**ANALISIS PREDIKSI HARGA SAHAM MANCHESTER UNITED MENGGUNAKAN METODE SUPPORT VECTORE MACHINE**

**PROPOSAL TUGAS AKHIR**

Diajukan untuk memenuhi salah satu syarat

memperoleh gelar Sarjana Komputer

Program Studi Informatika

****

Diajukan Oleh:

**Fransiskus Jremiegi Saputra**

**205314062**

PROGRAM STUDI INFORMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS SANATA DHARMA

YOGYAKARTA

2023

# **KATA PENGANTAR**

# **DAFTAR ISI**

[KATA PENGANTAR i](#_Toc160541077)

[DAFTAR ISI ii](#_Toc160541078)

[Daftar Gambar iii](#_Toc160541079)

[Daftar Tabel iv](#_Toc160541080)

[BAB I PENDAHULUAN 1](#_Toc160541081)

[1.1 Latar Belakang 1](#_Toc160541082)

[1.2 Rumusan Masalah 2](#_Toc160541083)

[1.3 Batasan Masalah 3](#_Toc160541084)

[1.4 Tujuan Penelitian 3](#_Toc160541085)

[1.5 Manfaat Penelitian 3](#_Toc160541086)

[BAB II TINJAUAN PUSTAKAN 4](#_Toc160541087)

[2.1 Kajian Teori 4](#_Toc160541088)

[2.1.1 Saham 4](#_Toc160541089)

[2.1.2 Manchester United 4](#_Toc160541090)

[2.1.3 Support Vector Machine (SVM) 5](#_Toc160541091)

[2.2 Kajian Hasil Riset 7](#_Toc160541092)

[BAB III METODE PENELITIAN 13](#_Toc160541093)

[3.1 Pengumpulan Data 14](#_Toc160541094)

[3.2 Preprocessing 15](#_Toc160541095)

[3.3 Ekstraksi Fitur 16](#_Toc160541096)

[3.4 Pembagian Data 17](#_Toc160541097)

[3.5 Modeling 18](#_Toc160541098)

[3.6 Evaluasi 18](#_Toc160541099)

[3.7 Skenario Pengujian 18](#_Toc160541100)

[DAFTAR PUSTAKA 20](#_Toc160541101)

# **DAFTAR GAMBAR**

[Gambar 3. 1 Gambaran umum penelitian 13](#_Toc160541427)

[Gambar 3. 2 Contoh data harga saham Manchester United 15](#_Toc160541428)

[Gambar 3. 3 Alur proses preprocessing 16](#_Toc160541429)

[Gambar 3. 4 Alur proses ekstraksi fitur 16](#_Toc160541430)

# **DAFTAR TABEL**

[Tabel 2. 1 Rumus fungsi karnel 6](#_Toc155821978)

[Tabel 2. 2 Hasil kajian teori 7](#_Toc155821979)

[Tabel 3. 1 Skenario pengujian 18](#_Toc155821966)

# **PENDAHULUAN**

## **Latar Belakang**

Dewasa ini, trend terhadap saham kian populer dikalangan masyarkat mulai dari iming-iming manis yang menjanjikan untuk cuan sampai pada penipuan yang merugikan bahkan tak jarang adu gengsian agar tidak jadi bahan gosipan. Adanya jaminan kemudahan dalam mengakses pasar saham juga menjadi pemikat agar masyarkat semakin tertarik untuk terjun ke dunia investasi. Lantas, fenomena ini menujukan bahwa pasar saham menjadi daya tarik utama dikalangan masyarakat umum sehingga mencerminkan minat yang meningkat terhadap investasi dan keuangan.

Penting untuk memahami bahwa investasi saham memerlukan pengetahuan yang cukup, riset yang matang, dan perencanaan yang hati-hati. Hal ini dikarenakan saham bergerak secara non linier dan perubahan harga bisa terjadi begitu cepat yang dipengaruhi oleh banyak faktor seperti kondisi keuangan perusahaan, laju inflasi, tingkat suku bunga, permintaan dan penawaran, dan masih banyak lagi (Fadilah, Agfiannisa, & Azhar, 2020) (Wulandari & Anubhakti, 2022) (Syahputra, Ramadhan, & Burha, 2022) (Maulana & Kumalasari, 2019) (Lestari & Pangaribuan, 2020). Oleh karena itu, penting untuk mengetahui dan mempelajari perilaku pergerakan harga saham dengan melakukan analisis prediksi harga saham sehingga berguna untuk mengetahui prospek dan alur investasi di masa mendatang serta membantu dalam mempertimbangkan pengambilan keputusan transaksi saham (Fadilah, Agfiannisa, & Azhar, 2020) (Wulandari & Anubhakti, 2022) (Syahputra, Ramadhan, & Burha, 2022) (Maulana & Kumalasari, 2019) (Lestari & Jasuni, 2023) . Dengan demikian, Adanya model prediksi pergerakan harga saham yang akurat dapat menjadi tools yang bisa membantu para investor dalam melakukan pengambilan keputusan terkait dengan saham dikarenakan pergerakan harga saham yang cenderung non linier ini akan menyulitkan investor dalam melakukan prediksi. Terdapat beberapa penelitian yang membahas terkait dengan prediksi saham dengan menggunakan beberapa model algoritma.

Penelitian yang dilakukan Wulandari & Anubhakti (2022) dalam memprediksi harga saham PT. Garuda Indonesia, Tbk menggunakan pengujian model SVM yakni menghasilkan performance prediction trend accuracy sebesar 0.545 dari 520 dataset yang ada (Wulandari & Anubhakti, 2022). Masih menggunakan metode yang sama yakni SVM namun dikombinasikan dengan melakukan perbandingan menggunakan KNN untuk memprediksi harga saham PT. Telekomunikasi Indonesia menghasilkan nilai akurasi tertinggi yakni 0.9653 dan nilai RMSE sebesar 0.0091 untuk algoritma SVM sedangkan algoritma KNN nilai akurasi sebesar 0.9456 dan RMSE sebesar 0.1162 (Fadilah, Agfiannisa, & Azhar, 2020). Tidak hanya itu, terdapat juga penelitian terhadap prediksi harga saham Bank BRI menggunakan algoritma Linier Regression dan memperoleh akurasi R-2 sebesar 0.9125 dengan nilai error pada prediksi yang diukur menggunakan MAPE dan menghasilkan persentase sebesar 13,751% untuk data pelatihan, 13,773% untuk data pengujian, dan 13,755% untuk data keseluruhan atau persentase error berada pada rentang 10% - 20% (Syahputra, Ramadhan, & Burha, 2022).

Berdasarkan latar belakang yang telah diuraikan tersebut maka dapat dilakukan penelitian dalam memprediksi harga saham dengan pendekatan yang serupa sehingga dapat membantu para investor dalam melihat prospek suatu saham tertentu. Tidak hanya itu, metode algoritma yang digunakan juga diharapkan mampu menjawab permasalahan yang telah diuraikan serta dapat mengetahui tingkat keakuratan dari suatu model algoritma dalam memprediksi harga saham.

## **Rumusan Masalah**

1. Bagaimanakah metode Support Vector Machine (SVM) dapat digunakan untuk memprediksi harga saham secara akurat ?
2. Bagaimana hasil dari tingkat keakuratan dari suatu model algoritma yang digunakan dalam meprediksi harga saham dapat menjawab permasalah seperti yang telah diuraikan pada bagian latar belakang ?

## **Batasan Masalah**

1. Data saham yang diambil ialah saham Manchester United dengan kode MANU pada bursa efek New York Stock Exchange (NYSC).
2. Data yang diambil yakni data pergerakan harian harga saham selama 5 tahun terakhir dengan rentang waktu yakni 31 Oktober 2018 sampai 30 Oktober 2023.
3. Algoritma yang digunakan hanya terbatas pada Support Vector Machine (SVM)

## **Tujuan Penelitian**

Pentingnya pengetahuan yang cukup, riset yang matang dan perencanaan yang hati-hati dalam investasi membutuhkan analisis prediksi saham yang akurat sehingga memberikan gambaran terkait dengan prospek saham di masa yang mendatang. Dengan demikian, hadirnya model dalam memprediksi pergerakan harga saham dapat menjadi tools untuk membantu dalam menganalisis serta mengambil keputusan sebagai bahan pertimbangan. Di sisi lain, metode yang digunakan juga menjadi tolak ukur untuk melihat sejauh mana keakuratan yang dihasilkan serta seberapa efektif model tersebut melahirkan prediksi yang akurat ditengah kondisi pasar saham yang bergerak secara non-linier.

## **Manfaat Penelitian**

1. Membantu para investor dalam memprediksi harga saham dan mengetahui prospek saham di masa yang akan datang sehingga dapat meminimalkan resiko dan memaksimalkan keuntungan.
2. Membantu investor dalam mempertimbangkan pengambilan keputusan sehingga menjadi panduan dalam menyusun starategi investasi baik jangka pendek maupun jangka Panjang.
3. Metode yang digunakan untuk memprediksi menjadi tolak ukur dalam mengetahui tingkat keakuratan dan efektifitas sehingga membantu sejauh mana model tersebut dapat diandalkan dan menghasilkan prediksi yang sesuai ditengah pasar saham yang bergerak secara non-linier.

# **TINJAUAN PUSTAKAN**

## **Kajian Teori**

### **Saham**

Saham merupakan bukti tanda kepemilikan dan penyertaan modal perseroan terbatas yang memberikan hak atas dividen dan lain-lain berdasarkan besar kecilnya modal (Rusmalawati, Furqon, & Indriati, 2018). Dengan membeli saham perusahaan menandakan bahwa adanya penanaman modal yang nantinya akan digunakan oleh pihak manejemen perusahaan untuk membiayai operasional perusahaan. Lantas, keuntungan yang diperoleh dari penanaman modal tersebut yakni *dividen* yang merujuk pada laba yang akan dibagian kepada pemegang saham sesuai dengan porsi kepemilikannya dan *capital gain* yang merujuk pada keuntungan yang didapatkan dengan memperdagangkan saham sehingga keuntungan yang diperoleh sangat bergantung pada harga saham di pasar modal sendiri (Rusmalawati, Furqon, & Indriati, 2018).

### **Manchester United**

Mendengar kata Manchester United tidak terlepas dari bayang-bayang klub sepak bola liga Inggris dengan segudang prestasi. Lantas, Manchester United bukan hanya sekedar klub sepak bola professional di Inggris melainkan juga perusahaan yang bergerak di bidang olahraga serta berinvestasi di bidang properti. Manchester United Ltd atau yang sekarang dikenal dengan Manchester United plc ini mengoperasikan Manchester United Football Club serta membangun hubungan pemasaran dan sponsorship dengan perusahaan-perusahaan internasional dan regional untuk meningkatkan merekanya (Manchester United, n.d.). Tidak hanya itu, Perusahaan ini juga menjual pakaian olahraga, pakaian rekreasi, pakaian latihan dan pakaian lainnya serta menjual produk berlisensi lainya dengan menampilkan merek dagang Manchester United. Produk-produk tersebut di distribusikan melalui pusat ritel bermerek Manchester United, platform *e-commerce* dan saluran distribusi grosir mitra perusahaan. Selain itu, perusahaan juga mendistribusikan konten sepak bola secara langsung dan bersama dengan mitra komersialnya serta menyiarkan hak siar televisi yang berkaitan dengan kompetisi klub bahkan menayangkan program-program Manchester United melalui saluran televisi MUTV ke berbagai belahan wilayah dunia (Manchester United, n.d.). Kini, Manchester United plc berkantor pusat di Manchester, Inggris.

### **Support Vector Machine (SVM)**

Dalam konteks analisis prediksi saham, sangat penting untuk memahami secara cermat setiap trend dan pergerakan yang terjadi. Hal ini memungkinkan untuk mengidentifikasi potensi cuan secara optimal. Namun hal tersebut tidak terlepas dari adanya analisis fundamental dan analisis teknikal yang telah cukup populer dikalangan para investor. Akan tetapi, pendekatan lain juga bisa dilakukan dalam melakukan analisis saham yakni dengan menggunakan metode algoritma support vector machine (SVM).

Support Vector Machine (SVM) pertama kali diperkenalkan oleh Vapnik pada tahun 1992 sebagai konsep unggulan dalam bidang pattern recognition dan memiliki kemampuan dalam memilih model otomatis serta tidak memiliki masalah overfitting. Sejauh ini, SVM tidak hanya mampu menyelesaikan kasus secara linier melainkan juga bisa berkerja pada kasus non-linier dengan menggunakan konsep karnel pada ruang kerja berdimensi tinggi, dengan mencari *hyperplane* yang dapat memaksimalkan jarak antar kelas data (Fadilah, Agfiannisa, & Azhar, 2020).

Tingkat keakurasian motode SVM sangat dipengaruhi oleh fungsi karnel dan parameter yang digunakan. Fungsi karnel ini digunakan untuk memetakan dimensi awal (dimensi yang lebih rendah) himpunan data ke dimensi baru (dimensi yang realtif tinggi). Lantas, fungsi karnel kelak akan menentukan fitur baru dimana *hyperplane* akan dicari serta menjadi garis pemisah beberapa kelompok data (Fadilah, Agfiannisa, & Azhar, 2020).

Dengan demikian, secara matematis *hyperplane* klasifikasi linier pada support vector machine (SVM) dapat dilihat melalui persamaan yakni:

|  |  |
| --- | --- |
|  | (2.1) |

atau dengan kata lain dapat juga menggunakan persamaan:

|  |  |
| --- | --- |
|  | (2.2) |

|  |  |
| --- | --- |
|  | (2.3) |

dimana untuk nilai merupakan himpunan bilangan traning, dan merupakan label dari kelas . Sedangkan untuk fungsi karnel sendiri digambarkan melalui persamaan sebagai berikut:

Tabel 2. 1 Rumus fungsi karnel

|  |  |  |
| --- | --- | --- |
| Karnel | Persamaan |  |
| Polynomial |  | (2.4) |
| RBF |  | (2.5) |

dimana untuk nilai dan merupakan pasang dari dua data traning dan parameter merupakan konstanta. Pemilihan fungsi karnel ini bergantung pada sifat data sehingga memainkan peran penting dalam mencari dan menentukan *hyperplane* yang terbaik. Oleh karena itu, tingkat akurasi dari metode support vector machine (SVM) sangat dipengaruhi sekali oleh faktor dari fungsi karnel dan parameter yang digunakan.

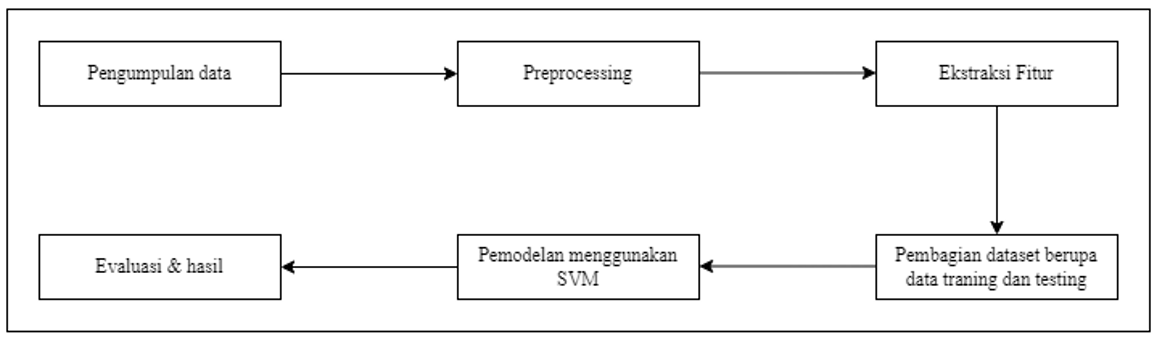
## **Kajian Hasil Riset**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **No** | Peneliti, Penerbit, dan Tahun Terbit | Judul | Algoritma | Hasil Penelitian |
| **1** | Peneliti (Wulandari & Anubhakti, 2022)  Ratih Febrilia Tri Wulandari & Dian Anubhakti  Penerbit dan Tahun Terbit  Indonesia Journal Information System (IDEALIS), vol 4 no 2 tahun 2021 | Implementasi Algoritma Support Vector Machine (SVM) Dalam Memprediksi Harga Saham PT. Garuda Indonesia Tbk | Support Vector Machine (SVM) | Hasil dari pengujian model SVM yakni menghasilkan performance prediction trend accuracy sebesar 0.545 dari 520 dataset yang ada dan dengan hasil nilai trend akurasi prediksi tersebut bahwa SVM dapat digunakan sebagai model untuk memprediksi harga saham penutupan pada PT. Garuda Indonesia, Tbk. |
| **2** | Peneliti (Syahputra, Ramadhan, & Burha, 2022)  Janur Syahputra, Rima Dias Ramadhan & Auliya Burhanudin  Penerbit dan Tahun Terbit  Journal of Dinda, vol 2 no 1 tahun 2022 | Prediksi Harga Saham Bank BRI Menggunakan Algoritma Linier Regression Sebagai Strategi Jual Beli Saham | Linier Regression | Penelitian ini melakukan pembagian data yakni 60:40, 65:35, 70:30, 75:25, dan 80:20 untuk mengetahui akurasi latih dan uji tertinggi. Rasio terbaik ada pada 80:20 yang menghasilkan akurasi train dan test sebesar 0,89 dan 0,91. Kemudian masing-masing data training dan data testing dimasukan kedalam model linear regression untuk kemudian dilakukan prediksi. Adapun hasil akurasi yang dihasilkan dari model yang sudah diuji dengan data prediksi dalam bentuk R¬2 yaitu sebesar 0.9127. Hasil error dari prediksi dihitung menggunakan MAPE dan menghasilkan persentase sebesar 13,751% untuk data pelatihan, 13,773% untuk data pengujian, dan 13,755% untuk data keseluruhan. Ketiga hasil tersebut mengidikasikan bahwa persentase error berada pada rentang 10% - 20%. |
| **3** | Peneliti (Maulana & Kumalasari, 2019)  Reza Maulana & Devy Kumalasari  Penerbit dan Tahun Terbit  Jurnal Informatika Kaputama (JIK), vol 3 no 1 tahun 2019 | Analisis Dan Perbandingan Algoritma Data Mining Dalam Prediksi Harga Saham GGRM | Algoritma Data Mining seperti Neural Network, Linear Regression, Support Vector Machine, Gaussian Process, dan Polynomial Regression. | Dalam penelitian ini akan dilakukan prediksi harga saham GGRM dengan membandingkan beberapa model algoritma yaitu Neural Network, Linear Regression, Support Vector Machine, Gaussian Process, dan Polynomial Regression. Untuk mengukur tingkat akurasi dari tiap model algoritma digunakan model validasi 10 Fold Cros Validation dan evaluasi menggunakan Root Mean Square Error (RMSE). Hasil penelitian diperoleh sebagai berikut:   * Algoritma NN memperoleh RMSE 612.472 +/- 89.402 (mikro 618.916 +/- 0.000) * Algoritma LR memperoleh RMSE 659.806 +/- 76.480 (mikro 664.360 +/- 0.000) * Algoritma SVM memperoleh RMSE 7668.182 +/- 627.179 (mikro 7694.784 +/- 0.000) * Algoritma GP memperoleh RMSE 52295.739 +/-448.393 (mikro 52297.714 +/- 0.000) * Algoritma PR memperoleh RMSE 13049082159319.207 +/-19076530613123.832 (mikro 23133462625506.945 +/- 0.000)   Dapat dilihat bahwa algoritma Neural Network memiliki hasil akurasi prediksi RMSE paling kecil yakni 612.474 +/- 89.402 (mikro: 618.916 +/- 0.000) dibandingkan dengan model algoritma lainnya. |
| **4** | Peneliti (Lestari & Pangaribuan, 2020)  Jefri Junifer Pangaribuan & Megawaty Lestari  Penerbit dan Tahun Terbit  Information System Development, vol 5 no 1 tahun 2020 | Perbandingan Metode Moving Average (MA) Dan Neural Network Yang Berbasis Algoritma Backpropagation Dalam Prediksi Harga Saham | Metode Moving Average (MA) Dan Neural Network Yang Berbasis Algoritma Backpropagation | Prediksi harga saham dengan metode Moving Average (MA) lebih akurat dibandingkan dengan Neural Network algoritma Backpropagation, dimana tingkat akurasi untuk Moving Average (MA) adalah 80,11% dan untuk Neural Network algoritma Backpropagation adalah 78,91%. |
| **5** | Peneliti (Lestari & Jasuni, 2023)  Siti Sarah Sobariah Lestari & Ana Yuliana Jasuni  Penerbit dan Tahun Terbit  Jurnal Bisnisman; Riset Bisnis dan Manajemen, vol 5 no 1 tahun 2023 | Analisis Rantai Makrov Lima Status Pada Return Harga Saham BBCA | Rantai Makrov | Berdasarkan hasil analisis yang telah dilakukan, Dalam jangka panjang peluang return harga saham harian penutupan Bank Central Asia Tbk. Terbesar terjadi pada status 2 dimana saaat return harga saham bernilai lebih besar dari -1 dan lebih kecil dari 0 dengan peluang sebesar 0,4186. |
| **6** | Peneliti (Fadilah, Agfiannisa, & Azhar, 2020)  Widya Rizka Ulul Fadilah, Dewi Agfiannisa & Yufis Azhar  Penerbit dan Tahun Terbit  Fountain of Informatics Journal, vol 5 no 2 tahun 2020 | Analisis Prediksi Harga Saham PT. Telekomunikasi Indonesia Menggunakan Metode Support Vector Machine | Support Vector Machine | Peneliti menggunakan Split Validation untuk melakukan pembagian data yakni 20% testing dan 80% training. Setelah itu, peneliti memilih karnel SVM yaitu kernel RBF untuk digunakan dalam algoritma SVM berdasarkan hasil pengujian dibandingkan dengan kernel lain seperti polynomial. Hal ini dikarenakan hasil pengujian yang diperoleh mendapatkan nilai akurasi tertinggi yakni 0.9653 dan nilai RMSE sebesar 0.0091. Peneliti juga melakukan pengujian dengan mencoba menggunakan algoritma KNN sebagai pembanding untuk menilai apakah SVM lebih unggul atau tidak dengan menggunakan dataset yang sama. Hasil yang diperoleh menggunakan algoritma KNN yakni nilai akurasi sebesar 0.9456 dan RMSE sebesar 0.1162. Dengan demikian, dapat diketahui bahwa algoritma SVM lebih baik ketimbang algoritma KNN dalam prediksi harga saham PT. Telekomunikasi. Hal ini dikarenakan algoritma SVM hampir medekati nilai yang sebenarnya dan tingkat error yang dihasilkan oleh metode KNN lebih besar dibandingkan oleh metode SVM. |

Tabel 2. 2 Hasil kajian teori

# **METODE PENELITIAN**

Dengan menggunakan data berupa histori pergerakan harga harian saham Manchester United (MANU) akan dilakukan analisis prediksi harga saham dengan menggunakan metode support vector machine (SVM) sehingga secara garis besar akan menghasilkan gambaran umum yakni sebagai berikut:



Gambar 3. 1 Gambaran umum penelitian

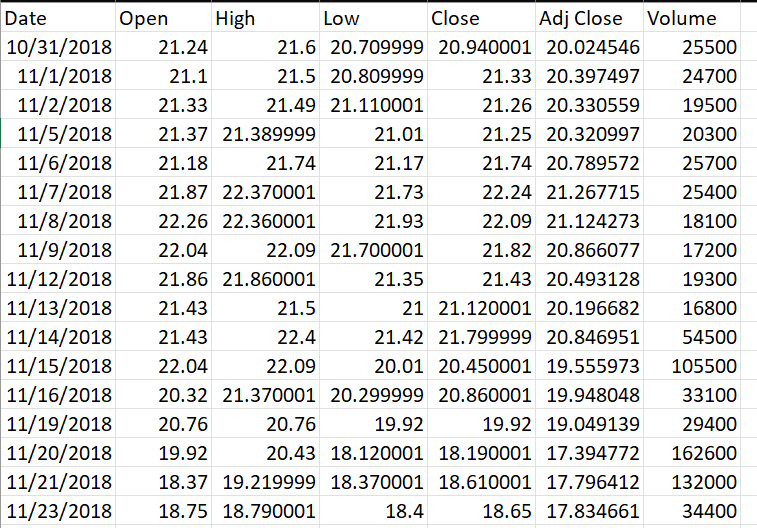
Gambaran umum tersebut merupakan proses atau tahapan yang akan dilakukan dalam penelitian sehingga dapat menjadi peta/alur dalam penelitian. Hal ini tentunya sangat membantu peneliti untuk mencapai pemahaman mendalam tentang langkah-langkah yang perlu diambil dalam mencapai tujuan penelitian. Oleh karena itu, jikalau merujuk pada gambaran umum maka peneliti akan melewati beberapa proses yakni pengumpulan data histori pergerakan harga harian saham Manchester United (MANU), preprocessing data yang bertujuan supaya data yang digunakan dapat lebih efektif dalam proses analisis lebih lanjut, ekstraksi fitur yang bertujuan untuk mengindentifikasi sekaligus mengekstraksi fitur yang relevan, pembagian dataset yang meliputi data traning dan data testing, pemodelan serta evaluasi dan hasil untuk mengetahui sejauh mana model berkerja dengan baik dan optimal dalam menjawab permasalahan yang dirumuskan. Detail dari setiap proses pada gambaran umum tersebut akan dijabarkan sebagai berikut:

## **Pengumpulan Data**

Pada penelitian ini data yang digunakan ialah data pergerakan harian harga saham Manchester United dengan kode saham yang terdaftar pada New York Stock Exchange (NYSC) ialah MANU. Data tersebut diperoleh melalui <https://finance.yahoo.com/> dalam rentang waktu 5 tahun terakhir yakni pada 31 Oktober 2018 sampai 30 Oktober 2023 dengan jumlah data 1257. Pada data tersebut mengandung 7 atribut yang dikemas dalam format file csv yakni sebagai berikut:

1. Date merupakan tanggal transaksi harga saham berlangsung.
2. Open merupakan harga buka saham pada waktu tertentu dengan tipe data bilangan desimal.
3. Close merupakan harga tutup saham pada waktu tertentu dengan tipe data bilangan desimal.
4. High merupakan harga tertinggi saham pada waktu tertentu dengan tipe data bilangan desimal.
5. Low merupakan harga terendah saham pada waktu tertentu dengan tipe data bilangan desimal.
6. Adj Close merupakan harga tutup yang disesuaikan pada waktu tertentu dengan tipe data bilangan desimal.
7. Volume merupakan jumlah saham yang diperdagangkan pada waktu tertentu dengan tipe data bilangan desimal.

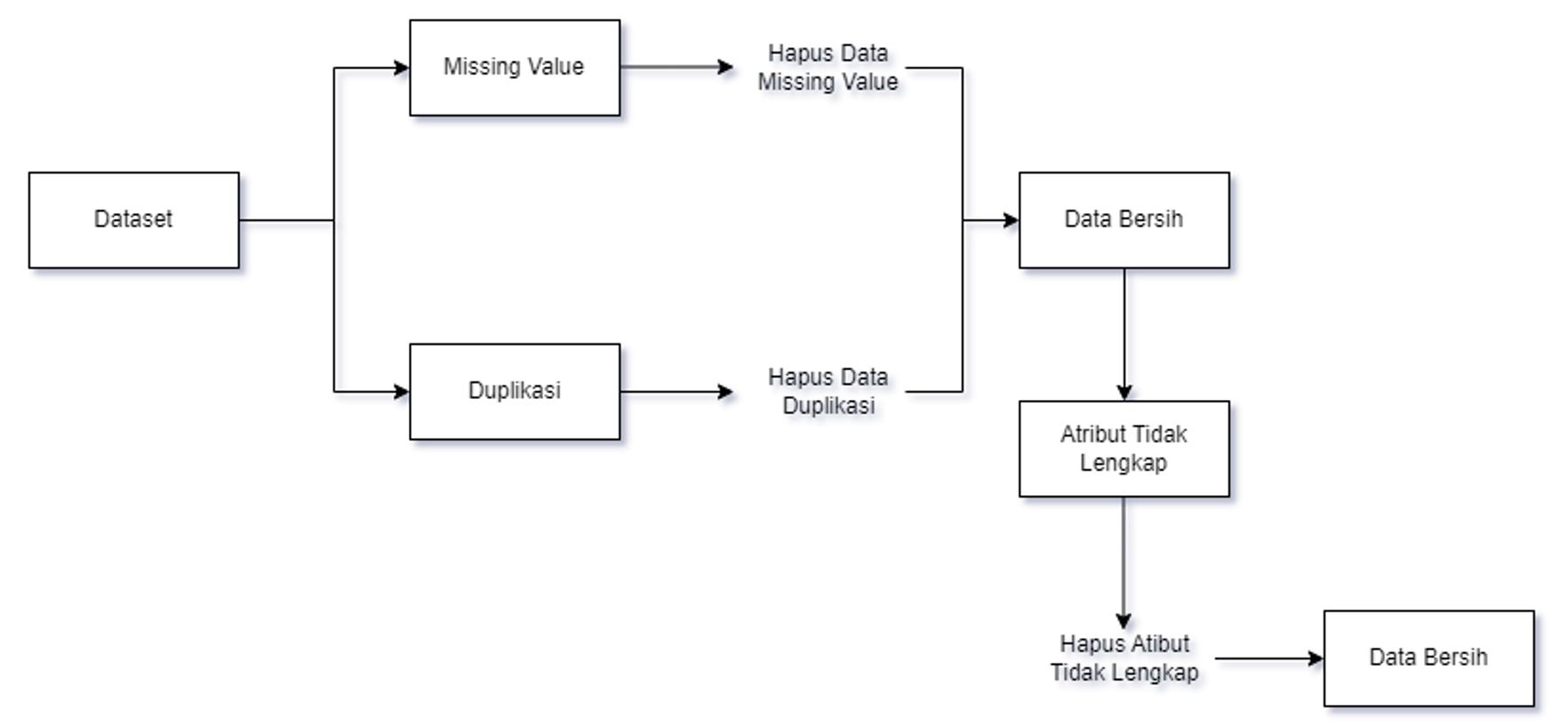
Berikut merupakan contoh data dari pergerakan harian harga saham Manchester United (MANU) dalam format CSV :



Gambar 3. 2 Contoh data harga saham Manchester United

## **Preprocessing**

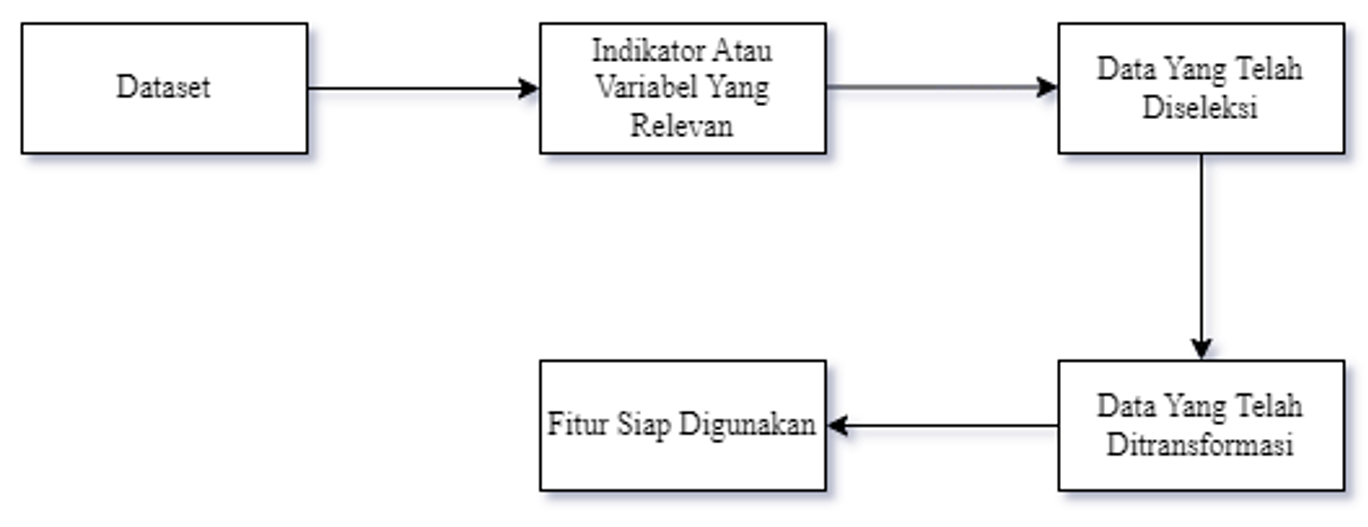
Pada tahap preprocessing akan dilakukan terlebih dahulu pengecekan untuk setiap nilai apakah mengandung *missing value* dan duplikasi. Jika terdapat nilai yang mengandung missing value maka akan dilakukan pendekatan berupa penghapusan terhadap nilai tersebut, begitu juga untuk nilai yang mengandung duplikasi. Akan tetapi, dikarenakan data yang dimiliki tidak mengandung *missing value* dan duplikasi maka tidak perlu melakukan tindakan tersebut namun tetap dilakukan pengecekan untuk memastikan data tersebut benar-benar tidak mengandung *missing value* dan duplikasi. Kemudian setelah itu, dikarenakan nilai pada dataset telah bertipe data float dan integer maka tidak perlu untuk melakukan konversi tipe data sehingga bisa dilakukan cleaning pada beberapa nilai atribut yang tidak lengkap untuk dihapus datanya. Hal ini bertujuan untuk meminimalisir tingkat error dalam proses pengklasifikasian data. Namun, hal ini tidak dilakukan pada atribut volume sebab atribut tersebut menggambarkan jumlah saham yang diperdagangkan sehingga bisa saja pada waktu tertentu volume transaksi bernilai 0 yang artinya tidak adanya penjualan saham yang terjadi pada hari tersebut. Berikut merupakan gambar flowchat dari proses preprocessing:



Gambar 3. 3 Alur proses preprocessing

## **Ekstraksi Fitur**

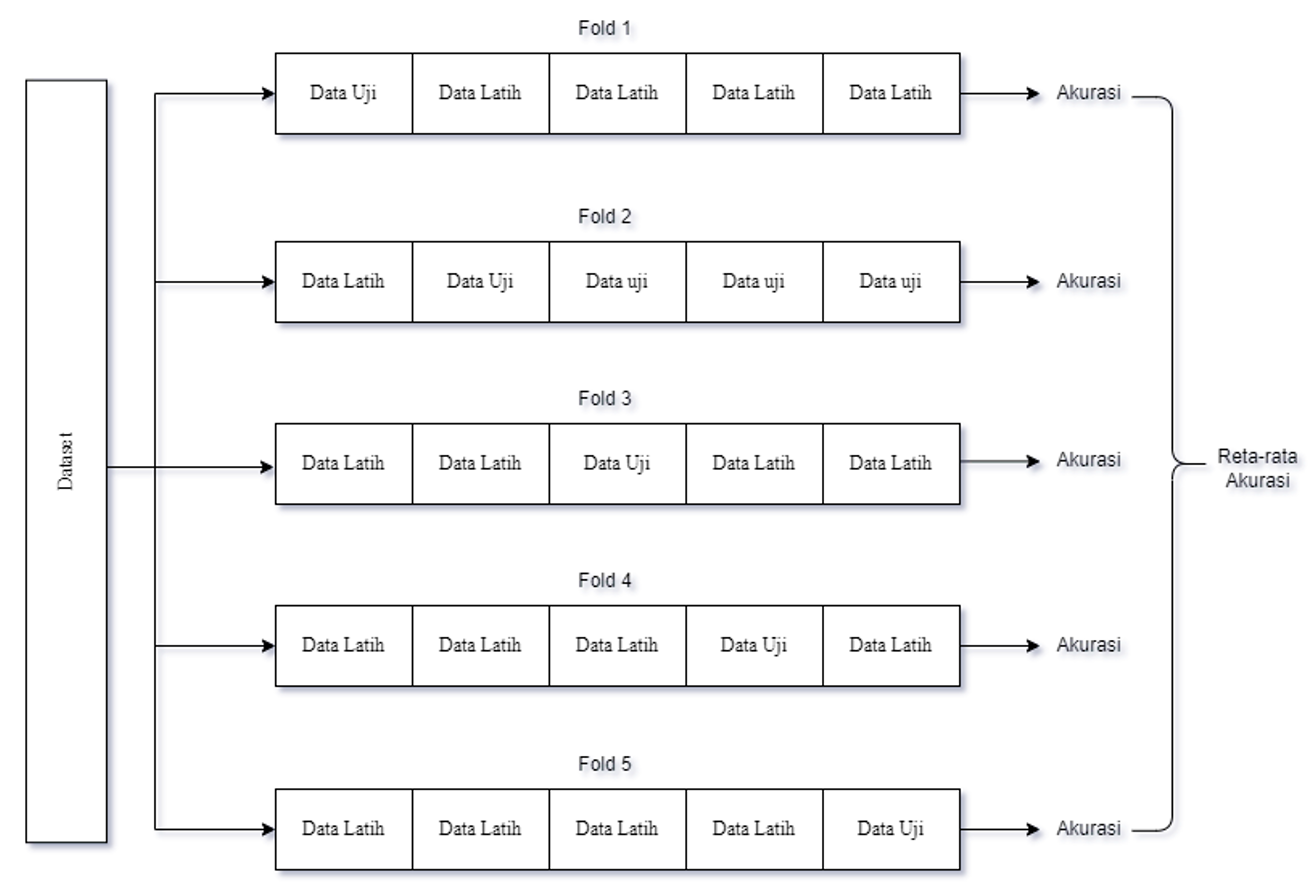
Ekstraksi fitur bertujuan untuk mengidentifikasi, memilih dan menggunakan berbagai indikator atau variabel yang dianggap relevan dalam mengembangkan model prediksi. Hal ini disebabkan karena tidak semua atribut data memiliki korelasi satu sama lainnya sehingga menyebabkan hadirnya data redundan sehingga bisa dihilangkan dalam kategorisasi atau klasifikasi. Sebagai contoh memilih dan menggunakan atribut close selama n periode dibagi dengan jumlah periode untuk menghitung simple moving average (SMA). Berikut merupakan gambar flowchat dari proses ekstraksi fitur:



Gambar 3. 4 Alur proses ekstraksi fitur

## **Pembagian Data**

Tujuan adanya pembagian antara data latih dan data uji yakni untuk mengetahui performa dari proses pelatihan model (Syahputra, Ramadhan, & Burha, 2022). Pada proses ini akan dilakukan pembagian data dengan persentasi yakni 80 % untuk data latih dan 20 % untuk data uji. Pembagian data tersebut menggunakan metode k-fold cross validation sehingga hal ini akan memungkinkan untuk memperoleh akurasi yang lebih baik. Berikut merupakan gambar flowchat dari proses k-fold cross validation:



Gambar 3. 5 Alur proses metode k-flod cross validation

Nilai k-fold yang digunakan dalam penelitian ini yakni mencoba semua kemungkinan nilai sampai mengetahui sejauh mana menemukan titik terbaik dari hasil yang diperoleh. Hal tersebut akan dibantu dengan grafik yang menujukkan berbagai nilai k-fold sehingga dapat digunakan dalam menentukan nilai k-fold yang optimal untuk digunakan.

## **Modeling**

Pada proses modeling terlebih dahulu menentukan karnel yang akan digunakan. Karnel merupakan salah satu metode non parametrik yang berfungsi untuk memetakan data dalam ruang berdimensi tinggi (Miraltamirus, Fitri, Vionanda, & Permana, 2023). Oleh sebab itu, karnel yang digunakan ialah polynomial kernel dan RBF karnel serta dengan nilai parameter C yakni 0.1, 0 dan 10 sehingga dari kedua karnel tersebut kelak akan dipilih yang menghasilkan hyperplane dengan nilai yang paling optimal.

## **Evaluasi**

Setelah melatih model maka kemudian melakukan evaluasi model dengan melihat hasil dari kinerja model yang telah dibangun. Evaluasi model tersebut dituangkan dalam confusion metrix atau metrix seperti akurasi, presisi, recall dan F1-score untuk melihat kemampuan prediksi model.

## **Skenario Pengujian**

Pada penelitian ini terdapat beberapa skenario pengujian yang akan dilakukan yakni seperti pembagian data menggunakan test split validation dan cross validation dengan nilai k-fold sama dengan 5 serta karnel yang digunakan berupa polynomial dan RBF dengan parameter C yakni 0.1, 0 dan 10. Berikut merupakan rincian terkait perihal diatas:

Tabel 3. 1 Skenario pengujian

|  |  |  |  |
| --- | --- | --- | --- |
| **No** | **Pembagian Data** | **C** | **Karnel** |
| 1 | Test Split | 0.1 | Polynomial |
| 2 | Test Split | 0 | Polynomial |
| 3 | Test Split | 10 | Polynomial |
| 4 | Cross validation | 0.1 | Polynomial |
| 5 | Cross validation | 0 | Polynomial |
| 6 | Cross validation | 10 | Polynomial |
| 7 | Test Split | 0.1 | RBF |
| 8 | Test Split | 0 | RBF |
| 9 | Test Split | 10 | RBF |
| 10 | Cross validation | 0.1 | RBF |
| 11 | Cross validation | 0 | RBF |
| 12 | Cross validation | 10 | RBF |

# **DAFTAR PUSTAKA**

Fadilah, W. R., Agfiannisa, D., & Azhar, Y. (2020). Analisis Prediksi Harga Saham PT. Telekomunikasi Indonesia Menggunakan Metode Support Vector Machine. *Fountain of Informatics Journal, 5*.

Lestari, M., & Pangaribuan, J. J. (2020). Perbandingan Metode Moving Average (MA) Dan Neural Network Yang Berbasis Algoritma Backpropagation Dalam Prediksi Harga Saham. *Information System Development, 5*.

Lestari, S. S., & Jasuni, A. Y. (2023). Analisis Rantai Makrov Lima Status Pada Return Harga Saham BBCA. *Jurnal Bisnisman; Riset Bisnis dan Manajemen, 5*.

Manchester United. (n.d.). *Manchester United plc*. (Yahoo Finance) Retrieved 11 25, 2023, from https://finance.yahoo.com/quote/MANU/profile?p=MANU

Maulana, R., & Kumalasari, D. (2019). Analisis Dan Perbandingan Algoritma Data Mining Dalam Prediksi Harga Saham GGRM. *Jurnal Informatika Kaputama (JIK), 3*.

Miraltamirus, I., Fitri, F., Vionanda, D., & Permana, D. (2023). Stock Price Prediction of PT Bank Syariah Indonesia Tbk Using Support Vector Regression. *UNP JOURNAL OF STATISTICS AND DATA SCIENCE, 1*, 112-119.

Rusmalawati, V., Furqon, M. T., & Indriati. (2018). Peramalan Harga Saham Menggunakan Metode Support Vector Regression (SVR) Dengan Particle Swarm Optimization (PSO). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, 2*, 1980-1990.

Syahputra, J., Ramadhan, R. D., & Burha, A. (2022). Prediksi Harga Saham Bank BRI Menggunakan Algoritma Linier Regression Sebagai Strategi Jual Beli Saham. *Journal of Dinda, 2*.

Wulandari, R. F., & Anubhakti, D. (2022). Implementasi Algoritma Support Vector Machine (SVM) Dalam Memprediksi Harga Saham PT. Garuda Indonesia Tbk. *Indonesia Journal Information System (IDEALIS), 4*.