

## **BAB II**

### **LANDASAN TEORI**

#### **2.1. Penelitian Terdahulu**

Penelitian yang dilakukan oleh Kandari Puteri dan Astried Silvanie dengan judul '*Machine Learning* Untuk Model Prediksi Harga Sembako Dengan Metode Regresi Linear Berganda' bertujuan untuk memprediksi harga sembako harian dengan penerapan metode regresi linear berganda, yang menggunakan sampel data sembako di DKI Jakarta. Himpunan data tersebut memiliki empat variabel, yaitu harga, komoditas, pasar, dan tanggal. Variabel harga merupakan variabel terikat, sedangkan variabel komoditas, pasar, dan tanggal merupakan variabel bebas. Hasilnya, harga sembako harian dapat diprediksi dengan bantuan sistem *machine learning* baik untuk prediksi harga di masa lalu, sekarang, maupun masa depan. Persentase kontribusi variabel bebas terhadap variabel terikat sebesar 84,2%, dengan 15,8% tambahan dipengaruhi oleh variabel yang tidak termasuk dalam penelitian ini [4].

Persamaan penelitian ini dengan penelitian yang dilakukan oleh Kandari Puteri dan Astried Silvanie adalah, peneliti sama-sama melakukan penerapan algoritma *machine learning* untuk prediksi harga sembako pada suatu daerah. Perbedaannya terletak pada karakteristik daerah yang diteliti. Penelitian sebelumnya berfokus pada prediksi harga sembako di DKI Jakarta, yang merupakan daerah dengan karakteristik ekonomi dan distribusi barang yang sangat berbeda dibandingkan dengan Kabupaten Pandeglang. DKI Jakarta sebagai ibu kota negara

memiliki pola distribusi sembako yang lebih kompleks dan dinamis, sementara Kabupaten Pandeglang, yang merupakan daerah agraris, memiliki karakteristik yang lebih dipengaruhi oleh faktor-faktor lokal seperti produksi dan musim tanam.

Penelitian terkait selanjutnya dilakukan oleh Moch Farryz Rizkilloh dan Sri Widiyanesti dengan judul ‘Prediksi Harga *Cryptocurrency* Menggunakan Algoritma *Long Short-Term Memory* (LSTM)’ bertujuan untuk membuat model prediksi yang dapat meramalkan harga pasar *cryptocurrency*. Hasilnya, model menunjukkan performa prediksi terbaik terjadi pada saat melakukan pengujian terhadap jenis koin DOGE dengan jumlah *Epoch* 20 yang mendapatkan nilai RMSE sebesar 0,0630 [3].

Persamaan penelitian ini dengan penelitian yang dilakukan oleh Moch Farryz Rizkilloh dan Sri Widiyanesti adalah, peneliti sama-sama melakukan penerapan algoritma LSTM untuk melakukan tugas *forecasting*. Selain itu, data yang digunakan juga sama-sama berbentuk *time series*. Perbedaannya terletak pada arsitektur model yang dirancang, di mana penelitian ini tidak hanya menggunakan LSTM sebagai dasar model, tetapi juga menggabungkannya dengan teknik dropout *layers* untuk mengurangi *overfitting*, serta mengoptimalkan model dengan algoritma optimasi lain yaitu Adam *optimizer*.

Penelitian terkait selanjutnya dilakukan oleh Agus Nursikuwagus dan Tono Hartono dengan judul “Implementasi Algoritma Apriori Untuk Analisis Penjualan Dengan Berbasis *Web*” bertujuan untuk memberikan laporan hasil penelitian mengenai implementasi algoritma apriori untuk analisis penjualan. Hasil penelitian ini adalah aplikasi yang didasarkan pada algoritma yang berasal dari penggunaan

*itemset* dan aturan asosiasi. Aplikasi ini digunakan untuk mengotomatisasi perhitungan analisis penjualan barang [5].

Persamaan penelitian ini dengan penelitian yang dilakukan oleh Agus Nursikuwagus dan Tono Hartono adalah, peneliti sama-sama melakukan penerapan algoritma berbasis *web*. Selain itu, aplikasi yang dibuat juga ditujukan untuk otomatisasi perhitungan menggunakan model yang dibangun. Perbedaannya terletak pada algoritma yang digunakan dan jenis tugas yang dilakukan, penelitian ini fokus pada tugas *forecasting*.

## **2.2. Teori Pendukung**

### **2.2.1. Machine Learning**

*Machine Learning* (ML) adalah salah satu bidang studi di bidang artificial intelligence dengan fokus pengembangan dan studi algoritma statistik yang dapat belajar dari data yang diberikan. Dengan demikian, *machine learning* dapat melakukan tugas tanpa instruksi eksplisit [6]. *Machine learning* sudah diterapkan pada berbagai aplikasi, diantaranya natural language processing, computer vision, speech recognition, email filtering, agrikultur, dan medis [7]. *Machine learning* dibagi menjadi tiga kategori, yaitu:

1. *Supervised Learning*, komputer diberikan contoh *input* dan *output* yang diinginkan dengan tujuan untuk mempelajari aturan atau pola dari data yang diberikan, sehingga komputer dapat memetakan *input* ke *output* secara otomatis [8].

2. *Unsupervised Learning*, komputer diberikan data tanpa label untuk menemukan struktur ataupun pola tersembunyi dalam data [8].
3. *Reinforcement Learning*, metode pembelajaran mesin berbasis umpan balik di mana agen belajar berperilaku di lingkungannya dengan melakukan tindakan dan melihat hasilnya. Untuk tindakan baik, agen diberi umpan balik positif, dan untuk tindakan buruk, mereka diberi umpan balik negatif atau penalti [8].

### **2.2.2. Analisis Peramalan**

Analisis peramalan (*forecasting analysis*) adalah suatu proses analisis yang digunakan untuk memproyeksikan atau meramalkan nilai-nilai masa depan berdasarkan data historis atau tren yang telah terjadi. Tujuannya tidak hanya sekadar memperkirakan nilai-nilai di masa depan, tetapi juga untuk mengidentifikasi pola-pola dan tren yang mendasari data historis tersebut. Dengan memahami pola-pola tersebut, analisis peramalan dapat memberikan estimasi yang lebih akurat tentang kemungkinan nilai-nilai di masa mendatang. Kemudian, informasi ini dapat digunakan untuk mengambil keputusan yang lebih baik dan lebih tepat, terutama dalam konteks ekonomi, keuangan, manufaktur, dan bidang lainnya. Dengan demikian, analisis peramalan tidak hanya menjadi alat prediktif, tetapi juga menjadi alat penting dalam mendukung pengambilan keputusan yang efektif berdasarkan informasi yang tersedia [9].

### 2.2.3. Bahasa Pemrograman Python

Python ialah sebuah bahasa pemrograman yang terkenal dan memiliki fleksibilitas yang tinggi. Dibuat dengan prinsip-prinsip yang memprioritaskan kejelasan dan keterbacaan kode, Python telah menjadi pilihan utama bagi berbagai keperluan pengembangan perangkat lunak. Selain digunakan dalam pengembangan *web*, Python juga sangat berguna dalam analisis data, pembuatan model kecerdasan buatan, pemrosesan bahasa alami, dan bidang-bidang lainnya yang membutuhkan pemrograman. Kemampuan Python yang luas dan kemudahan penggunaannya membuatnya menjadi salah satu bahasa pemrograman yang paling diminati dan digunakan oleh para pengembang di seluruh dunia [10].

```
def my_function(food):
    for x in food:
        print(x)

fruits = ["apple", "banana", "cherry"]

my_function(fruits)
```

**Gambar 2. 1 Contoh Code Python**

(Sumber: <https://w3schools.com>)

### 2.2.4. Google Colab

Google Colab adalah platform komputasi *cloud* yang diberikan secara cuma-cuma oleh Google. Platform ini memungkinkan pengguna untuk menulis dan menjalankan kode Python di lingkungan *cloud* tanpa memerlukan persiapan atau konfigurasi tambahan. Colab menggunakan dasar dari proyek

*open-source* yang disebut Jupyter Notebook, yang memfasilitasi pembuatan dan berbagi notebook interaktif yang menggabungkan kode, teks, gambar, serta hasil eksekusi [11].



**Gambar 2. 2 Logo Google Colab**

(Sumber: <https://colab.google>)

## **2.2.5. Algoritma yang Digunakan**

### **2.2.5.1. Long Short-Term Memory**

*Long Short-Term Memory* (LSTM) adalah jenis arsitektur jaringan rekurensi yang dikembangkan khusus untuk menangani masalah pembelajaran berbasis urutan, seperti prediksi *time series*. LSTM memiliki kemampuan untuk mengatasi masalah vanishing gradient yang sering terjadi pada jaringan rekurensi tradisional, yang memungkinkannya untuk menangkap ketergantungan jarak jauh dalam data urutan [12].

Arsitektur LSTM terdiri dari sel-sel memori yang kompleks dengan tiga gerbang (*input*, *output*, dan *forget gate*) yang mengatur aliran informasi dalam sel. Sel-sel memori ini memungkinkan LSTM untuk menyimpan informasi dalam jangka waktu yang lama dan memilih informasi mana yang harus diingat atau dilupakan.

Proses pelatihan LSTM melibatkan optimisasi bobot dan bias dalam sel-sel memori, serupa dengan algoritma yang digunakan dalam pelatihan ANN. Namun, LSTM menggunakan mekanisme khusus yang disebut *backpropagation through time* (BPTT) untuk menyesuaikan bobot dan bias dengan memperhitungkan urutan waktu data *input* [12][13].

#### **2.2.5.2. Artificial Neural Network**

*Artificial Neural Network* (ANN) adalah sebuah model komputasi yang terinspirasi dari struktur dan fungsi otak manusia. ANN terdiri dari jaringan neuron buatan yang terdiri dari beberapa lapisan, termasuk lapisan *input*, lapisan tersembunyi (*hidden layer*), dan lapisan *output*. Setiap neuron dalam jaringan terhubung dengan neuron di lapisan sebelumnya dan sesudahnya dengan bobot yang mempengaruhi informasi yang dikirimkan.

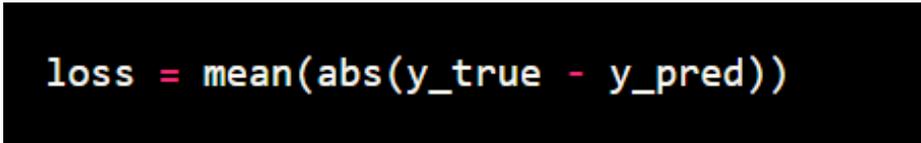
Proses pelatihan ANN melibatkan pengoptimalan bobot dan bias dalam jaringan untuk menghasilkan prediksi yang akurat. Ini dilakukan dengan menggunakan algoritma pembelajaran seperti *backpropagation*, di mana galat antara prediksi dan nilai sebenarnya digunakan untuk menyesuaikan bobot agar menghasilkan prediksi yang lebih baik [14][15].

## 2.2.6. Evaluation Metrics yang digunakan

Average Absolute Error (MAE), Mean Squared Error (MSE), dan Root Mean Squared Error (RMSE), dan Mean Absolute Percentage Error (MAPE) adalah beberapa metrik evaluasi yang umum digunakan untuk mengevaluasi seberapa baik model dalam melakukan tugas peramalan. Ketiga metrik ini memberikan informasi yang bermanfaat tentang seberapa baik model memprediksi harga harian sembako.

### 2.2.6.1. Mean Absolute Error (MAE)

MAE adalah metrik evaluasi yang mengukur rata-rata dari selisih absolut antara prediksi dan nilai sebenarnya. MAE menunjukkan seberapa dekat rata-rata prediksi dengan nilai sebenarnya tanpa memperhatikan arah kesalahan [16]. Formula untuk MAE adalah sebagai berikut:



```
loss = mean(abs(y_true - y_pred))
```

Gambar 2. 3 Formula MAE

(Sumber: [https://keras.io/api/metrics/regression\\_metrics/](https://keras.io/api/metrics/regression_metrics/))

di mana:

1.  $y_{pred}$  adalah nilai yang diprediksi,
2.  $y_{true}$  adalah nilai sebenarnya

### 2.2.6.2. Mean Squared Error (MSE)

MSE adalah salah satu metrik evaluasi yang umum digunakan dalam analisis peramalan. MSE mengukur rata-rata dari kuadrat selisih antara nilai prediksi dengan nilai sebenarnya dari data yang diamati. Semakin rendah nilai MSE, semakin baik kinerja model peramalan [17]. Adapun formula untuk perhitungan MSE adalah:

```
loss = mean(square(y_true - y_pred))
```

**Gambar 2. 4 Formula MSE**

(Sumber: [https://keras.io/api/metrics/regression\\_metrics/](https://keras.io/api/metrics/regression_metrics/))

di mana:

1.  $y_{pred}$  adalah nilai yang diprediksi,
2.  $y_{true}$  adalah nilai sebenarnya

### 2.2.6.3. Root Mean Squared Error (RMSE)

RMSE merupakan akar kuadrat dari MSE dan memberikan gambaran tentang seberapa besar kesalahan prediksi dalam unit yang sama dengan variabel dependen. RMSE lebih sensitif terhadap kesalahan besar daripada MSE karena mengecilkan efek dari kesalahan kecil [18]. Adapun formula untuk perhitungan RMSE adalah:

```
loss = sqrt(mean((y_pred - y_true) ** 2))
```

**Gambar 2. 5 Formula RMSE**

(Sumber: [https://keras.io/api/metrics/regression\\_metrics/](https://keras.io/api/metrics/regression_metrics/))

di mana:

1.  $y_{pred}$  adalah nilai yang diprediksi
2.  $y_{true}$  adalah nilai sebenarnya

#### 2.2.6.4. Mean Absolute Percentage Error (MAPE)

*Mean Absolute Percentage Error* (MAPE) adalah salah satu metrik evaluasi yang umum digunakan dalam analisis peramalan. MAPE mengukur persentase rata-rata dari selisih absolut antara nilai prediksi dengan nilai sebenarnya dari data yang diamati, dibagi dengan nilai sebenarnya. Semakin rendah nilai MAPE, semakin baik kinerja model peramalan [9]. Adapun formula untuk perhitungan MAPE adalah:

```
loss = 100 * mean(abs((y_true - y_pred) / y_true))
```

**Gambar 2. 6 Formula MAPE**

(Sumber: [https://keras.io/api/metrics/regression\\_metrics/](https://keras.io/api/metrics/regression_metrics/))

di mana:

1.  $y_{pred}$  adalah nilai yang diprediksi,
2.  $y_{true}$  adalah nilai sebenarnya

## **2.2.7. Implementasi Web**

### **2.2.7.1. Aplikasi Web**

Aplikasi *web* merupakan perangkat lunak yang diprogram untuk diakses melalui browser *web* dan bekerja dalam lingkungan internet. Aplikasi *web* menggunakan teknologi *web* seperti HTML, CSS, dan JavaScript untuk menampilkan konten yang interaktif kepada pengguna. Aplikasi *web* ini dapat diakses dari berbagai perangkat yang terhubung ke internet [19].

### **2.2.7.2. HTML**

HTML (*Hypertext Markup Language*) adalah bahasa markah yang digunakan untuk membuat dan mengatur konten halaman *web*. HTML digunakan untuk mendefinisikan struktur dan elemen-elemen dasar dari sebuah halaman *web*, seperti teks, gambar, tautan, dan elemen-elemen lainnya. Dengan HTML, penulis dapat membuat halaman *web* yang terstruktur dan dapat diakses oleh berbagai perangkat dan browser *web* [19].

```
<!DOCTYPE html>
<html>
<head>
<title>Page Title</title>
</head>
<body>

<h1>This is a Heading</h1>
<p>This is a paragraph.</p>

</body>
</html>
```

**Gambar 2. 7 Contoh Code HTML**

(Sumber: [w3school.com](http://w3school.com))

### 2.2.7.3. CSS

CSS (Cascading Style Sheets) adalah bahasa stylesheet yang digunakan oleh pengembang untuk mengatur tampilan dan gaya elemen-elemen HTML dalam suatu halaman *web*. Dengan CSS, pengembang dapat mengendalikan berbagai aspek penampilan elemen, seperti warna, ukuran, jenis font, tata letak, dan animasi. Penggunaan CSS memungkinkan pengembang untuk memisahkan struktur konten (HTML) dari gaya visualnya, sehingga memungkinkan pembuatan halaman *web* yang lebih dinamis, konsisten, dan mudah dikelola. Prinsip kaskade dan pewarisan dalam CSS memungkinkan aturan gaya dapat ditentukan pada berbagai tingkatan, dan elemen-elemen yang lebih spesifik dapat mewarisi gaya dari elemen yang lebih umum [19].

```
body {  
  background-color: lightblue;  
}  
  
h1 {  
  color: white;  
  text-align: center;  
}  
  
p {  
  font-family: verdana;  
  font-size: 20px;  
}
```

**Gambar 2. 8 Contoh Code CSS**  
(Sumber: w3school.com)

#### **2.2.7.4. PHP**

PHP (*Hypertext Preprocessor*) adalah bahasa pemrograman yang umumnya dipakai dalam pembuatan aplikasi *web* yang dinamis. PHP beroperasi di sisi server, artinya skrip PHP dieksekusi di server *web* sebelum hasilnya disampaikan ke peramban pengguna. PHP berguna untuk berbagai keperluan, termasuk pengolahan formulir, akses basis data, penciptaan konten yang dinamis, serta interaksi dengan pengguna [20].

```
<!DOCTYPE html>
<html>
<body>

<?php
echo "My first PHP script!";
?>

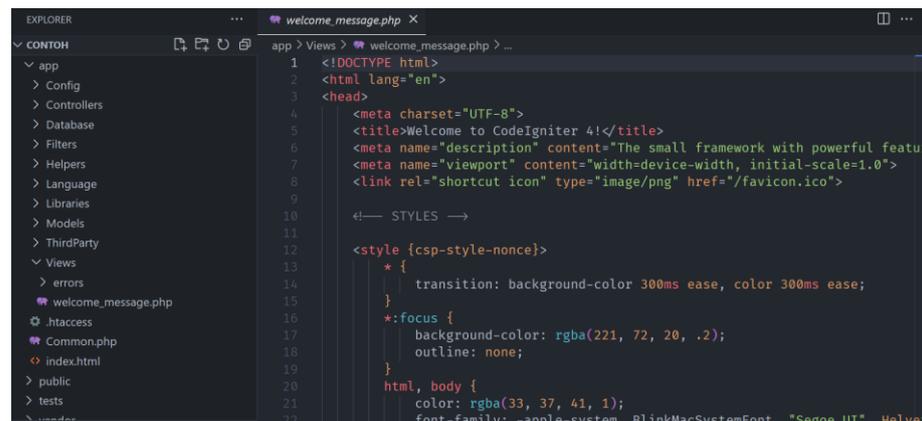
</body>
</html>
```

**Gambar 2. 9 Contoh Code PHP**

(Sumber: [w3school.com](http://w3school.com))

#### 2.2.7.5. Framework CodeIgniter

CodeIgniter adalah salah satu kerangka kerja (*framework*) PHP yang dikenal karena ringan dan responsif, diciptakan untuk menyederhanakan proses pengembangan aplikasi *web*. Framework ini dilengkapi dengan beragam fitur yang handal untuk mempercepat pembangunan aplikasi *web* secara efisien. CodeIgniter juga terkenal dengan dokumentasinya yang lengkap dan komunitas pengguna yang aktif, menjadikannya sebagai pilihan utama di kalangan para pengembang *web* [21].



The image shows a code editor window with a file explorer on the left and a code editor on the right. The file explorer shows a folder named 'CONTOH' containing a sub-folder 'app'. Inside 'app', there are folders for 'Views' and 'welcome\_message.php'. The code editor shows the content of 'welcome\_message.php', which is an HTML file with a head section containing meta tags for charset, title, description, viewport, and a link to a favicon. It also includes a style section with a transition and a focus effect, and a body section with a color and font-family.

```

1 <!DOCTYPE html>
2 <html lang="en">
3 <head>
4   <meta charset="UTF-8">
5   <title>Welcome to CodeIgniter 4!</title>
6   <meta name="description" content="The small framework with powerful featur
7   <meta name="viewport" content="width=device-width, initial-scale=1.0">
8   <link rel="shortcut icon" type="image/png" href="/favicon.ico">
9
10  <!-- STYLES -->
11
12  <style {csp-style-nonce}>
13    * {
14      transition: background-color 300ms ease, color 300ms ease;
15    }
16    *:focus {
17      background-color: rgba(221, 72, 20, .2);
18      outline: none;
19    }
20  }
21  html, body {
22    color: rgba(33, 37, 41, 1);
23    font-family: -apple-system, BlinkMacSystemFont, "Segoe UI", Helvetica

```

**Gambar 2. 10 Contoh Struktur Folder Framework CodeIgniter**

#### 2.2.7.6. *Web Scraping*

*Web scraping* merupakan proses pengambilan data secara otomatis dari halaman *web*. Proses ini menggunakan program komputer atau bot untuk menjelajahi situs *web*, mengekstraksi informasi yang diinginkan dari kode HTML, dan menyimpannya dalam format yang dapat diproses, seperti file CSV atau *database* [22].

#### 2.2.7.7. **Flask**

Flask adalah kerangka kerja *web* ringan yang sering diterapkan oleh pengembang untuk menyajikan model *machine learning*. Melalui Flask, pengembang dapat membuat antarmuka pengguna yang mudah diakses melalui permintaan HTTP, memungkinkan pengguna untuk berinteraksi dengan model *machine learning* secara langsung [23].

## 2.2.8. Pustaka yang digunakan Dalam Python

### 2.2.8.1. NumPy

NumPy, yang merupakan singkatan dari Numerical Python, ialah sebuah modul Python yang berguna untuk komputasi numerik. Modul ini menawarkan array multidimensi yang hemat memori dan beragam fungsi untuk melakukan operasi matematika pada array tersebut. NumPy menjadi fondasi bagi sebagian besar pustaka ilmu data lainnya di Python [24].

```
import numpy as np

arr = np.array([1, 2, 3, 4, 5])

print(arr)

print(type(arr))
```

**Gambar 2. 11 Contoh Code Pustaka Numpy**

(Sumber: [w3school.com](http://w3school.com))

### 2.2.8.2. Pandas

Pandas adalah sebuah pustaka Python yang penting dalam analisis data. Pustaka ini menyediakan berbagai struktur data yang intuitif, terutama DataFrame, yang dapat dibandingkan dengan lembar kerja dalam spreadsheet. Selain itu, Pandas juga menyediakan berbagai fungsi untuk membaca, menulis, dan melakukan manipulasi data tabular

dengan mudah. Dengan demikian, Pandas memungkinkan pengguna untuk melakukan analisis data dengan lebih efisien dan efektif [24].

```
import pandas

mydataset = {
    'cars': ["BMW", "Volvo", "Ford"],
    'passings': [3, 7, 2]
}

myvar = pandas.DataFrame(mydataset)

print(myvar)
```

**Gambar 2. 12 Contoh Code Pustaka Pandas**  
(Sumber: w3school.com)

### 2.2.8.3. Matplotlib

Matplotlib merupakan sebuah pustaka Python yang berguna untuk membuat visualisasi data. Dengan menggunakan Matplotlib, pengguna dapat menghasilkan berbagai jenis plot seperti grafik garis, histogram, scatter plot, dan lainnya. Pustaka ini sering dipasangkan dengan NumPy untuk melakukan visualisasi data [24].

```
import matplotlib.pyplot as plt
import numpy as np

xpoints = np.array([0, 6])
ypoints = np.array([0, 250])

plt.plot(xpoints, ypoints)
plt.show()
```

**Gambar 2. 13 Contoh Code Pustaka Matplotlib**  
(Sumber: w3school.com)

#### 2.2.8.4. Scikit Learn

*Scikit-learn* adalah salah satu pustaka Python yang paling populer untuk *machine learning* dan *data mining*. Pustaka ini menawarkan beragam algoritma *machine learning* yang dapat diterapkan untuk berbagai tujuan seperti klasifikasi, regresi, pengelompokan, dan sebagainya. Selain itu, *scikit-learn* juga menyediakan berbagai fungsi penting untuk evaluasi model dan pemrosesan data. Dengan koleksi algoritma yang luas dan fitur-fitur bawaan yang kuat, *scikit-learn* menjadi pilihan utama bagi para praktisi dan peneliti dalam melakukan analisis data dan pengembangan model *machine learning* [24].