

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Kebutuhan finansial di masa depan adalah masalah yang akan dihadapi oleh banyak orang. Berinvestasi adalah kegiatan menanam modal pada sarana investasi dengan harapan dapat menunda konsumsi saat ini dan memperoleh manfaat ekonomi. Investasi merupakan penanaman modal untuk satu atau lebih aktiva yang dimiliki dan biasanya berjangka waktu lama dengan harapan menghasilkan keuntungan di masa-masa yang akan datang (Sunariyah, 2011). Salah satu investasi yang paling banyak diminati adalah saham.

Di Indonesia, untuk mengetahui pergerakan harga saham rata-rata di dalam negeri dapat diperoleh dari Indeks Harga Saham Gabungan (IHSG). IHSG merupakan gabungan dari beberapa saham yang ada di Indonesia dan tercatat di bursa efek Indonesia yang dihitung nilai rata-rata dari beberapa saham tersebut (Tandelilin, 2010). IHSG digunakan untuk mengetahui kinerja jumlah saham yang berada di Bursa Efek dengan cara menjumlahkan seluruh harga saham dan membaginya dengan total jumlah saham tersebut yang tercatat di Bursa Efek Indonesia. IHSG menggambarkan seluruh kombinasi harga saham pasar modal, sehingga kondisi pasar modal domestik dapat dibaca dari IHSG. Karena volatilitas IHSG yang tinggi berdasarkan waktu (mingguan), diperlukan metodologi peramalan untuk membantu investor mengambil keputusan.

Data IHSG yang diperoleh dari suatu penelitian seringkali berupa fungsi atas waktu, dan ada hubungan antar pengamatannya (autokorelasi), sehingga analisis regresi sederhana tidak bisa digunakan untuk menganalisis hubungan fungsional antara pengamatan dengan waktunya. Jenis data seperti ini disebut data deret waktu (*times series*) dan untuk menganalisisnya harus menggunakan metode analisis deret waktu. Metode *K-Nearest Neighbor* merupakan suatu metode yang bisa diterapkan dalam menganalisis data deret waktu. Metode ini juga bisa digunakan pada data yang tidak memenuhi asumsi klasik dan jika karakteristik data yang tidak linear.

Kelebihan dari metode ini juga bisa melakukan prediksi pada data deret waktu dengan ukuran data yang besar dan peubah penjelas yang banyak.

Metode *K-Nearest Neighbor* (KNN) merupakan metode yang menggunakan algoritma *supervised learning*. Algoritma *supervised learning* ini tujuannya untuk mendapatkan pola baru berdasarkan data-data yang sudah diketahui sebelumnya. Metode ini dianggap sebagai suatu metode analisis data yang mudah untuk diterapkan dalam analisis data yang besar dengan peubah penjelas yang banyak (Alkhatib, Najadat, Hmeidi, dan Shatnawi, 2013). Sebelumnya, metode KNN merupakan metode untuk analisis klasifikasi, tetapi beberapa tahun terakhir metode KNN bisa digunakan untuk prediksi.

Metode ini memiliki kekurangan dalam melakukan prediksi, yaitu tidak dapat secara pasti melakukan perhitungan jarak ( $k$ ), yang mana memiliki hasil yang mendekati nilai sebenarnya. Oleh karena itu, untuk mendapatkan hasil yang konsisten dan baik perlu dilakukan optimasi dengan cara menambahkan suatu teknik. Teknik *Ensemble* adalah teknik yang akan digunakan untuk optimasi atau meningkatkan secara optimal kinerja metode KNN dalam melakukan prediksi (Sinta, Wijayanto, dan Sartono, 2014).

Teknik *Ensemble* adalah salah satu teknik penting untuk meningkatkan kemampuan prediksi dari berbagai hasil prediksi yang berbeda. Teknik ini menggabungkan berbagai hasil prediksi terbaik diantara banyak kandidat prediksi yang ada menjadi suatu satu buah prediksi akhir. Dengan kata lain, metode *Ensemble* dalam metode KNN tidak perlu mencari nilai  $k$  (jumlah tetangga terdekat) yang optimal (Satriya, Santoso, dan Sutrisno, 2018).

Berdasarkan uraian latar belakang yang telah diuraikan, maka pada penelitian ini akan dilakukan prediksi nilai Indeks Harga Saham Gabungan (IHSG) menggunakan metode *Ensemble K-Nearest Neighbor*. Tujuan dari penelitian ini adalah untuk menganalisis hasil akurasi metode *K-Nearest Neighbor* yang dioptimasi dengan teknik *Ensemble* dalam melakukan prediksi nilai Indeks Harga Saham Gabungan (IHSG) di Indonesia.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang dikemukakan, rumusan masalah dalam penelitian ini adalah apakah prediksi nilai Indeks Harga Saham Gabungan (IHSG) menggunakan metode *K-Nearest Neighbor* yang dioptimasi dengan teknik *Ensemble* menghasilkan akurasi yang lebih akurat dibandingkan dengan KNN tunggal?

## 1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang dikemukakan, tujuan penelitian dalam penelitian ini adalah menganalisis hasil akurasi dari prediksi nilai Indeks Harga Saham Gabungan (IHSG) menggunakan metode *K-Nearest Neighbor* yang dioptimasi dengan teknik *Ensemble*.

## 1.4 Batasan Masalah

Batasan masalah yang didefinisikan dalam penelitian ini adalah sebagai berikut:

1. Data yang digunakan dalam penelitian ini adalah data yang berasal dari website *yahoo.finance.com* dan *id.investing.com*.
2. Data yang digunakan adalah data mingguan nilai IHSG Indonesia pada Januari 2021 sampai Maret 2022 dengan total periode  $(t) = 64$ .
3. Ada 2 variabel dalam penelitian ini yaitu variabel respon ( $Y$ ) dan variabel prediktor ( $X$ ). Variabel prediktor terdiri dari harga emas ( $X_1$ ), Nikkei 225 ( $X_2$ ), dan nilai tukar rupiah terhadap dollar ( $X_3$ ) dan variabel responnya yaitu jumlah nilai Indeks Harga Saham Gabungan (IHSG) ( $Y$ ).
4. Tingkat akurasi diukur berdasarkan *Mean Absolute Percentage Error* (MAPE).

## 1.5 Tinjauan Pustaka

Penelitian sebelumnya yang dilakukan oleh Satriya, Santoso dan Sutrisno (2018) yang memprediksi nilai tukar rupiah terhadap dollar Amerika dengan metode *Ensemble K-Nearest Neighbor*, dengan melakukan beberapa skenario

pengujian, salah satunya dengan melihat nilai MAE, MAPE dan RMSEP. Penelitian tersebut disimpulkan bahwa semakin besar nilai  $k$  maka nilai MAE, MAPE, dan RMSEP semakin besar. Nilai terkecil terletak pada  $k=3$  dengan nilai MAE = 360,8, MAPE = 2,74%, dan RMSEP = 421,72, sedangkan pada metode *Ensemble* nilai MAE = 456,56, MAPE = 3,47%, dan RMSEP = 534,88. Hasil dari metode *Ensemble* KNN tidak berbeda jauh dari metode KNN yang memiliki nilai terkecil karena pada metode *Ensemble* KNN menggunakan teknik *Weighted Mean* yang mana data yang digunakan untuk menghitung *Ensemble* KNN berasal dari hasil prediksi metode KNN pada setiap  $k$ . Karena hal tersebut hasil dari metode *Ensemble* KNN tidak akan berada di bawah nilai terkecil dan di atas nilai terbesar dari metode KNN.

Alfin, Murtopo dan Fadilah (2022) melakukan penelitian tentang prediksi produksi bawang merah menggunakan metode *Ensemble K-Nearest Neighbor*. Hasil dari penelitian ini menunjukkan bahwa metode *Ensemble* KNN memiliki hasil yang lebih baik dibandingkan KNN Tunggal. Hasil prediksi harga bawang merah pada bulan Januari 2022 Rp 23.450 per kilogram dengan nilai MAPE sebesar 1,46%, MAE sebesar 123,10 dan RMSE sebesar 152,75 yang mana masuk pada akurasi yang tinggi. Nilai MAPE, MAE dan RMSE kecil ketika ( $k$ ) tetangga terdekat besar

Jusman, Nur'eni dan Handayani (2022) melakukan penelitian tentang prediksi IHSG menggunakan metode *Ensemble K-Nearest Neighbor*. Penelitian menyimpulkan bahwa hasil dari metode *Ensemble* KNN memiliki hasil yang lebih baik dibandingkan KNN Tunggal. Hasil prediksi dari metode *Ensemble* KNN rata-rata sebesar 6078, 634 dengan nilai MAPE sebesar 7,16% yang mana masuk pada akurasi yang tinggi.

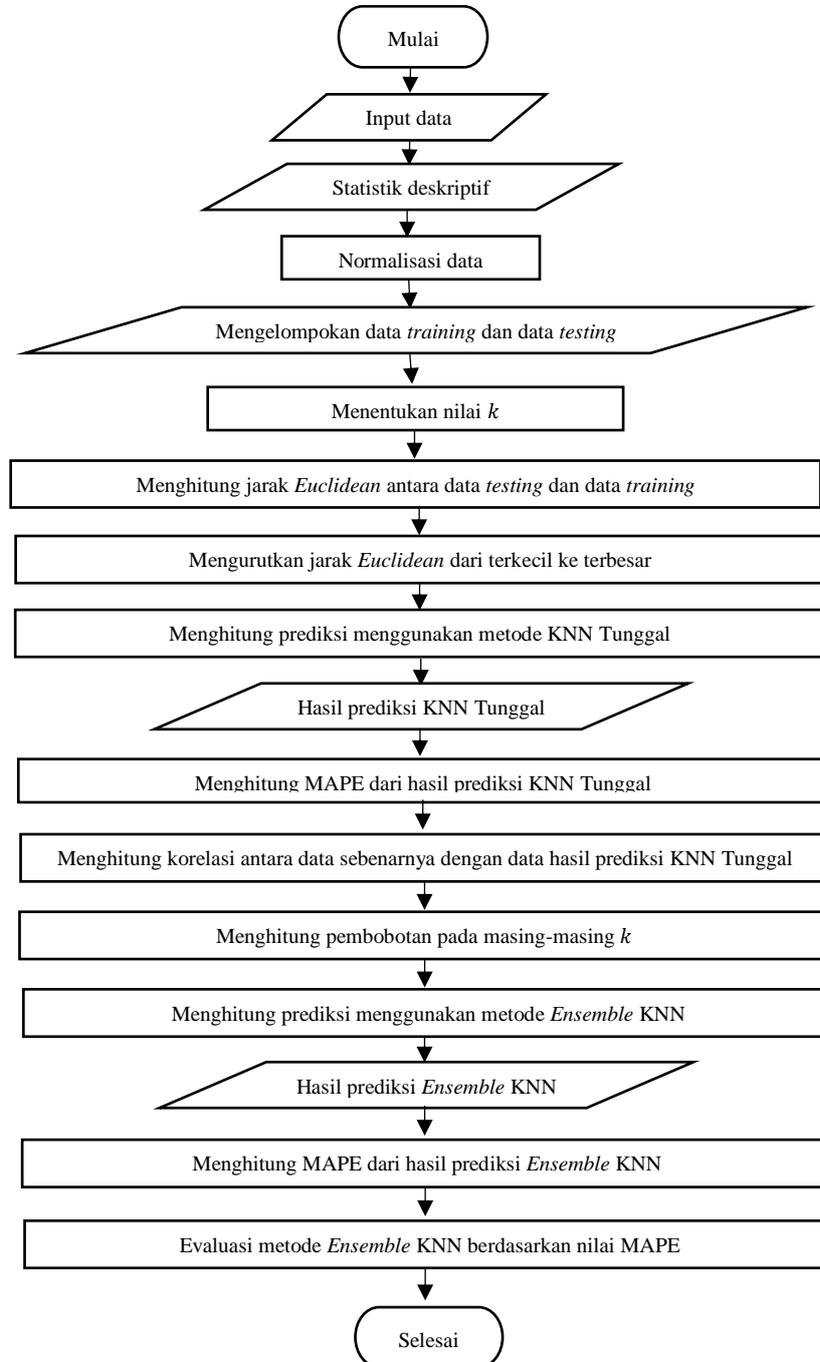
Sutanto, Murhadi, dan Ernawati (2013) melakukan penelitian tentang analisis pengaruh ekonomi makro, indeks Dow Jones, dan indeks Nikkei 225 terhadap IHSG di BEI menggunakan teknik analisis Regresi Linier Berganda. Penelitian menyimpulkan bahwa harga emas dunia, indeks Nikkei 225 dan indeks Dow Jones berpengaruh positif signifikan terhadap IHSG.

## 1.6 Metodologi Penelitian

Penelitian ini termasuk dalam kategori penelitian deskriptif kuantitatif. Penelitian deskriptif kuantitatif adalah penelitian yang menggambarkan suatu kejadian berdasarkan karakteristik dan fakta-fakta yang ditemukan. Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari website *yahoo.finance.com* dan *id.investing.com*, yang terdiri dari data Indeks Harga Saham Gabungan (IHSG) sebagai variabel respon, harga emas, indeks Nikkei 225 dan nilai tukar rupiah terhadap dolar sebagai variabel prediktor. Data yang digunakan merupakan data mingguan nilai IHSG Indonesia pada Januari 2021 sampai Maret 2022 dengan total periode ( $t$ ) = 64. Adapun metode yang digunakan untuk data ini menggunakan metode *K-Nearest Neighbor* yang di optimasi dengan menggunakan teknik *Ensemble*. Tahapan-tahapan analisis dalam melakukan prediksi nilai Indeks Harga Saham Gabungan (IHSG) dengan metode *Ensemble K-Nearest Neighbor* adalah sebagai berikut:

1. Pengumpulan data.
2. Statistik deskriptif dari data.
3. Normalisasi data.
4. Setelah data sudah dinormalisasi dilanjutkan dengan membagi data *training* dan data *testing*.
5. Menentukan nilai  $k$ .
6. Menghitung jarak *Euclidean* antara data *training* dan data *testing*.
7. Mengurutkan jarak *Euclidean* dari jarak yang terkecil hingga terbesar.
8. Menghitung prediksi dengan menggunakan metode KNN.
9. Menghitung MAPE dari hasil prediksi KNN.
10. Menghitung korelasi antara data sebenarnya dengan data hasil prediksi KNN.
11. Menghitung pembobotan pada masing-masing  $k$ .
12. Menghitung prediksi menggunakan metode *Ensemble* KNN.
13. Menghitung MAPE dari hasil prediksi *Ensemble* KNN
14. Melakukan evaluasi pada metode *Ensemble* KNN berdasarkan nilai akurasi dengan melihat *Mean Absolute Percentage Error* (MAPE).
15. Selesai.

Langkah-langkah analisis dalam penelitian ini dapat dilakukan dengan diagram alir yang ditampilkan pada Gambar 1.1.



**Gambar 1.1** Diagram alir *ensemble K-Nearest Neighbor*