digunakan dalam data *time series* ini yaitu *Date*, $x(n-5)$, $x(n-4)$, $x(n-3)$, $x(n-2)$, $x(n-1)$, $x(n)$, dan $x(n+1)$. Dimana masing – masing fitur memiliki penjelasan sebagai berikut.

- a. *Date*, merupakan tanggal dari suatu harga saham.
- b. $x(n-5)$, merupakan harga saham 5 hari sebelumnya.
- c. $x(n-4)$, merupakan harga saham 4 hari sebelumnya.
- d. $x(n-3)$, merupakan harga saham 3 hari sebelumnya.
- e. $x(n-2)$, merupakan harga saham 2 hari sebelumnya.
- f. $x(n-1)$, merupakan harga saham 1 hari sebelumnya.
- g. $x(n)$, merupakan harga saham pada hari itu.
- h. $x(n+1)$, merupakan prediksi harga saham 1 hari berikutnya.

Dengan adanya empat indikator lainnya, terdapat perbedaan variabel yang digunakan dalam penamaan label yang dibuat untuk membedakan antara fitur satu dengan fitur lainnya didalam *dataset* dalam melakukan prediksi menggunakan satu sampai empat indikator dimana masing – masing dijelaskan sebagai berikut

- a. variabel x untuk fitur *close*
- b. variabel o untuk fitur *open*
- c. variabel h untuk fitur *high*
- d. variabel l untuk fitur *low*

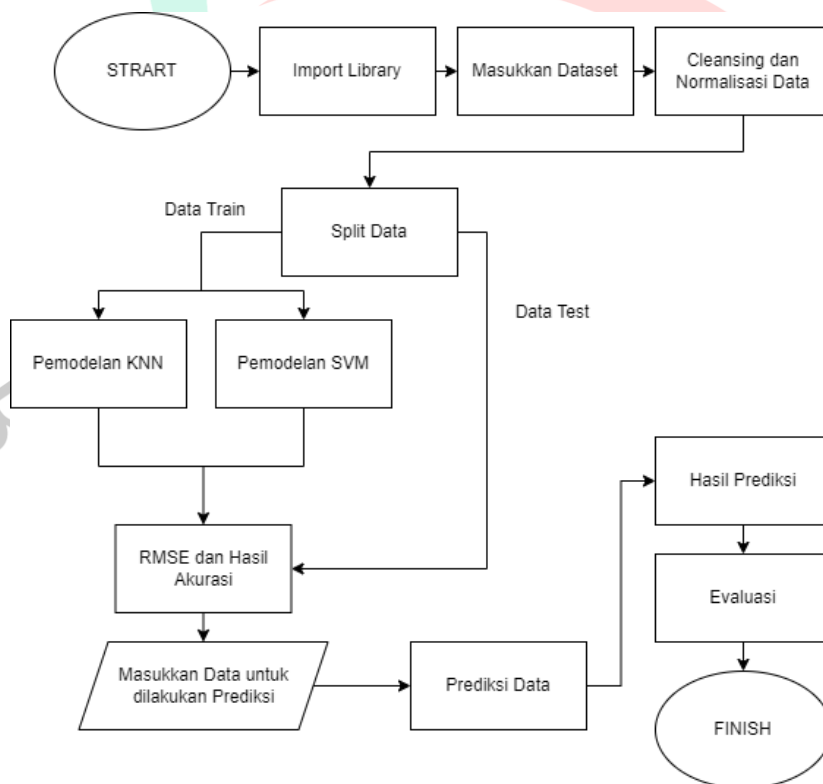
4.2.2 Machine Learning

Dalam penelitian ini dilakukan perancangan sebuah model *machine learning* untuk memprediksi harga sebuah saham di hari berikutnya. Dimana *machine learning* yang dirancang melakukan proses regresi data. Metode yang digunakan dalam merancang *machine learning* ini termasuk ke dalam metode *supervised learning* karena dataset dalam memprediksi harga saham memiliki label tetap yang sudah dilakukan perubahan data sebelumnya menjadi data *time series*. Label yang terdapat pada dataset yaitu $x(n-5)$, $x(n-4)$, $x(n-3)$, $x(n-2)$, $x(n-1)$, $x(n)$, dan $x(n+1)$ untuk satu indikator fitur *close*, label $o(n-5)$, $o(n-4)$, $o(n-3)$, $o(n-2)$, $o(n-1)$ $o(n)$, $x(n-$

5), $x(n-4)$, $x(n-3)$, $x(n-2)$, $x(n-1)$, $x(n)$, dan $x(n+1)$ untuk dua indikator fitur *open* dan *close*, label $o(n-5)$, $o(n-4)$, $o(n-3)$, $o(n-2)$, $o(n-1)$ $o(n)$, $h(n-5)$, $h(n-4)$, $h(n-3)$, $h(n-2)$, $h(n-1)$ $h(n)$, $x(n-5)$, $x(n-4)$, $x(n-3)$, $x(n-2)$, $x(n-1)$, $x(n)$, dan $x(n+1)$ untuk tiga indikator fitur *open*, *high*, dan *close*, serta label $o(n-5)$, $o(n-4)$, $o(n-3)$, $o(n-2)$, $o(n-1)$ $o(n)$, $h(n-5)$, $h(n-4)$, $h(n-3)$, $h(n-2)$, $h(n-1)$ $h(n)$, $l(n-5)$, $l(n-4)$, $l(n-3)$, $l(n-2)$, $l(n-1)$, $l(n)$, $x(n-5)$, $x(n-4)$, $x(n-3)$, $x(n-2)$, $x(n-1)$, $x(n)$, dan $x(n+1)$ untuk empat indikator fitur *open*, *high*, *low* dan *close* dimana nantinya *machine learning* akan selalu membaca dan mengeksplor dataset tersebut berdasarkan label yang sudah dibuat dan ditetapkan berdasarkan fungsi dari masing – masing label yang sudah dijelaskan pada konsep data.

4.3 Mekanisme Melakukan Prediksi

Mekanisme yang dilakukan *machine learning* dalam melakukan prediksi harga saham PT Bank Mandiri dilakukan secara bertahap agar berjalan secara terurut agar lebih mekanis.



Gambar 4. 2 Flowchart Mekanisme

Berdasarkan grafik *flowchart* diatas, mekanisme *machine learning* dalam melakukan prediksi diuraikan sebagai berikut.

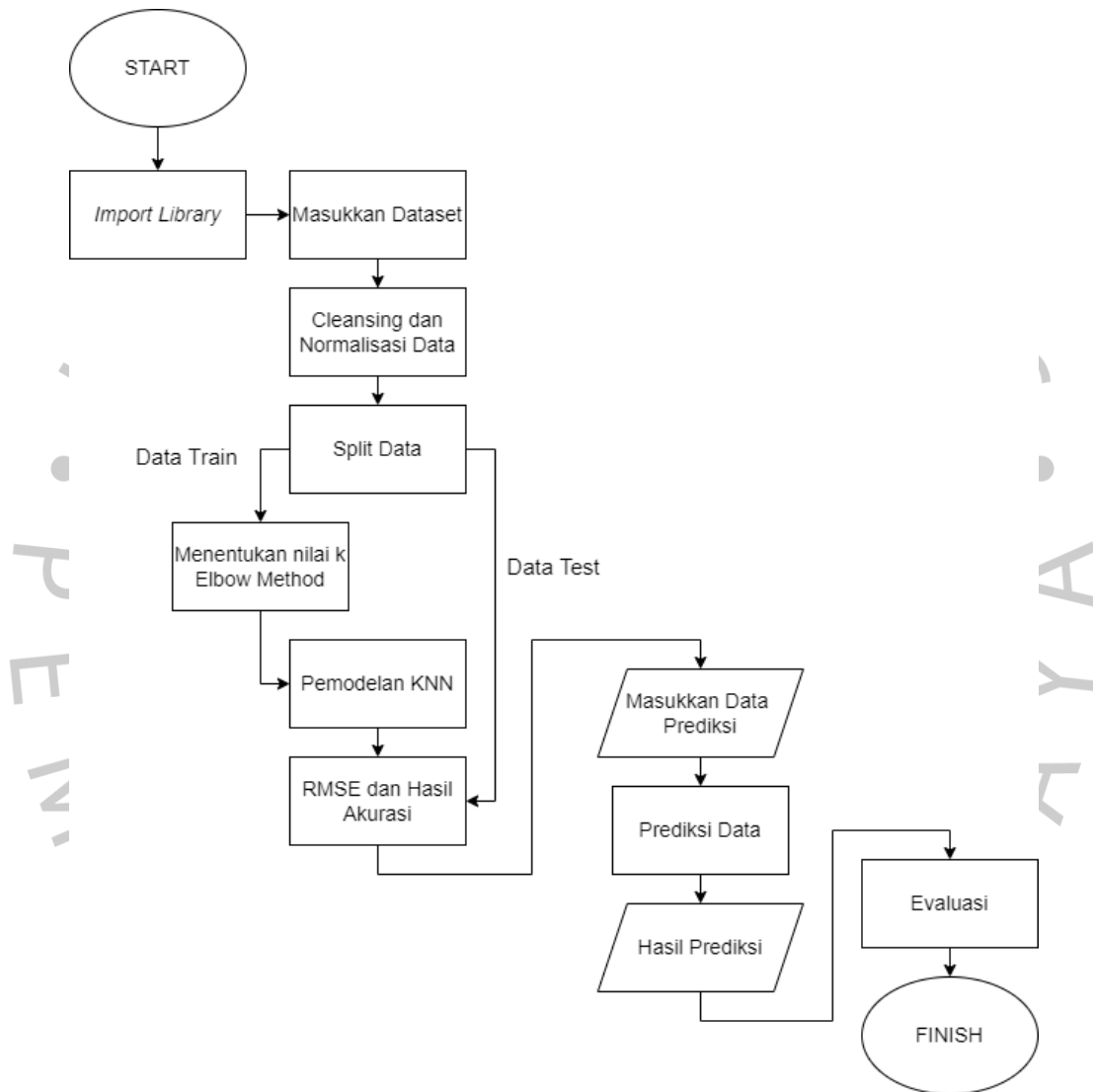
- (1) *Import library* dibutuhkan dalam komputasi algoritma. *Library* yang digunakan yaitu *pandas* (analisis data), *numpy* (komputasi dan numerik), *matplotlib, pyplot* (visualisasi data dalam bentuk grafik) dan *math* (mengelola perhitungan matematika).
- (2) Memasukkan *dataset* harga saham PT Bank Mandiri Tbk yang telah diperoleh dari situs *Yahoo Finance* dengan periode sesuai batasan penelitian dan disimpan pada Google Drive.
- (3) Melakukan *cleansing* dan normalisasi data untuk melakukan perbaikan data yang memiliki kesalahan dengan adanya NaN dan membersihkan kolom data yang bernilai nol atau kosong pada data.
- (4) Melakukan Split data dengan membagi menjadi data latih dan data uji yang diinisialisasikan dalam variabel x dan y .
- (5) Membangun pemodelan KNN dengan menggunakan nilai K-optimal yang sudah didapatkan.
- (6) Melakukan pemodelan menggunakan algoritma SVM dengan memasukan data *train* yang nantinya akan diuji.
- (7) Melakukan pemodelan RMSE dan hasil akurasi untuk mengetahui nilai *error* dan berapa besar akurasi yang dihasilkan pada hasil pemodelan *machine learning*.
- (8) Memasukkan data untuk dilakukan prediksi harga saham pada hari berikutnya
- (9) Hasil dari pemodelan selanjutnya dievaluasi apakah sudah baik atau belum.

4.4 Perancangan Pengujian

Perancangan pengujian dilakukan untuk memberi gambaran terkait model yang akan dirancang. Dalam penelitian ini melakukan pembuatan suatu pemodelan pergerakan harga saham PT Bank Mandiri Tbk menggunakan *machine learning* dengan algoritma KNN dan SVM. Rancangan pengujian dilakukan dengan membuat *flowchart* yang digunakan sebagai gambaran alur dari kedua algoritma yang telah ditentukan. Rancangan diperlukan agar perencanaan dalam membangun sebuah model dapat dilakukan secara terarah sehingga penelitian berjalan dengan baik.

4.4.1 Perancangan *K-Nearest Neighbors*

K-Nearest Neighbors merupakan algoritma *machine learning* yang digunakan untuk memprediksi suatu data dengan menggunakan nilai *K* optimal. Berikut merupakan rancangan algoritma KNN yang digambarkan dalam bentuk *flowchart*.



Gambar 4. 3 Flowchart Algoritma KNN

Berdasarkan diagram *flowchart* algoritma KNN diatas, dapat dijelaskan pada uraian berikut.

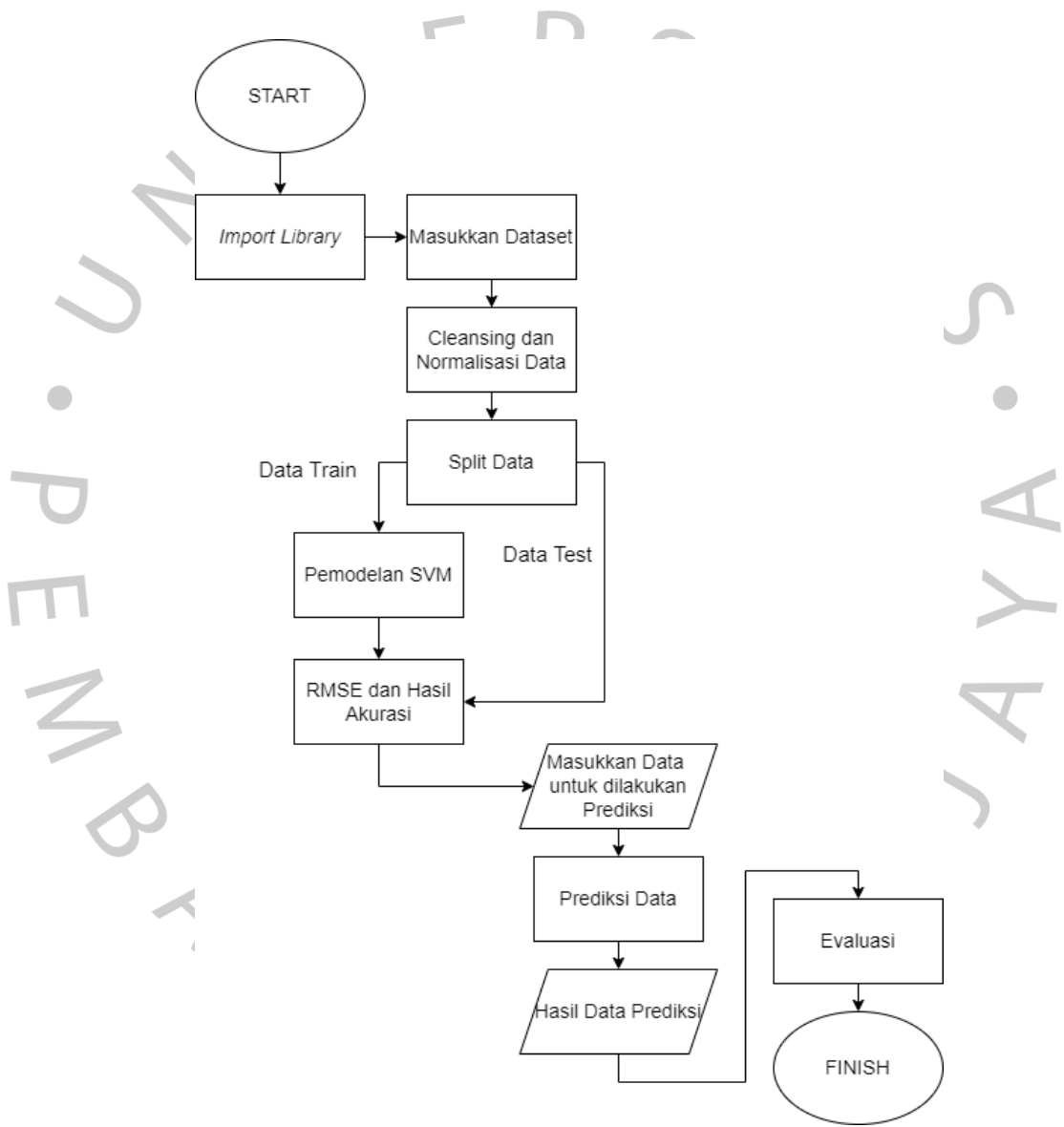
- (1) *Import library* dibutuhkan dalam komputasi algoritma KNN. *Library* yang digunakan yaitu *pandas* (analisis data), *numpy* (komputasi dan numerik),

matplotlib, pyplot (visualisasi data dalam bentuk grafik) dan *math* (mengelola perhitungan matematika).

- (2) Memasukkan *dataset* harga saham PT Bank Mandiri Tbk yang telah diperoleh dari situs *Yahoo Finance* dengan periode sesuai batasan penelitian dan disimpan pada Google Drive.
- (3) Melakukan *cleansing* untuk melakukan perbaikan data yang memiliki kesalahan dengan adanya NaN pada data dan melakukan normalisasi data dengan membersihkan kolom data yang bernilai nol atau data kosong. Normalisasi dilakukan dengan menggunakan metode *Standard scaler* dimana berfungsi untuk menormalkan seluruh data pada baris memiliki nilai yang sama.
- (4) Melakukan Split data dengan membagi menjadi data latih dan data uji yang diinisialisasikan dalam variabel *x* dan *y*. Variabel *x* digunakan untuk data dari
 - 5 hari sebelumnya sampai data pada hari utama, dan untuk variabel *y* digunakan untuk data pada hari berikutnya.
- (5) Menentukan nilai K-optimal dengan menggunakan metode Elbow dimana nilai K-optimal merupakan jumlah nilai tetangga yang paling terdekat dalam pemodelan KNN.
- (6) Menerapkan pemodelan KNN dengan menggunakan nilai K-optimal yang sudah didapatkan.
- (7) Melakukan pemodelan RMSE untuk mengetahui nilai *error* yang terjadi pada hasil pemodelan KNN.
- (8) Hasil yang dikeluarkan berupa akurasi presentase dalam memprediksi harga saham menggunakan pemodelan KNN.
- (9) Memasukkan data untuk dilakukan prediksi harga saham pada hari berikutnya.
- (10) Menampilkan hasil prediksi harga saham pada hari berikutnya dan melakukan perbandingan dengan harga aktual.
- (11) Melakukan evaluasi model yang dibutuhkan untuk mengetahui apakah pemodelan sudah dilakukan dengan baik atau belum.

4.4.2 Perancangan *Support Vector Machine*

Support Vector Machine merupakan algoritma *machine learning* yang digunakan dalam melakukan pemodelan *machine learning* prediksi harga saham PT Bank Mandiri Tbk. Berikut merupakan rancangan algoritma SVM yang digambarkan dalam bentuk *flowchart*.



Gambar 4. 4 Flowchart Algoritma SVM

Berdasarkan diagram *flowchart* algoritma KNN diatas, dapat dijelaskan pada uraian berikut.

- (1) *Import library* dibutuhkan dalam komputasi algoritma SVM. *Library* yang digunakan yaitu *pandas* (analisis data), *numpy* (komputasi dan numerik), *matplotlib, pyplot* (visualisasi data dalam bentuk grafik) dan *math* (mengelola perhitungan matematika).
- (2) Memasukkan *dataset* harga saham PT Bank Mandiri Tbk yang telah diperoleh dari situs *Yahoo Finance* dengan periode sesuai batasan penelitian dan disimpan pada Google Drive.
- (3) Melakukan *cleansing* untuk melakukan perbaikan data yang memiliki kesalahan dengan adanya NaN pada data dan melakukan normalisasi data dengan membersihkan kolom data yang bernilai nol atau data kosong. Normalisasi dilakukan dengan menggunakan metode *Standard scaler* dimana berfungsi untuk menormalkan seluruh data pada baris memiliki nilai yang sama.
- (4) Melakukan Split data dengan membagi menjadi data latih dan data uji yang diinisialisasikan dalam variabel *x* dan *y*. Variabel *x* digunakan untuk data dari 5 hari sebelumnya sampai data pada hari utama, dan untuk variabel *y* digunakan untuk data pada hari berikutnya.
- (5) Melakukan pemodelan menggunakan algoritma SVM dengan memasukan data *train* yang nantinya akan diuji.
- (6) Melakukan pemodelan RMSE untuk mengetahui nilai *error* yang terjadi pada hasil pemodelan KNN.
- (7) Hasil yang dikeluarkan berupa akurasi presentase dalam memprediksi harga saham menggunakan pemodelan SVM.
- (8) Memasukkan data untuk dilakukan prediksi harga saham pada hari berikutnya.
- (9) Menampilkan hasil prediksi harga saham pada hari berikutnya dan melakukan perbandingan dengan harga aktual.
- (10) Melakukan evaluasi model yang dibutuhkan untuk mengetahui apakah pemodelan sudah dilakukan dengan baik atau belum.