

**PERAMALAN DATA *TIME SERIES* PADA HARGA KOMODITAS PANGAN  
MENGUNAKAN METODE *LONG SHORT TERM MEMORY* (LSTM)  
(Studi Kasus: Harga Penutupan Kedelai Amerika Serikat)**

**(Skripsi)**

**Oleh**

**BERLIAN SASYA DEVI PRADANA  
NPM 1917031085**



**JURUSAN MATEMATIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMPUNG  
BANDAR LAMPUNG  
2023**

## **ABSTRACT**

### **FORECASTING TIME SERIES DATA ON FOOD COMMODITY PRICES USING THE LONG SHORT TERM MEMORY (LSTM) METHOD (Case Study: Closing Prices of United States Soybeans)**

**By**

**Berlian Sasya Devi Pradana**

Forecasting the closing price of US soybeans plays a crucial role in preventing price fluctuations at a specific time and can be used as a reference for tomorrow's opening price. Therefore, we need a method that can be used for forecasting the closing price of US soybeans. This study discusses the analysis of forecasting time series data on food commodity prices, namely the closing price of US soybeans, using the Long Short Term Memory (LSTM) method with several parameters needed, namely hidden layer, epoch, hidden neurons, batch size and activation function. This study identified the optimal LSTM model from several models formed based on the results of a combination of hidden layer parameters, number of epochs, number of hidden neurons, and number of batch sizes. The model chosen is the one that predicts the closing price of US soybeans with the smallest Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). The best model is given by the composition of hidden layer parameter values of 1, epoch of 100, hidden neurons of 64, and batch size of 16, which can predict the closing price of US soybeans with an RMSE value of 23.6674 and a MAPE value of 1.2152%.

**Keywords:** Soybean Closing Price, Forecasting, Prediction, Time Series, LSTM, RMSE and MAPE.

## ABSTRAK

### PERAMALAN DATA *TIME SERIES* PADA HARGA KOMODITAS PANGAN MENGUNAKAN METODE *LONG SHORT TERM MEMORY* (LSTM) (Studi Kasus: Harga Penutupan Kedelai Amerika Serikat)

Oleh

**Berlian Sasya Devi Pradana**

Peramalan harga penutupan kedelai Amerika Serikat berperan sangat penting dalam mencegah fluktuasi harga pada waktu tertentu dan dapat menjadi acuan untuk pembukaan harga esok hari. Oleh karena itu, diperlukan metode yang dapat digunakan untuk peramalan harga penutupan kedelai Amerika Serikat. Penelitian ini membahas tentang analisis peramalan data *time series* pada harga komoditas pangan yaitu harga penutupan kedelai Amerika Serikat menggunakan metode *Long Short Term Memory* (LSTM) dengan sejumlah parameter yang dibutuhkan yakni *hidden layer*, *max epoch*, *neuron hidden*, *batch size* dan fungsi aktivasi. Penelitian ini mengidentifikasi model LSTM yang optimal dari sejumlah model yang terbentuk berdasarkan hasil kombinasi parameter *hidden layer*, jumlah *epoch*, jumlah *neuron hidden*, dan jumlah *batch size*. Model yang dipilih adalah model yang memberikan hasil prediksi harga penutupan kedelai Amerika Serikat dengan dengan nilai *Root Mean Square Error* (RMSE) dan *Mean Absolute Percentage Error* (MAPE) paling kecil. Model terbaik diberikan oleh komposisi nilai parameter *hidden layer* sebanyak 1, *epoch* sebanyak 100, *neuron hidden* sebanyak 64, *batch size* sebanyak 16 yang mampu meramalkan harga penutupan kedelai Amerika Serikat dengan nilai RMSE sebesar 23,6674 dan nilai MAPE sebesar 1,2152%.

**Kata kunci :** Harga Penutupan Kedelai, Peramalan, Prediksi, *Time Series*, LSTM, RMSE dan MAPE.

**PERAMALAN DATA *TIME SERIES* PADA HARGA KOMODITAS PANGAN  
MENGUNAKAN METODE *LONG SHORT TERM MEMORY (LSTM)*  
(Studi Kasus: Harga Penutupan Kedelai Amerika Serikat)**

Oleh

**BERLIAN SASYA DEVI PRADANA**

(Skripsi)

**Sebagai Salah Satu Syarat untuk Mencapai Gelar  
SARJANA MATEMATIKA**

Pada

**Jurusan Matematika  
Fakultas Matematika dan Ilmu Pengetahuan Alam**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMPUNG  
BANDAR LAMPUNG  
2023**

Judul Skripsi : **PERAMALAN DATA *TIME SERIES* PADA HARGA KOMODITAS PANGAN MENGGUNAKAN METODE *LONG SHORT TERM MEMORY (LSTM)* (STUDI KASUS: HARGA PENUTUPAN KEDELAI AMERIKA SERIKAT)**

Nama Mahasiwa : **Berlian Sasya Devi Pradana**

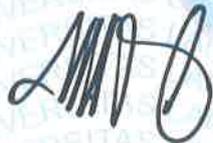
Nomor Pokok Mahasiswa : **1917031085**

Jurusan : **Matematika**

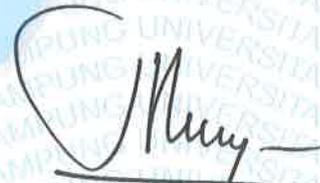
Fakultas : **Matematika dan Ilmu Pengetahuan Alam**

**MENYETUJUI**

**1. Komisi Pembimbing**

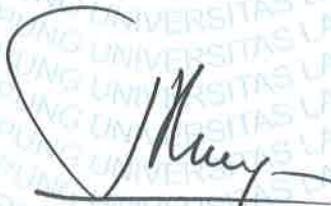


**Dr. Ahmad Faisol, S.Si, M.Sc.**  
NIP 19800206 200312 1 003



**Dr. Aang Nuryaman, S.Si., M.Si.**  
NIP 19740316 200501 1 001

**2. Ketua Jurusan Matematika**



**Dr. Aang Nuryaman, S.Si., M.Si.**  
NIP 19740316 200501 1 001

**MENGESAHKAN**

**1. Tim Penguji**

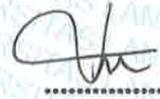
**Ketua : Dr. Ahmad Faisal, S.Si., M.Sc.**



**Sekretaris : Dr. Aang Nuryaman, S.Si., M.Si**



**Penguji  
Bukan Pembimbing : Drs. Nusyirwan, M.Si.**



**2. Plt. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam  
Universitas Lampung,**



**Dr. Eng. Heri Satria, S.Si., M.Si.**  
NIP. 197110012005011002

**Tanggal Lulus Ujian Skripsi: 28 Februari 2023**

## PERNYATAAN SKRIPSI MAHASISWA

Yang bertanda tangan di bawah ini :

Nama : **Berlian Sasya Devi Pradana**  
Nomor Pokok Mahasiswa : **1917031085**  
Jurusan : **Matematika**  
Judul Skripsi : **PERAMALAN DATA *TIME SERIES* PADA HARGA KOMODITAS PANGAN MENGGUNAKAN METODE *LONG SHORT TERM MEMORY (LSTM)* (STUDI KASUS: HARGA PENUTUPAN KEDELAI AMERIKA SERIKAT)**

Dengan ini menyatakan bahwa skripsi ini adalah hasil pekerjaan saya sendiri. Apabila kemudian hari terbukti bahwa skripsi ini merupakan hasil salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku.

Bandar Lampung, 7 Maret 2023

Yang menyatakan,



**Berlian Sasya Devi Pradana**

**NPM. 1917031085**

## **RIWAYAT HIDUP**

Penulis bernama lengkap Berlian Sasya Devi Pradana lahir di Wonosobo, pada tanggal 27 Maret 2002, merupakan anak pertama dari lima bersaudara dari pasangan Bapak Supratigno, S.P. dan Ibu Nur Hayati, S.TP.

Penulis mengawali pendidikan di Taman Kanak-Kanak (TK) Tunas Ceria pada tahun 2005-2007. Kemudian menempuh pendidikan Sekolah Dasar (SD) di SD Negeri 1 Jati Baru pada tahun 2007-2013. Kemudian melanjutkan ke Sekolah Menengah Pertama di MTS Al-Ikhlas Tanjung Bintang pada tahun 2013-2016. Selanjutnya penulis melanjutkan lagi jenjang pendidikannya ke Sekolah Menengah Atas di SMA Negeri 1 Tanjung Bintang pada tahun 2016-2019. Pada tahun 2019, penulis terdaftar sebagai mahasiswa S1 Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA) Universitas Lampung melalui jalur Seleksi Bersama Masuk Perguruan Tinggi Negeri (SBMPTN).

Selama menjadi mahasiswa penulis aktif dalam mengikuti kegiatan UKM Sains dan Teknologi (SAINTEK) Universitas Lampung pada periode 2021. Pada Tahun 2022 penulis melaksanakan Kuliah Kerja Praktik (KP) di Dinas Ketahanan Pangan, Tanaman Pangan dan Hortikultura Provinsi Lampung serta mengikuti kegiatan Kuliah Kerja Nyata (KKN ) di Desa Karya Makmur, Lampung Timur. Selama menjadi mahasiswa penulis juga mengikuti kegiatan Magang dan Studi Independen Bersertifikat (MSIB) yaitu Studi Independen bidang *Data Analytics for Business* di Bitlabs Academy pada bulan Agustus-Desember di PT.Indobit Digital Raya.

## **KATA INSPIRASI**

*“Allah tidak akan membebani seseorang melainkan sesuai dengan kadar kesanggupannya.”*

*(QS Al-Baqarah : 286)*

*“Dan hanya kepada Tuhanmulah hendaknya kamu berharap.”*

*(QS Al-Insyirah : 8)*

*“Boleh jadi kamu membenci sesuatu padahal ia amat baik bagimu, dan boleh jadi pula kamu menyukai sesuatu padahal ia amat buruk bagimu, Allah mengetahui sedang kamu tidak mengetahui”*

*(QS Al – Baqarah : 216)*

*“Satu-satunya batasan untuk meraih mimpi adalah keragu-raguan kita akan hari ini. Marilah kita maju dengan keyakinan yang aktif dan kuat”*

*(Franklin Roosevelt)*

## **PERSEMBAHAN**

*Alhamdulillah, puji dan syukur kepada Allah SWT atas nikmat serta hidayahnya sehingga skripsi ini dapat terselesaikan dengan baik dan tepat pada waktunya. Oleh karena itu, dengan rasa syukur dan bahagia penulis persembahkan rasa karya sederhana kepada:*

### ***Ibu dan Ayah***

*Sebagai tanda terima kasih kepada kedua orang tuaku yang membuat segalanya menjadi mungkin hingga penulis bisa sampai pada tahap di mana skripsi ini akhirnya selesai atas segala pengorbanan, waktu, dukungan, motivasi, serta doa baik yang tidak pernah berhenti kalian berikan, karena atas doa dan ridho kalian Allah memudahkan perjalanan hidup ini.*

### ***Dosen Pembimbing dan Pembahas***

*Terimakasih kepada dosen pembimbing dan pembahas yang sudah sangat membantu, memberikan motivasi, memberikan masukan, arahan serta ilmu yang bermanfaat bagi penulis.*

***Almamater Tercinta , Universitas Lampung***

## SANWACANA

Segala puji dan syukur penulis ucapkan kepada Allah SWT atas segala nikmat dan karunia-Nya yang tak terhingga sehingga penulis dapat menyelesaikan skripsi yang berjudul **“Peramalan Data *Time Series* Pada Harga Komoditas Pangan Menggunakan Metode *Long Short Term Memory (LSTM)* (Studi Kasus: Harga Penutupan Kedelai Amerika Serikat)”**. Dalam penulisan skripsi ini tidak dapat terselesaikan tanpa adanya bimbingan, bantuan dan dukungan dari berbagai pihak. Sehingga, dalam kesempatan ini penulis mengucapkan terima kasih kepada:

1. Bapak Dr. Ahmad Faisol, S.Si., M.Sc., selaku dosen pembimbing I yang telah banyak memberikan kemudahan, nasihat, masukan dan saran, serta dorongan kepada penulis dalam menyelesaikan skripsi ini.
2. Bapak Dr. Aang Nuryaman, S.Si., M.Si., selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Univeristas Lampung dan dosen pembimbing II yang telah memberikan bimbingan, pengarahan, serta saran sehingga penulis dapat menyelesaikan skripsi ini.
3. Bapak Drs. Nusyirwan, M.Si. selaku dosen penguji yang telah memberikan kritik dan saran yang membangun sehingga skripsi ini dapat terselesaikan.
4. Ibu Dr. Fitriani, S.Si. ,M.Sc., selaku dosen pembimbing akademik saya, yang telah banyak membantu dalam urusan akademik dan memberikan bimbingan serta arahan selama masa perkuliahan.
5. Bapak Dr. Eng. Heri Satria, S.Si., M.Si. selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
6. Seluruh Dosen, staf dan karyawan jurusan Matematika FMIPA Universitas Lampung.
7. Teruntuk kedua orang tuaku tercinta, Ayah Supartigno S.P dan Ibu Nur Hayati S.TP, terimakasih atas doa-doa yang tak henti-hentinya kalian

curahkan, dukungan serta semangat yang selalu diberikan, cinta kasih, perhatian dan pengorbanan yang diberikan tiada terbalas demi kesuksesan penulis. Semoga dikemudian hari dapat membahagiakan dan menjadi kebanggaan kalian.

8. Adik-adikku tersayang Adelia, Raihan, Ghibran, Iqbal, Akmal dan Zidna, terimakasih atas dukungan dan doa serta pengertian untuk penulis sebagai kakak. Semoga dapat menjadi contoh yang baik untuk kalian semua.
9. Eyang kakung, Eyang Uti serta Om dan Tante, terimakasih atas doa, dukungan dan motivasi serta jasa-jasa kalian sehingga penulis dapat menyelesaikan skripsi ini dengan lancar.
10. Sahabat terbaikku Dewi Ratna Fahmi Arifin, yang senantiasa memberikan waktu untuk mendengarkan keluh kesah dalam setiap perjalanan hidupku dari SMP hingga saat ini serta doa, bantuan dan motivasi dalam menyelesaikan skripsi ini.
11. Teman-teman seperjuanganku yaitu Zida, Tika, Linda, Siti dan Mega yang telah banyak membantu selama perkuliahan dan selalu memberikan semangat dan saran dalam pengerjaan skripsi ini.
12. Keluarga KKN Karya Makmur Yollanda Nababan, Raineven S.V.C., dan Putri Ramadhani A., yang selalu hadir menyempatkan waktu dalam setiap acara, memberikan bantuan, doa, dukungannya dan semangat kepada penulis dalam menyelesaikan skripsi ini.
13. Rekan mahasiswa Maematika Universitas Lampung 2019.
14. Semua pihak yang tidak dapat disebutkan satu persatu yang telah membantu penulis dalam menyelesaikan skripsi ini.

Penulis menyadari bahwa masih banyak kekurangan dalam penulisan skripsi ini. Oleh karena itu, penulis mengharapkan masukan serta saran untuk dijadikan pelajaran kedepannya.

Bandar Lampung, 7 Maret 2023  
Penulis,

Berlian Sasya Devi Pradana

## DAFTAR ISI

	Halaman
<b>ABSTRAK</b> .....	ii
<b>DAFTAR GAMBAR</b> .....	xv
<b>DAFTAR TABEL</b> .....	xvi
<b>I. PENDAHULUAN</b> .....	<b>1</b>
1.1 Latar Belakang dan Masalah .....	1
1.2 Tujuan Penelitian .....	4
1.3 Manfaat Penelitian .....	4
<b>II. TINJAUAN PUSTAKA</b> .....	<b>5</b>
2.1 Peramalan ( <i>forecasting</i> ) .....	5
2.2 Data <i>Time Series</i> (Deret Waktu) .....	6
2.3 Data <i>Mining</i> .....	7
2.3.1 <i>Preprocessing Data</i> .....	7
2.4 <i>Machine Learning</i> .....	9
2.5 <i>Artifical Neural Network</i> .....	10
2.6 Fungsi Aktivasi .....	11
2.7 <i>Long Short Term Memory</i> (LSTM).....	13
2.7.1 Membangun Model LSTM.....	22
2.8 Denormalisasi Data .....	23
2.9 Evaluasi Model.....	23
<b>III. METODOLOGI PENELITIAN</b> .....	<b>25</b>
3.1 Waktu dan Tempat Penelitian.....	25
3.2 Data Penelitian .....	25
3.3 Metode Penelitian .....	25
<b>IV. HASIL DAN PEMBAHASAN</b> .....	<b>28</b>
4.1 <i>Dataset</i> .....	28
4.2 <i>Preprocessing Data</i> .....	30
4.2.1 <i>Ascending</i> .....	31
4.2.2 Segmentasi.....	31

4.2.3	Normalisasi Data Dengan <i>Min Max Normalization</i> .....	31
4.3	Pembagian Data <i>Train</i> dan Data <i>Testing</i> .....	33
4.4	Konstruksi Model LSTM .....	34
4.5	Perhitungan Manual .....	37
4.6	Pengujian Model .....	39
4.7	Denormalisasi Data .....	41
4.8	Evaluasi Model LSTM .....	42
4.9	Peramalan Harga Penutupan Kedelai Amerika Serikat.....	43
<b>V.</b>	<b>KESIMPULAN</b> .....	<b>46</b>
	<b>DAFTAR PUSTAKA</b> .....	<b>47</b>
	<b>LAMPIRAN</b>	

## DAFTAR TABEL

Tabel	Halaman
1. Contoh Perhitungan Manual pada $X = 1$ dan $t = 1$ .....	15
2. Kriteria Nilai MAPE .....	24
3. Data Awal.....	28
4. Data <i>Input</i> .....	29
5. Statistik Deskriptif Data Harga Penutupan Kedelai Amerika Serikat .....	30
6. Normalisasi Data.....	32
7. Pembagian Data .....	33
8. Perhitungan Manual RMSE .....	35
9. Perhitungan Manual MAPE .....	35
10. Perpaduan Parameter Model LSTM.....	36
11. Contoh Perhitungan Manual pada $X = 1$ dan $t = 1$ .....	37
12. Hasil Prediksi .....	40
13. Evaluasi Model LSTM.....	42
14. Hasil Peramalan Harga.....	43

## DAFTAR GAMBAR

Gambar	Halaman
1. Pola Data <i>Time Series</i> .....	6
2. Ilustrasi Pola Data <i>Time Series</i> .....	8
3. Model Arsitektur <i>Artificial Neural Network</i> .....	11
4. Grafik Fungsi Sigmoid .....	12
5. Grafik Fungsi Aktivasi <i>Tangen Hiperbolik</i> .....	13
6. Arsitektur <i>Long Short Term Memory</i> .....	14
7. Sel Memori <i>Long Short Term Memory</i> .....	14
8. Alur <i>Forget Gate</i> .....	16
9. Alur <i>Input Gate</i> .....	17
10. Alur Untuk <i>Cell State</i> Baru.....	20
11. Alur <i>Output Gate</i> .....	20
12. Diagram Alir Metode Penelitian.....	20
13. Plot Data Historis Kedelai Amerika Serikat.....	29
14. Ilustrasi Proses Segmentasi Pola <i>Time Series</i> .....	31
15. Plot Hasil Prediksi Harga Kedelai Amerika Serikat.....	40
16. Plot Hasil Peramalan Harga Penutupan Kedelai Amerika Serikat .....	45

## I. PENDAHULUAN

### 1.1 Latar Belakang dan Masalah

Pada era modern saat ini kemudahan dalam mengakses berbagai macam data, khususnya dalam mengakses data harga komoditas pangan. Komoditas pangan yang harganya sering mengalami fluktuasi yaitu beras, jagung, kedelai, tepung terigu, gula pasir, minyak goreng, bawang merah, bawang putih, cabai, telur, daging dan susu. Perubahan harga komoditas bahan pangan dapat menjadi penyumbang terbesar laju inflasi (Sumaryanto, 2009). Salah satu komoditas pangan yaitu kedelai yang merupakan komoditas pangan terpenting setelah padi dan jagung.

Menurut Badan Pusat Statistik (2021), Indonesia mengimpor kedelai dari Amerika Serikat sebesar 2.152.633 Ton kedelai. Ketergantungan impor kedelai ini dikenal dengan istilah *dekedelainisasi* (Muslim, 2014). Tujuan dari impor kedelai ini adalah untuk memenuhi kebutuhan permintaan konsumsi dalam negeri, karena produksi dalam negeri tidak dapat memenuhi kebutuhan dalam negeri, dan karena produksi kedelai dalam negeri cenderung menurun sehingga mengakibatkan semakin melebarnya kesenjangan antara penawaran dan permintaan dalam negeri, maka harus dipenuhi dengan impor. Oleh karena itu, pasokan kedelai Indonesia sangat tergantung pada impor (Nuhung, 2013). Hingga saat ini produsen kedelai didominasi oleh negara-negara di Benua Amerika salah satunya adalah Amerika Serikat (Mursidah, 2005). Pergerakan harga komoditas pangan terutama kedelai Amerika Serikat (AS) selalu mengalami fluktuasi dari waktu ke waktu, serta harga penutupan yang sangat penting karena menjadi acuan untuk harga pembukaan esok harinya, sehingga diperlukan suatu metode peramalan pada data *time series*. Fakta ini didukung berdasarkan Data Histori Kedelai Amerika Serikat Periode Januari

2017-2022 yang diperoleh dari *website* investing.com bahwa harga kedelai menunjukkan ketidak stabilan. Fluktuasi harga kedelai dapat diprakirakan dengan peramalan harga kedelai. Peramalan harga kedelai yang akurat penting guna memastikan harga kedelai dimasa depan, selain itu juga untuk membantu pengambilan keputusan. Oleh karena itu sangat penting untuk mengkaji harga kedelai dari periode waktu ke waktu, sehingga fluktuasi harga dapat diantisipasi dengan baik.

Peramalan adalah suatu teknik untuk memperkirakan suatu nilai pada masa yang akan datang dengan mempertimbangkan data masa lalu serta data saat ini. Sangat sulit memprediksi atau meramalkan suatu keadaan di masa mendatang karena faktor ketidakpastian sangat besar pengaruhnya (Gunaryati, dkk. 2018). Salah satu metode peramalan yang paling berkembang saat ini adalah metode berbasis runtun waktu (*time series*). Data runtun waktu (*time series*) adalah nilai-nilai suatu variabel yang berurutan menurut waktu (misal: hari, minggu, bulan, tahun).

Pengembangan metode *time series* adalah *deep learning*, salah satunya adalah *Recurrent Neural Network* (RNN) yang melakukan pembelajaran dengan lapisan yang lebih kompleks untuk mencapai akurasi dan lebih efisien yang tinggi, serta memberikan hasil akurat daripada metode peramalan tradisional karena mampu memodelkan data yang linear dan nonlinear (Arissinta, dkk. 2022). Ketidakpastian harga kedelai merupakan data non stasioner dan non linear. RNN masih memiliki kelemahan yaitu tidak mampu menampung memori jangka panjang sehingga sulit untuk mengingat informasi sebelumnya, mengakibatkan informasi penting dari awal akan tertinggal (Wiranda dan Sadikin, 2019). Untuk mengatasi masalah tersebut terdapat variasi RNN yaitu *Long Short Term Memory* (LSTM) dapat digunakan untuk pemodelan data *time series* dan yang bertujuan membuat peramalan akurat terhadap suatu variabel berdasarkan pada tingkat kesalahan peramalan semakin kecil, maka semakin tepat sebuah metode dalam meramalkan (Yanti, dkk. 2016). Maka digunakan metode *Long Short Term Memory* (LSTM) pada peramalan harga kedelai yang dirancang untuk mengenali pola sebagai urutan data dan berguna dalam prediksi dan prakiraan. Keunggulain lain dari LSTM

adalah memiliki memori untuk mengatasi masalah ketergantungan temporal jangka panjang dengan masukan *sequences*, juga memproses *input* dan *output* tiap langkah waktu, memungkinkan *input* dan *output* yang panjang.

Berdasarkan penelitian terkait dengan peramalan harga komoditas kedelai adalah sebagai berikut. Wang dan Gao (2018), melakukan prediksi pada data harga tertinggi dan terendah kedelai berjangka dari data *Dalian Commodity Exchange* dengan volatilitas. Berdasarkan volatilitas berbeda menghasilkan model LSTM dengan prediksi pada harga cukup bagus sebesar 70%. Rokhimah, dkk. (2022), meramalkan jumlah produksi, luas panen dan harga kedelai dengan metode ARIMA, diperoleh hasil peramalan harga kedelai di Provinsi Jawa Tengah tahun 2021-2022 dengan menggunakan metode ARIMA menunjukkan adanya kenaikan harga kedelai setiap tahunnya. Sen, dkk. (2020). Melakukan penelitian mengenai Komparasi metode *Multilayer Perceptron* (MLP) dan *Long Short Term Memory* (LSTM) dalam peramalan harga beras dari 2016 -2019 yang menghasilkan kesimpulan bahwa hasil LSTM lebih unggul dari MLP dengan nilai RMSE-nya yang lebih kecil dibandingkan dengan MLP. Setiawan, dkk. (2022). melakukan prediksi harga jual kakao dengan metode *Long Short Term Memory* menggunakan optimasi *root mean square propagation* dan *adaptive moment estimation* dilengkapi *GUI Rshiny* diperoleh akurasi terbaik yaitu nilai MSE 491505,1 dan nilai MAPE 1,739155% dan perkiraan kakao cenderung menurun.

Berdasarkan pemaparan di atas, peneliti tertarik untuk mengetahui seberapa akurat metode LSTM dalam meramalkan harga kedelai Amerika Serikat. Oleh karena itu, peneliti mengangkat judul untuk penelitian ini yaitu “Peramalan Data *Time Series* Pada Harga Komoditas Pangan dengan Menggunakan Metode *Long Short Term Memory* (LSTM) Studi Kasus: Harga Penutupan Kedelai Amerika Serikat”.

## 1.2 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah:

1. Memperoleh model LSTM terbaik pada Data Histori Kedelai Amerika Serikat untuk meramalkan harga terakhir atau penutupan pada data Histori Kedelai Amerika Serikat.
2. Mengetahui nilai RMSE dan MAPE yang dihasilkan oleh metode LSTM yang optimal dalam melakukan peramalan harga terakhir atau penutupan pada data Histori Kedelai Amerika Serikat.
3. Memperoleh hasil peramalan harga penutupan pada data Histori Kedelai Amerika Serikat yang dihasilkan oleh metode LSTM.

## 1.3 Manfaat Penelitian

Adapun manfaat dari dilakukannya penelitian ini adalah:

1. Menambah pengetahuan bagi penulis dan pembaca pada umumnya tentang *Long Short Term Memory (LSTM)*.
2. Menjadi sarana pengembangan keilmuan bagi mahasiswa dan masyarakat umum khususnya di bidang peramalan harga komoditas terutama Kedelai menggunakan metode LSTM.
3. Penelitian ini dapat dijadikan bahan referensi pemerintah dalam mengambil keputusan ketika terjadi lonjakan pada impor kedelai serta bagi importir kedelai menjaga stok di gudang importir.
4. Membantu para pedagang untuk memperkirakan harga Kedelai yang akan dibeli.

## II. TINJAUAN PUSTAKA

### 2.1 Peramalan (*Forecasting*)

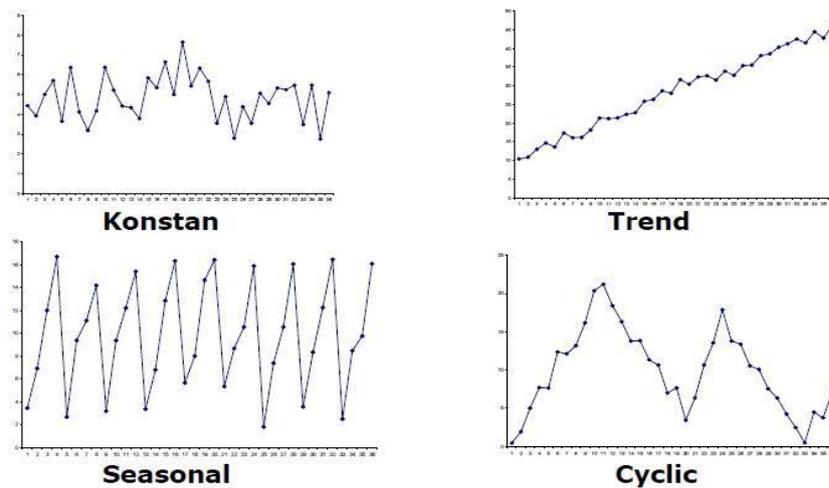
Peramalan atau *forecasting* adalah proses dalam memperkirakan suatu kejadian masa depan yang mungkin terjadi dengan menggunakan informasi data masa lalu dan data masa sekarang. Peramalan tidak selalu memberikan jawaban yang pasti melainkan mencoba memberikan jawaban yang sedekat mungkin dengan kejadiannya (Kafil, 2019). Terdapat dua kategori metode peramalan yaitu metode kuantitatif dan metode kualitatif. Menurut Makridakis, dkk. (1999), metode peramalan kualitatif didasarkan opini atau pendapat dari pihak yang bersangkutan sedangkan metode peramalan kuantitatif menggunakan data historis untuk model peramalan seperti model kausal dan *time series*.

Menurut Heizer dan Render (2015), dari segi jangka waktu peramalannya dapat dikelompokkan menjadi tiga jenis sebagai berikut:

1. Peramalan jangka pendek ialah peramalan dengan jangka waktu kurang dari 3 bulan.
2. Peramalan jangka menengah ialah peramalan mencakup waktu antara 3 hingga 18 bulan.
3. Peramalan jangka panjang ialah peramalan dengan jangka waktu lebih dari 18 bulan.

## 2.2 Data Time Series

Menurut Wei (2006), data *time series* (deret waktu) ialah serangkaian data yang diperoleh dari pengamatan atau kegiatan yang dilakukan selama periode waktu tertentu dan direkam secara berurutan dalam interval waktu yang tetap. Data *time series* merupakan data yang meliputi beberapa objek periode waktu, yaitu harian, mingguan, bulanan, tahunan dan lain-lain yang terdiri dari satu objek. Tujuan peramalan adalah *error* atau kesalahan. Karena semakin sedikit kesalahan maka hasil peramalan akan semakin baik atau mendekati data aktualnya. *Time series* mampu mengatasi satu variabel yang bergantung pada waktu. Pola data *time series* terdapat empat macam pola yaitu horizontal, musiman, siklis dan *trend* (Makridakis, dkk., 1999). Pola horizontal terjadi jika nilai data fluktuasi disekitar rata-rata yang konstan dan membentuk suatu garis horizontal. Pola musiman terjadi ketika suatu deret dipengaruhi oleh faktor musiman (misalnya kuartal tahun tertentu, bulanan, atau hari-hari pada minggu tertentu). Pola siklis terjadi ketika data dipengaruhi oleh fluktuasi jangka panjang. Sementara itu, pola *trend* terjadi ketika mengalami kenaikan atau penurunan selama periode jangka panjang dalam data.



Gambar 1. Pola Data *Time Series*

## 2.3 Data Mining

*Data mining* adalah proses pencarian hubungan, pola, dan tren yang bermakna dengan memeriksa kumpulan data besar yang disimpan dalam *repository* sehingga memerlukan teknik dalam pengenalan pola seperti teknik statistik dan matematika (Larose, 2006). *Data mining* adalah kombinasi yang menggabungkan beberapa disiplin ilmu seperti pengenalan pola, menggunakan pembelajaran mesin, statistik, *database* dan visualisasi untuk memecahkan masalah. Proses tersebut meliputi pembersihan data (*data cleaning*), pemilihan data, penemuan pola, evaluasi pola, dan kesimpulan.

### 2.3.1 Preprocessing Data

Data yang telah diperoleh selanjutnya dilakukan langkah *preprocessing*. Data mentah sebelum digunakan akan diproses terlebih dahulu agar data yang digunakan dapat bekerja secara optimal. Berikut ini tahap dalam *preprocessing data*.

#### 1. Seleksi data

Proses pengumpulan data yang akan digunakan.

#### 2. *Ascending*

Dalam proses *preprocessing data* tahap *ascending* ini memiliki fungsi untuk mengurutkan data dari tanggal terendah hingga terbaru. Data historis kedelai Amerika Serikat merupakan data *time series* sehingga sudah dipastikan sudah sesuai urutan.

#### 3. Normalisasi

Pada langkah ini terjadi proses penskalaan, dimana data akan diubah dalam rentang 0 hingga 1. Normalisasi dilakukan untuk menyamaratakan data. Nilai normalisasi diperoleh dari *min-max normalization*:

$$x' = \frac{(X - X_{min})}{(X_{max} - X_{min})} \quad (2.1)$$

dengan:

$x'$  = data hasil normalisasi

$X$  = data asli

$X_{max}$  = nilai maksimum dari  $x$

$X_{min}$  = nilai minimum dari  $x$

Contoh pada data penutupan harga kedelai Amerika Serikat berdasarkan persamaan (2.1) sebagai berikut:

$$\text{Hari ke-1 : } x'_{1norm} = \frac{(986,75-791)}{(1769-791)} = 0.20015$$

$$\text{Hari ke-2 : } x'_{2norm} = \frac{(1006,25-791)}{(1769-791)} = 0.220092$$

$$\text{Hari ke-3 : } x'_{3norm} = \frac{(1003,50-791)}{(1769-791)} = 0.21728$$

...

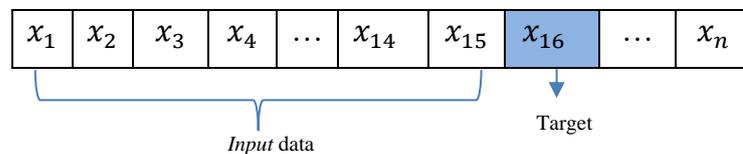
$$\text{Hari ke-1522 : } x'_{1522norm} = \frac{(1469,5-791)}{(1769-791)} = 0.69376$$

$$\text{Hari ke-1523 : } x'_{1523norm} = \frac{(1431,5-791)}{(1769-791)} = 0.65490$$

$$\text{Hari ke-1524 : } x'_{1524norm} = \frac{(1440,38-791)}{(1769-791)} = 0.663987$$

#### 4. Segmentasi

Dalam proses ini mengubah data mentah dengan pengelompokan data dan pemisahan data untuk mengubah data mentah menjadi data yang dibutuhkan sistem.



Gambar 2. Ilustrasi pola data *time series*

Pada Gambar 1, apabila dilakukan pengelompokan yang terdiri dari 15 *time step* maka *time step* 1 hingga *time step* 15 merupakan bagian dari *input*. Sedangkan bagian target merupakan pada *time step* 16.

#### 5. Penentuan data *training* dan data *testing*

Pada proses ini data dibagi menjadi dua bagian yaitu data *training* (latih) dan data *testing* (uji). Proporsi pembagian data bersifat subjektif tergantung peneliti serta persentase data *training* lebih besar dibanding dengan data *testing* (Aisyah, dkk., 2021).

### 2.4 *Machine Learning*

*Machine Learning* adalah sekumpulan teknik yang dapat membantu dalam memproses dan memprediksi data dengan jumlah yang sangat besar dengan mempresentasikan data tersebut melalui algoritma pembelajaran. *Machine learning* memiliki kegunaan yaitu klasifikasi dan prediksi. Proses pelatihan dan pembelajaran, atau *training* merupakan ciri khas dari pembelajaran mesin. Dengan demikian, pembelajaran mesin memerlukan data untuk dipelajari yang dikenal sebagai data *training*. Prediksi dipakai oleh mesin guna menyimpulkan *output* dari suatu data *input* berdasarkan data yang telah dipelajari pada pelatihan.

Fausett (1994), menyatakan bahwa secara umum terdapat dua metode dalam pembelajaran yakni, yaitu metode *supervised learning* (pembelajaran terawasi) dan metode pembelajaran *unsupervised learning* (tidak terawasi).

#### 1. *Supervised Learning*

Metode ini dapat dipakai apabila nilai *output* yang diinginkan telah diketahui sebelumnya. Nilai *output* dari jaringan yang akan dibandingkan dengan nilai *output* targetnya yang disebut galat atau eror. Ketika kesalahan atau *error*nya masih cukup besar berarti perlu dilakukan pembelajaran kembali hingga diperoleh nilai eror yang kecil atau nilai *output* dari jaringan dengan nilai *output* target memiliki nilai yang tidak berbeda jauh.

## 2. *Unsupervised Learning*

Metode *unsupervised learning* (pembelajaran tak terawasi) adalah metode pembelajaran yang tidak memerlukan acuan nilai *output* yang diharapkan. Dalam metode ini tidak diketahui berapa nilai *output* yang diharapkan, tetapi hanya ditentukan nilai bobot pada suatu range sesuai nilai *input* yang digunakan.

### 2.5 *Artificial Neural Network*

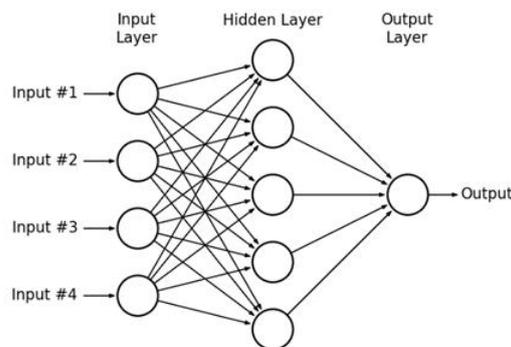
Menurut Prathama, dkk. (2017), *artificial neural network* (ANN) adalah suatu pendekatan dengan cara kerja dalam mengolah suatu informasi seperti cara kerja otak manusia. Pada otak manusia setiap *neuron* saling terhubung dan informasi akan mengalir pada kesetiap *neuron*. Sebuah model ANN memiliki *layer* yang merupakan elemen penting dalam model. *Layer* ini dapat dikategorikan sebagai berikut:

1. *Input layer* adalah *layer* yang berperan sebagai lokasi informasi dari data-data yang bertindak sesuai dengan *output* yang diinginkan. Pada *input layer* terdapat beberapa *neuron* yang mampu mempresentasikan parameter-parameter penting untuk dapat penyelesaian masalah. Melalui *input layer* ini data-data akan yang akan diteruskan menuju *hidden layer* maupun *output layer*.
2. *Hidden layer* terletak di antara *input layer* dan *output layer*, yang bertugas menerima data dari *input layer* dan meneruskan menuju *output layer*.
3. *Output layer* adalah *layer* yang menerima data dan memberikan hasil perhitungan dari *input layer* menggunakan fungsi aktivasi. Nilai dari *output layer* melambangkan hasil keluaran dari  $X$  menjadi nilai  $Y$ .

Menurut Kusumadewi (2003), komponen-komponen artificial neural network sebagai berikut.

1. *Input* : sama seperti dendrit yang terdapat pada otak manusia memiliki fungsi sebagai penerima informasi masukan dari *neuron* lain.

2. *Neuron* : komponen yang bertugas untuk memproses informasi.
3. *Bobot* : sama seperti sinopsis yang terdapat pada otak manusia berfungsi untuk menunjukkan kekuatan hubungan antara *neuron* satu dengan yang lainnya.
4. *Fungsi aktivasi* : suatu nilai tertentu yang memetakan fungsi hasil penjumlahan yang diterima oleh semua *input* dari suatu *neuron*.
5. *Output* : pada otak manusia seperti axon pada otak manusia yang memiliki fungsi untuk memproses pembelajaran atau proses perhitungan suatu fungsi aktivasi yang akan menghasilkan suatu *output* dari jaringan yang telah di *input* atau bahkan akan menjadi *input*-an bagi *neuron* yang lain.



Gambar 3. Model Arsitektur *Artificial Neural Network*  
(Sumber: Tian, dkk., 2018)

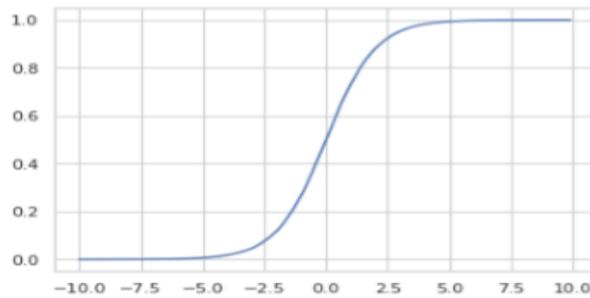
## 2.6 Fungsi Aktivasi

Fungsi aktivasi adalah fungsi yang digunakan dalam jaringan saraf untuk mengaktifkan atau menonaktifkan *neuron* (Julpan, dkk., 2015). Fungsi aktivasi dapat berupa linier atau non linear tergantung pada fungsi yang mempresentasikannya dan mengontrol *output* dari jaringan saraf luar. Fungsi aktivasi yang sering digunakan sebagai berikut:

1. Fungsi aktivasi sigmoid merupakan fungsi aktivasi nonlinear yang dipakai pada *Neural Network*. Fungsi Sigmoid mengubah nilai *input*  $x$  menjadi nilai antara 0 hingga 1 (Pradasari, dkk. 2013).

$$\sigma(x) = \frac{1}{(1+e^{-x})} \quad (2.2)$$

Adapun fungsi sigmoid ini apabila digambarkan dalam bentuk grafik akan menghasilkan kurva yang berbentuk seperti huruf S seperti berikut:



Gambar 4. Grafik Fungsi Sigmoid

2. Fungsi aktivasi alternatif untuk sigmoid adalah fungsi aktivasi tanh. Seperti fungsi sigmoid, fungsi tanh berbentuk “S”, tetapi fungsi ini menghasilkan nilai antara -1 dan 1. Oleh karena itu, fungsi tanh mirip dengan fungsi sigmoid, tetapi rentang nilai yang luas membuatnya efektif untuk pemodelan nonlinear yang kompleks. Bentuk dari fungsi tanh adalah:

$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)} \quad (2.3)$$

Fungsi tanh atau tangen hiperbolik dapat dirumuskan sebagai berikut:

Dimana  $\sinh x = \frac{e^x - e^{-x}}{2}$  dan  $\cosh x = \frac{e^x + e^{-x}}{2}$ , sehingga:

$$\tanh(x) = \frac{\frac{e^x - e^{-x}}{2}}{\frac{e^x + e^{-x}}{2}}$$

$$\tanh(x) = \frac{e^x - e^{-x}}{2} \cdot \frac{2}{e^x + e^{-x}}$$

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

$$\tanh(x) = \frac{e^x(1 - e^{-2x})}{e^x(1 + e^{-2x})}$$

$$\tanh(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}}$$

$$\tanh(x) = \frac{2}{1 + e^{-2x}} - \frac{(1 + e^{-2x})}{1 + e^{-2x}}$$

$$\tanh(x) = \frac{2}{1 + e^{-2x}} - 1$$

Berdasarkan persamaan (2.2), maka diperoleh:

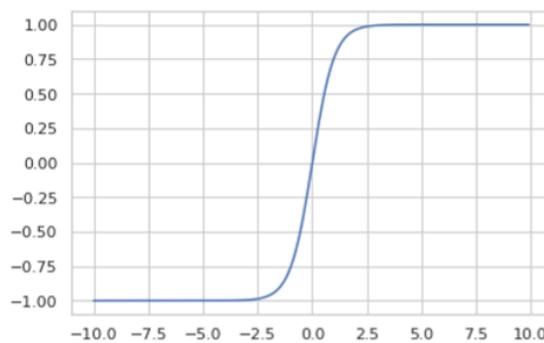
$$\tanh(x) = 2\sigma(2x) - 1 \quad (2.4)$$

dengan:

$\sigma$  = fungsi sigmoid

$x$  = data *input*

Adapun grafik dari fungsi tanh sebagai berikut:



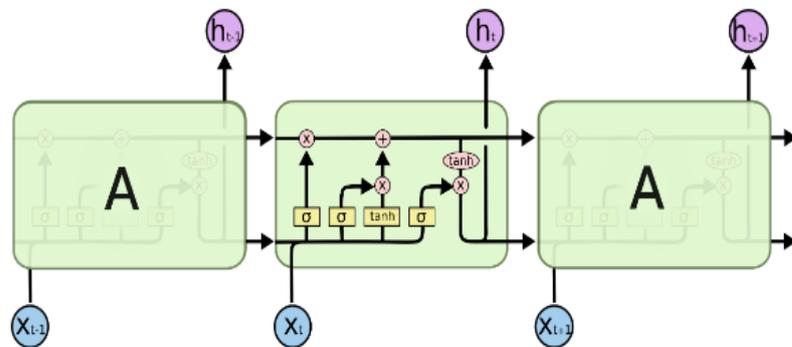
Gambar 5. Grafik Fungsi Aktivasi *Tangen Hiperbolik*

## 2.7 Long Short Term Memory (LSTM)

*Long Short Term Memory* adalah modifikasi dari *Recurrent Neural Network* (RNN) dengan menambahkan *memory cell* yang mampu menyimpan informasi dalam waktu yang cukup lama. LSTM mempunyai *gated cell* yang merupakan tambahan jalur penyimpanan informasi diluar dari aliran informasi RNN. Informasi yang melewati jalur ini akan hanya mengalami perubahan linear sehingga nilai yang terkandung tidak mudah berubah. Informasi pada sel hanya mengalami perubahan

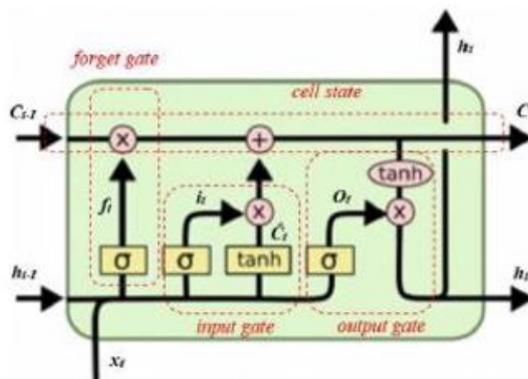
ketika melalui struktur gerbang yang disebut dengan *gates* (Olah, 2015). Sel pada LSTM dapat menghubungkan informasi sebelumnya dengan informasi selanjutnya dan keefektifan dalam menyimpan informasi yang panjang sangat diperlukan dalam mengolah data *time series*. LSTM mampu mengatasi permasalahan pada menghilangnya atau meledaknya gradient.

Model LSTM melakukan penyaringan informasi melalui struktur gerbang untuk mempertahankan serta memperbarui keadaan sel memori. Terdapat tiga gerbang yang mengatur alur informasi dalam sel LSTM terdiri dari *input gate*, *forget gate*, dan *output gate*. Berikut merupakan arsitektur LSTM.



Gambar 6. Arsitektur *Long Short Term Memory*  
(Sumber: Olah, 2015)

Pada LSTM setiap sel memori memiliki tiga lapisan sigmoid dan satu lapisan tanh serta lapisan tersembunyi terdiri dari sel memori, satu sel memori tersusun atas tiga gates, yaitu *forget gate*, *output gate*, dan *input gate* (Vinayukumar, dkk., 2017).



Gambar 7. Sel Memori *Long-Short Term Memory*  
(Sumber: Olah, 2015)

Gambar 8 menunjukkan bagian dari lapisan tersembunyi LSTM, yaitu sel-sel memori yang mampu menyimpan nilai atau status (keadaan sel) untuk jangka waktu lama atau singkat. Berikut ini langkah dan contoh pada setiap *time step* dengan sampel perhitungan. Nilai weight dan biases ditentukan secara acak oleh program.

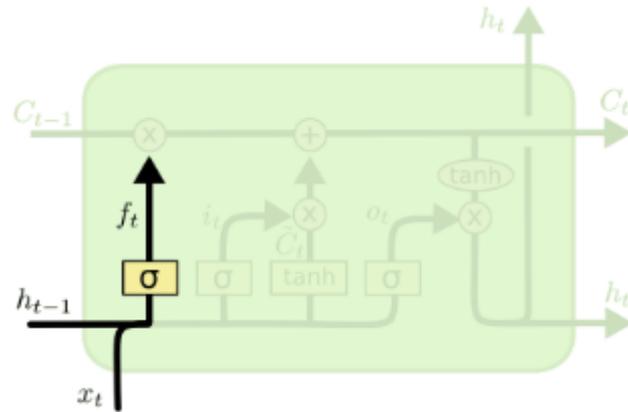
Tabel 1. Contoh perhitungan manual pada  $X_1$  dan  $t = 1$

$X_1$	$t = 1$
$x_t$	0.20015
$W_f$	$-5.91199808E - 02$
$W_i$	$5.97282536E - 02$
$W_c$	$1.61976200E - 02$
$W_o$	$7.33344676E - 03$
$b_f$	-0.000993
$b_i$	-0.09135693
$b_c$	0.05770339
$b_o$	-0.03704451

Bobot atau *weight* adalah hubungan antar node dengan suatu nilai. *Output* adalah keluaran dari suara node. Error adalah tingkat kesalahan yang terdapat dalam suatu node dari proses yang dilakukan. *Weight* merupakan bobot dari node tersebut ke node yang lain pada layer yang berbeda. Bobot awal dalam suatu jaringan syaraf tiruan biasanya diperoleh secara random dan sebaiknya diinisialisasi dengan nilai yang relatif kecil, yaitu berkisar antara -1 sampai 1. Pada tahap pelatihan, bobot tersebut akan mengalami penyesuaian melalui suatu proses perhitungan matematik agar tercapai nilai bobot yang sesuai.

Berikut adalah gerbang dalam satu sel LSTM.

1. *Forget gate* merupakan gerbang pertama yang dilalui. Gerbang ini menentukan apakah informasi dari input sel  $x_t$  dan *output*  $h_{t-1}$  pantas lewat atau tidak. *Layer neuron* sigmoid akan membuat keputusan nilai mana yang harus dilewati dan menghasilkan  $f_t$ . Pada outputnya, mendekati 1 artinya “biarkan lewat” sedangkan mendekati 0 artinya “lupakan/abaikan informasi ini”. Berikut pada gambar 8 adalah forget gate LSTM:

Gambar 8. Alur *Forget Gate*

(Sumber: Olah, 2015)

Persamaan *forget gate* diuraikan sebagai berikut :

$$f_t = \sigma(W_f \cdot x_t + W_f \cdot h_{t-1} + b_f) \quad (2.5)$$

dengan:

$f_t$  = *forget gate*

$\sigma$  = fungsi sigmoid

$W_f$  = nilai *weight* untuk *forget gate*

$h_{t-1}$  = nilai *output* sebelum orde ke  $t$

$x_t$  = nilai *input* pada order ke  $t$

$b_f$  = nilai bias pada *forget gate*

*Forget gate* berfungsi untuk merekam seberapa banyak status sel  $C_{t-1}$  dari waktu sebelumnya dicadangkan ke status sel  $C_t$  dari waktu saat ini. Gerbang akan memunculkan nilai antara 0 hingga 1 berdasarkan pada  $h_{t-1}$  dan  $x_t$ , dimana *outputnya* bernilai 0, maka informasi dianggap tidak lagi berguna dan bisa dihapus. Begitu sebaliknya, apabila *outputnya* bernilai 1 maka informasi tersebut disimpan untuk penggunaan di masa mendatang.

Time step  $t = 1$

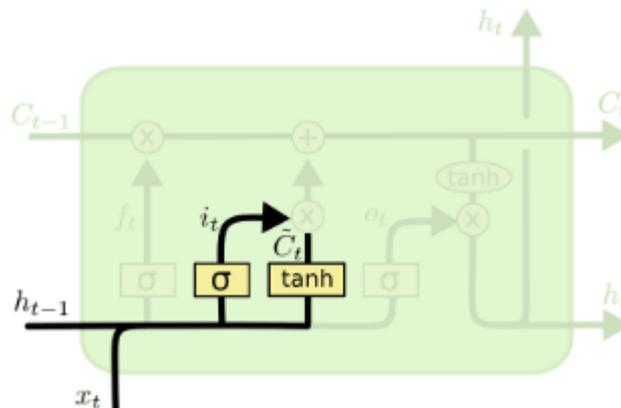
$h_{t-1} = 0$

$c_{t-1} = 0$

Menghitung forget gate ( $f_t$ ) berdasarkan persamaan (2.5) sebagai berikut:

$$\begin{aligned}
 f_t &= \sigma(W_f \cdot x_t + W_f \cdot h_{t-1} + b_f) \\
 &= \sigma((-5.91199808E - 02 * 0.20015) + (-5.91199808E - 02 * 0) - \\
 &\quad 0.000993)) \\
 &= \sigma(-16.47177) \\
 &= \frac{1}{(1+e^{-(-16.47177)})} \\
 &= 7,02104 \times 10^{-8}
 \end{aligned}$$

2. *Input gate* adalah gerbang yang memiliki masukan dengan dua fungsi aktivasi (sigmoid dan tanh) dengan fungsi memilih bagian yang akan diperbaharui. Pertama, *layer neuron* sigmoid memutuskan nilai mana yang akan diperbarui yang menghasilkan  $i_t$  dan fungsi aktivasi tanh memberikan bobot pada nilai-nilai yang dilewati dan menentukan tingkat kepentingannya menghasilkan  $\hat{C}_t$  dengan alur, yaitu:



Gambar 9. Alur *Input Gate*

(Sumber: Olah, 2015)

Tujuan dari *input gate* adalah untuk menentukan berapa banyak *input* jaringan waktu saat ini ( $x_t$ ) dicadangkan ke dalam status sel  $C_t$ , yang mencegah konten tidak signifikan dari memasuki sel memori. *Input gate* mempunyai dua fungsi yaitu untuk menemukan keadaan sel yang harus diperbarui. Persamaan *input gate* diuraikan sebagai berikut.

$$i_t = \sigma(W_i \cdot x_t + W_i \cdot h_{t-1} + b_i) \quad (2.6)$$

dengan:

$i_t$  = *input gate*

$\sigma$  = fungsi sigmoid

$W_f$  = nilai *weight* untuk *input gate*

$h_{t-1}$  = nilai *output* sebelum orde ke  $t$

$x_t$  = nilai *input* pada order ke  $t$

$b_i$  = nilai bias pada *input gate*

Menghitung *input gate* berdasarkan persamaan (2.6) sebagai berikut:

$$\begin{aligned}
 i_t &= \sigma(W_i \cdot x_t + W_i \cdot h_{t-1} + b_i) \\
 &= \sigma((5.97282536E - 02 * 0.20015) + (5.97282536E - 02 * 0) + \\
 &\quad (-0.09135693)) \\
 &= \sigma(15,74416) \\
 &= \frac{1}{(1+e^{-(15,74416)})} \\
 &= 1
 \end{aligned}$$

Selanjutnya, untuk fungsi kedua dari *input gate* ialah untuk memperbarui informasi ke keadaan sel. Kandidat baru  $\hat{C}_t$  dibuat melalui lapisan tanh agar dapat mengontrol berapa banyak informasi baru ditambahkan. Persamaan kandidat baru sebagai berikut:

$$\hat{C}_t = \tanh(W_c \cdot x_t + W_c \cdot h_{t-1} + b_c) \quad (2.7)$$

dengan:

$\hat{C}_t$  = nilai baru yang ditambahkan ke *cell state*

$\tanh$  = fungsi tanh

$W_c$  = nilai *weight* untuk *cell state*

$h_{t-1}$  = nilai *output* sebelum orde ke  $t$

$x_t$  = nilai *input* pada order ke  $t$

$b_i$  = nilai bias pada *cell state*

Menghitung  $\hat{C}_t$  berdasarkan persamaan (2.7) sebagai berikut

$$\begin{aligned}
 \hat{C}_t &= \tanh(W_c \cdot x_t + W_c \cdot h_{t-1} + b_c) \\
 &= \tanh((1.61976200E - 02 * 0.20015) + (1.61976200E - 02 * 0) + \\
 &\quad (0.05770339)) \\
 &= \tanh(3,94497) \\
 &= 2 \sigma(2 * 3,94497) - 1 \\
 &= \frac{2}{(1+e^{-(3,94497)})} - 1 \\
 &= 0.962031
 \end{aligned}$$

Setelah itu *cell state* yang lama akan diperbaharui menjadi *cell state* yang baru dengan mengalikan *state* lama dengan *forget gate* ( $f_t$ ) untuk menghapus informasi yang telah ditentukan pada *layer forget gate* kemudian ditambahkan dengan  $i_t * \hat{C}_t$  yang merupakan nilai baru untuk memperbaharui *state*, sehingga menghasilkan persamaan *cell state* sebagai berikut.

$$C_t = f_t * C_{t-1} + i_t * \hat{C}_t \quad (2.8)$$

dengan:

$C_t$  = *cell state*

$f_t$  = *forget gate*

$C_{t-1}$  = *cell state* sebelum orde ke  $t$

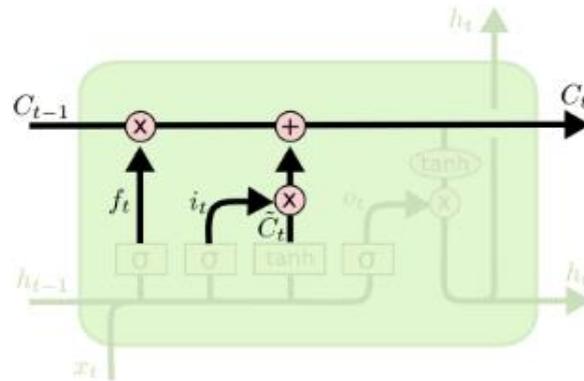
$i_t$  = nilai *output* sebelum orde ke  $t$

$x_t$  = *input gate*

$b_i$  = nilai baru yang dapat ditambahkan ke *cell state*

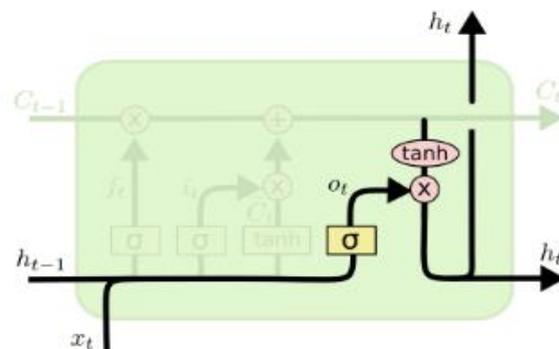
Menghitung *cell state* ( $C_t$ ) berdasarkan persamaan (2.8) sebagai berikut:

$$\begin{aligned}
 C_t &= f_t * C_{t-1} + i_t * \hat{C}_t \\
 &= 7,02104 \times 10^{-8} * 0 + 1 * 0.962031 \\
 &= 0.962031
 \end{aligned}$$



Gambar 10. Alur untuk *Cell State* Baru  
(Sumber: Olah, 2015)

3. *Output gate* berfungsi untuk memutuskan apa yang akan dihasilkan. Pertama menjalankan *layer neuron* sigmoid yang menentukan bagian sel apa yang akan dihasilkan pada  $O_t$ . Langkah berikutnya mengalikan dengan *cell state* yang telah melewati *layer neuron* tanh dan hasilnya akan menjadi  $h_t$  dan meneruskan menuju *cell state*, tetapi *gate* inilah yang membedakan *cell state*  $C_t$  dan  $h_t$  yang sebenarnya, dengan alur yaitu:



Gambar 11. Alur *Output Gate*  
(Sumber: Olah, 2015)

Persamaan *output gate* sebagai berikut.

$$O_t = \sigma(W_o * x_t + W_o * h_{t-1} + b_o) \quad (2.9)$$

dengan:

$O_t$  = *output gate*

$\sigma$  = fungsi sigmoid

$W_o$  = nilai *weight* untuk *output gate*

$h_{t-1}$  = nilai *output* sebelum orde ke  $t$

$x_t$  = nilai *input* pada orde ke  $t$

$b_o$  = nilai bias pada *output gate*

Menghitung *output gate* berdasarkan persamaan (2.9) sebagai berikut:

$$\begin{aligned}
 O_t &= \sigma(W_o * x_t + W_o * h_{t-1} + b_o) \\
 &= \sigma((7.33344676E - 03 * 0.20015) + (7.33344676E - 03 * 0) + \\
 &\quad (-0.03704451)) \\
 &= \sigma(19,29688) \\
 &= \frac{1}{(1+e^{-(19,29688)})} \\
 &= 1
 \end{aligned}$$

Setelah nilai dari *output gate* didapatkan maka *cell state* ditempatkan melalui tanh. Lalu mengalikan dengan *output gate* dan sigmoid layer. Persamaan nilai *output* orde ke  $t$  pada persamaan sebagai berikut.

$$h_t = O_t * \tanh(C_t) \quad (2.10)$$

dengan:

$h_t$  = nilai *output* order ke  $t$

$O_t$  = *output gate*

tanh = fungsi tanh

$C_t$  = *cell state*

Menghitung *output gate* berdasarkan persamaan (2.10) sebagai berikut:

$$\begin{aligned}
 h_t &= O_t * \tanh(C_t) \\
 &= 1 * (2 \sigma(2 * 0.962031) - 1)
 \end{aligned}$$

$$\begin{aligned}
&= 1 * \left( \frac{2}{(1+e^{-(0.962031)})} - 1 \right) \\
&= 1 * (0,447057) \\
&= 0,447057
\end{aligned}$$

Hasil dari nilai  $h_t$  dan  $c_t$  yang diperoleh akan dibawa ke *time step* selanjutnya.

### 2.7.1 Membangun Model *Long Short Term Memory*

#### 1. *Hidden dan Nodes Layer*

Heaton (2008), berpendapat bahwa jumlah *hidden layer* yang digunakan akan berpengaruh terhadap perubahan hasil dari pelatihan. Jumlah *neuron* yang terlalu sedikit akan menghasilkan *output* yang kurang baik, sedangkan apabila *neuron* terlalu banyak akan memperlambat proses pelatihan dan mungkin akan terjadi pelatihan yang tak hingga. Menurut Fausett (1994), penentuan jumlah *hidden layer* untuk jaringan *backpropagation* satu *hidden layer* cukup untuk memperkirakan setiap pemetaan.

#### 2. *Epoch dan Batch Size*

Menurut Vijayalakshmi dan Venkatachalapathy (2019), *epoch* merupakan jumlah iterasi selama proses pelatihan yang menghasilkan *input* dari jaringan dan juga memperbarui bobot jaringan. Dalam proses pelatihan, satu *epoch* membutuhkan waktu yang lama. Dengan demikian, untuk mempercepat proses pelatihan dilakukan pembagian per *batch* yang disebut dengan *batch size*. Ukuran *batch* adalah istilah yang digunakan dalam pembelajaran mesin dan mengacu pada jumlah contoh pelatihan yang digunakan dalam satu iterasi. (Rochmawati, dkk., 2021). Nilai yang biasa digunakan antara lain adalah 16, 32, 64 dengan mempertimbangkan angka kelipatan 2 yang dapat memudahkan proses pembagian dataset dalam pelatihan model.

### 3. Fungsi Aktivasi

Terlebih dulu menentukan jumlah node pada *neural network*, khususnya pada *hidden layer*, karena *hidden layer* berperan penting dalam menghitung hasil akhir dari *neural network* maka penentuan ini menjadi hal yang sangat penting sebelum menjalankan *training*. Fungsi aktivasi merupakan fungsi yang digunakan pada jaringan syaraf untuk mengaktifkan atau tidak mengaktifkan *neuron*.

## 2.8 Denormalisasi Data

Sebelum dilakukan perhitungan akurasi maka dari hasil prediksi yang didapatkan perlu dilakukan denormalisasi, yaitu data diubah kembali menjadi bentuk aslinya karena data hasil prediksi masih berupa interval pada saat normalisasi data (Ashar, dkk., 2018). Tujuan denormalisasi adalah agar *output* mudah dibaca dan mudah dipahami. Dibawah ini adalah rumus untuk denormalisasi.

$$X_t = x'_t(X_{max} - X_{min}) + X_{min} \quad (2.11)$$

dengan:

- $X_t$  = nilai hasil denormalisasi
- $x'_t$  = nilai *output* data setelah dihasilkan
- $X_{max}$  = nilai maksimal pada data aktual
- $X_{min}$  = nilai minimal pada data aktual

## 2.9 Evaluasi Model

Menurut Makridakis, dkk. (1999), bahwa ketepatan atau akurasi menunjukkan seberapa jauh model peramalan tersebut mampu memproduksi data yang telah diketahui. Dua alasan utama melihat tingkat akurasi pada prediksi model *time*

series yaitu melihat nilai *Root Mean Square Error* (RMSE) dan *Mean Absolute Percentage Error* (MAPE) (Vercellis, 2009).

### 1. *Root Mean Square Error* (RMSE)

*Root Mean Square Error* digunakan dalam menghitung perbedaan atau kesalahan diantara data prediksi dan aktual dibagi dengan banyaknya data. Dengan persamaan sebagai berikut:

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (\hat{Y}_t - Y_t)^2}{n}} \quad (2.12)$$

dengan:

$Y_t$  = nilai data aktual

$\hat{Y}_t$  = nilai akhir data peramalan

$n$  = banyaknya data

### 2. *Mean Absolute Percentage Error* (MAPE)

Berdasarkan Riyadi (2005), MAPE adalah rata-rata diferensiasi absolut antara nilai peramalan dan aktual, yang dinyatakan sebagai persentase nilai aktual. MAPE digunakan untuk menghitung persentase kesalahan antara nilai aktual dan nilai prediksi. MAPE dirumuskan sebagai berikut:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|\hat{Y}_t - Y_t|}{Y_t} \times 100\% \quad (2.13)$$

dengan:

$Y_t$  = nilai data aktual

$\hat{Y}_t$  = nilai akhir data peramalan

$n$  = banyaknya data

Tabel 2. Kriteria nilai MAPE

No	MAPE	Penjelsan Nilai
1	<10 %	Sangat Baik
2	10-20 %	Baik
3	20-50 %	Sedang
4	>50 %	Buruk

### **III. METODOLOGI PENELITIAN**

#### **3.1 Waktu dan Tempat Penelitian**

Penelitian ini dilakukan pada semester ganjil tahun akademik 2022/2023 di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

#### **3.2 Data Penelitian**

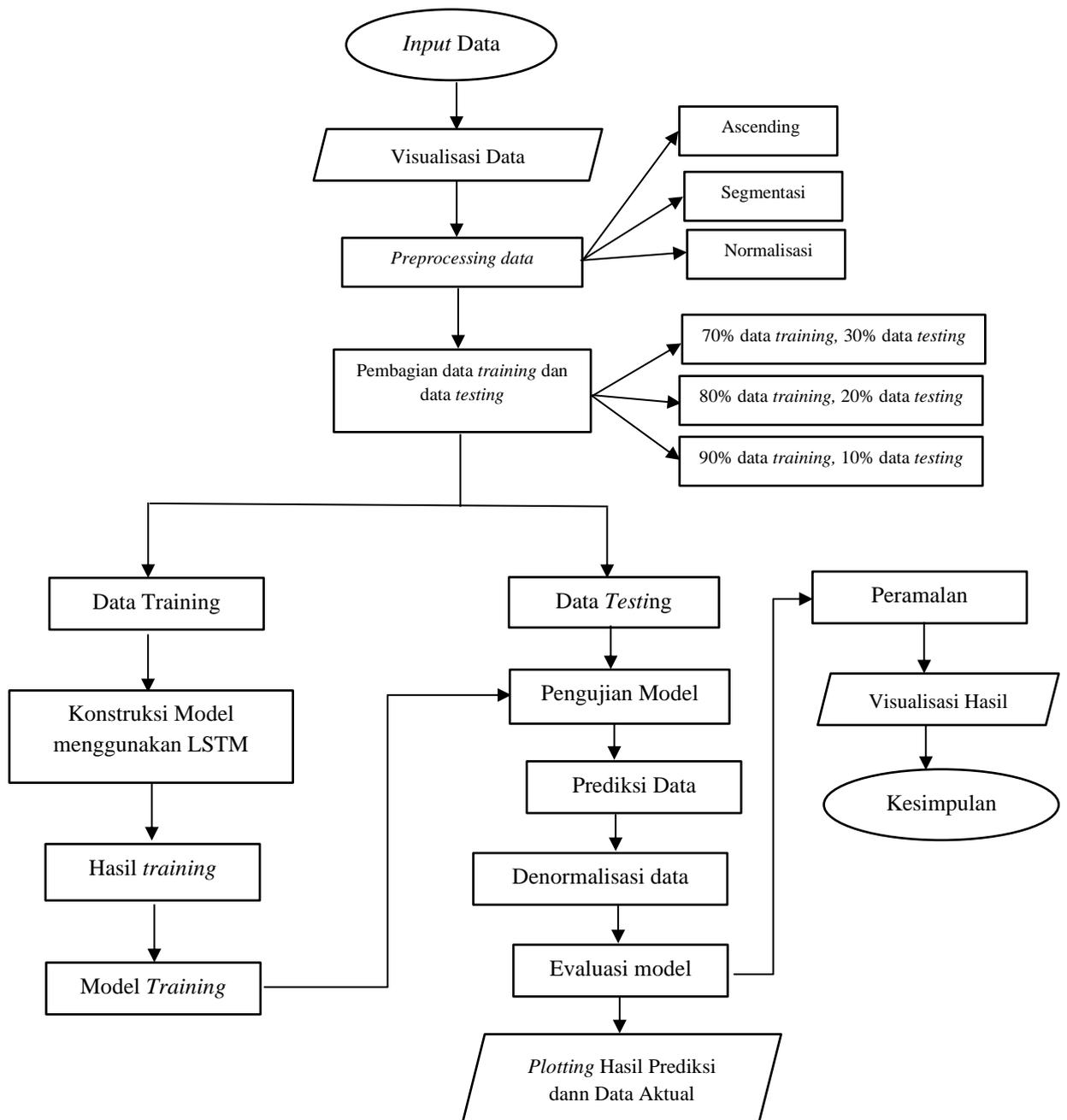
Data yang digunakan pada penelitian ini adalah data sekunder yang diperoleh dari <https://id.investing.com/> mengenai data histori kedelai AS berjangka selama 6 tahun terhitung sejak 03 Januari 2017 sampai dengan 02 Desember 2022 dalam skala harian. Data berjumlah 1524 berbentuk tabel yang memiliki kolom Tanggal, Terakhir, Pembukaan, Tertinggi, dan Terendah. Kolom yang digunakan adalah Tanggal dan harga Terakhir (penutupan).

#### **3.3 Metode Penelitian**

Dalam penelitian ini, peneliti akan menampilkan model terbaik LSTM untuk meramalkan harga terakhir atau penutupan pada Kedelai AS menggunakan metode *Long Short Term Memory* (LSTM) dengan bantuan *software* Python yang didukung oleh Google Colab. Lalu, model yang dihasilkan akan dievaluasi berdasarkan nilai RMSE dan MAPE.

Adapun langkah-langkah yang dilakukan pada metode penelitian ini sebagai berikut:

1. Melakukan visualisasi data dengan membuat grafik *time series* dan analisis deskriptif data.
2. Melakukan *preprocessing* data yaitu mengurutkan data dari tanggal terlama hingga tanggal terbaru dan melakukan normalisasi data menggunakan *min-max normalization* berdasarkan variabel yang sudah ditentukan.
3. Membagi data menjadi data *training* untuk membentuk model dan data *testing* untuk menguji model dengan pembagian data sebesar
  - 70% data *training* dan 30% data *testing*,
  - 80% data *training* dan 20% data *testing*,
  - 90% data *training* dan 10% data *testing*.
4. Melakukan konstruksi model dengan LSTM berdasarkan masing-masing pembagian data menggunakan data latih dengan menentukan inisialisasi parameter-parameter yang dibutuhkan yaitu jumlah *epoch*, jumlah *neuron hidden* dan jumlah *batch size*. Simpan model *training*.
5. Menguji model *training* dengan data uji.
6. Melakukan prediksi untuk melihat prediksi harga pada indeks “Terakhir” atau harga penutupan kedelai Amerika Serikat.
7. Melakukan evaluasi model menggunakan RMSE dan MAPE. Dengan model yang tersimpan akan dilanjutkan untuk peramalan harga penutupan.
8. Melakukan peramalan untuk dengan data *testing* untuk mendapatkan harga penutupan kedelai Amerika Serikat 2 bulan ke depan.
9. Melihat visualisasi yang dihasilkan dari data yang diramalkan.
10. Memaparkan kesimpulan berdasarkan hasil yang diperoleh.



Gambar 12. Diagram Alir Metode Penelitian

## V. KESIMPULAN

Berdasarkan hasil dari pembahasan, maka diperoleh kesimpulan sebagai berikut:

1. Model LSTM terbaik untuk data harga penutupan kedelai Amerika Serikat adalah model dengan data *training* 80% dan 20% data *testing*, dan parameter terbaik meliputi 1 *hidden layer*, 100 *epoch*, 64 *neuron hidden*, dan 16 *batch size*.
2. Berdasarkan nilai RMSE dan MAPE yang dihasilkan dari model LSTM yakni sebesar 23,6674 dan 1,2152%. Nilai MAPE dalam model ini kurang dari 10% di mana menandakan kemampuan prediksi data harga penutupan kedelai Amerika Serikat sangat baik. Oleh karena itu, metode LSTM dari parameter yang optimal ini bisa dimanfaatkan sebaik mungkin untuk meramalkan harga terakhir atau penutupan kedelai Amerika Serikat atau komoditas pangan yang lain, dikarenakan metode ini hanya menghasilkan nilai *error* yang kecil.
3. Berdasarkan hasil peramalan harga penutupan kedelai Amerika Serikat menggunakan metode LSTM untuk 2 bulan ke depan sejak tanggal 05 Desember 2022 hingga 27 Januari 2023 menunjukkan harga kedelai Amerika Serikat cenderung meningkat.

## DAFTAR PUSTAKA

- Aisyah, S, Wahyuningsih. S., dan Amijaya., F.D.T. 2021. Peramalan Jumlah Titik Panas Provinsi Kalimantan Timur Menggunakan Metode Radial Basis Function Neural Network. *Jambura Journal of Probability And Statistics*. **2**(2):64-74.
- Arissinta, I.O., Sulistiyawati, I.D., dan Kurnianto, D. 2022. Pemodelan *Time Series* untuk Peramalan *Web Traffic* Menggunakan Algoritma Arima, LSTM, dan GRU, hlm. 693-700. Prosiding Seminar Nasional Matematika, Semarang.
- Ashar, N.M., Cholissodin, I., dan Dewi, C. 2018. Penerapan Metode Extreme Learning Machine (ELM) untuk Memprediksi Jumlah Produksi Pipa yang Layak (Studi Kasus pada PT. KHI Pipe Industries). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*. **2**(11): 4621–4628.
- Badan Pusat Statistik (BPS). 2021. Impor Kedelai menurut Negara Asal Utama, 2017-2021. <https://www.bps.go.id/statictable/2019/02/14/2015/imporkedelai-menurut-negara-asal-utama-2017-2021.html>. Diakses pada 20 November 2022.
- Deng, L. dan Yu, D. 2014. Deep Learning : Methods and Applications. *Foundations and Trends in Signal Processing*. **7**(3-4):197 - 387.
- Fausett, L. 1994. *Fundamentals of Neural Networks: Achitectures, Algorithms, and Applications*. Prentice Hall. New Jersey.
- Gunaryati, A., Fauziah, Andryan, S. 2018. Perbandingan Metode-Metode Peramalan Statistika Untuk Data Indeks Harga Pangan. *Jurnal String*. **2**(3): 241-248.
- Heizer, Jay, dan Render, B. 2015. *Manajemen Operasi : Manajemen Keberlangsungan dan Rantai Pasokan*. Ed ke-11. Salemba Empat, Jakarta.

- Heaton, J. 2008. *Introduction to Neural Network for Java*. 2nd Ed. Heaton Research Inc. Florida.
- Julpan, Budhiarti, E., dan Zarlis, M. 2015. Analisis Fungsi Aktivasi *Sigmoid* Biner dan *Sigmoid* Bipolar Dalam Algoritma Backpropagation pada Prediksi Kemampuan Mahasiswa. *Jurnal Teknik dan Inovasi Mesin Otomotif, Komputer, Industri dan Elektronika*. **2**(1): 103-116.
- Kafil, M. 2019. Penerapan Metode K-Nearest Neighbors Untuk Prediksi Penjualan Berbasis Web Pada Boutiq Dealove Bondowoso. *Jurnal Mahasiswa Teknik Informatika*. **3**(2): 59-66.
- Kusumadewi, S. 2003. *Artificial Intelligence (Teknik dan Aplikasinya)*. Graha Ilmu. Yogyakarta.
- Larose, D.T. 2006. *Discovering Knowledge in Data: An Introduction to Data Mining*. New Jersey John Wiley & Sons, Inc, USA.
- Makridakis, S., Wheelwright, S.C., dan McGee, V.E. 1999. *Metode dan Aplikasi Peramalan*. Edisi Ke-2. Jakarta : Binarupa Aksara.
- Muslim, A. 2014. Peramalan Harga Paritas Kedelai Model Anfi. *Jurnal Widyariset*. **17**(1): 13-24.
- Mursidah. 2005. Perkembangan Produksi Kedelai Nasional dan Upaya Pengembangannya di Provinsi Kalimantan Timur. *EPP*, **2**(1):39-44.
- Nuhung, I., A. 2013. Kendali dan Politik Pangan. *Forum Penelitian Agro Ekonomi*. **31**(2): 123-135.
- Olah, C. 2015. Understanding LSTM Networks. <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>. Diakses pada 10 November 2022.
- Pradasari, N.I., Pontia F.T., Triyanto, D. 2013, Jaringan Syaraf Tiruan untuk Memprediksi Penyakit Saluran Pernafasan dengan Metode Backpropagation. *Jurnal Coding* **1**(1): 20-30.
- Prathama, A.Y., Aminullah, A., dan Saputra, A. 2017. Pendekatan ANN (Artificial Neural Network) Untuk Penentuan Prosentase Bobot Pekerjaan Dan Estimasi Nilai Pekerjaan Struktur Pada Rumah Sakit Pratama. *Teknosains*. **7**(1): 14-25.
- Puspitaningrum, D. 2006. *Pengantar Jaringan Syaraf Tiruan*. Penerbit Andi, Yogyakarta.

- Riyadi, S. 2015. Aplikasi Peramalan Penjualan Obat Menggunakan Metode Pemulusan (Studi Kasus: Instalasi Farmasi RSUD Dr Murjani). *Open Journal System Universitas AMIKOM*. (1): 1-6.
- Rochmawati, N., Hidayati, H.B., Yamasari, Y., Tjahyaningtjas, H.P.A., Yustanti, W., dan Prihanto, A. 2021. Analisa Learning Rate dan Batch size Pada Klasifikasi Covid Menggunakan Deep learning dengan Optimizer Adam. *Journal Information Engineering and Educational Technology*. 5(2):44-48.
- Rokhimah, S., Widjoko, T., dan Mandamdari, A.N. 2022. Analisis Peramalan Produksi, Luas Panen, dan Harga Kedelai di Provinsi Jawa Tengah, hlm. 124-130. Prosiding Seminar Nasional Hasil Penelitian Agribisnis VI, Purwokerto.
- Sen, S., Sugiarto, D., dan Rochman, A. 2020. Komparasi Metode Multilayer Perceptron (MLP) dan Long Short Term Memory (LSTM) dalam Peramalan Harga Beras. *Jurnal Teknik Informatika*. 12(1): 35-41.
- Setiawan, Y., Tarno dan Kartikasari, P. 2022. Prediksi Harga Jual Kakao dengan Metode Long Short Term Memory Menggunakan Metode Optimasi Root Mean Square Propagation dan Adaptive Moment Estimation Dilengkapi GUI Rshiny. *Jurnal Gaussian*. 11(1): 99-107.
- Strivastava. N, Hinton. G, Krizhevsky. A, Sutskever, dan Salakhutdinov. R. 2014. Dropout: A Simple Way to Prevent Neural Networks Overfitting. *Journal of Machine Learning Research*. 15(56): 1929-1934.
- Sumaryanto. 2009. Analisis Volatilitas Harga Eceran Beberapa Komoditas Pangan Utama dengan MODEL ARCH/GARCH. *Jurnal Agro Ekonomi*. 27 (2):135-16.
- Sze, V. Chen, Y.H, dan Yang T.J. 2017, Efficient Processing of Deep Neural Networks: A Tutorial and Survey. Tersedia di: <http://arxiv.org/abs/1703.09039> Diakses: 10 November 2022.
- Tian, C., Ma, J., Zhang, C., dan Zhan, P. 2018. A Deep Neural Network Model for Short-Term Load Forecast Based on Long-Short Term Memory Network and Convolution Neural Network. *Energies Journal*. 11: 1-13.
- Vijayalakshmi, V., dan Venkatachalapathy, K. 2019. Deep Neural Network for Multi-Class Prediction of Student Performance in Educational Data. *International Journal of Recent Technology and Engineering (IJRTE)*. 8(2): 5073-5081.

- Vercellis, C. 2009. *Business Intelligence: Data Mining and Optimization for Decision Making*. John Wiley & Sons. United Kingdom.
- Wang, C., dan Gao, Q. 2018. High and Low Prices Prediction of Soybean Futures with LSTM Neural Network. *IEEE 9<sup>th</sup> International Conferences on Software Engineering and Service Sciences (ICSESS)*, Beijing.
- Wei, W.W.S. 2006. *Time series Analysis: Univariate and Multivariate Methods*. 2nd ed. Pearson Education Hall. New Jersey.
- Wiranda, L., dan Sadikin. M. 2019. Penerapan Long Short Term Memory Pada Data Time Series Untuk Memprediksi Penjualan Produk PT. Metiska Farma. *Jurnal Nasional Pendidikan Teknik Informatika*. **8**(3): 184-196.
- Yanti, N.P.L.P., Tuningrat, I.A.M., dan Wiranatha, A.A.P.A.S. 2016. Analisis Peramalan Penjualan Produksi Kecap Pada Perusahaan Kecap Manalagi Denpasar Bali. *Jurnal Rekayasa dan Manajemen Agroindustri*. **4**(1): 72-81.