

Prediksi Nilai Ekspor Pulp di Indonesia Menggunakan Metode *Long Short Term Memory*

Dinda Tamara Silaen¹⁾, Aldowad Alles Sandro Hamonangan Simanjuntak²⁾,
Kurniawan Tarigan³⁾, Evta Indra^{4)*}

¹⁾²⁾³⁾⁴⁾ Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Prima Indonesia

^{*)}Correspondence Author: evtaindra@unprimdn.ac.id, Medan, Indonesia

DOI: <https://doi.org/10.37012/jtik.v9i2.1599>

Abstrak

Ekspor pulp merupakan kegiatan ekonomi penting bagi perusahaan dan pemerintah, yang membutuhkan informasi akurat mengenai permintaan pasar dan strategi bisnis yang tepat. Metode *Long Short Term Memory (LSTM)* digunakan untuk memprediksi hasil produksi pulp di masa depan dengan memanfaatkan data historis dan faktor-faktor yang mempengaruhinya. Tahapan penelitian meliputi studi literatur untuk memahami metode *LSTM*, pengumpulan data, seleksi dan transformasi data untuk mempersiapkan dataset yang akan digunakan, serta visualisasi data ekspor untuk mendapatkan wawasan yang lebih baik. Selanjutnya, metode *LSTM* diterapkan dengan langkah-langkah pembentukan model, pelatihan model, prediksi nilai ekspor, dan evaluasi hasil prediksi. Hasil penelitian ini untuk nilai *Root Mean Squared Error (RMSE)* terhadap produk *Jumbo Roll Tissue, Napkin Tissue, Multi Purpose Tissue, dan Facial Tissue* berturut-turut adalah 2.52, 1.88, 2.77, dan 2.67. Semakin kecil nilai *RMSE*, semakin baik performa model. Nilai *RMSE* yang kecil pada setiap produk menunjukkan bahwa model memiliki performa yang baik dalam memprediksi semua produk. Dengan prediksi yang lebih akurat, perencanaan produksi dan persediaan bahan baku dapat dilakukan dengan lebih efisien dan efektif, sehingga mengoptimalkan produktivitas dan mengurangi biaya produksi. Hasil penelitian ini diharapkan dapat membantu masyarakat dan pemerintah dalam pengambilan keputusan strategis terkait produksi pulp dan kebijakan ekspor. Selain itu, penelitian ini juga diharapkan dapat meningkatkan efisiensi pemasaran dan distribusi produk, meningkatkan pengetahuan tentang pasar dan produk, serta membuka peluang pasar baru untuk produk pulp.

Kata Kunci: Prediksi nilai ekspor, Pulp, *Long Short Term Memory (LSTM)*

Abstract

Pulp exports are an important economic activity for companies and governments, which require accurate information about market demand and appropriate business strategies. The Long Short Term Memory (LSTM) method is used to predict future pulp production results by utilizing historical data and the factors that influence it. The research stages include literature study to understand the LSTM method, data collection, data selection and transformation to prepare the dataset to be used, as well as export data visualization to gain better insight. Next, the LSTM method is applied with the steps of model formation, model training, export value prediction, and evaluation of prediction results. The results of this research for the Root Mean Squared Error (RMSE) values for Jumbo Roll Tissue, Napkin Tissue, Multi Purpose Tissue, and Facial Tissue products were 2.52, 1.88, 2.77, and 2.67, respectively. The smaller the RMSE value, the better the model performance. A small RMSE value for each product indicates that the model has good performance in predicting all products. With more accurate predictions, production planning and raw material supplies can be carried out more efficiently and effectively, thereby optimizing productivity and reducing production costs. It is hoped that the results of this research can help society and the government in making strategic decisions regarding pulp production and export policies. Apart from that, this research is also expected to increase product marketing and distribution efficiency, increase knowledge about markets and products, and open up new market opportunities for pulp products.

Keywords: Prediction of export value, Pulp, *Long Short Term Memory (LSTM)*

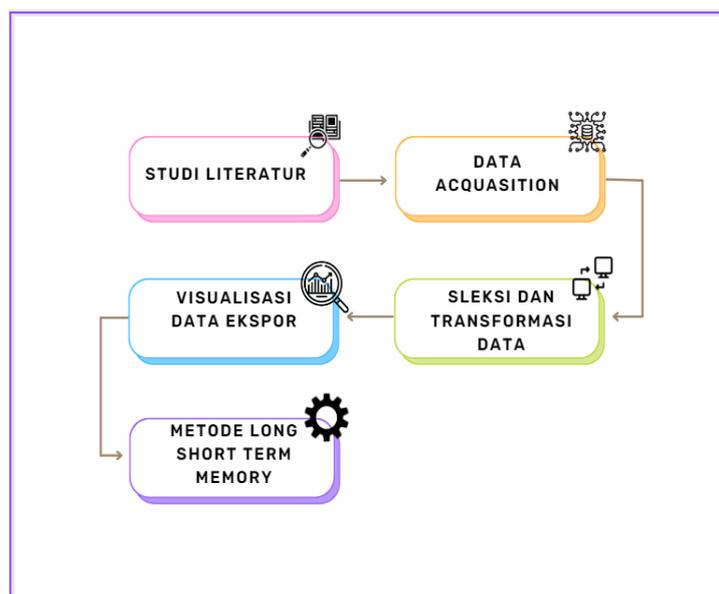
PENDAHULUAN

Ekspor merupakan suatu kegiatan ekonomi yang melibatkan penjualan produk-produk yang berasal dari dalam negeri ke pasar-pasar di luar negeri (Julian, 2021). Sementara itu, prediksi (*forecasting*) adalah suatu seni dan ilmu yang digunakan untuk mengestimasi atau memperkirakan kejadian-kejadian yang akan terjadi di masa depan (Amrustian, Afrizal, Widayat, & Wirawan, 2022). Untuk melakukan prediksi ini, biasanya melibatkan pengumpulan data historis yang relevan dan menganalisisnya untuk memproyeksikan ke masa yang akan datang (Wiko Putra, Albar, and Irmawati 2019). Metode yang digunakan dapat berupa pembentukan model matematis yang mempertimbangkan aspek-aspek penting, atau juga menggunakan pendekatan prediksi yang bersifat subyektif berdasarkan intuisi yang baik dan pengalaman peneliti (Alfandi & Hatuaon Sihite, 2022). Telah banyak penelitian mengenai prediksi suatu produk, contohnya, dataset yang digunakan antara tahun 2017 sampai dengan 2019, dengan hasil perhitungan *Root Mean Squared Error* (RMSE) dalam bentuk rupiah sebesar 13,762,154 dan hasil menggunakan *Mean Absolute Percentage Error* (MAPE) adalah 12% (Lattifia, Wira Buana, and Rusjyanthi 2022).

Di masa kini, hampir semua perusahaan yang bergerak di sektor industri ekspor menghadapi tantangan yang semakin ketat dalam persaingan. Oleh karena itu, perusahaan-perusahaan tersebut perlu melakukan perencanaan yang matang terkait semua parameter produksi, termasuk kapasitas produksi, agar dapat memenuhi permintaan pasar dengan tepat waktu dan dalam jumlah yang sesuai (Aisyah, Diah, Purboyo, & Kallista, 2023). Melalui prediksi nilai ekspor produk, baik perusahaan maupun pemerintah dapat memperoleh pemahaman tentang prospek di masa depan dan menjadi lebih siap dalam menghadapi tantangan pasar (Akbar, Rahmatul, Santoso, & Warsito, 2023). Hal ini memungkinkan pengambilan keputusan strategis yang tepat dalam menetapkan target pasar yang sesuai dan efisien (Sabar Sautomo and Hilman Ferdinandus Pardede 2021). Oleh karena itu, sangat penting untuk memahami kebutuhan pasar berdasarkan hasil prediksi nilai ekspor produk, sehingga strategi bisnis yang diambil dapat membidik dengan akurat dan tepat sasaran (Aprian, Aji, Azhar, & Nastiti, 2020). Dari kegiatan ekspor pulp, *stakeholder* dapat memperkirakan permintaan dan pasokan produk, sehingga dapat membuat keputusan produksi dan strategi pasar yang tepat untuk meningkatkan pertumbuhan ekonomi.

METODE

Penelitian adalah sebuah studi atau observasi yang bertujuan untuk menemukan jawaban atas suatu permasalahan dengan menggunakan pendekatan ilmiah dan mengumpulkan fakta-fakta. Dalam penelitian ini, peneliti menggunakan metode *Long Short Term Memory* (LSTM). LSTM, sebuah varian dari Jaringan Saraf Rekuren (RNN), telah dimodifikasi dengan menggabungkan komponen memori (memory cell) yang memungkinkan penyimpanan informasi dalam jangka waktu yang lebih lama (Wiranda and Sadikin 2019).



Gambar 1. Proses Penelitian

Sistem akuisisi data adalah suatu sistem yang bertugas mengambil, mengumpulkan, dan menyusun data, serta memprosesnya agar menghasilkan data yang diinginkan (Husni, 2022). Dataset pada penelitian ini diambil dari PT. Pulp Papi yang merupakan perusahaan yang memproduksi berbagai olahan dari bahan mentah pulp, menyediakan data ekspor pulp mengenai nilai ekspor, jumlah, negara penerima, dan lain sebagainya. Pengambilan dataset ini untuk mengumpulkan informasi atau bertujuan untuk menganalisa pertumbuhan produksi PT. Pulp Papi di tahun mendatang. Dataset ini terdiri 5055 baris dan 14 kolom, tetapi yang akan diolah pada penelitian ini sebanyak 7 kolom. Dataset yang digunakan adalah dataset dalam kurun waktu 1 tahun, yaitu 2022. Dataset dapat dilihat pada tabel 1 di bawah ini.

Tabel 1. Dataset PT. Pulp Papi

| DATA SET EKSPOR PT.PULP | | | | | | | | | | | | | |
|-------------------------|-----------|-----------|-------------|---|----------|----------------|---------------|-----------|--------------|---------------|----------------|-------------------|-----------------|
| Nomor PEB | Tahun PEB | Bulan PEB | Tanggal PEB | Uraian Barang | Hs Code | Jumlah Kms Brg | Jenis Kms Brg | Netto Brg | Jenis Satuan | Jml Container | Pelabuhan Muat | Pelabuhan Bongkar | Negara Penerima |
| 1 | 2022 | January | 2 | NAPKIN TISSUE BASIS WEIGHT 18 GSM +/- 0.5 NAPKIN TISSUE BASIS WEIGHT 20 GSM +/-0.5 | 48030090 | 136 | RO | 16.885 | TNE | 1 | IDPWG | TWKHH | TW |
| 2 | 2022 | January | 5 | JUMBO TOILET TISSUE | 48030090 | 85 | RO | 71.92 | TNE | 5 | IDPWG | TTPOS | TT |
| 3 | 2022 | January | 5 | TISSUE FINISHED GOODS | 48181000 | 4.39 | BL | 27.42 | TNE | 3 | IDPWG | AUBNE | AU |
| 4 | 2022 | January | 5 | TISSUE PRODUCT | 48030090 | 908 | RO | 239.093 | TNE | 16 | IDPWG | GRSKG | GR |
| 5 | 2022 | January | 5 | TISSUE PRODUCT | 48030090 | 221 | RO | 198.025 | TNE | 13 | IDPWG | GRPIR | GR |
| 6 | 2022 | January | 5 | TISSUE PRODUCT | 48030090 | 290 | RO | 52.588 | TNE | 3 | IDPWG | INMAA | IN |
| 7 | 2022 | January | 5 | TISSUE FINISHED PRODUCT | 48182000 | 1.362 | CT | 11.728 | TNE | 1 | IDPWG | USMA | US |
| 8 | 2022 | January | 5 | JUMBO HAND TOWEL TISSUE PP ORIGINAL 000 | 48030090 | 55 | RO | 76.833 | TNE | 4 | IDPWG | USRLO | US |
| 9 | 2022 | January | 5 | JUMBO TOILET TISSUE PP ORIGINAL 000 | 48030090 | 54 | RO | 95.065 | TNE | 9 | IDPWG | USRLO | US |
| 10 | 2022 | January | 5 | JUMBO TOILET TISSUE PP ORIGINAL 000 | 48030090 | 54 | RO | 96.963 | TNE | 9 | IDPWG | USRLO | US |
| 11 | 2022 | January | 5 | JUMBO TOILET TISSUE PP ORIGINAL 000 | 48030090 | 30 | RO | 51.685 | TNE | 5 | IDPWG | USRLO | US |
| 12 | 2022 | January | 5 | JUMBO TOILET TISSUE PP ORIGINAL 000 | 48030090 | 36 | RO | 65.28 | TNE | 6 | IDPWG | USRLO | US |
| 13 | 2022 | January | 5 | VIRGIN PULP KTCHEN TOWEL TISSUEJUMBO ROLL | 48030090 | 32 | RO | 32.462 | TNE | 2 | IDPWG | MYPKG | MY |
| 14 | 2022 | January | 5 | NAPKIN TISSUE BASIS WEIGHT 18 GSM +/- 0.5 NAPKIN TISSUE BASIS WEIGHT 20 GSM +/-0.5 | 48030090 | 124 | RO | 17.869 | TNE | 1 | IDPWG | TWTXG | TW |
| 15 | 2022 | January | 5 | JUMBO TOILET TISSUE PP ORIGINAL 000 | 48030090 | 44 | RO | 99.607 | TNE | 11 | IDPWG | USFRR | US |

Seleksi data bertujuan untuk menghasilkan himpunan data target. Transformasi data untuk mengubah data kedalam bentuk yang sesuai agar dapat diproses dengan baik (Owen et al. 2022). Pada tahapan ini, peneliti melakukan seleksi dengan memilih dan mengelompokkan produk yang diekspor ke 6 kategori komoditas ekspor, yaitu hand towel tissue, facial tissue, kitchen tissue, toilet tissue, napkin tissue, dan bath tissue.

Visualisasi data adalah proses pengolahan data yang mengubahnya menjadi bentuk yang lebih mudah dipahami, seringkali dalam bentuk diagram seperti Line Chart dan Histogram (Informatika, 2023). Tujuan dari visualisasi data ini adalah untuk memberikan gambaran yang jelas dan intuitif tentang pola, tren, dan distribusi data sehingga memudahkan pemahaman dan analisis data secara visual (Saragih, Saragih, and Wanto 2018).

Long Short Term Memory adalah sistem penyimpanan data yang memiliki kemampuan untuk memprediksi dan memproses informasi dalam jangka waktu yang panjang (Mukhlis, Kustiyo, and Suharso 2021). *Long Short Term Memory* (LSTM) merupakan modifikasi dari *Recurrent Neural Network* (RNN) yang dirancang untuk melakukan prediksi yang akurat terhadap suatu variabel (Aulia, Alwi, 2021). Kemampuan

LSTM dalam melakukan peramalan didasarkan pada tingkat kesalahan prediksi, dimana semakin kecil tingkat kesalahannya, semakin akurat pula hasil prediksinya.

Berikut adalah langkah-langkah metode *Long Short Term Memory*:

1. Persiapan Data: Kumpulkan dan siapkan data beserta faktor-faktor yang mempengaruhi ekspor.
2. Preprocessing Data: Preprocessing data termasuk normalisasi data, memisahkan data latih dan data uji, dan mengubah data menjadi format tensor yang dapat diterima oleh jaringan saraf tiruan(Maygirtasari, Yulianto, and Mawardi 2015).
3. Pembentukan Model: Buat arsitektur model LSTM dan kompilasikan model dengan metode optimasi dan *loss function* yang sesuai (Ivan Dwi Nugraha, 2022).
4. Latih Model: Latih model dengan data latih dan evaluasi model dengan data uji untuk menentukan akurasi model(Laksamana, Utami, and Fatta 2021).
5. *Hyperparameter Tuning*: Sesuaikan *hyper parameter model* seperti jumlah layer, jumlah neuron, *dropout rate*, dll untuk meningkatkan akurasi model(Sanjaya and Heksaputra 2020).
6. Prediksi: Gunakan model yang terlatih untuk memprediksi nilai ekspor produk di masa depan dengan menggunakan data masa lalu dan faktor-faktor yang mempengaruhi ekspor(Rumapea 2020).
7. Evaluasi Hasil: Evaluasi hasil prediksi dan bandingkan dengan data aktual untuk menentukan tingkat akurasi model(Karno 2020).

HASIL DAN PEMBAHASAN

Untuk identifikasi permasalahan yang ada, masalah yang dihadapi adalah kesulitan dalam menentukan jumlah stok barang yang harus disiapkan untuk keperluan ekspor, terutama ketika terdapat banyak jenis barang yang perlu dipertimbangkan. Penyediaan stok barang yang tidak memadai dapat mengakibatkan ketidaktersediaan produk saat penjualan, sementara terlalu banyak stok pada jenis barang tertentu dapat menyebabkan penumpukan yang tidak efisien. Keterbatasan dalam pengolahan data secara manual menghambat pengecekan stok barang, yang berpotensi mempengaruhi kualitas perusahaan. Oleh karena itu, rumusan masalah dalam penelitian ini adalah bagaimana menerapkan metode *Long Short*

Term Memory (LSTM) dalam sistem untuk memperkirakan jumlah produk di masa depan, sehingga dapat digunakan sebagai pedoman dalam pengadaan stok bahan produksi.

Data yang bisa diberikan oleh PT. Pulp Papi untuk penelitian, yakni data tahun 2022. Diperoleh data penjualan dan pembelian antara tanggal 1 Januari 2022 hingga 31 Desember 2022. Data tersebut pada awalnya masih berbentuk catatan ekspor pulp. Data tersebut kemudian dimasukkan ke dalam file penjualan dan kemudian dipindahkan ke dalam spreadsheet.

Pada tahap pengolahan data, dibahas tentang tahapan penelitian mulai dari Persiapan data, Preprocessing data, Pembentukan model, Melatih data dan melakukan Prediksi serta Memvisualisasikan data untuk mendapatkan hasil prediksi yang baik. Peneliti menggunakan bahasa pemrograman Python melalui Google Collab dengan beberapa library seperti Numpy, Pandas dan Keras dari Tensorflow.

Langkah pertama, peneliti melakukan persiapan untuk mengolah dataset Pulp berupa angka yang akan divisualisasikan dalam bentuk grafik. Dalam tahap ini, peneliti menggunakan Platform Google Collab untuk menjalankan Program Python dan mengimport library yang digunakan.

```
✓ [23] 0s import pandas as pd
import re
import string
import math
import numpy as np
import matplotlib.pyplot as plt

from keras.models import Sequential
from keras.layers import Dense
from keras.layers import LSTM
from sklearn.preprocessing import MinMaxScaler
from sklearn.metrics import mean_squared_error
```

Gambar 2. Potongan Kode Python Untuk Import Library

Dataset dalam bentuk excel atau csv kemudian dimasukkan kedalam google collab dan kemudian tampilkan dataset.

```
[24] from google.colab import files
      uploaded = files.upload()
```

Choose Files DatasetPulp.xlsx
• DatasetPulp.xlsx(application/vnd.openxmlformats-officedocument.spreadsheetml.sheet) - 254215 bytes, last modified: 3/8/2023 - 100% done
Saving DatasetPulp.xlsx to DatasetPulp.xlsx

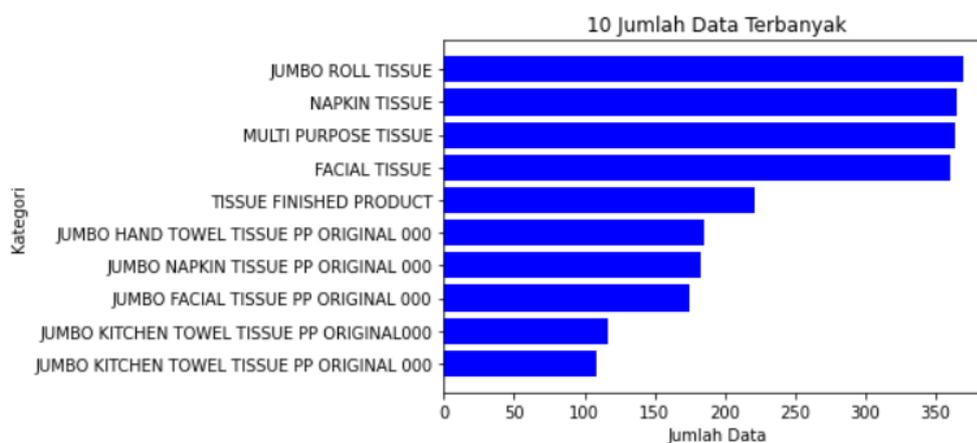
```
[25] df=pd.read_excel('DatasetPulp.xlsx')
      df
```

| | Nomor PEB | Tanggal | Barang | Kode Barang | Jumlah Kms Brg | Jenis Kms Brg | Produksi | Jenis Satuan | Jml Container | Pelabuhan Muat | Pelabuhan Bongkar | Negara Penerima |
|------|-----------|------------|---|-------------|----------------|---------------|----------|--------------|---------------|----------------|-------------------|-----------------|
| 0 | 1 | 2022-01-01 | NAPKIN TISSUE BASIS WEIGHT 18 GSM +/-0.5NAPKIN... | 48183000 | 136.00 | RO | 16.885 | TNE | 1 | IDPWG | TWKHH | TW |
| 1 | 2 | 2022-01-01 | NAPKIN TISSUE | 48183000 | 908.00 | RO | 155.093 | TNE | 16 | IDPWG | GRSKG | GR |
| 2 | 3 | 2022-01-01 | JUMBO ROLL TISSUE | 48183000 | 484.00 | RO | 156.500 | TNE | 21 | IDPWG | KRPUS | KR |
| 3 | 4 | 2022-01-01 | FACIAL TISSUE | 48181000 | 4.39 | BL | 160.148 | TNE | 3 | IDPWG | AUBNE | AU |
| 4 | 5 | 2022-01-01 | MULTI PURPOSE TISSUE | 48183000 | 244.00 | RO | 153.148 | TNE | 3 | IDPWG | VNCLI | VN |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 3776 | 3777 | 2022-12-31 | JUMBO KITCHEN TOWEL TISSUE PP ORIGINAL000 | 48183000 | 54.00 | RO | 150.035 | TNE | 3 | IDPWG | USLGB | US |

Gambar 3. Tampilan Dataset di Google Colab

Kode tersebut merupakan bagian dari program Python yang berjalan di Google Colaboratory. Kode ini digunakan untuk mengimpor dataset dari komputer lokal ke Google. Modul "files" diimpor dari pustaka "google.colab". Modul ini digunakan untuk mengunggah file dari komputer lokal ke Google Colab.

```
[26] df = pd.read_excel('DatasetPulp.xlsx')
      count_df = pd.DataFrame({'count': df.groupby(['Barang']).size()}).reset_index()
      top_10_counts = count_df.sort_values('count', ascending=False).head(10)
      plt.barh(top_10_counts['Barang'], top_10_counts['count'], color='blue')
      plt.xlabel('Jumlah Data')
      plt.ylabel('Kategori')
      plt.title('10 Jumlah Data Terbanyak')
      plt.gca().invert_yaxis()
      plt.show()
```



Gambar 4. Tampilan Jumlah Produksi terbanyak

Dengan menggunakan metode `.query` (`Barang == "JUMBO ROLL TISSUE"`), menghasilkan DataFrame baru yang hanya berisi baris-baris dimana nilai kolom 'Barang' sama dengan "JUMBO ROLL TISSUE". Dalam Penelitian ini, akan diolah 4 jenis produk, yaitu Jumbo Roll Tissue, Napkin Tissue, Multi Purpose Tissue, dan Facial Tissue.

```
[28] df=df.query('Barang == "JUMBO ROLL TISSUE"')
df
```

| | Nomor PEB | Tanggal | Barang | Kode Barang | Jumlah Kms Brg | Jenis Kms Brg | Produksi | Jenis Satuan | Jml container | Pelabuhan Muat | Pelabuhan Bongkar | Negara Penerima |
|------|-----------|------------|-------------------|-------------|----------------|---------------|----------|--------------|---------------|----------------|-------------------|-----------------|
| 2 | 3 | 2022-01-01 | JUMBO ROLL TISSUE | 48183000 | 484.0 | RO | 156.500 | TNE | 21 | IDPVG | KRPUS | KR |
| 6 | 7 | 2022-01-02 | JUMBO ROLL TISSUE | 48183000 | 320.0 | RO | 155.500 | TNE | 8 | IDPVG | KRICH | KR |
| 10 | 11 | 2022-01-03 | JUMBO ROLL TISSUE | 48183000 | 480.0 | RO | 156.500 | TNE | 12 | IDPVG | KRICH | KR |
| 15 | 16 | 2022-01-04 | JUMBO ROLL TISSUE | 48183000 | 320.0 | RO | 157.500 | TNE | 8 | IDPVG | KRICH | KR |
| 33 | 34 | 2022-01-05 | JUMBO ROLL TISSUE | 48183000 | 120.0 | RO | 156.500 | TNE | 3 | IDPVG | KRICH | KR |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 3696 | 3697 | 2022-12-27 | JUMBO ROLL TISSUE | 48183000 | 120.0 | RO | 169.626 | TNE | 7 | IDPVG | KRICH | KR |
| 3702 | 3703 | 2022-12-28 | JUMBO ROLL TISSUE | 48183000 | 300.0 | RO | 170.097 | TNE | 5 | IDPVG | KRICH | KR |
| 3718 | 3719 | 2022-12-29 | JUMBO ROLL TISSUE | 48183000 | 240.0 | RO | 170.783 | TNE | 2 | IDPVG | KRICH | KR |
| 3728 | 3729 | 2022-12-30 | JUMBO ROLL TISSUE | 48183000 | 200.0 | RO | 171.809 | TNE | 5 | IDPVG | KRPUS | KR |

Gambar 5. Tampilan produk yang akan diolah

Proses selanjutnya adalah normalisasi data, mengubah tipe dataframe agar bisa diolah.

```
[30] df['Tanggal'] = pd.to_datetime(df.Tanggal)
df['Produksi'] = df['Produksi'].astype(float)

<ipython-input-30-61c2c543ce37>:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
df['Tanggal'] = pd.to_datetime(df.Tanggal)
<ipython-input-30-61c2c543ce37>:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
df['Produksi'] = df['Produksi'].astype(float)

[31] df.dtypes

Tanggal    datetime64[ns]
Produksi   float64
dtype: object
```

Gambar 6. Tampilan potongan Kode Python untuk Import Library

Kode `df['Tanggal']= pd.to_datetime(df.Tanggal)`, digunakan untuk mengubah tipe data kolom 'Tanggal' pada dataframe 'df' menjadi datetime. Dalam banyak kasus, kolom tanggal atau waktu pada dataframe memiliki tipe data string atau objek yang harus diubah menjadi datetime agar dapat diolah lebih lanjut. Baris kode kedua `df['Produksi'] = df['Produksi'].astype(float)` digunakan untuk mengubah tipe data kolom 'Produksi' pada dataframe 'df' menjadi float. Tipe data float digunakan untuk merepresentasikan bilangan pecahan pada python.

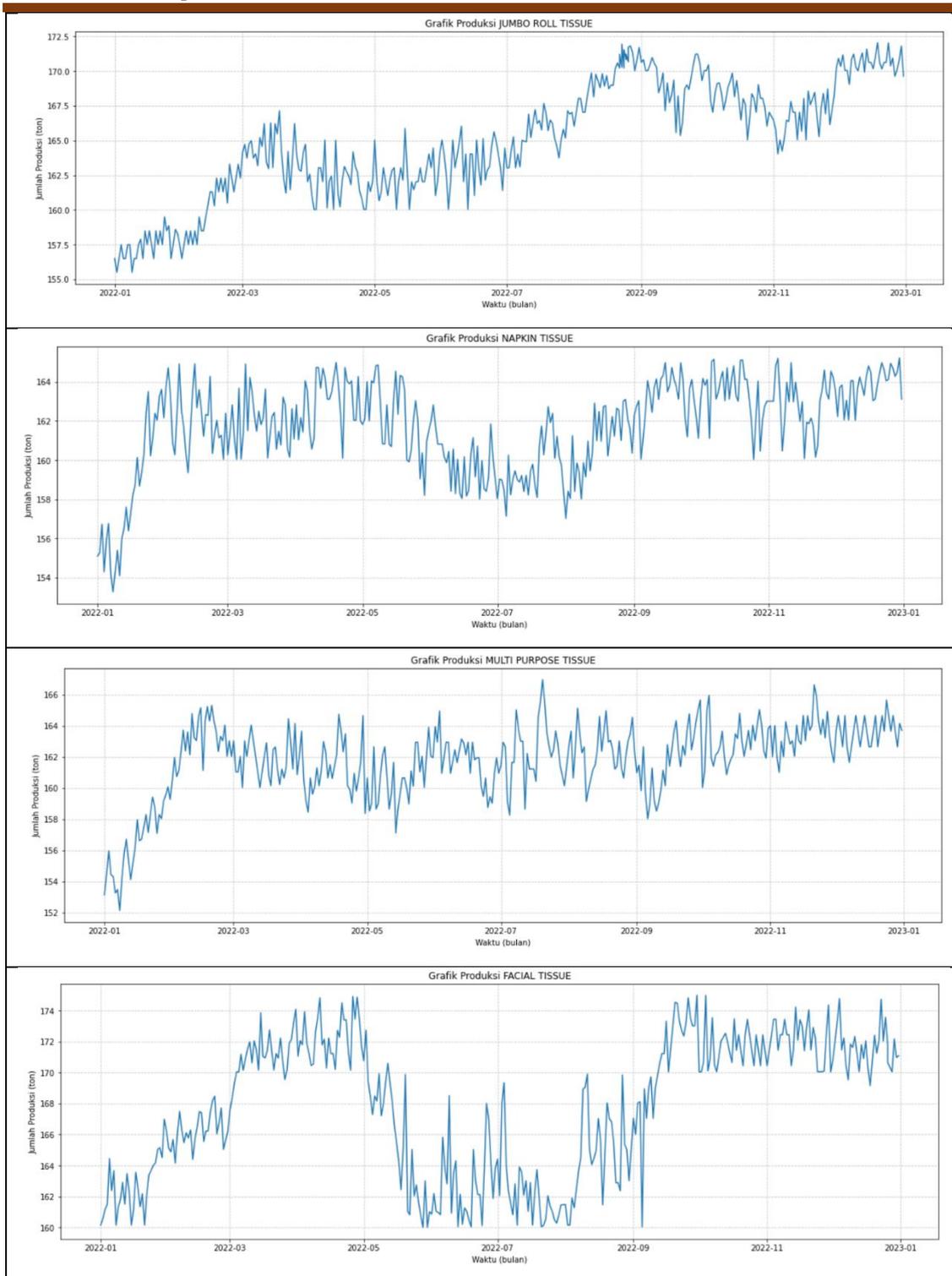
```
✓ [33] fig, ax = plt.subplots(figsize=(20, 6))  
0s ax.plot(df.index, df['Produksi'])  
ax.set_xlabel('Waktu (bulan)')  
ax.set_ylabel('Jumlah Produksi (ton)')  
ax.set_title('Grafik Produksi JUMBO ROLL TISSUE')  
ax.grid(linestyle='--', alpha=0.7)  
plt.show()
```

Gambar 7. Potongan Kode Python untuk menampilkan dataset kedalam grafik

Codingan tersebut menggunakan library matplotlib untuk membuat sebuah grafik garis (line plot) dari data produksi pada DataFrame df dengan sumbu x (ax.plot(df.index, df['Produksi'])) menggunakan indeks DataFrame sebagai nilai x dan kolom 'Produksi' sebagai nilai y. Selanjutnya, ax.set_xlabel ('Waktu') dan ax.set_ylabel ('Jumlah Produksi') digunakan untuk menetapkan label pada sumbu x dan sumbu y, sedangkan ax.set_title ('Grafik Produksi NAPKIN TISSUE') digunakan untuk menetapkan judul grafik. Visualisasi data set dalam LSTM (*Long Short-Term Memory*) adalah proses menampilkan data sesuai urutan waktu pada plot atau grafik untuk membantu memahami dan menganalisis data dengan lebih baik. Visualisasi data dapat memberikan wawasan yang berguna tentang tren, pola, anomali, atau kecenderungan dalam data urutan waktu.

Berikut merupakan hasil visualisasi dari dataset, grafik tersebut menjelaskan jumlah produksi produk Jumbo Roll Tissue, Napkin Tissue, Multi Purpose Tissue, dan Facial Tissue per Januari 2022 - Desember 2022.

Preprocessing Data pada LSTM atau *Long Short-Term Memory* adalah serangkaian tahapan persiapan data sebelum data tersebut dimasukkan ke dalam model LSTM. Selanjutnya dilakukan pengembangan dengan menambahkan Layer RNN menggunakan Simple RNN yang dilengkapi dengan blok memori. Selanjutnya disertakan lapisan dense untuk meningkatkan kompleksitas jaringan saraf (NN) yang digunakan. Diterapkan *dropout* untuk mengatasi masalah *overfitting* yang mungkin terjadi selama pelatihan model. Selanjutnya dihasilkan keluaran data yang direpresentasikan dalam bentuk nilai biner, yaitu 0 dan 1. Nilai-nilai ini mewakili probabilitas bahwa data tersebut tergolong dalam kategori normal atau depresi. Label ini kami gunakan untuk mengklasifikasikan kalimat dalam pemrosesan data yang penting untuk membantu model LSTM mempelajari pola dan melakukan prediksi data dengan tingkat akurasi yang tinggi.



Gambar 8. Tampilan grafik dataset produk yang akan diolah

Kemudian dataset tersebut *discaling* menggunakan `MinMaxScaler()` dari modul `sklearn.preprocessing`. Scaling pada data bertujuan untuk menormalkan rentang nilai pada

setiap fitur (kolom) dalam dataset sehingga nilainya memiliki skala yang seragam. Proses scaling ini dapat membantu meningkatkan performa model *machine learning Long Short Term Memory*. Setelah itu, menentukan *train size* sebesar 75 persen dari dataset, dan sisanya 25 persen sebagai *test size*.

```
[34] df.shape
(370, 1)

[35] scaler = MinMaxScaler(feature_range=(0, 1))
df = scaler.fit_transform(df)

[36] train_size = int(len(df) * 0.75)
test_size = len(df) - train_size
train = df[0:train_size,:]
test = df[train_size:len(df),:]
print("train size: {}, test size: {}".format(len(train), len(test)))

train size: 277, test size: 93
```

Gambar 9. Potongan Kode Python untuk membagi data train dengan test

Dalam mengolah data menggunakan LSTM, timestamp adalah informasi waktu yang terkait dengan setiap data dalam urutan waktu. Ini memungkinkan jaringan LSTM untuk melacak perubahan dalam data dari waktu ke waktu, sehingga dapat mempelajari pola dan tren dalam data tersebut. Setelah itu, *trainX* dan *testX* masing-masing diubah dimensinya menjadi (jumlah baris data latih/test, 1, jumlah timestep) menggunakan `np.reshape()`. Hal ini dilakukan karena model LSTM yang akan dibuat membutuhkan input data berupa array 3 dimensi dengan bentuk (jumlah baris, 1, jumlah timestep), yaitu berisi jumlah sampel (baris data), jumlah fitur (dalam hal ini hanya satu), dan jumlah timestep (dalam hal ini 5).

```
[37] time_stemp = 1
dataX = []
dataY = []
for i in range(len(train)-time_stemp-1):
    a = train[i:(i+time_stemp), 0]
    dataX.append(a)
    dataY.append(train[i + time_stemp, 0])
trainX = np.array(dataX)
trainY = np.array(dataY)

dataX = []
dataY = []
for i in range(len(test)-time_stemp-1):
    a = test[i:(i+time_stemp), 0]
    dataX.append(a)
    dataY.append(test[i + time_stemp, 0])
testX = np.array(dataX)
testY = np.array(dataY)

[39] trainX = np.reshape(trainX, (trainX.shape[0], 1, trainX.shape[1]))
testX = np.reshape(testX, (testX.shape[0], 1, testX.shape[1]))
```

Gambar 10. Potongan Kode Python untuk Persiapan data

Dengan cara ini, dimensi dari array input data menjadi sesuai dengan format yang dibutuhkan oleh model LSTM yang akan digunakan. Dalam hal ini, *trainX* dan *testX* akan digunakan sebagai input pada saat melakukan pelatihan dan pengujian model.

Proses Train data set menggunakan model LSTM dilakukan dengan mengkompilasi model menggunakan fungsi `compile()`, yang diberi parameter `loss='mean_squared_error'` untuk menentukan jenis loss function yang digunakan, `optimizer='adam'` untuk menentukan jenis optimizer yang digunakan (dalam hal ini Adam optimizer), dan `metrics=['mean_squared_error']` untuk menentukan metrik evaluasi model yang digunakan. Selanjutnya, model dilatih menggunakan fungsi `fit()`. Pada fungsi ini, parameter `trainX` dan `trainY` digunakan untuk menyediakan input dan output data yang akan digunakan pada pelatihan model. Parameter `epochs=50` menunjukkan bahwa model akan dilatih sebanyak 50 epoch, sedangkan `batch_size=1` menunjukkan bahwa model akan mengambil satu sampel data pada setiap iterasinya. Proses pelatihan model akan menghasilkan model yang dapat digunakan untuk memprediksi data pada waktu berikutnya.

```
[18] model = Sequential()
      model.add(LSTM(10, input_shape=(1, time_step)))
      model.add(Dense(1))
      model.compile(loss='mean_squared_error', optimizer='adam', metrics=['mean_squared_error'])
      model.fit(trainX, trainY, epochs=50, batch_size=1)

Epoch 1/50
275/275 [=====] - 6s 5ms/step - loss: 0.1126 - mean_squared_error: 0.1126
Epoch 2/50
275/275 [=====] - 1s 5ms/step - loss: 0.0273 - mean_squared_error: 0.0273
Epoch 3/50
275/275 [=====] - 1s 3ms/step - loss: 0.0199 - mean_squared_error: 0.0199
Epoch 4/50
275/275 [=====] - 1s 2ms/step - loss: 0.0146 - mean_squared_error: 0.0146
Epoch 5/50
275/275 [=====] - 1s 2ms/step - loss: 0.0114 - mean_squared_error: 0.0114
Epoch 6/50
275/275 [=====] - 1s 3ms/step - loss: 0.0101 - mean_squared_error: 0.0101
Epoch 7/50
275/275 [=====] - 1s 3ms/step - loss: 0.0093 - mean_squared_error: 0.0093
Epoch 8/50
275/275 [=====] - 1s 3ms/step - loss: 0.0092 - mean_squared_error: 0.0092
Epoch 9/50
275/275 [=====] - 1s 3ms/step - loss: 0.0093 - mean_squared_error: 0.0093
Epoch 10/50
275/275 [=====] - 1s 2ms/step - loss: 0.0093 - mean_squared_error: 0.0093

[20] trainPredict = model.predict(trainX)
      testPredict = model.predict(testX)
      trainPredict = scaler.inverse_transform(trainPredict)
      trainY = scaler.inverse_transform([trainY])
      testPredict = scaler.inverse_transform(testPredict)
      testY = scaler.inverse_transform([testY])
      trainScore = math.sqrt(mean_squared_error(trainY[0], trainPredict[:,0]))
      print('Train Score JUMBO ROLL TISSUE : %.2f RMSE' % (trainScore))
      testScore = math.sqrt(mean_squared_error(testY[0], testPredict[:,0]))
      print('Test Score JUMBO ROLL TISSUE : %.2f RMSE' % (testScore))
```

Gambar 11. Kode Python untuk Melatih model menggunakan model LSTM dan Dense

Setelah melakukan train dengan model LSTM, dapat dievaluasi hasil model yang didapatkan. Pertama-tama, prediksi dilakukan pada data latih menggunakan fungsi `predict()`, yang mengambil input `trainX` sebagai masukan dan menghasilkan prediksi output pada

trainPredict. Selanjutnya, prediksi dilakukan pada data uji menggunakan fungsi yang sama dengan input test X sebagai masukan dan menghasilkan prediksi output pada testPredict. Selanjutnya, nilai prediksi yang telah dinormalisasi (dalam rentang 0 hingga 1) diubah kembali menjadi nilai semula dengan menggunakan fungsi inverse_transform() dari objek scaler. Hal ini dilakukan untuk mengembalikan nilai prediksi ke dalam satuan aslinya. Proses selanjutnya menentukan RMSE Berdasarkan Hasil Data Train Dan Test

```

9/9 [=====] - 1s 2ms/step
3/3 [=====] - 0s 4ms/step
Train Score JUMBO ROLL TISSUE : 1.58 RMSE
Test Score JUMBO ROLL TISSUE : 1.30 RMSE
    
```

Gambar 12. Hasil RMSE produk Jumbo Roll Tissue

Berdasarkan keempat produk yang diuji, nilai RMSE yang rendah menunjukkan bahwa kualitas model *Machine Learning* cukup baik dalam melakukan prediksi nilai produksi pada data latih dan data uji, karena semakin sedikit tingkat *error* hasil prediksi dengan data aktualnya.

Tabel 2. Nilai RMSE Hasil prediksi Produk Tissue

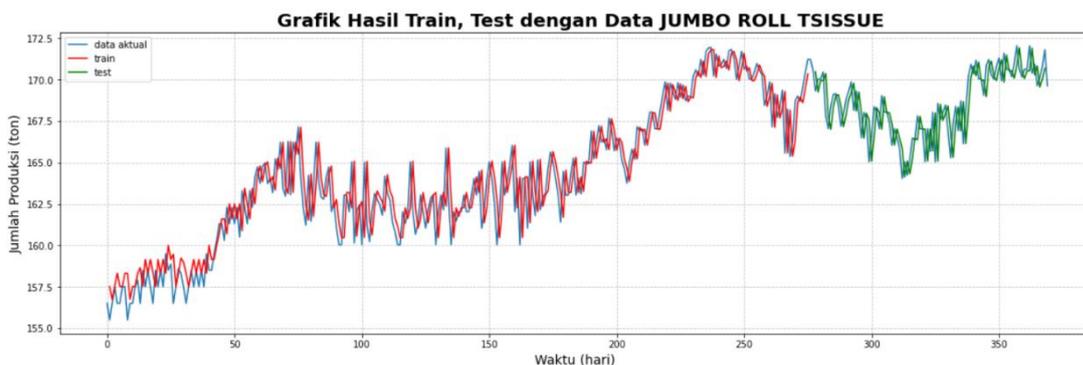
| Jenis Produk | Train Score (RMSE) | Test Score (RMSE) |
|----------------------|--------------------|-------------------|
| Jumbo Roll Tissue | 1.58 | 1.30 |
| Napkin Tissue | 1.47 | 1.29 |
| Multi Purpose Tissue | 1.51 | 1.29 |
| Facial Tissue | 2.09 | 1.94 |

```

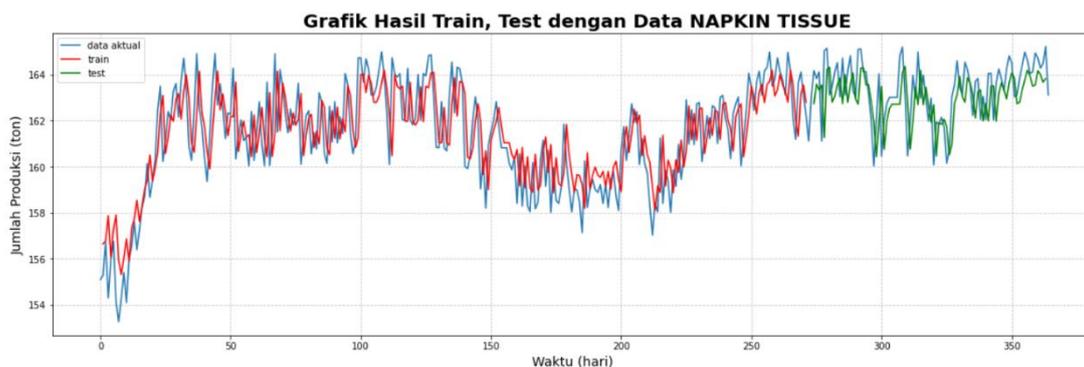
[113] # shifting train
trainPredictPlot = np.empty_like(df)
trainPredictPlot[:, :] = np.nan
trainPredictPlot[time_stemp:len(trainPredict)+time_stemp, :] = trainPredict
# shifting test predictions for plotting
testPredictPlot = np.empty_like(df)
testPredictPlot[:, :] = np.nan
testPredictPlot[len(trainPredict)+(time_stemp*2)+1:len(df)-1, :] = testPredict
# plot baseline and predictions
plt.figure(figsize=(20,6))
plt.plot(scaler.inverse_transform(df),label='data aktual')
plt.plot(trainPredictPlot,color='red',label='train')
plt.plot(testPredictPlot,color='green',label='test')
plt.title(' Grafik Hasil Train, Test dengan Data JUMBO ROLL TSISSUE')
plt.xlabel('Waktu')
plt.ylabel('Jumlah Produksi')
plt.grid(linestyle='--', alpha=0.7)
plt.legend()
plt.show()
    
```

Gambar 13. Potongan Kode Python menampilkan Grafik hasil prediksi

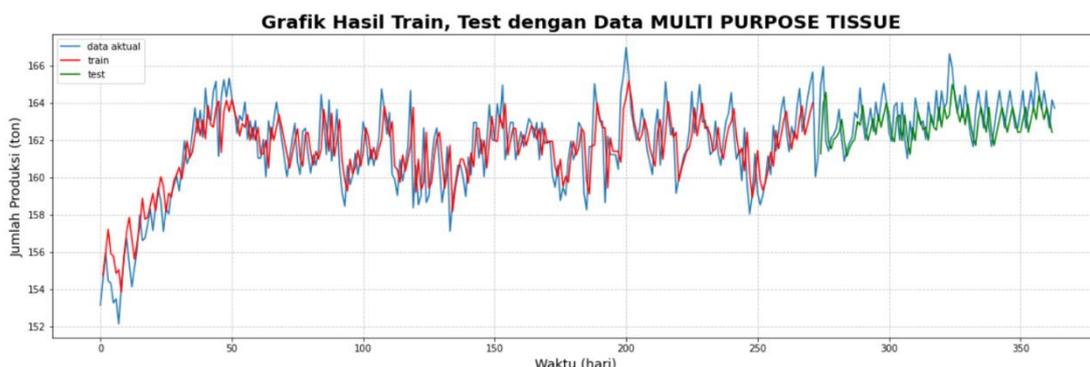
Proses berikutnya memvisualisasikan Grafik Hasil Prediksi Model LSTM Dengan Data Aktual. Berikut merupakan grafik, hasil gabungan dari hasil train, test, dan data real. Grafik ini bertujuan agar mempermudah kita dalam membaca dan menemukan perbedaannya. Hasil train ditandai dengan warna merah, sedangkan hasil test ditandai dengan warna hijau, dan data aktual ditandai dengan warna biru.



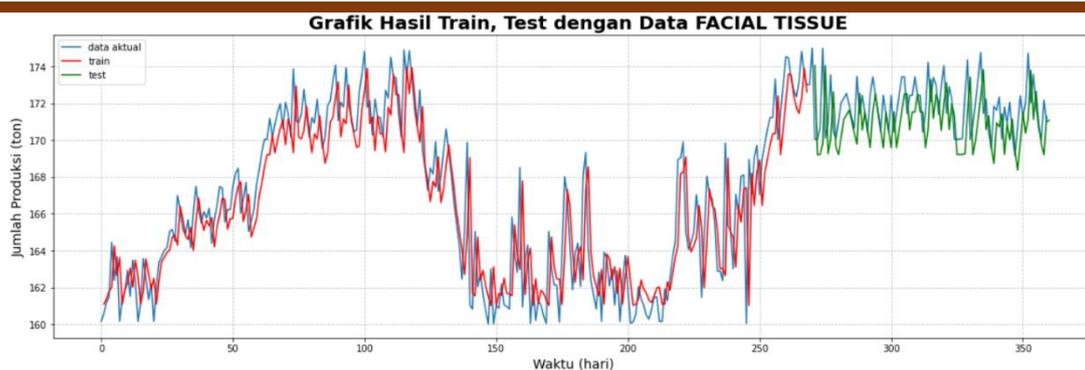
Gambar 14. Hasil Prediksi Produk Jumbo Roll Tissue



Gambar 15. Hasil Prediksi Produk Napkin Tissue



Gambar 16. Hasil Prediksi Produk Multi Purpose Tissue



Gambar 17. Hasil Prediksi Produk Facial Tissue

KESIMPULAN DAN REKOMENDASI

Berdasarkan hasil pelatihan dan evaluasi model *LSTM*, dapat disimpulkan bahwa model tersebut baik dalam memprediksi produksi Pulp menggunakan data historis selama 365 hari. Evaluasi menunjukkan bahwa nilai *RMSE* untuk produk *Jumbo Roll Tissue*, *Napkin Tissue*, *Multi Purpose Tissue*, dan *Facial Tissue* berturut-turut adalah 2.52, 1.88, 2.77, dan 2.67. Semakin kecil nilai *RMSE*, semakin baik performa model. Nilai *RMSE* yang kecil pada setiap produk menunjukkan bahwa model memiliki performa yang baik dalam memprediksi semua produk. Dengan prediksi yang lebih akurat, perencanaan produksi dan persediaan bahan baku dapat dilakukan dengan lebih efisien dan efektif, sehingga mengoptimalkan produktivitas dan mengurangi biaya produksi. Namun, hasil prediksi dari model hanya sebatas perkiraan dan tidak selalu akurat, terutama jika terdapat faktor-faktor eksternal yang tidak dapat diprediksi oleh model seperti perubahan permintaan pasar atau masalah produksi internal. Oleh karena itu, hasil prediksi sebaiknya dijadikan sebagai bahan pertimbangan atau referensi untuk mengetahui pola produksi Pulp.

Setelah melatih model *LSTM* ini, ada beberapa saran yang dapat diberikan: 1) Melakukan evaluasi dan peningkatan model: Meskipun model *LSTM* sudah dilatih, evaluasi terhadap performa model masih diperlukan. Peningkatan model juga bisa dilakukan dengan memodifikasi arsitektur model, menambah data yang lebih banyak atau berkualitas, dan meningkatkan hyperparameter. 2) Menjaga kualitas data: Kualitas data sangat penting untuk memperoleh hasil prediksi yang akurat. Oleh karena itu, diperlukan pengecekan kualitas data secara berkala untuk memastikan bahwa data yang digunakan untuk melatih model benar-benar mewakili kondisi produksi pabrik.

REFERENSI

- Aisyah, D, Purboyo, TW, & ... (2023). Prediksi Penderita Tuberkulosis Dengan Algoritma Long Short-Term Memory. *eProceedings ...*,telkomuniversity.ac.id, <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/19439>
- Akbar, Rahmatul, Santoso, R., & Warsito, B. (2023). Prediksi Tingkat Temperatur Kota Semarang Menggunakan Metode Long Short Term Memory (LSTM). *Jurnal Gaussian*, 11(4), 572-579.
- Alfandi, M., & Hatuaon Sihite, A. (2022). Penerapan Metode CNN-LSTM dalam Memprediksi Hujan pada Wilayah Medan. *Jurnal Nasional Teknologi Informasi dan Komputer*, 6(1), 490-499. Retrieved from <http://ejurnal.stmik-budidarma.ac.id/index.php/komik>
- Amrustian, Afrizal, M., Widayat, W., & Wirawan, A. M. (2022). Analisis Sentimen Evaluasi terhadap Pengajaran Dosen di Perguruan Tinggi Menggunakan Metode LSTM. *Jurnal Media Informatika Budidarma*, 6(1), 535.
- Aprian, Aji, B., Azhar, Y., & Nastiti, V. R. (2020). Prediksi Pendapatan Kargo Menggunakan Arsitektur Long Short Term Memory. *Jurnal Komputer Terapan*, 6(2), 148-157.
- Aulia, Alwi. (2021). Implementasi Algoritma Gated Recurrent Unit dalam Melakukan Prediksi Harga Kelapa Sawit dengan Memanfaatkan Model Recurrent Neural Network (RNN). *Prosiding SNASTIKOM: Seminar Nasional Teknologi Informasi & Komunikasi*, (pp. 288-294).
- Firdaus, R, & Mukhtar, H (2023). Prediksi Indeks Harga Produsen Pertanian Karet Di Indonesia Menggunakan Metode LSTM. *JURNAL FASILKOM*, ejurnal.umri.ac.id, <https://ejurnal.umri.ac.id/index.php/JIK/article/view/4851>
- Husni, D. T. (2022). Analisis Big data Penjualan Video Games Menggunakan Eda. *Jurnal Teknik Informasi dan Komputer (Tekinkom)*, 5(1), 43.
- Informatika, T. (2023). Prediksi Indeks Harga Produsen Pertanian Karet di Indonesia Menggunakan Metode LSTM. *Jurnal FASILKOM*, 13(1), 1-6.

- Ivan Dwi Nugraha, Y. A. (2022). Deteksi Depresi Pengguna Twitter Indonesia Menggunakan LSTM-RNN. *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, 11(3), 320-329.
- Julian, R. M. (2021). Peramalan Harga Saham Pertambangan Pada Bursa Efek Indonesia (BEI). *Jurnal Teknik Informatika dan Sistem Informasi*, 8(3), 1570-1580.
- Karno, Adhitho Satyo Bayangkari. (2020). “Prediksi Data Time Series Saham Bank BRI Dengan Mesin Belajar LSTM (Long ShortTerm Memory).” *Journal of Informatic and Information Security* 1(1): 1–8.
- Laksamana, M Imam Budi, Ema Utami, and Hanif Al Fatta. (2021). “Prediksi Jumlah Kunjungan Wisatawan Kabupaten Lombok Barat Menggunakan Algoritma Long Short Term Memory Magister Teknik Informatika , Universitas Amikom Yogyakarta Pendahuluan.” 6(2): 81–95.
- Lattifia, Tita, Putu Wira Buana, and Ni Kadek Dwi Rusjyanthi. (2022). “Model Prediksi Cuaca Menggunakan Metode LSTM.” *JITTER Jurnal Ilmiah Teknologi dan Komputer* 3(1): 994–1000.
<https://ojs.unud.ac.id/index.php/jitter/article/view/85000/43781>.
- Maygirtasari, Tyanma, Edy Yulianto, and Kholid Mukhammad Mawardi. 2015. “Faktor-Faktor Yang Mempengaruhi Volume Ekspor Crude Palm Oil (CPO) Indonesia.” *Jurnal Administrasi Bisnis (JAB)* 25(2): 1–8.
administrasibisnis.studentjournal.ub.ac.id.
- Meizar, Abdul, Wirhan Fahrozi, Evta Indra, and Muhardi Saputra. (2022). “Analisis Trend Moment Pada Datamining Forecasting Dalam Memprediksi Jumlah Persediaan Obat Herbal.” *Jurnal Sistem Informasi dan Ilmu Komputer Prima(JUSIKOM PRIMA)* 5(2): 103–6.
- Mukhlis, Mukhlis, Aziz Kustiyo, and Aries Suharso. (2021). “Peramalan Produksi Pertanian Menggunakan Model Long Short-Term Memory.” *Bina Insani Ict Journal* 8(1): 22.
- Owen, Michael, Vincent Vincent, Riama Br Ambarita, and Evta Indra. (2022). “Implementasi Metode Long Short Term Memory Untuk Memprediksi Pergerakan Nilai Harga Emas.” *Jurnal Teknik Informasi dan Komputer (Tekinkom)* 5(1): 96.

-
- Rumapea, Enjelica. (2020). “Analisis Jaringan Syaraf Tiruan Dalam Memprediksi Data Ekspor Ikan Tongkol / Tuna Menurut Negara Tujuan Utama.” *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS) SAINTEKS 2020 ISBN:: 232–35*.
- Sabar Sautomo, and Hilman Ferdinandus Pardede. (2021). “Prediksi Belanja Pemerintah Indonesia Menggunakan Long Short-Term Memory (LSTM).” *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)* 5(1): 99–106.
- Sanjaya, Fadil Indra, and Dadang Heksaputra. (2020). “Prediksi Rerata Harga Beras Tingkat Grosir Indonesia Dengan Long Short Term Memory.” *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)* 7(2): 163–74.
- Saragih, Jonas Rayandi, Mhd. Billy Sandi Saragih, and Anjar Wanto. (2018). “Analisis Algoritma Backpropagation Dalam Prediksi Nilai Ekspor (Juta Usd).” *Jurnal Pendidikan Teknologi dan Kejuruan* 15(2): 254–64.
- Wiko Putra, Biondi Bagasta, Moh. Ali Albar, and Budi Irmawati. (2019). “Penerapan Algoritma Jaringan Saraf Tiruan Metode Backpropagation Untuk Memprediksi Jumlah Nilai Ekspor Di Provinsi NTB.” *Jurnal Teknologi Informasi, Komputer, dan Aplikasinya (JTIKA)* 1(2).
- Wiranda, Laras, and Mujiono Sadikin. (2019). “Penerapan Long Short Term Memory Pada Data Time Series Untuk Memprediksi Penjualan Produk Pt. Metiska Farma.” *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)* 8(3): 184–96.