

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Prediksi adalah kegiatan yang memperkirakan suatu nilai pada masa yang akan datang dengan menggunakan data masa lalu. Tujuan dari diadakannya prediksi sendiri yaitu untuk memperoleh informasi sebagai bentuk usaha dalam pengambilan keputusan di masa yang akan datang [1]. Maka dari itu, saat ini banyak lembaga ataupun perusahaan yang menerapkan metode prediksi untuk membantu dalam pengambilan keputusan.

Berdasarkan jenisnya, prediksi terbagi menjadi dua yaitu prediksi kualitatif dan prediksi kuantitatif. Prediksi kualitatif merujuk pada penilaian data dari masa lalu berupa pendapat, pengetahuan, intuisi, dan pengalaman. Sedangkan prediksi kuantitatif merujuk pada data historis numerik seperti deret waktu (*time series*) dari masa lalu [2]. Pada prediksi kuantitatif, banyak metode yang tersedia untuk memprediksi data *time series*, baik metode tradisional seperti *Autoregressive Integrated Moving Average* (ARIMA), *Exponential Smoothing*, analisis *trend*, maupun dengan metode *machine learning* seperti *Convolutional Neural Network* (CNN), *K-Nearest Neighbor* (KNN), *Multilayer Perceptron* (MLP), *Long Short Term Memory* (LSTM), dll.

Untuk prediksi dengan metode *machine learning*, saat ini sudah banyak penelitian yang menerapkannya. Misalnya penelitian dalam memprediksi harga cabai rawit yang dilakukan oleh Sebastianus Reczy (2020) dengan algoritma *K-Nearest Neighbor* [3]. Di mana tingkat akurasi yang dihasilkan sebesar 72,22% dengan nilai  $k$  yang menghasilkan nilai tertinggi yaitu  $k=13$ . Penelitian ini

menggunakan atribut *dataset* nilai luas panen, produksi panen, dan curah hujan untuk pelatihan dan pengujian sistem. Nilai *output* yang dikeluarkan dari hasil prediksi pun berupa kelas, yaitu kelas murah, mahal, dan sangat mahal dari sebuah harga yang telah diprediksikan. Karena nilai prediksi harga cabai rawit yang dihasilkan hanya berupa kelas, Maka peneliti pun penasaran bagaimana bila nilai prediksi harga cabai rawit yang dihasilkan berupa *time series* dan dalam bentuk grafik agar terlihat lebih jelas. Serta mencari atribut *dataset* lain yang mempengaruhi harga cabai rawit dan tentunya mudah didapat.

Maka dari itu, peneliti pun tertarik pada suatu algoritma *deep learning* yang bernama *Long Short Term Memory* (LSTM) untuk diimplementasikan pada penelitian kali ini dan sebagai pembaruan algoritma pada penelitian prediksi harga cabai rawit sebelumnya. Alasan peneliti memilih algoritma ini dikarenakan nilai *output* yang dihasilkan bisa berupa nilai deret waktu (*time series*) dan disajikan dalam bentuk grafik agar lebih spesifik dalam menunjukkan hasil prediksi seperti yang sudah dijelaskan sebelumnya [4]. Sehingga algoritma ini cocok digunakan untuk memprediksi harga cabai rawit yang bersifat *time series* dan kontinyu.

*Long Short Term Memory* (LSTM) adalah salah satu algoritma regresi pada *deep learning* yang merupakan evolusi dari arsitektur jaringan *Recurrent Neural Network* (RNN). LSTM diciptakan oleh Sepp Hochreiter dan Jürgen Schmidhuber pada tahun 1997. Di mana pada LSTM nanti akan dilakukan sebuah modifikasi pada RNN dengan menambahkan *memory cells* yang dapat menyimpan pola dalam jangka waktu lebih lama, dengan kemampuan yang selektif yang dapat memanggil kembali data [5]. Sehingga algoritma LSTM lebih cocok untuk memprediksi data *time series* yang memiliki dependensi terhadap data masa lalu dan dinilai lebih

mampu dalam melatih lebih banyak data-data sebelumnya untuk dijadikan bahan dalam menemukan pola tersembunyi. Sehingga pola-pola yang sudah ditemukan bisa memprediksi data pada masa yang akan datang yang bahkan dalam rentang waktu panjang.

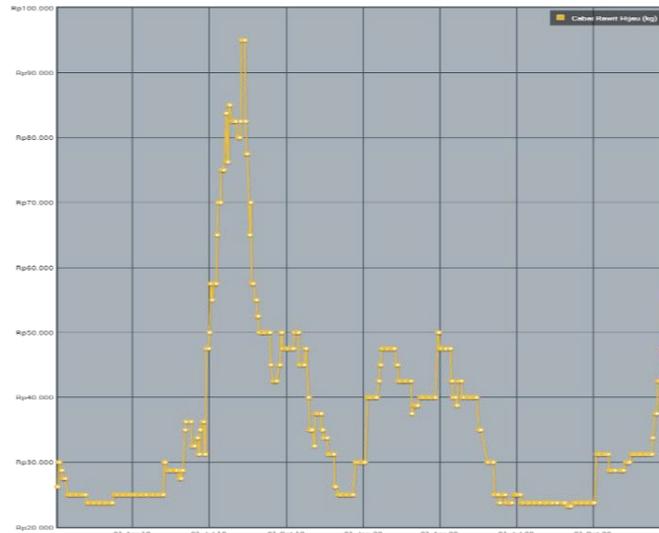
Untuk penelitian ini, peneliti menggunakan model *vanilla* LSTM atau LSTM konvensional, *stack* LSTM dan *bidirectional* LSTM sebagai varian dari algoritma LSTM. Pada *vanilla* LSTM, selama proses pelatihan model akan dilakukan pembelajaran pola dari data yang diberikan yang kemudian akan disimpan ke dalam *memory cells*. Sehingga proses pelatihan model pun menjadi lebih efektif walaupun hanya dengan *single hidden layer* [5]. Misalnya penelitian mengenai perbandingan penggunaan algoritma *machine learning* pada prediksi tren harga saham Netflix oleh Harry Akbar Al Hakim dan Dhomas Hatta Fudholi (2021) membuktikan bahwa algoritma LSTM lebih baik dibandingkan dengan algoritma *Linear Regression* dan *Decision Tree Regression* [6]. Dari penelitian lain yang dilakukan oleh Shahzad Muzaffar dan Afshin Afshari (2018) menunjukkan bahwa algoritma LSTM lebih baik dibandingkan dengan metode *Autoregressive Moving Average with Exogenous Inputs* (ARMAX), dan *Seasonal Autoregressive Integrated Moving Average* (SARIMA) [7].

Kemudian selanjutnya peneliti memilih model *stack* LSTM dikarenakan model ini dapat mempelajari lebih dalam dengan banyaknya *hidden layer*. Sehingga tingkat akurasi yang dihasilkan pun semakin tinggi dari hasil pembelajaran yang dilakukan model [8]. Misalnya penelitian yang dilakukan oleh Sangyeon Kim dan Myungjoo Kang (2019) yang berjudul *Financial Series Prediction Using Attention*

*LSTM* membuktikan bahwa model *stack LSTM* lebih baik dalam memprediksi data sekuensial daripada MLP dan 1D CNN [9].

Terakhir peneliti memilih juga model *bidirectional LSTM* dikarenakan model ini dapat menggabungkan lintasan informasi sebelumnya (*input forward*) dan lintasan informasi sesudahnya (*input backward*) yang digunakan untuk pelatihan model. Sehingga proses pelatihan model pun menjadi lebih optimal [10]. Seperti penelitian yang dilakukan oleh Fikri Masri, Deni Saepudin, dan Didit Adytia (2020) yang berjudul *Forecasting of Sea Level Time Series using Deep Learning RNN, LSTM, and BiLSTM, Case Study in Jakarta Bay, Indonesia* membuktikan bahwa model *bidirectional LSTM* memberikan hasil yang lebih baik daripada RNN dan LSTM [11].

Dengan mempelajari karakteristik algoritma LSTM, maka algoritma ini cocok digunakan untuk memprediksi harga cabai rawit dengan data berjenis *time series*. Karena perlu diketahui pula bahwa harga cabai rawit ini bersifat naik turun secara tidak tetap dan pergerakannya sulit untuk ditebak. Seperti data historis harga cabai rawit di kota Bandung dari tahun 2019 – 2020 yang diambil dari situs Pusat Informasi Harga Pangan Strategis (PIHPS) Nasional berikut ini.



Gambar 1.1 Grafik Harga Cabai Rawit di Kota Bandung 2019 -2020

Pada gambar 1.1 dapat dilihat harga eceran cabai rawit di kota Bandung pada tanggal 1 Januari 2019 seharga Rp.77.500/Kg. Lalu pada tanggal 1 Januari 2020 seharga Rp.52.500/Kg. Kemudian terakhir pada tanggal 1 Januari 2021 harga cabai rawit sebesar Rp.67.500/Kg [12].

Dari permasalahan dan referensi penelitian di atas, maka penelitian ini akan melakukan prediksi harga cabai rawit dengan algoritma *Long Short Term Memory* dengan tempatnya di kota Bandung sebagai studi kasus penelitian agar menghasilkan nilai prediksi yang yang bisa dilihat secara rinci. Dengan begitu maka tingkat evaluasi berdasarkan nilai *loss* yang diperoleh juga bisa diketahui. Maka dari latar belakang permasalahan tersebut akan dijadikan objek tugas akhir yang berjudul **“Penerapan Algoritma *Long Short Term Memory* (LSTM) untuk memprediksi Harga Cabai Rawit”**.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang dijelaskan, maka diperoleh rumusan permasalahan sebagai berikut :

1. Bagaimana cara menerapkan algoritma *Long Short Term Memory* (LSTM) untuk memprediksi harga cabai rawit?
2. Bagaimana hasil evaluasi berdasarkan nilai *loss* dengan *Root Mean Squared Error* pada prediksi harga cabai rawit dengan model *vanilla* LSTM, *stack* LSTM dan *bidirectional* LSTM pada algoritma LSTM?

## 1.3 Tujuan Penelitian

Tujuan yang diperoleh dari penelitian ini yaitu :

1. Dapat menerapkan algoritma *Long Short Term Memory* untuk memprediksi harga cabai rawit.
2. Dapat mengetahui nilai *loss* sebagai hasil evaluasi dengan *Root Mean Squared Error* pada hasil prediksi harga cabai rawit dengan model *vanilla* LSTM, *stack* LSTM, dan *bidirectional* LSTM pada algoritma LSTM.

## 1.4 Manfaat Penelitian

Manfaat yang diperoleh dari penelitian ini yaitu :

1. Mampu memperoleh pengetahuan dalam penerapan algoritma *Long Short Term Memory*.
2. Mampu menyelesaikan masalah dari hasil prediksi harga cabai rawit dengan tepat melalui pengujian nilai *loss* sebagai hasil evaluasi.

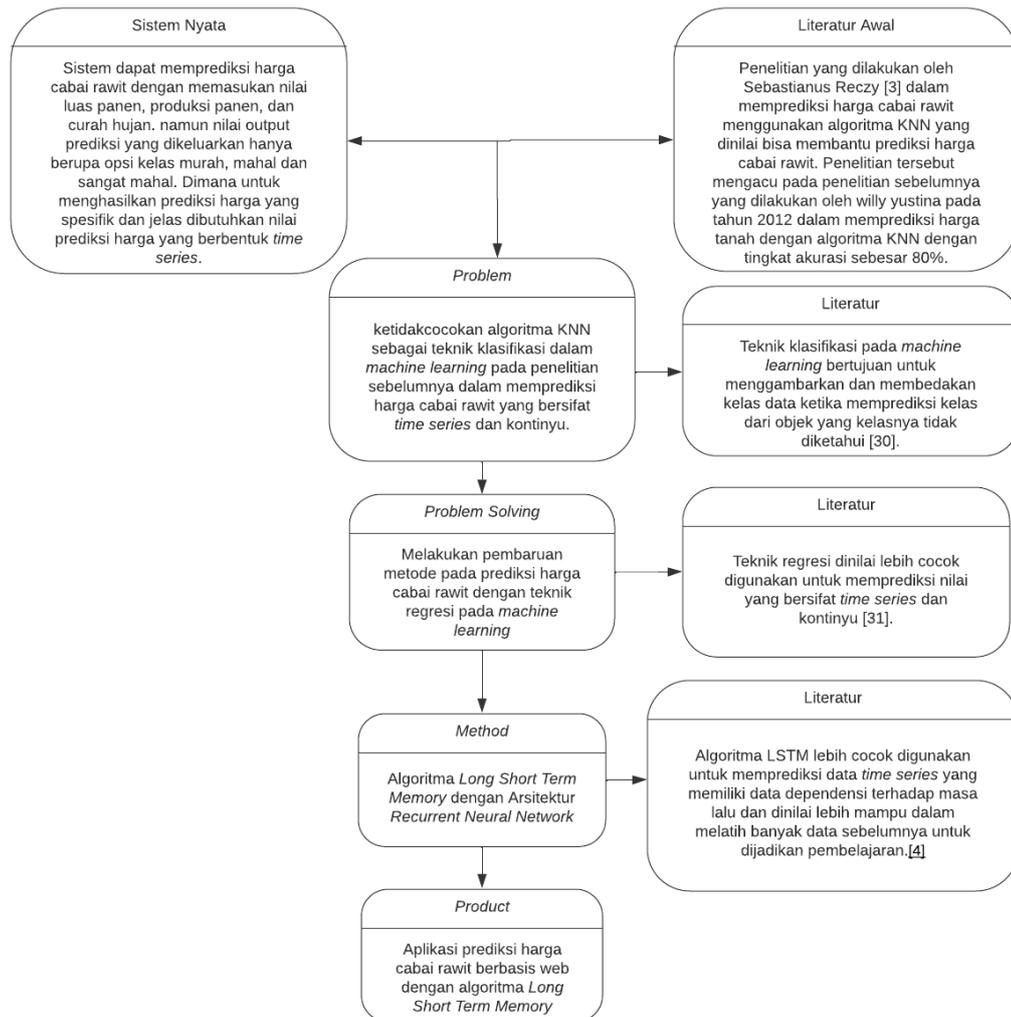
## 1.5 Batasan Masalah

Berdasarkan penelitian ini, batasan masalah yang terdapat pada penelitian ini yaitu sebagai berikut :

1. Sistem yang dibuat menggunakan bahasa pemrograman Javascript, PHP dan *database MySql*.
2. Data yang digunakan bersumber dari website Pusat Informasi Harga Pangan Strategis Nasional (PIHPS) dengan link :  
<https://hargapangan.id/tabel-harga/pasar-tradisional/daerah>.
3. Data yang digunakan adalah data harga cabai rawit di kota Bandung dari tanggal 16 Agustus 2017 sampai 20 Juli 2021 yang berjumlah 966 *record* (baris) yang meliputi tanggal dan harga cabai rawit per kilogram.
4. Prediksi dilakukan berdasarkan parameter *epoch* dan *memory cells*.
5. Model algoritma *Long Short Term Memory* yang digunakan yaitu *vanilla LSTM*, *stack LSTM*, dan *bidirectional LSTM*.

## 1.6 Kerangka Pemikiran

Kerangka pemikiran dari penelitian tugas akhir ini dapat dilihat pada gambar berikut :



Gambar 1.2 Kerangka Pemikiran

## 1.7 Metodologi Penelitian

### 1.7.1 Teknik Pengumpulan Data

Teknik dalam proses pengumpulan data pada penelitian ini dilakukan dengan menggunakan data *time series* pada harga cabai rawit di kota Bandung. Sehingga didapatkan harga cabai rawit di kota Bandung setiap hari dari tanggal 6 Agustus 2017 sampai 19 Juli 2021 yang berjumlah 966 *record*. *Dataset* yang didapatkan bersumber dari situs website <https://hargapangan.id/tabel-harga/pasar-tradisional/daerah>.

### 1.7.2 Metode Pengembangan

Metode pengembangan yang digunakan pada penelitian ini yaitu *Cross-Industry Standard Process for Data Mining* atau CRISP-DM. Berikut ini adalah tahapan – tahapannya :

#### 1. *Business Understanding*

Pada tahap pertama ini akan dilakukan pemahaman terlebih dahulu dari tujuan dan kegiatan apa saja yang akan dilakukan dari perspektif bisnis.

#### 2. *Data Understanding*

Pada tahap ini, peneliti akan melakukan pengumpulan data dan pendeskripsian data.

#### 3. *Data Preparation*

Kemudian pada tahap ini akan dilakukan persiapan data agar data siap untuk diuji.

#### 4. *Modelling*

Tahap ini dilakukannya pemodelan sesuai dengan teknik yang telah ditentukan.

## **5. Evaluation**

Setelah melakukan pemodelan maka akan dilakukan pengujian dan meninjau kembali apakah data yang dihasilkan sudah sesuai dengan tujuan yang sudah dirancang sebelumnya.

## **6. Deployment**

Bila sistem sudah diuji secara keseluruhan maka akan dibuat laporan yang berisi tentang dokumentasi keseluruhan sistem yang sudah dibuat.

### **1.8 Sistematika Penulisan**

Sistematika penulisan ini dibuat untuk mempermudah pemahaman dalam penyusunan penelitian yang akan dilakukan. Sistematika penulisan ini dibagi menjadi beberapa bagian antara lain :

#### **BAB I PENDAHULUAN**

Pada bab ini terdapat latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, metode penelitian, kerangka pemikiran dan sistematika penulisan.

#### **BAB II STUDI PUSTAKA**

Pada bab ini berisi uraian penjelasan teori yang akan digunakan dalam menganalisa permasalahan pada penelitian yang dilakukan. Studi pustaka dapat diperoleh dari jurnal, buku, serta berbagai penelitian yang dilakukan sebelumnya.

### **BAB III METODOLOGI PENELITIAN**

Pada bab ini berisi penjelasan alur metodologi penelitian yang dilakukan untuk pembuatan sistem mulai dari *business understanding*, *data understanding*, *data preparation*, dan *modelling*.

### **BAB IV HASIL DAN PEMBAHASAN**

Pada bab ini menjelaskan mengenai hasil dari sistem yang dibuat yang meliputi proses evaluasi dan *deployment*.

### **BAB V PENUTUP**

Pada bab ini berisi kesimpulan dan saran pada aplikasi yang telah dibuat dalam upaya mengevaluasi kekurangan yang bertujuan mendapatkan hasil yang lebih baik.

