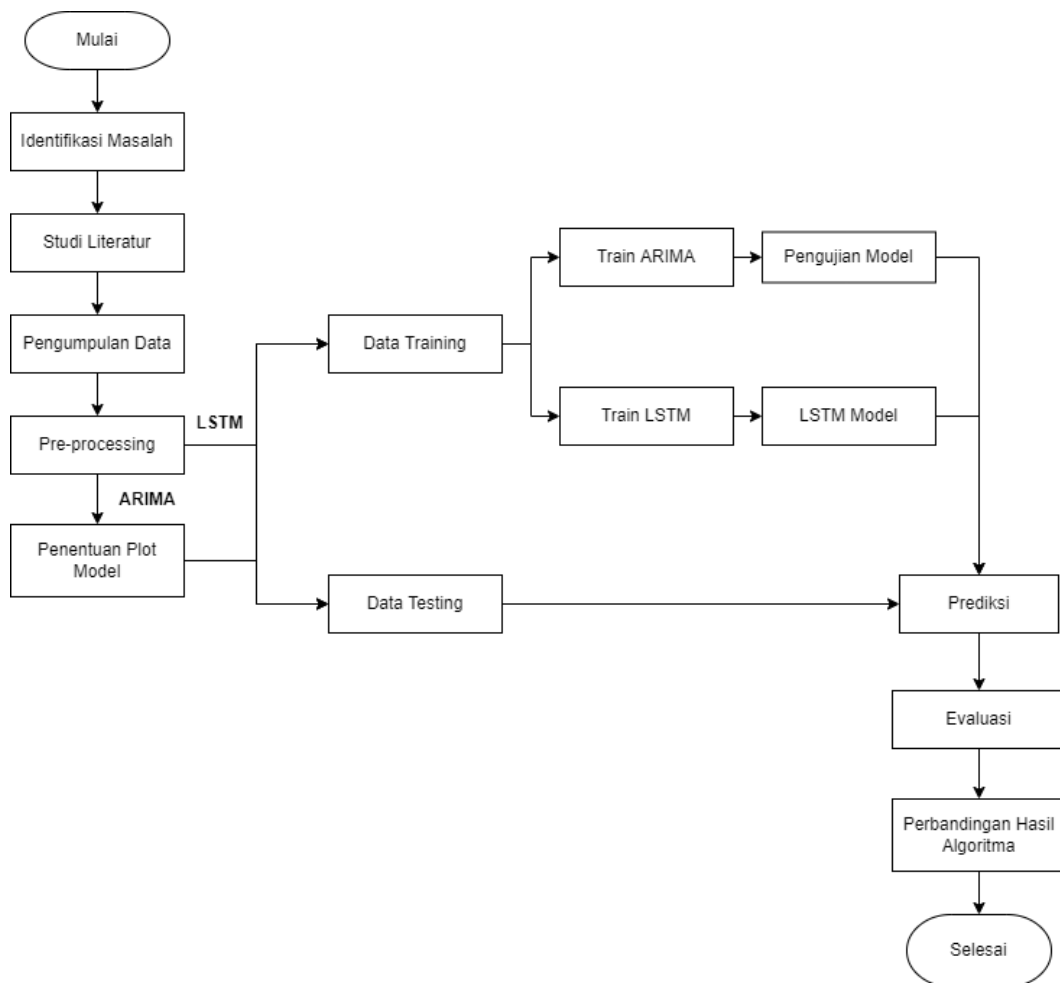


BAB III

METODE PENELITIAN

3.1. Tahapan Penelitian

Berdasarkan berbagai penelitian terkait sebelumnya, maka untuk dapat mencapai tujuan penelitian, penulis menyusun prosedur penelitian yang berisikan gambaran umum dari tahapan penelitian yang akan dilakukan. Tahapan penelitian dapat dilihat pada Gambar 3.1.



Gambar 3.1 Diagram Alir Tahapan Penelitian

3.2. Identifikasi Masalah

Tahapan yang pertama dilakukan oleh peneliti pada penelitian ini adalah identifikasi dan perumusan masalah. Hal ini dilakukan untuk menemukan

permasalahan yang relevan dalam lingkungan penelitian. Pada tahap ini penulis mendapatkan gambaran ruang lingkup permasalahan.

3.3. Studi Literatur

Pada tahap ini, peneliti mencari teori pendukung dan informasi yang sudah ada sebelumnya. Sumber didapatkan dari buku ataupun jurnal penelitian yang memiliki materi berkaitan dengan peramalan harga saham, metode ARIMA, model LSTM serta informasi lainnya yang berhubungan dengan kegiatan penelitian ini. Studi literatur menjadi landasan dasar yang penting dalam pembuatan penelitian yang dilakukan.

3.4. Pengumpulan Data

Pada tahap ini, data awal yang digunakan dalam penelitian adalah data time series. Data deret waktu memiliki nilai variabel numerik untuk setiap interval harian. Data saham ANTM, ADRO, ICBP, KLBF dan TLKM dari *website* Yahoo Finance, termasuk tanggal (date), buka (buka), tutup (close), tinggi (tinggi), rendah (rendah), harga penutupan yang disesuaikan (harga penutupan yang disesuaikan) dan volume (jumlah saham yang diperdagangkan).

3.5. Pra-proses Data

Tahapan ini melakukan beberapa tahapan proses data yang menghapus beberapa atribut yang sekiranya tidak diperlukan atau tidak memengaruhi hasil akhir (*attribute removal*), cek *missing value*, uji stasioner, mengubah data menjadi stasioner (*differencing*), normalisasi *MinMax*, pembagian data *train* dan data *test*.

3.5.1. Attribute Removal

Attribute Removal dilakukan seperti seleksi atribut yang tidak dibutuhkan pada proses peramalan yang akan datang. Atribut tidak akan digunakan karena tidak akan mempengaruhi hasil akhir. Pemilihan atribut yang digunakan disesuaikan dengan kebutuhan data dari metode ARIMA dan juga LSTM.

3.5.2. Missing Value

Data yang tidak lengkap atau data yang hilang dapat mempengaruhi hasil analisis data yang diolah. Data yang hilang tersebut biasanya disebabkan terdapat kesalahan ketika mengisi data. Jika *missing value* data hanya sedikit, misalkan berjumlah 1% dari total keseluruhan data, maka tidak terlalu bermasalah terhadap pemodelan atau analisis nantinya, tetapi jika data tersebut banyak yang hilang maka perlu diatasi untuk diproses lebih lanjut.

3.5.3. Uji Stasioner

Uji stasioner dilakukan untuk mengetahui apakah data sudah stasioner atau belum, hal tersebut dilakukan khusus untuk metode ARIMA. Uji stasioner dapat dilakukan dengan melihat langsung plot grafik dari dataset yang dipakai, jika grafik berada pada garis lurus maka data sudah berbentuk stasioner. Cara lain yang dapat dilakukan dengan mudah adalah dengan melakukan uji *Augmented Dickey-Fuller* (ADF). Uji ADF ini membandingkan nilai statistik ADF dengan nilai kritis dan juga melihat apakah p-value lebih kecil atau lebih besar dari 0,05. Jika nilai statistik ADF < nilai kritis dan p-value < 0,05 maka dinyatakan stasioner dan jika keadaan menunjukkan sebaliknya maka data tersebut tidak stasioner atau jika salah satu tidak memenuhi syarat untuk kestasioneran maka dilihat terlebih dahulu dari plot ACF dan PACF. Kemudian jika belum stasioner maka akan dilakukan *differencing*.

3.5.4. Differencing

Untuk melakukan proses pada metode ARIMA, suatu deret waktu yang tidak stasioner harus diubah menjadi data stasioner dengan melalui proses *differencing*. Proses *differencing* adalah menghitung perubahan atau selisih nilai observasi. Nilai selisih yang diperoleh dilihat lagi apakah sudah stasioner atau belum. Jika belum stasioner maka dilakukan *differencing* lagi. Proses *differencing* tersebut dapat dilakukan dengan cara mengurangi nilai satu periode dengan nilai periode sebelumnya.

3.6. Penentuan Plot Model

ARIMA memiliki tiga buah ordo yaitu p , d , q . Penentuan tiga ordo tersebut dilakukan dengan menggunakan metode *correlogram* yang dapat dianalisis melalui *Autocorrelation Function* (ACF) dan *Partial Autocorrelation Function* (PACF). Plot ACF digunakan untuk menjelaskan korelasi antara data time series yang berdekatan dengan selisih waktu (*time lag*). Parameter p dan q dapat dilihat jika dalam plot tersebut terdapat lag yang mengalami kondisi *cut off*. Kondisi tersebut terjadi jika terdapat lag yang menembus area batas signifikan.

3.7. Proses Model LSTM

Setelah melalui tahapan pra-proses khusus metode LSTM maka dilakukan tahapan selanjutnya yaitu membagi data menjadi data *train* dan juga data *test*, masing-masing memiliki porsi data sebesar 80% dan 20%.

Model LSTM ini terdapat inisialisasi untuk parameter-parameter yang dibutuhkan dalam pemodelan yaitu *hidden layer*, *units*, *epoch*, dan juga *batch_size*. Selain itu proses pemodelan LSTM juga akan dilakukan dengan fungsi optimasi dan *dropout*. Dalam penelitian ini akan menggunakan fungsi optimasi *adam* yang berguna untuk memaksimalkan tingkat akurasi yang dihasilkan model dan juga menentukan bobot yang optimal. Pertama, model LSTM akan melalui proses *train* dengan melewati beberapa *gates* seperti yang sudah dijelaskan sebelumnya, data akan berhenti dilatih jika sudah mencapai batas parameter yang sudah ditentukan. Kemudian model tersebut akan dilakukan *testing* dengan data *test* yang sudah dibagi sebelumnya.

3.8. Pengujian Model

Pengujian model tersebut dilakukan dengan prediksi dan menghitung error rate hasil prediksi. Prediksi dilakukan untuk memperoleh harga saham pembuka selama satu minggu kedepan. Kemudian diperoleh hasil prediksi selanjutnya dihitung error rate setiap hasil prediksi menggunakan MAPE.

3.9. Evaluasi Hasil Perbandingan Algoritma

Pada tahapan terakhir ini evaluasi untuk menguji performa peramalan menggunakan Mean Absolute Percent Error (MAPE). MAPE adalah rata-rata dari total persen kesalahan atau selisih antara data aktual dan prediksi. Hasil prediksi

yang telah didapatkan dari model akan dilihat perbandingan metode ARIMA dan LSTM yang memiliki nilai MAPE terkecil dari 2 algoritma yang menjadi perbandingan untuk harga saham sehingga dapat memvalidasi performa model.

3.10. Alat Bantu Penelitian

Dalam melakukan penelitian ini, diperlukan perangkat keras serta perangkat lunak yang dapat menunjang penulis dalam penelitian. Berikut adalah uraian penunjang yang digunakan pada penelitian ini.

3.9.1. Perangkat Keras

Spesifikasi perangkat keras yang digunakan berupa laptop dengan spesifikasi sebagai berikut:

- a. Model : Lenovo ideapad Slim 3
- b. Processor : Intel ® Core™ i5-1035G1 (4C / 8T, 1.0 / 3.6GHz, 6MB)
- c. Memory : 8 GB RAM
- d. Hard Disk : 512 SSD
- e. System Type : 64-bit Operating System

3.9.2. Perangkat Lunak

Perangkat lunak yang digunakan selama proses penelitian adalah sebagai berikut:

- a. Sistem Operasi : Windows 10 Pro 64-bit
- b. Google Colab
- c. Spyder
- d. Microsoft Office 2019