

**IMPLEMENTASI *MACHINE LEARNING* DALAM MEMREDIKSI
PERMINTAAN MODEL *BUSINESS TO BUSINESS* DENGAN
MENGUNAKAN ALGORITMA *AUTOREGRESSIVE INTEGRATED MOVING
AVERAGE (ARIMA)* DAN *LONG SHORT TERM MEMORY (LSTM)* GUNA
MENGURANGI *FOOD WASTE* (STUDI KASUS : PT TANIHUB INDONESIA)**

TUGAS AKHIR

**Diajukan Sebagai Salah Satu Syarat Untuk Mendapatkan Gelar Sarjana Strata-1
Pada Jurusan Teknik Industri Fakultas Teknologi Industri**



Nama : Dinda Meilasari

No Mahasiswa : 18522210

**PROGRAM STUDI TEKNIK INDUSTRI
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS ISLAM INDONESIA**

YOGYAKARTA

2022

PERNYATAAN KEASLIAN

Demi Allah, saya akui karya ini adalah karya hasil kerja saya sendiri kecuali nukilan dan ringkasan yang setiap satunya telah saya jelaskan sumbernya. Jika dikemudian hari ternyata terbukti pengakuan saya ini tidak benar dan melanggar peraturan yang sah dalam karya tulis dan hak kekayaan intelektual maka saya bersedia ijazah yang telah saya terima untuk ditarik kembali oleh Universitas Islam Indonesia.

Yogyakarta, 8 Agustus 2022



Dinda Meilasari

الجامعة الإسلامية
الاستدراك

LEMBAR PENGESAHAN PEMBIMBING

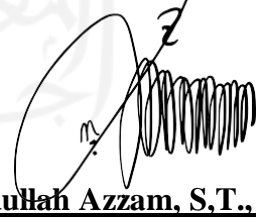
**IMPLEMENTASI *MACHINE LEARNING* DALAM MEMPREDIKSI
PERMINTAAN MODEL *BUSINESS TO BUSINESS* DENGAN
MENGUNAKAN ALGORITMA *AUTOREGRESSIVE INTEGRATED MOVING
AVERAGE (ARIMA)* DAN *LONG SHORT TERM MEMORY (LSTM)* GUNA
MENGURANGI *FOOD WASTE* (STUDI KASUS : PT TANIHUB INDONESIA)**

TUGAS AKHIR**Disusun Oleh :**

Nama : Dinda Meilasari

NIM : 18522210

Yogyakarta, 8 Agustus 2022

Pembimbing I**Ir.Ira Promasanti Rachmadewi, M.Eng.****Pembimbing II****Abdullah Azzam, S.T., M.T**

LEMBAR PENGESAHAN PENGUJI

**IMPLEMENTASI *MACHINE LEARNING* DALAM MEMPREDIKSI PERMINTAAN
MODEL *BUSINESS TO BUSINESS* DENGAN MENGGUNAKAN ALGORITMA
AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) DAN *LONG SHORT
TERM MEMORY (LSTM)* GUNA MENGURANGI *FOOD WASTE* (STUDI KASUS : PT
TANIHUB INDONESIA)**

TUGAS AKHIR

Oleh :

Nama : Dinda Meilasari
NIM : 18522210

Telah dipertahankan di depan sidang penguji sebagai salah satu syarat untuk
memperoleh gelar Sarjana Strata-1 Teknik Industri
Yogyakarta, 18 Agustus 2022

Tim Penguji

Ir. Ira Promasanti Rachmadewi, M.Eng.

Ketua

Yuli Agusti Rochman, S.T., M.Eng.

Anggota I

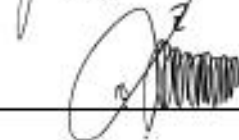
Andrie Pasca Hendradewa, S.T., M.T.

Anggota II

Abdullah Azzam, S.T., M.T.

Anggota III



Mengetahui,

Ketua Program Studi Teknik Industri

Fakultas Teknologi Industri

Universitas Islam Indonesia



Ir. Muhammad Ridwan Andi Purgomo, S.T., M.Sc., Ph.D., IPM.

HALAMAN PERSEMBAHAN

Alhamdulillah *rabbil 'alamin* atas izin dan ridha Allah SWT, saya ingin mempersembahkan karya tulis ini untuk kedua orang tua saya Bapak Paimin dan Ibu Susiati yang selalu tiada hentinya untuk mendoakan dan memberikan dukungan agar anaknya ini dapat selalu diberikan kemudahan dalam menyelesaikan karya tulis ini. Selain itu, kepada para sahabat sedari semester satu dan teman-teman saya semasa perkuliahan, terimakasih sudah selalu berjuang bersama hingga kita dapat berada di titik ini dan selalu mengingatkan untuk dapat menyelesaikan kuliah dengan baik dan dalam waktu yang tepat.



MOTTO

أَخْطَأْنَا أَوْ يِنَا نَسْ إِنْ تُؤَاخِذْنَا لَا رَبَّنَا ۗ أَكْتَسَبْتُمْ مَا وَعَلَيْهَا كَسَبْتُمْ مَا لَهَا ۗ وَسَعَهَا إِلَّا نَفْسًا اللَّهُ يُكَلِّفُ لَا
وَأَعْفُ بِهِ ۗ لَنَا طَاقَةٌ لَا مَا نُحْمِلُنَا وَلَا رَبَّنَا ۗ قَبْلَنَا مِنَ الَّذِينَ عَلَى حَمَلَتُهُ كَمَا إِصْرًا عَلَيْنَا تَحْمِيلٌ وَلَا رَبَّنَا
الْكُفْرِينَ وَمَالِقٍ عَلَى فَأَنْصُرْنَا مَوْلَانَا أَنْتَ ۗ وَأَرْحَمْنَا لَنَا وَأَعْفِرْ عَنَّا

“Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya. Ia mendapat pahala (dari kebajikan) yang diusahakannya dan ia mendapat siksa (dari kejahatan) yang dikerjakannya. (Mereka berdoa): "Ya Tuhan kami, janganlah Engkau hukum kami jika kami lupa atau kami tersalah. Ya Tuhan kami, janganlah Engkau bebankan kepada kami beban yang berat sebagaimana Engkau bebankan kepada orang-orang sebelum kami. Ya Tuhan kami, janganlah Engkau pikulkan kepada kami apa yang tak sanggup kami memikulnya. Beri maafilah kami; ampunilah kami; dan rahmatilah kami. Engkaulah Penolong kami, maka tolonglah kami terhadap kaum yang kafir".

(QS. Al-Baqarah : 286)

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Puji dan syukur penulis panjatkan kepada Allah SWT yang telah melimpahkan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan karya tulis ini. Shalawat serta salam tidak lupa saya ucapkan kepada junjungan kita, Nabi Muhammad SAW beserta keluarga dan sahabat beliau karena telah menuntun dari masa jahiliyah hingga ke jalan yang telah di ridhai oleh Allah SWT.

Alhamdulillah rabbil 'alamin atas izin dan karunia-Nya penulis dapat menyelesaikan karya tulis ini. Karya tulis ini dibuat sebagai salah satu syarat untuk memperoleh gelar Strata-1. Tentunya dalam Menyusun karya tulis ini, penulis menyadari banyak peran yang ikut serta membantu dalam proses penyelesaian. Oleh karena itu penulis mengucapkan terima kasih kepada :

1. Bapak Prof. Dr. Ir. Hari Purnomo, M.T., selaku Dekan Fakultas Teknologi Industri, Universitas Islam Indonesia.
2. Bapak Muhammad Ridwan Andi Purnomo, S.T., M.Sc., PhD selaku Ketua Jurusan Teknik Industri Fakultas Teknologi Industri, Universitas Islam Indonesia.
3. Bapak Dr. Taufiq Immawan, S.T., M.M. selaku Ketua Program Studi Teknik Industri Fakultas Teknologi Industri, Universitas Islam Indonesia.
4. Ibu Ira Promasanti Rachmadewi, Ir., M.Eng., dan Bapak Abdullah Azzam, S.T., M.T., selaku Dosen Pembimbing 1 dan 2 yang selalu membimbing dalam penyelesaian laporan ini.
5. Bapak Abdul Wahid Santoso selaku Manager Commercial Banten PT TaniHub Indonesia dan Mentor Magang Merdeka serta teman-teman Tim Commercial Banten yang tidak bisa disebutkan satu persatu yang sudah memberikan bimbingan dan pengalaman yang sangat berkesan

selama periode magang merdeka serta telah memberikan izin untuk dapat mengangkat kasus permasalahan di PT TaniHub Indonesia.

6. Kedua orang tua penulis, Bapak Paimin dan Ibu Susiati yang tiada hentinya untuk mendoakan dan memberikan dukungan untuk dapat selalu semangat dan yakin dapat menyelesaikan karya tulis ini.
7. Kakak dan Adik penulis, Kak Disa Widiastrika dan Muhammad Rasyad Arhaburrizqi yang selalu setia menemani dan mendukung dalam proses pembuatan hingga penyelesaian.
8. Rania Abdul Aziz Baraba selaku sahabat sedari TK hingga masa perkuliahan yang selalu meyakinkan diri penulis untuk yakin dalam menyelesaikan karya tulis ini.
9. Sahabat semasa perkuliahan, Putri Shafira Carolina, Indah Nur Nailiya, dan Bagas Zikri Afdila serta teman-teman angkatan 2018 lainnya yang selalu memberikan semangat satu sama lain untuk mendapatkan gelar Strata-1.

Penulis menyadari bahwa karya tulis ini masih terdapat banyak kekurangan. Oleh karena itu, kritik dan saran sangat diharapkan sehingga membuat karya tulis menjadi lebih baik lagi. Semoga kebaikan dan bantuan yang telah diberikan akan mendapat balasan dari Allah SWT dan semoga karya tulis ini dapat bermanfaat bagi pembaca ataupun penelitian selanjutnya. Aamiin.

Wassalamu'alaikum Warahmatullahi Wabarakatuh.

Yogyakarta, 8 Agustus 2022



Dinda Meilasari

ABSTRAK

Food waste merupakan pangan yang terbuang pada tahap distribusi dan retail, serta konsumsi. PT TaniHub Indonesia merupakan perusahaan berbasis *startup* yang bergerak di bidang agroteknologi yang berfokus dalam penjualan hasil tani kepada *customer*. Permasalahan *food waste* sering sekali terjadi di PT TaniHub Indonesia terkhususnya untuk *perishable product*. Hal ini dikarenakan tidak adanya jumlah permintaan stok produk hanya didasari intuisi *commercial specialist*. Pada penelitian ini akan melakukan pemodelan prediksi pada sektor *business to business* dengan menggunakan ARIMA dan LSTM yang akan diproses menggunakan bahasa pemrograman *Python*. Pengujian tingkat error akan dilakukan menggunakan *Root Mean Squared Error* (RMSE) dan *Mean Absolute Percentage Error* (MAPE). Hasil pemodelan menunjukkan bahwa performa metode LSTM memiliki akurasi yang lebih tinggi dengan nilai RMSE lebih kecil dibandingkan ARIMA yaitu 358.633 sedangkan untuk hasil peramalan memiliki nilai rata-rata MAPE 20% yang dinilai memiliki hasil akurasi yang baik.

Kata Kunci : *Food Waste, Perishable Product, ARIMA, LSTM, RMSE, MAPE.*

DAFTAR ISI

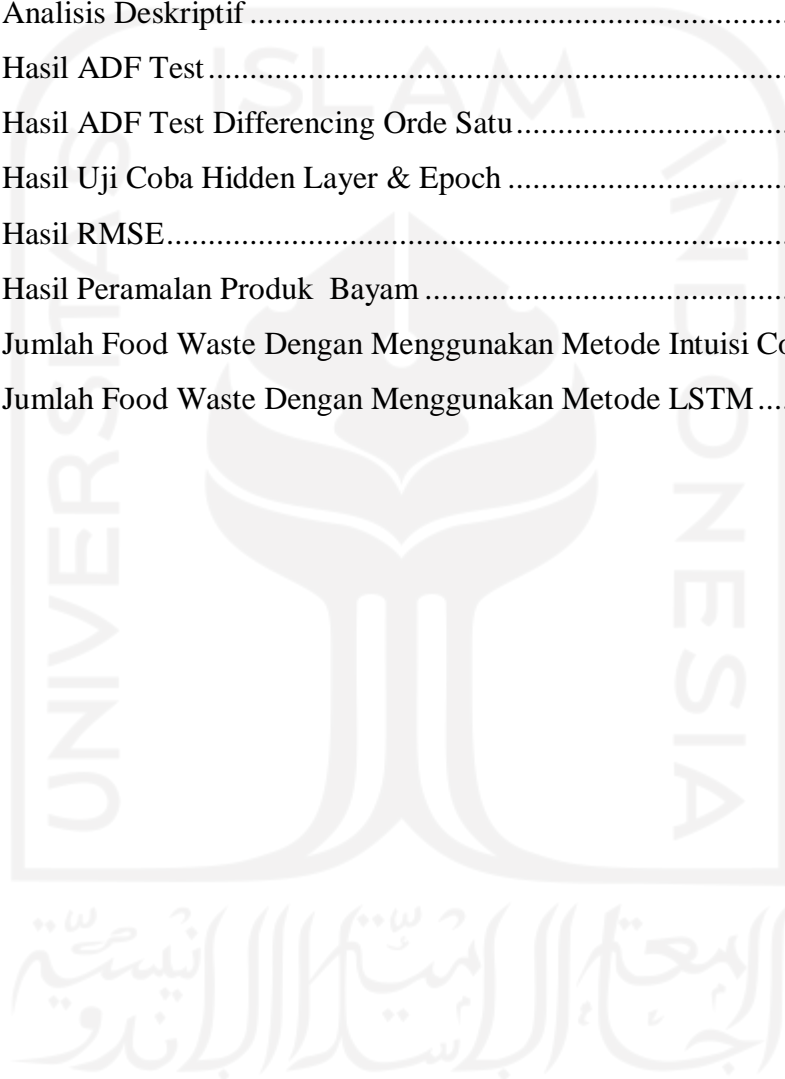
PERNYATAAN KEASLIAN	ii
LEMBAR PENGESAHAN PEMBIMBING.....	iii
LEMBAR PENGESAHAN PENGUJI.....	iv
HALAMAN PERSEMBAHAN	v
MOTTO.....	vi
KATA PENGANTAR.....	vii
ABSTRAK.....	ix
DAFTAR ISI.....	x
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR	xiv
BAB I PENDAHULUAN.....	15
1.1 Latar Belakang	15
1.2 Rumusan Masalah	18
1.3 Batasan Masalah	18
1.4 Tujuan Penelitian	18
1.5 Manfaat Penelitian	19
1.6 Sistematika Penulisan.....	19
BAB II KAJIAN LITERATUR	21
2.1 Kajian Deduktif.....	21
2.1.1 <i>Food Waste</i>	21
2.1.2 <i>Data</i>	22
2.1.3 <i>Time Series</i>	23
2.1.4 <i>Stasioneritas</i>	23
2.1.5 <i>Autoregressive Integrated Moving Average (ARIMA)</i>	24

2.1.6	<i>Long Short Term Memory (LSTM)</i>	25
2.1.7	<i>Machine Learning</i>	26
2.1.8	Pengujian Tingkat Error	26
2.2	Kajian Induktif	29
BAB III METODE PENELITIAN.....		41
3.1	Diagram Alur Penelitian.....	41
3.2	Identifikasi Masalah	42
3.3	Perumusan Masalah.....	42
3.4	Penentuan Tujuan dan Batasan Masalah	42
3.5	Pengumpulan Data	43
3.6	Pengolahan Data	43
3.6.1	<i>Pre-processing Data</i>	43
3.6.2	Pemodelan ARIMA dan LSTM.....	43
3.6.3	Pengujian Tingkat Error	44
3.7	Hasil dan Pembahasan.....	44
3.8	Kesimpulan dan Saran.....	45
BAB IV PENGUMPULAN DAN PENGOLAHAN DATA.....		46
4.1	Pengumpulan Data	46
4.2	<i>Pre-Processing Data</i>	49
4.2.1	<i>Data Cleaning</i>	50
4.2.2	<i>Data Selection</i>	50
4.2.3	<i>Data Transformation</i>	51
4.2.4	Penjumlahan Harian (Berat Produk).....	52
4.2.5	Persiapan Data	53
4.3	Pemodelan LSTM	53
4.3.1	Mendefinisikan Model	53
4.3.2	<i>Training Model LSTM</i>	54

4.3.3	Evaluasi Model LSTM	54
4.4	Pemodelan ARIMA.....	55
4.4.1	Identifikasi Model.....	55
4.4.2	<i>Differencing</i>	56
4.4.3	Estimasi Parameter Model.....	56
4.4.4	<i>Training Model</i> ARIMA.....	57
4.4.5	Evaluasi Model ARIMA.....	57
BAB V HASIL DAN PEMBAHASAN		58
5.1	<i>Analisis Pre-Processing Data</i>	58
5.1.1	<i>Analisis Data Cleaning</i>	58
5.1.2	<i>Analisis Data Selection</i>	58
5.1.3	<i>Analisis Data Transformation</i>	59
5.1.4	Analisis Penjumlahan Harian	63
5.1.5	Analisis Persiapan Data.....	64
5.2	Analisis Pemodelan ARIMA	64
5.3	Analisis Pemodelan LSTM.....	69
5.4	Analisis Evaluasi Model ARIMA & LSTM.....	71
5.5	<i>Analisis Food Waste</i>	74
BAB VI PENUTUP		78
6.1	Kesimpulan.....	78
6.2	Saran.....	78
DAFTAR PUSTAKA.....		79

DAFTAR TABEL

Tabel 2. 1 Kajian Induktif.....	36
Tabel 4. 1 Entitas Data	46
Tabel 5. 1 Hasil Transformasi Product Name.....	60
Tabel 5. 2 Analisis Deskriptif	64
Tabel 5. 3 Hasil ADF Test.....	66
Tabel 5. 4 Hasil ADF Test Differencing Orde Satu.....	67
Tabel 5. 5 Hasil Uji Coba Hidden Layer & Epoch	70
Tabel 5. 6 Hasil RMSE.....	72
Tabel 5. 7 Hasil Peramalan Produk Bayam	72
Tabel 5. 8 Jumlah Food Waste Dengan Menggunakan Metode Intuisi Commercial	74
Tabel 5. 9 Jumlah Food Waste Dengan Menggunakan Metode LSTM.....	76



DAFTAR GAMBAR

Gambar 1. 1 Data Food Loss and Waste di Indonesia	16
Gambar 2. 1 Sel LSTM	25
Gambar 3. 1 Alur Penelitian.....	41
Gambar 4. 1 Dataframe Historis Permintaan.....	49
Gambar 4. 2 Data Cleaning.....	50
Gambar 4. 3 Data Selection Farm Produce	51
Gambar 4. 4 Data Selection Vegetable.....	51
Gambar 4. 5 Data Transformation Unit Name.....	52
Gambar 4. 6 Data Transformation Product Name	52
Gambar 4. 7 Penjumlahan Harian	52
Gambar 4. 8 Persiapan Data	53
Gambar 4. 9 Mendefinisikan Model	54
Gambar 4. 10 Training Model LSTM	54
Gambar 4. 11 Evaluasi Model LSTM	55
Gambar 4. 12 Identifikasi Model	55
Gambar 4. 13 Differencing	56
Gambar 4. 14 Mengestimasi Parameter Model.....	56
Gambar 4. 15 Training Model ARIMA.....	57
Gambar 4. 16 Evaluasi Model ARIMA.....	57
Gambar 5. 1 Plot Data	65
Gambar 5. 2 Hasil Differencing	67
Gambar 5. 3 Hasil Auto ARIMA	68
Gambar 5. 4 Plot Hasil ARIMA.....	69
Gambar 5. 5 Plot Hasil LSTM	71

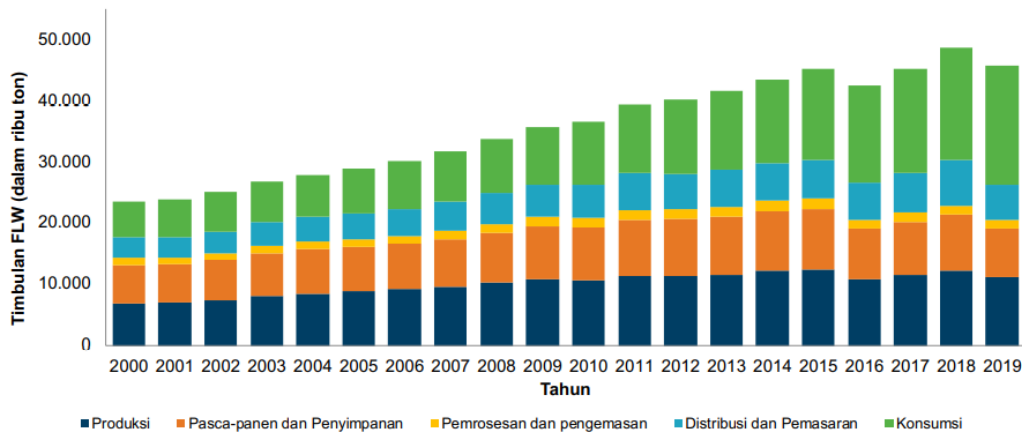
BAB I

PENDAHULUAN

1.1 Latar Belakang

Trend penjualan produk-produk *farm* yang meliputi daging, sayur, buah, ikan & *seafood* yang saat ini sering kita temui dijual di beberapa *platform e-commerce* memiliki peningkatan penjualan yang cukup tinggi. Hal ini dikarenakan kondisi pandemik yang saat ini terjadi menyebabkan masyarakat untuk mengurangi mobilitas di tempat umum seperti pasar maupun supermarket. Tapi pada nyatanya produk *farm* ini termasuk dalam *perishable product* yaitu produk yang memiliki jangka siklus hidup yang pendek. Umur simpan produk segar harus diasumsikan satu hari sedemikian rupa sehingga jumlah permintaan harian harus dipesan, disimpan, dan dijual setiap hari. Dampak dari penjualan produk *farm* ini dapat menyebabkan *deterioration*, yaitu penurunan nilai terhadap suatu produk berdasarkan kerusakan suatu produk termasuk dari jangka umur mendekati kadaluarsa. Hal ini dapat menyebabkan penurunan *revenue* yang diperoleh oleh perusahaan. Pada *perishable product*, pengendalian persediaannya dinilai cukup sulit dikarenakan permintaan produk yang fluktuatif dan tidak pasti (Wulansari, 2017). Tidak adanya peramalan yang akurat dalam penjualan *perishable product* dapat menyebabkan produk-produk ini menjadi *food waste*.

Food waste merupakan pangan yang terbuang pada tahap distribusi dan retail, serta konsumsi. Berdasarkan *The Economist Intelligence Unit* (2017), Indonesia menduduki peringkat kedua dengan *food loss* dan *food waste* terbesar di dunia setelah Arab Saudi dengan menghasilkan 13 juta ton *food waste* di setiap tahunnya. Pada dasarnya timbulan *food loss* dan *food waste* dibagi menjadi beberapa tahap yaitu : tahap produksi, tahap pascapanen dan penyimpanan, tahap pemrosesan dan pengemasan, tahap distribusi dan pemasaran, dan tahap konsumsi. Namun, untuk *food waste* sendiri ada ditahap distribusi dan pemasaran serta tahap konsumsi. Berikut merupakan sebaran data *food waste* di berbagai tahap di Indonesia :



Gambar 1. 1 Data *Food Waste* di Indonesia

Sumber : Laporan Kajian Bappenas

Peramalan *time series* merupakan suatu peramalan dengan menggunakan pendekatan kuantitatif dengan menjadikan data masa lampau sebagai objek peramalan untuk di masa yang akan datang. Salah satu metode peramalan *time series* yang dapat digunakan dan memiliki akurasi yang tinggi adalah *machine learning*. Menurut Shwartz dan David (2014) *Machine Learning* merupakan sebuah studi tentang algoritma untuk mempelajari sesuatu dalam melakukan beberapa hal tertentu yang dilakukan oleh manusia secara otomatis. Algoritma yang berbasis *machine learning* merupakan algoritma yang sangat baik yang telah dipelajari secara ekstensif dan memiliki potensi yang akurat dalam melakukan peramalan (Shen, 2012). Oleh karena itu, *machine learning* merupakan metode yang tepat untuk melakukan peramalan suatu permintaan.

TaniHub merupakan perusahaan berbasis *startup* yang bergerak di bidang agroteknologi yang berfokus dalam penjualan hasil tani kepada *customer*. Tak hanya menjual hasil tani saja namun tanihub juga menjual produk-produk rumah tangga lainnya yang diklasifikasikan menjadi produk *farm* dan produk *non farm*. Beberapa produk farm yang dijual dari hasil tani tersebut tergolong dalam *perishable product* yang memiliki jangka hidup yang relatif singkat. Ritchie Goenawan (2021) selaku *Chief Marketing Officer* TaniHub dalam wawancaranya mengatakan bahwa, TaniHub masih mengalami minim

informasi dan akses yang berdampak pada harga dan *food waste*. *Food waste* yang terjadi di TaniHub berlangsung pada tahap pemasaran. Hal ini dikarenakan banyaknya produk yang tidak terjual hingga waktu umur simpan produk. Walaupun pada *warehouse* TaniHub sudah menerapkan beberapa aturan penyimpanan produk yang menyesuaikan pada suhu ketahanan produk namun seringkali masih ditemukan produk-produk yang sudah tidak layak jual dan konsumsi sehingga digolongkan menjadi *food waste*. Hal ini dikarenakan jumlah perkiraan penjualan di TaniHub hanya di dasari oleh intuisi dari setiap *commercial specialist* untuk memprediksi jenis dan kuantitas produk yang akan terjual dalam periode bulan tersebut. Dikarenakan tidak adanya peramalan permintaan terhadap suatu produk yang akurat sehingga menyebabkan *perishable product* terkhususnya *produk farm* di TaniHub menumpuk di *warehouse* dan menyebabkan *food waste*. Dengan menggunakan sistem peramalan penjualan yang efisien yang memperhitungkan akurasi dan ketidakpastian merupakan langkah dasar yang signifikan dalam mengurangi *food waste* yang terjadi pada *retail industry* (Arunraj, 2014).

Berdasarkan permasalahan diatas, maka diperlukan prediksi permintaan terhadap produk-produk *farm* untuk menanggulangi isu *food waste*. Peramalan yang digunakan menggunakan dua metode yaitu *Autoregression Integrated Moving Average* (ARIMA) dan *Long Short-term Memory* (LSTM). Metode peramalan ARIMA memiliki ketepatan yang sangat baik untuk peramalan dengan jangka pendek dan untuk data *time series* non stasioner pada saat linier (Munarsih, 2011). ARIMA mengkombinasikan model *autoregressive* dan *moving average*. Kelebihan dari ARIMA adalah memiliki sifat yang fleksibel (mengikuti pola data), memiliki tingkat akurasi peramalan yang cukup tinggi dan cocok digunakan untuk meramal sejumlah variabel yang cepat, sederhana, akurat dan murah (Kolker, 2011). Sedangkan LSTM adalah sebuah algoritma yang didasarkan oleh *Recurrent Neural Network* (RNN) yang memiliki tingkat keakuratan yang tinggi dalam melakukan peramalan data dalam bentuk *time series* (Robby, 2021). Harapannya melalui kedua metode termasuk nantinya akan dicari metode yang memiliki tingkat error yang paling rendah untuk dapat dijadikan sebagai metode alternatif sehingga dapat direkomendasikan kepada PT. TaniHub Indonesia untuk menanggulangi isu *food waste*.

1.2 Rumusan Masalah

Rumusan masalah pada penelitian ini adalah sebagai berikut :

1. Bagaimana model prediksi permintaan yang berbasis *machine learning* dengan menggunakan metode *Autoregression Integrated Moving Average* (ARIMA) dan *Long Short-term Memory* (LSTM) untuk item *vegetable* pada produk *farm* dalam sektor *business-to-business* di PT. TaniHub Indonesia ?
2. Metode apakah yang memiliki tingkat error yang rendah sehingga dapat dijadikan sebagai rekomendasi metode untuk dapat mengurangi *food waste* ?

1.3 Batasan Masalah

Adapun batasan-batasan pada penelitian yang dilakukan di PT. TaniHub Indonesia adalah sebagai berikut :

1. Objek penelitian yang digunakan adalah *farm product*.
2. Objek penelitian pada *farm product* hanya berfokus pada item *vegetable*.
3. Data historis yang digunakan adalah data penjualan pada 29 Desember 2020 – 30 Desember 2021.

1.4 Tujuan Penelitian

Adapun tujuan pada penelitian yang dilakukan di PT. TaniHub Indonesia adalah sebagai berikut :

1. Mengetahui model prediksi permintaan yang berbasis *machine learning* dengan menggunakan metode *Autoregression Integrated Moving Average* (ARIMA) dan *Long Short-term Memory* (LSTM) untuk item *vegetable* pada produk *farm* dalam sektor *business-to-business* di PT. TaniHub Indonesia.
2. Mengetahui metode yang memiliki tingkat error yang rendah sehingga dapat dijadikan sebagai rekomendasi metode untuk dapat mengurangi *food waste*.

1.5 Manfaat Penelitian

Adapun batasan-batasan pada penelitian yang dilakukan di PT. TaniHub Indonesia adalah sebagai berikut:

1. Bagi Keilmuan

Membantu pengembangan ilmu pengetahuan dalam kajian keilmuan dan perkembangan teknologi, serta hasil penelitian ini diharapkan dapat menjadi pembelajaran yang berguna bagi penelitian lainnya.

2. Bagi Perusahaan

Membantu perusahaan dalam menyelesaikan permasalahan yang ada yaitu dalam melakukan peramalan permintaan dan menjadi faktor pertimbangan dalam pengambilan keputusan.

1.6 Sistematika Penulisan

Berikut merupakan sistematika penulisan tugas akhir yang dibagi menjadi enam bab :

BAB 1 PENDAHULUAN,

Pada bab 1 terdiri dari kajian singkat yang memuat penentuan masalah sebagai dasar dilakukannya penelitian ini sehingga nantinya dituangkan dalam latar belakang. Selanjutnya melakukan penentuan terhadap rumusan dan batasan masalah serta tujuan dan manfaat penelitian.

BAB II KAJIAN LITERATUR

Pada bab 2 ini memuat landasan teoritik yang dijadikan pedoman dalam melakukan penelitian berdasarkan dari teori dan penelian-penelitian yang telah dilakukan sebelumnya. Pada kajian literatur ini terdapat dua pokok pembahasan yaitu kajian deduktif dan kajian induktif.

BAB III METODOLOGI PENELITIAN

Pada bab 3 ini memuat bagan alur penelitian dari awal hingga akhir penelitian. Dalam bagan tersebut juga berisi terkait metode yang akan digunakan. Tahapan dimulai dari identifikasi masalah, perumusan masalah, penentuan tujuan dan batasan penelitian, pengolahan data (*pre-processing data*, pemodelan ARIMA & LSTM), hasil dan pembahasan, dan kesimpulan dan saran.

BAB IV PENGUMPULAN DAN PENGOLAHAN DATA

Pada bab 4 ini dimulai dari pengumpulan data yang digunakan untuk pemodelan. Selanjutnya akan dilakukan pengolahan terhadap data-data tersebut yang dimulai dari *pre-processing data*. Setelah data seragam maka dilakukan proses persiapan data untuk menentukan proporsi perbandingan jumlah data *training* dan data *test*. Terakhir akan dilakukan pemodelan dan peramalan menggunakan ARIMA dan LSTM serta dilakukan pengujian tingkat error dengan menggunakan *Root Mean Squared Error* (RMSE) dan (MAPE).

BAB V HASIL DAN PEMBAHASAN

Pada bab 5 ini berisi hasil dari pengolahan data yang telah dilakukan pada bab sebelumnya dan memberikan analisis pembahasan terkait hasil pengolahan data tersebut. Selain itu, pada bab ini juga akan membahas terkait perbandingan hasil prediksi yang paling baik berdasarkan hasil pengujian tingkat error yang paling rendah. Hasil akhir akan berisi contoh peramalan pada salah satu produk *farm* yang diikuti dengan persentase error antara hasil prediksi dengan data aktual.

BAB VI KESIMPULAN DAN SARAN

Pada bab 6 berisi kesimpulan yang menjawab dari rumusan masalah dan berisi saran ataupun rekomendasi terkait penelitian yang telah dilakukan sehingga dapat menjadi bahan evaluasi untuk penelitian selanjutnya.

BAB II

KAJIAN LITERATUR

2.1 Kajian Deduktif

Kajian deduktif merupakan kajian yang membahas teori-teori konseptual yang dapat mendukung dalam penelitian ini.

2.1.1 *Food Waste*

Dalam rantai pasok makanan terdiri dari 5 tahap yaitu : Produksi, Pasca panen dan penyimpanan, Pemrosesan dan pengemasan, Distribusi dan pemasaran, dan konsumsi. *Food waste* merupakan pangan yang terbuang pada tahap distribusi dan retail, serta konsumsi. Menurut Legrand (2017), *food waste* diklasifikasikan menjadi 2 macam berdasarkan waktu dan tingkat kemungkinannya. Berdasarkan waktunya, *food waste* dibagi menjadi 3 macam, yaitu :

a. *Pre-consumer Waste*

Segala pemangkasan makanan basi ataupun bahan baku di dapur yang belum sempat untuk diolah menjadi menu yang akan dikonsumsi.

b. *Post-consumer Waste*

Sampah yang terkumpul hasil sisa dari mengonsumsi makanan.

c. *Packaging waste dan Operation Supplies*

Sampah makanan yang berada dalam plastik yang tidak terdekomposisi alami. Sedangkan *operation supplies* adalah bahan-bahan yang digunakan dan menjadi *waste* dalam mengolah makanan.

2.1.2 Data

Menurut Willian dan Swyer (2007) data merupakan sekumpulan informasi yang berisi fakta dan angka. Selain itu menurut Mulyanto (2009) data merupakan suatu bahan mentah yang berisi peristiwa, konsep, manusia, hewan dan lain sebagainya yang berbentuk kata, angka, gambar ataupun media lainnya sebagai wujud representasi suatu peristiwa yang perlu diolah lebih lanjut untuk memperoleh informasi. Data dapat digolongkan menjadi 2 jenis data yaitu data kualitatif dan kuantitatif. Data kualitatif merupakan data yang berbentuk gambar ataupun kata. Sedangkan data kuantitatif merupakan data yang berbentuk angka. Selain itu, data juga dapat diklasifikasikan menjadi beberapa golongan diantaranya :

2.1.2.1 Klasifikasi Berdasarkan Sifat Data

a. Data Diskrit

Data diskrit adalah data yang nilainya memiliki keterbatasan antara nilai satu dengan nilai lainnya. Pada umumnya data diskrit diperoleh dari hasil menghitung. Adapun contohnya : jumlah pulpen yang dimiliki siswa, jumlah karyawan di perusahaan X, jumlah penjualan suatu barang.

b. Data Kontinu

Data kontinu adalah data yang nilainya tidak terbatas dalam kisaran tertentu. Pada umumnya data kontinu diperoleh dari hasil mengukur. Adapun contohnya : data suhu badan (menggunakan thermometer), data diameter cincin (menggunakan jangka sorong), data berat badan (menggunakan timbangan)

2.1.2.2 Klasifikasi Berdasarkan Waktu Pengumpulannya

a. Data *Cross Section*

Data *cross section* adalah data yang dikumpulkan pada titik waktu tertentu. Dalam pengumpulannya data ini tidak mengumpulkan pada satu titik persis pada suatu periode waktu. Adapun contohnya : Data konsumen PT XYZ pada Bulan Mei 2020.

b. Data *Time Series*

Data time series adalah data yang dikumpulkan dari waktu ke waktu atau periode secara historis. Adapun contohnya : Data penjualan pada Bulan Mei 2020 – Mei 2021

2.1.3 Time Series

Time series merupakan metode peramalan yang memiliki asumsi bahwa masa depan merupakan fungsi dari masa lalu (Heizer & Render, 2009). Hal ini selaras dengan tujuan dari metode dari *time series* yaitu menganalisis pola berdasarkan dari deret data historis. Dalam pengamatan *time series* menggunakan data yang berasal dari satu tempat yang sama yang berurutan dan memiliki interval yang tetap. Interval yang digunakan dapat berupa minggu, bulan maupun tahun. Pada umumnya terdapat beberapa jenis *time series*, yaitu (Hanke & Wichren, 2005) :

- a. Pola Siklus
Pola yang mengalami fluktuasi dari data yang memiliki kurun waktu lebih dari satu tahun.
- b. Pola Horizontal
Pola yang berasal dari kejadian tidak terduga dan bersifat acak namun kemunculannya dapat mempengaruhi fluktuasi data *time series*.
- c. Pola *Trend*
Pola yang memiliki kecenderungan terhadap arah data dalam jangka Panjang dapat berupa kenaikan maupun penurunan.
- d. Pola Musiman
Pola yang mengalami fluktuasi dikarenakan data yang periodik dalam kurun waktu kurang dari satu tahun.

2.1.4 Stasioneritas

Menurut Brillinger (2005), salah satu hipotesis dalam melakukan prediksi data *time series* adalah data harus stasioner. Stasioneritas adalah suatu peristiwa ketika data mengalami

kenaikan maupun penurunan. Jika pola data tidak bersifat trend, tidak terdapat fluktuasi periodik dan memiliki keragaman yang konstan, maka data tersebut stasioner (Marvilla, 2013). Untuk mengetahui stasioneritas data dapat melihatnya dengan plot data. Jika pada plot data terdapat fenomena naik turunnya lag pertama ke lag berikutnya yang jauh dan langsung terpotong (*cut off*), maka data runtun waktu stasioner. Namun pemeriksaan stasioneritas dengan melihat plot data termasuk dalam subyektif. Untuk mengetahui stasioneritas data secara pasti dapat menggunakan *Augmented Dickey-Fuller Test* (ADF Test). Apabila data mengandung non stasioner, cara untuk mengubah data yang bersifat non stasioner menjadi stasioner adalah dengan cara melakukan *differencing* (Marvilla, 2013).

2.1.5 *Autoregressive Integrated Moving Average* (ARIMA)

Metode *Autoregressive Integrated Moving Average* (ARIMA) merupakan metode yang dikembangkan oleh Box dan Jenkins pada tahun 1976. Pada sebuah data historis, model ini mengabaikan variabel independennya dan model ini tidak mengasumsikan pola tertentu. Dalam penggunaannya model ini mampu untuk digunakan pada *time series* yang stasioner. Metode ARIMA merupakan gabungan antara *autoregressive* (AR) dan *moving average* (MA) (Hanke, dkk., 2019).

a. *Autoregressive* (AR)

Bentuk umum dari model *autoregressive* dengan AR(p) atau model ARIMA (p,0,0) dinyatakan sebagai berikut:

$$Y_t = \theta_0 + \theta_1 Y_{t-1} + \dots + \theta_p Y_{t-p} + \varepsilon_t \quad (1)$$

b. *Moving Average* (MA)

Bentuk umum dari model *Moving Average* dengan MA(q) atau model ARIMA(0,0,q) dinyatakan sebagai berikut:

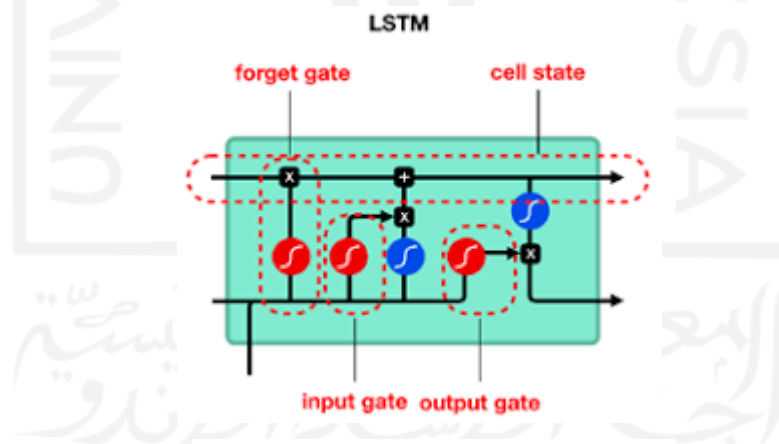
$$Y_t = \mu + \varepsilon_t - \omega_1 \varepsilon_{t-1} - \dots - \omega_q \varepsilon_{t-q} \quad (2)$$

c. *Autoregressive dan Moving Average (ARMA)*

$$Y_t = \theta_0 + \theta_1 Y_{t-1} + \dots + \theta_p Y_{t-p} + \varepsilon_t - \omega_1 \varepsilon_{t-1} - \dots - \omega_q \varepsilon_t \quad (3)$$

2.1.6 Long Short Term Memory (LSTM)

Algoritma *Long Short Term Memory* merupakan algoritma yang dikembangkan dari algoritma *Recurral Neural Network (RNN)* yang memiliki tingkat keakuratan yang tinggi untuk melakukan peramalan pada data *time series* (Roby, Rizki 2020). Model ini dikenalkan oleh Hochreiter dan Schmidhuber pada tahun 1997. Pada dasarnya struktur dari LSTM hampir sama dengan RNN yang memiliki bentuk rangkaian modul jaringan yang berulang. Namun, pada LSTM dilakukan modifikasi dengan menambahkan *memory cell* agar dapat menyimpan informasi dalam waktu yang lama. Pada LSTM terdapat *input gate*, *output gate* dan *forget gate*. *Input gate* berfungsi untuk melakukan pembaharuan sel, sedangkan *forget gate* digunakan untuk menentukan informasi yang akan disimpan ataupun dibuang dan untuk *output gate* berfungsi untuk menentukan *hidden state* berikutnya (Nguyen,2018).



Gambar 2. 1 Sel LSTM

2.1.7 Machine Learning

Menurut Shwartz dan David (2004), *machine learning* merupakan sebuah studi algoritma yang difungsikan untuk mempelajari sesuatu dalam melakukan beberapa *task* tertentu yang ditentukan oleh manusia secara otomatis. Tujuan dari *machine learning* adalah agar membuat sistem yang dapat memahami algoritma untuk dapat melakukan pekerjaan mandiri secara otomatis yang pada umumnya dilakukan oleh manusia. Implementasi dari *machine learning* sudah banyak diterapkan dalam kehidupan sehari-hari, seperti : *Face Detection* (sistem yang dapat mengenali wajah seseorang), *Speech Understanding* (sistem pemrosesan untuk mengubah suara menjadi suatu teks tulisan) dan *Customer Segmentation* (sistem yang menganalisis pola perilaku konsumen untuk membantu produsen dalam bidang pemasaran). Berdasarkan konsep pembelajarannya *machine learning* dibagi menjadi 2 jenis, yaitu :

1. *Supervised Learning*

Supervised Learning merupakan teknik pembelajaran mesin untuk memprediksi peristiwa di masa depan berdasarkan data yang diperoleh dari masa lampau dengan membuat suatu fungsi berdasarkan data tersebut.

2. *Unsupervised Learning*

Unsupervised Learning merupakan teknik pembelajaran mesin dengan membuat representasi pola berdasarkan dari data yang sudah ada. Pada teknik ini tidak adanya pengklasifikasian data namun mencari struktur tersembunyi dari data input tersebut.

2.1.8 Pengujian Tingkat Error

Pengukuran kesalahan dalam peramalan terjadi karena adanya perbedaan dengan peramalan aktual dengan peramalan permintaan (Taylor, 2004). Menurut Preswitch dkk (2014) Dalam pengukuran kesalahan banyak sekali metode-metode pengukuran yang dapat diterapkan untuk mengukur kesalahan peramalan, diantaranya :

2.1.7.1 *Scale Dependent Measures*

1. *Mean [Signed] Error (ME)*

$$\begin{aligned} & \text{Mean} (e_t) \\ & e_t = y_t - \hat{y}_t \end{aligned} \quad (4)$$

2. *Mean Square Error (MSE)*

$$\text{Mean} (e_t^2) \quad (5)$$

3. *Root-Mean Square Error (RMSE)*

$$\sqrt{\text{MSE}} \quad (6)$$

4. *Mean Absolute Error (MAE)*

$$\text{Mean} (|e_t|) \quad (7)$$

5. *Median Absolute Error (MdAE)*

$$\text{Median} (|e_t|) \quad (8)$$

2.1.7.2 *Percentage Measures*

1. *Mean Absolute Percentage Error (MAPE)*

$$\begin{aligned} & \text{Mean} (|p_t|) \\ & p_t = \frac{100 e_t}{y_t} \end{aligned} \quad (9)$$

2. *Median Absolute Percentage Error (MdAPE)*

$$\text{Median} (|p_t|) \quad (10)$$

3. *Root-Mean Square Percentage Error (RMSPE)*

$$\sqrt{\text{mean} (p_t^2)} \quad (11)$$

4. *Root-Median Square Percentage Error (RMdSPE)*

$$\sqrt{\text{median} (p_t^2)} \quad (12)$$

5. *Symmetric Mean Absolute Percentage Error (sMAPE)*

$$\text{Mean} \left(\frac{200 |e_t|}{y_t - \hat{y}_t} \right) \quad (13)$$

6. *Symmetric Median Absolute Percentage Error (sMdAPE)*

$$\text{Median} \left(\frac{200 |e_t|}{y_t - \hat{y}_t} \right) \quad (14)$$

2.1.7.3 *Relative Error Based Measurement*

1. *Mean Relative Absolute Error (MRAE)*

$$\text{Mean} (|r_t|)$$

$$p_t = \frac{e_t}{e_t^*} \quad (15)$$

2. *Median Relative Absolute Error (MdRAE)*

$$\text{Median} (|r_t|) \quad (16)$$

2.1.7.4 *Relative Measurement*

1. *Relative Mean Absolute Error (RelMAE)*

$$\frac{MAE}{MAE_b} \quad (17)$$

2. *Relative Mean Squared Error (RelMSE)*

$$\frac{MSE}{MSE_b} \quad (18)$$

3. *Relative Root Mean Squared Error (RelRMSE)*

$$\frac{RMSE}{RMSE_b} \quad (19)$$

2.2 Kajian Induktif

Kajian induktif merupakan sekumpulan penelitian terdahulu yang memiliki kaitan dengan penelitian saat ini sehingga dijadikan acuan dan referensi dalam penelitian yang dilakukan.

2.2.1 *Agribusiness Time Series Forecasting Using Wavelet Neural Networks and Metaheuristic Optimization: An Analysis of The Soybean Sack Price and Perishable Products Demand.*

Penelitian yang dilakukan oleh Weslly Puchalskya, Gabriel Trierweiler Ribeiro, Claudimar Pereira da Veigac, Roberto Zanetti Freirea, Leandro dos Santos Coelho yang dipublikasikan pada tahun 2018 bertujuan untuk mengevaluasi kinerja *Wavelet Neural Networks* (WNNs) yang dikombinasikan dengan lima teknik optimasi untuk mendapatkan peramalan *time series* terbaik dengan mempertimbangkan dua studi kasus di sektor agribisnis. Objek dari penelitian ini adalah Harga sekarung kedelai dan permintaan terhadap *perishable product* yang nantinya akan menghasilkan prediksi harga sekarung kedelai dalam 30 hari kedepan dan permintaan *perishable product* dalam 12 bulan kedepan. Alasan dari penggunaan objek tersebut ialah karena merupakan salah satu budaya agribisnis di brasil dan harga sekarung kedelai itu menjadi komoditas penting di seluruh dunia. Selain itu alasan penggunaan objek permintaan *perishable product* adalah dengan mempertimbangkan bahwa perusahaan dengan strategi peramalan permintaan yang tepat mampu meningkatkan efisiensi, dan meningkatkan tingkat layanan pelanggan. Hasil dari penelitian tersebut adalah diperlukan pengujian metode pelatihan yang berbeda untuk *time series forecasting*. Hal ini dikarenakan WNN memberikan suatu pendekatan yang sangat konsisten.

2.2.2 *A Machine Learning-Based Forecasting System of Perishable Cargo Flow in Maritime Transport*

Penelitian yang dilakukan oleh José Antonio Moscoso-López, Daniel Urda, Juan Jesús Ruiz-Aguilar, Javier González-Enrique, Ignacio J. Turias yang dipublikasikan pada tahun 2019 bertujuan untuk memprediksi nilai masa depan terhadap aliran kargo perishable. Data yang digunakan ialah catatan harian buah dan sayur dari tahun 2010 hingga 2017.

Sedangkan untuk penelitian ini digunakan metode SVM (*Support Vector Machine*) dan LSTM (*Long Short Term Memory*). Hasil analisis menunjukkan bahwa sistem berbasis *machine learning* ini mencapai tingkat kinerja 14,83% pada dataset buah segar dan 11,3% pada sayuran. Selain itu, SVM dan LSTM menjadi metode yang memiliki tingkat kinerja yang lebih baik untuk dataset buah dan sayuran.

2.2.3 Demand Forecasting for Perishable Asset in Importer Company (PT. TMM)

Pada penelitian yang dilakukan oleh Mahaning Indrawaty Wijaya, Ratih Hendayani, Sri Widiyanesti pada tahun 2014 bertujuan untuk mengetahui jumlah permintaan dimasa yang akan datang sehingga dapat membuat target yang tepat. Penelitian ini dilakukan karena adanya gap antara target penjualan dengan realita pada *perishable asset* sehingga menyebabkan kelebihan stok dan kekurangan stok. Data yang digunakan pada penelitian ini adalah data historis target dan permintaan perishable asset di PT. TMM. Berdasarkan data yang terkumpul, dipilih tiga jenis produk yang memiliki tingkat penjualan tertinggi yang selanjutnya akan di analisis menggunakan Microsoft Excel, Minitab dan Zaitun Time Series. Metode yang digunakan pada penelitian ini adalah SES, Holt, dan ETS. Hasil penelitian menunjukkan bahwa ETS memiliki tingkat kesalahan terendah untuk tiga jenis produk, yang berarti ini adalah metode terbaik untuk meramalkan permintaan masa depan terhadap *perishable asset* di PT. TMM.

2.2.4 Demand Forecasting in Food Retail : A Comparison Between The Holt Winters and ARIMA Models.

Penelitian yang dilakukan oleh Claudimar Pereira Da Veiga, Cassia Rita Pereira Da Veiga, Anderson Catapan, Ubirata Tortato, dan Wesley Vieira Da Silva yang dipublikasikan pada tahun 2014 bertujuan untuk membandingkan dua model peramalan permintaan linier yaitu Model *Autoregressive Integrated Moving Average* (ARIMA) model dan Model Holt-Winters (HW). Penelitian ini dilakukan dikarenakan pada perusahaan retail makanan yang memiliki produk makanan dengan jangka hidup yang pendek dan mudah rusak sangat penting untuk melakukan *forecasting* dikarenakan permintaan yang sangat fluktuatif.

Selain itu, pada penelitian ini juga menggunakan metode MAPE dan U-Theil untuk pengujian tingkat eror. Objek dari penelitian ini adalah produk *perishable dairy*. Hasil dari penelitian ini adalah Model Holt-Winters merupakan model yang memiliki tingkat kinerja yang baik dibandingkan model ARIMA karena dapat memberikan hasil peramalan yang akurat dengan teknik yang lebih kompleks.

2.2.5 Decision Support System for Customer Demand Forecasting and Inventory Management of Perishable Goods

Penelitian yang dilakukan oleh Prapai Sridama dan Chopaka Siribut yang dipublikasikan pada tahun 2018 yang bertujuan untuk mempelajari sistem pengambilan keputusan untuk peramalan permintaan pelanggan dan manajemen persediaan terhadap *perishable goods*. Nantinya model peramalan dibuat untuk diterapkan dalam perencanaan jumlah barang ke gudang untuk volume penyimpanan terendah. Objek dari penelitian ini adalah penjualan daging babi secara grosir maupun eceran di Provinsi Bungkan, Thailand. Data yang digunakan adalah data historis 350 hari yang lalu. Sedangkan untuk metode yang digunakan adalah *Single Exponential Smoothing* (SES) dan Holt. Hasil dari penelitian ini adalah algoritma SES memiliki tingkat efisiensi yang lebih tinggi dibandingkan metode yang lainnya.

2.2.6 Forecasting The Daily Sales of Perishable Food to Reduce Spoilage in Hypermarket

Penelitian ini dilakukan oleh Prof. Swapnali Kurhade, Ashwini Kadam, Dipali Warekar, dan Nagnath Kamble yang dipublikasikan pada tahun 2015 yang bertujuan untuk mengurangi pembusukan *perishable foods* di hypermarket secara akurat dengan memperkirakan jumlah penjualan setiap harinya. Alasan dari adanya penelitian ini didasari oleh fenomena ketika terdapat pesanan bahan makanan di supermarket dalam jumlah besar yang mengakibatkan stok bahan makanan lebih dari yang dibutuhkan dan akhirnya menyebabkan pemborosan. Nantinya penelitian ini akan memperkirakan permintaan harian

selama 30 hari untuk Sayuran , Daging , Buah dan Produk Susu. Metode yang digunakan adalah Holt-Winter.

2.2.7 Pemodelan *Time Series* Dengan Proses ARIMA Untuk Prediksi Indeks Harga Konsumen (IHK) Di Palu – Sulawesi Tengah

Penelitian yang dilakukan oleh Y. Wigati, Rais dan T. Utami yang dipublikasikan pada tahun 2016 bertujuan untuk memprediksi pergerakan data Indeks Harga Konsumen (IHK) di Palu – Sulawesi Tengah. Penelitian ini menggunakan metode ARIMA untuk melakukan prediksi kedepannya. Data yang digunakan adalah data IHK yang pergerakannya fluktuatif. Hasil dari penelitian ini adalah model dengan proses ARIMA (1,1,1) sebagai model terbaik.

2.2.8 Peramalan Harga Saham Pertambangan Pada Bursa Efek Indonesia (BEI) Menggunakan *Long Short Term Memory* (LSTM)

Penelitian yang dilakukan oleh Roby Julian dan Muhammad Rizky Pribadi yang dipublikasikan pada tahun 2021 bertujuan untuk melakukan prediksi harga saham agar mendapatkan tingkat akurasi yang tinggi. Hal ini didasari oleh pergerakan perusahaan yang sangat dinamis yang dapat menentukan harga saham. Dalam penelitian ini menggunakan metode *Long Short Term Memory* (LSTM) dengan metode pengukuran tingkat eror *Root Mean Square Error* (RMSE). Adapun objek dari penelitian ini adalah Bursa Efek Indonesia (BEI) dengan menggunakan data dari emiten ataupun perusahaan yang terdaftar di Bursa Efek Indonesia (BEI) pada sektor pertambangan dan yang mengalami pertumbuhan harga saham yang meningkat drastis serta terdapat fenomena yang kuat sehingga emiten saham tersebut layak untuk dilakukan penelitian. Hasil dari penelitian ini yaitu dengan menggunakan *Long Short Term Memory* (LSTM) penelitian ini menghasilkan nilai RMSE yang cukup baik dengan adanya peningkatan nilai RMSE berdasarkan penambahan jumlah variasi *epoch*.

2.2.9 Comparative Study of Wavelet-ARIMA and Wavelet-ANN Models for Temperature Time Series Data in Northeastern Bangladesh

Penelitian yang dilakukan oleh Ahmad Hasan Nury, Khairul Hasan dan Md. Jahir Bin Alam yang dipublikasikan pada tahun 2015 bertujuan untuk menganalisis tren dan pola suhu untuk melihat variasi transien. Metode yang digunakan pada penelitian ini ada metode kombinasi antara Wavelet-ANN dan Wavelet-ARIMA yang diterapkan pada data suhu maksimum dan data suhu minimum bulanan. Data *training set* yang digunakan adalah data pada tahun 1957-2000 untuk menyusun model dan untuk data *testing set* menggunakan data pada tahun 2001-2012 untuk memperkirakan kinerja. Pada penelitian ini juga menggunakan metode pengukuran tingkat eror RMSE (*Root Mean Square Error*). Hasil dari penelitian ini menunjukkan bahwa model wavelet-ARIMA lebih baik untuk memprediksi suhu di wilayah *Northeastern Bangladesh*.

2.2.10 The Use of LSTM-Based RNN and SVM Models to Detect Ludian Coseismic Landslides in Time Series Images

Penelitian yang dilakukan oleh Zeng-Guang Zhou, Binbin Chen, Ziyang Li dan Chuanrong Li yang dipublikasikan pada tahun 2020 bertujuan untuk mendeteksi tanah longsor yang dipicu oleh gempa Ms.6.5 di Ludian, China pada tahun 2014. Metode yang digunakan adalah *Long Short Term Memory* (LSTM), *Recurrent Neural Network* (RNN) dan *Support Vector Machine* (SVM). Data yang digunakan adalah 72 gambar yang tersedia dengan resolusi spasial 16/30m dari satelit Landsat-7/8, GaoFen-1 dan HJ-1A/1B. Hasil dari penelitian ini adalah adanya potensi kombinasi model LSTM dan SVM yang dapat digunakan untuk mendeteksi longsor dengan resolusi 30 m.

2.2.11 Web Traffic Time Series Forecasting Using ARIMA and LSTM RNN

Penelitian yang dilakukan oleh Tejas Shelatkar, Stephen Tondale, Swaraj Yadav, dan Sheetal Ahir yang dipublikasikan pada tahun 2020 yang bertujuan memprediksi lalu lintas pada sebuah website. Data yang digunakan adalah halaman wikipedia yang berisi

kunjungan halaman harian sebagai rangkaian waktu untuk setiap posting. Metode yang digunakan pada penelitian ini adalah ARIMA, LSTM dan RNN. Hasil dari penelitian ini adalah metode LSTM dan RNN membawa lebih banyak efisiensi ke sistem web tersebut.

2.2.12 Bitcoin Price Prediction Using ARIMA and LSTM

Penelitian yang dilakukan oleh Yiqing Hua pada tahun 2020 bertujuan untuk membandingkan prediksi harga bitcoin dalam USD berdasarkan model *Long Short Term Memory* (LSTM) dan *Autoregressive Integrated Moving Average* (ARIMA). Data yang digunakan adalah harga bitcoin yang diunduh dalam bitfinex. Hasil dari penelitian ini adalah ARIMA dan LSTM dapat bekerja dengan baik untuk memprediksi harga bitcoin. Namun, LSTM dapat membuat prediksi lebih efisien dan tingkat presisi juga lebih tinggi. Sedangkan ARIMA cukup efisien dalam membuat prediksi dalam waktu singkat tetapi seiring berjalannya waktu tingkat presisi akan menurun.

2.2.13 Time Series Investigation of Milk Production in Major States of India Using ARIMA Modelin

Penelitian yang dilakukan oleh P. Mishra, Chellai Fatih, G.K. Vani, Shiwani Tiwari, D. Ramesh dan Anurag Dubey yang dipublikasikan pada tahun 2020 bertujuan untuk membantu dalam merumuskan kebijakan pertanian nasional serta perencanaan tepat untuk produk susu. Metode yang digunakan adalah *Autoregressive Integrated Moving Average* (ARIMA) dengan menggunakan data *time series* produksi susu pada tahun 2017. Hasil dari penelitian ini yaitu perlu adanya peningkatan produksi susu dengan melakukan penyadaran kepada pemilik dan peternak sapi perah tentang program pemuliaan hewan dan praktik kesehatan.

2.2.14 Implementasi Machine Learning Dalam Memprediksi Permintaan Model Business to Business Dengan Menggunakan Algoritma Autoregressive Integrated Moving Average (ARIMA) dan Long Short Term Memory (LSTM) Guna Mengurangi Food Waste (Studi Kasus : PT TaniHub Indonesia)

Penelitian yang dilakukan oleh Dinda Meilasari pada tahun 2022 bertujuan untuk membuat model prediksi permintaan pada *perishable product* di PT TaniHub Indonesia. Model yang digunakan adalah *Autoregressive Integrated Moving Average* (ARIMA) dan *Long Short Term Memory* (LSTM) dengan menggunakan data *time series* permintaan produk *farm* di sektor *business to business*. Hasil dari penelitian ini adalah untuk membandingkan nilai tingkat uji error paling rendah dari dua metode tersebut untuk dijadikan referensi dalam melakukan pemodelan permintaan.





Tabel 2. 1 Kajian Induktif

No	Penulis, Tahun	Objek Penelitian	Metode											
			ARIMA	Holt	SES	TES	SVM	HW	LSTM	Wavelet	RNN	ANN		
1.	(Wellsy, Gabriel, Claudimar, Leandro, Roberto, 2018)	Harga Sekarung Kedelai dan Permintaan <i>Perishable Product</i> .									✓			
2.	(Jose, Daniel, Juan, Javier, Ignacio, 2019)	Aliran Kargo <i>Perishable</i>					✓			✓				

3. (Mahaning, *Perishable*
Ratih, Sri, *Asset*
2014)
4. (Claudimar, *Perishable*
Cassia, *Dairy*
Anderson, *Product*
Ubirata dan
Wesley
2014)
5. (Prapai dan *Perishable*
Chopaka, *Goods*
2018)
6. (Swapnali, *(Perishable*
Ashwini, *foods* :
Dipali dan Sayuran,
Nagnath, Daging,
2015) Buah dan
Produk
Susu)



- | | | | | | | |
|--|--|---|---|--|---|---|
| 7. (Wigati, Rais dan Utami, 2016) | Indeks Harga Konsumen | ✓ |  | | | |
| 8. (Roby dan Rizky, 2021) | Harga Saham | | | ✓ | | |
| 9. (Ahmad, Khairul dan Jahir, 2015) | Suhu di <i>Northeastern</i> Bangladesh | ✓ | | ✓ | ✓ | |
| 10. (Zeng, Binbin, Ziyang dan Chuanrong, 2020) | Gambar Tanah Longsor | ✓ | | ✓ | ✓ | |
| 11. (Tejas, Stephen, Swaraj dan | Halaman Wikipedia | ✓ | |  | ✓ | ✓ |

- Sheetan,
2020)
12. (Yiqing,
2020) Harga
 Bitcoin
13. (Mishra,
Chellai,
Vani,
Shiwani,
Ramesh
dan
Anurag,
2020) Produk Susu
 di India
- 14 (Dinda,
2022) *Perishable*
 product
 (i produk
 farm)
 TaniHub



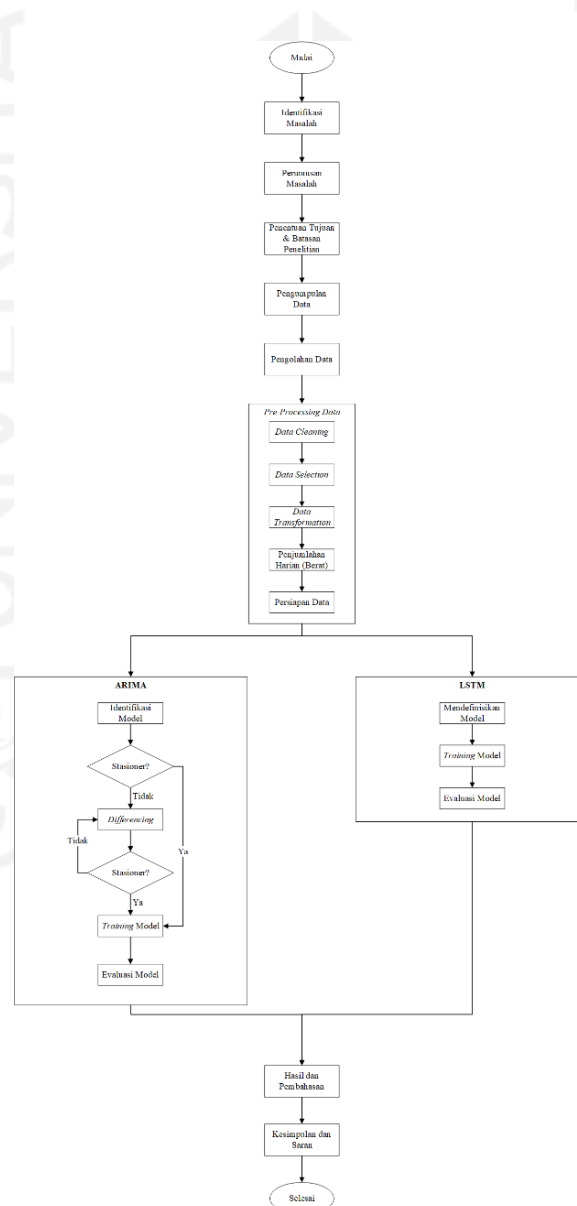


BAB III

METODE PENELITIAN

3.1 Diagram Alur Penelitian

Berikut merupakan *flowchart* yang menunjukkan tahapan penelitian dari awal hingga akhir:



Gambar 3. 1 Alur Penelitian

3.2 Identifikasi Masalah

Permasalahan *foodwaste* yang terjadi di PT TaniHub Indonesia yang terjadi pada tahap distribusi dan pemasaran menyebabkan kerugian dikarenakan tidak terjualnya produk. Identifikasi masalah tersebut dilakukan berdasarkan dari tugas harian yang dilakukan saat melakukan penelitian yang selanjutnya di analisis permasalahan yang ada pada perusahaan tersebut sehingga diperoleh masalah yang dapat dikorelasikan dengan keilmuan Teknik Industri. Dalam hal ini peneliti akan melakukan peramalan terhadap permintaan produk *farm* pada item *vegetable* di PT TaniHub Indonesia. Metode peramalan akan menggunakan data *time series*.

3.3 Perumusan Masalah

Dalam perumusan masalah bertujuan agar dapat membantu peneliti dalam menyelesaikan masalah dan dapat menentukan acuan yang akan digunakan untuk penyelesaian masalah tersebut. Berdasarkan masalah yang diangkat, peneliti akan membentuk model prediksi permintaan menggunakan metode *Autoregressive Integrated Moving Average* (ARIMA) dan *Long Short Term Memory* (LSTM).

3.4 Penentuan Tujuan dan Batasan Masalah

Pada tahap ini dilakukan penentuan tujuan dan batasan masalah penelitian. Tujuan penelitian ini berdasarkan permasalahan yang diangkat dengan berlandaskan dari perumusan masalah yaitu untuk mengetahui jumlah permintaan yang akurat terhadap produk *farm* dengan indikasi kategori *vegetable*. Dalam penyusunan batasan masalah ditentukan melalui studi lapangan dan studi literatur dengan topik terkait. Pada studi lapangan diperoleh melalui tahap diskusi dengan mentor terkait produk-produk yang sering mengalami permasalahan terkait *foodwaste* yaitu pada produk *vegetable* dengan data historis yang digunakan pada 29 Desember 2020 – 30 Desember 2021 dengan model transaksi *business to business*.

3.5 Pengumpulan Data

Metode pengumpulan data adalah dengan menggunakan data sekunder, data sekunder merupakan data yang diperoleh oleh peneliti dari sumber yang telah tersedia. Pengumpulan data dilakukan dengan melakukan pengunduhan data historis permintaan pada Tim *Business Intelligence* yang memiliki tanggung jawab dalam pengelolaan terkait *big data* perusahaan dengan format xlsx untuk periode 29 Desember 2020 – 30 Desember 2021 yang terakumulasi sebanyak 235.662 data.

3.6 Pengolahan Data

Pengolahan data dilakukan dengan menggunakan bahasa pemrograman python. Python merupakan bahasa pemrograman yang interpretatif dan multiguna dengan berdasarkan filosofi perancangan yang berfokus pada tingkat keterbacaan kode. Dalam penggunaan python juga akan menggunakan beberapa *library* yang akan membantu dalam pengolahan data, seperti : Pandas, Numpy, Sklearn, Seaborn, Statsmodel, Matplotlib, Pmdarima, dan Keras.

3.6.1 *Pre-processing Data*

Pre-processing data adalah langkah-langkah yang digunakan sebagai bentuk proses penemuan pengetahuan yang di dasarkan data yang berkualitas (Kumar & Cadha, 2012). *Pre-processing data* dilakukan untuk mempersiapkan data sebelum dilakukan pemodelan prediksi. Pada tahap ini dilakukan *data cleaning*, *data selection*, *data transformation*, Penjumlahan Harian terhadap berat produk, dan persiapan data.

3.6.2 Pemodelan ARIMA dan LSTM

Pada penelitian ini akan menggunakan metode *Autoregressive Integrated Moving Average* (ARIMA) dan *Long Short Term Memory* (LSTM). Model *Autoregressive Integrated Moving Average* (ARIMA) merupakan model yang menggambarkan tren dan musiman dalam suatu deret waktu sebagai fungsi dari nilai-nilai yang tertinggal (*Autoregressive*) dan rata-rata yang berubah dari waktu ke waktu (*Moving Averages*). ARIMA memiliki

ketepatan yang sangat baik untuk peramalan dengan jangka pendek dan untuk data *time series* non stasioner pada saat linier (Munarsih, 2011). Pada penelitian ini akan dilakukan beberapa tahap sebelum dilakukan peramalan dengan ARIMA diantaranya, identifikasi model, *differencing*, estimasi parameter model dan training model.

LSTM merupakan sebuah metode pengembangan dari *neural network*. Dalam LSTM terdapat sebuah *cell state* yang menyimpan nilai atau keadaan untuk periode waktu yang panjang ataupun singkat. LSTM memiliki 3 struktur utama yang memiliki tugasnya masing-masing yaitu *forget gate*, *input gate* dan *output gate*. Pada penelitian ini akan dilakukan penentuan model yang cocok terlebih dahulu dengan mengestimasi beberapa parameter yang diperlukan. Selanjutnya akan dilakukan training model untuk dilakukan peramalan.

3.6.3 Pengujian Tingkat Error

Pengujian tingkat error bertujuan untuk membandingkan, mengawasi, dan memastikan model peramalan berjalan dengan baik. Metode yang digunakan untuk pengujian tingkat error adalah *Root Mean Squared Error* (RMSE) dan *Mean Absolute Percentage Error* (MAPE). RMSE merupakan metode pengujian tingkat *error* yang digunakan untuk mengambil keputusan dalam penentuan metode peramalan yang akan digunakan. Sedangkan MAPE merupakan persentase absolut kesalahan antara nilai aktual dan nilai prediksi.

3.7 Hasil dan Pembahasan

Pada hasil dan pembahasan ini berisi analisis terkait hasil dari pengolahan data yang sudah dilakukan. Hasil dari analisis akan menentukan metode yang paling tepat yang berdasarkan dari pengujian tingkat error yang telah dilakukan. Selain itu, dilakukan analisis prediksi model terhadap data harian pada salah satu item dalam kategori *vegetable* untuk mengetahui persentase akurasi hasil peramalan dengan aktual.

3.8 Kesimpulan dan Saran

Pada bagian kesimpulan merupakan jawaban dari rumusan masalah yang sudah dituliskan diatas. Selain itu, terdapat saran yang berisi terkait rekomendasi yang dapat digunakan untuk perusahaan maupun penelitian selanjutnya.



BAB IV

PENGUMPULAN DAN PENGOLAHAN DATA

4.1 Pengumpulan Data

Pada penelitian ini menggunakan data sekunder, yaitu data historis permintaan pada sektor *Business to Business* PT TaniHub Indonesia. Pengumpulan data dilakukan dengan menghubungi pada tim terkait yang bertanggung jawab pada perekapan data historis transaksi klien yaitu Tim *Business Intelligence* dengan cara mengunduh pada *e-retail* yang merupakan *database* yang digunakan oleh perusahaan untuk melakukan transaksi penjualan. Pengunduhan data dilakukan pada data historis 29 Desember 2020 – 30 Desember 2021 dalam format *xlsx*. Total data yang dikumpulkan adalah 235.620 data historis permintaan. Terdapat 6 entitas data dengan 36 atribut yang terdapat dalam *database* tersebut, diantaranya adalah :

Tabel 4. 1 Entitas Data

No	Entitas	Atribut	Deskripsi
1	Date	Order Date	Tanggal ketika klien melakukan order.
		Delivered Date	Tanggal pesanan diantarkan ke klien.
2	Invoice Info	No Purchase Order	Nomor order klien.
		No Invoice	Nomor tagihan nota klien.
3	Client Info	Channel	Saluran penjualan

No	Entitas	Atribut	Deskripsi
			yang digunakan oleh klien.
		<i>Client Code</i>	Nomor kode klien.
		<i>Region Name</i>	Lokasi wilayah klien.
		<i>Client Province</i>	Provinsi tempat tinggal klien.
		<i>Client City</i>	Kota tempat tinggal klien.
		<i>Client District</i>	Kecamatan tempat tinggal klien.
		<i>Client Village</i>	Kelurahan tempat tinggal klien.
		<i>Client Type</i>	Tipe bisnis klien.
		<i>Client Subtype</i>	Subtipe bisnis klien
		<i>Client Business Name</i>	Nama bisnis klien
		<i>Client Store Name</i>	Nama toko klien.
		<i>Commercial in Charge</i>	Nama <i>commercial</i> yang bertanggung jawab
		<i>Payment Status</i>	Status pembelian klien.
		<i>Client First Order</i>	Keterangan tanggal dan waktu klien ketika melakukan transaksi pertama kali.
		<i>First Repeat</i>	Keterangan pembelian pertama

No	Entitas	Atribut	Deskripsi
			atau <i>repeat order</i> .
		<i>MAP</i>	Akumulasi nilai pembelian klien.
		<i>Order Province</i>	Provinsi pengantaran pesanan.
		<i>Order City</i>	Kota pengantaran pesanan.
		<i>Order District</i>	Kecamatan pengantaran pesanan.
4	<i>Order Info</i>	<i>Order Village</i>	Kelurahan pengantaran pesanan.
		<i>Qty Ordered</i>	Jumlah produk yang dipesan.
		<i>Qty Received</i>	Jumlah produk yang diterima.
		<i>Product Code</i>	Kode produk.
		<i>Product Name</i>	Nama produk.
		<i>Product Department</i>	Departemen produk.
		<i>Product Category</i>	Kategori produk.
		<i>Product Type</i>	Tipe produk.
5	<i>Product Info</i>	<i>Farm Category</i>	Kategori jenis hasil tani.
		<i>Unit Name</i>	Satuan produk.
		<i>Selling Price</i>	Harga jual produk.
6	<i>Profit Info</i>	<i>Revenue</i>	Keuntungan yang

No	Entitas	Atribut	Deskripsi
			diperoleh.
		<i>Margin</i>	Persentase keuntungan yang diperoleh.

	order_dt	delivered_dt	purchase_order_no	invoice_no	Channel	client_code	region_name	client_province	client_city	client_district	...	unit
0	2021-01-30	2021-01-31	CK200	INV/CK/202101000001	Offline	201023054141034	Cikarang	DKI Jakarta	Kota Jakarta Selatan	Jagakarsa	...	
1	2021-01-30	2021-01-31	CK200	INV/CK/202101000001	Offline	201023054141034	Cikarang	DKI Jakarta	Kota Jakarta Selatan	Jagakarsa	...	
2	2021-01-30	2021-01-31	CK200	INV/CK/202101000001	Offline	201023054141034	Cikarang	DKI Jakarta	Kota Jakarta Selatan	Jagakarsa	...	
3	2021-01-30	2021-01-31	CK200	INV/CK/202101000001	Offline	201023054141034	Cikarang	DKI Jakarta	Kota Jakarta Selatan	Jagakarsa	...	
4	2021-01-30	2021-01-31	CK200	INV/CK/202101000001	Offline	201023054141034	Cikarang	DKI Jakarta	Kota Jakarta Selatan	Jagakarsa	...	

Gambar 4. 1 *Dataframe* Historis Permintaan

Berdasarkan data sampel diatas, dapat diketahui bahwa item yang dijual oleh PT TaniHub Indonesia dikategorikan berdasarkan *farm category* dan *product department*. Pada *farm category* produk-produk tersebut dikategorikan dalam *farm produce* dan *non farm produce*. Sedangkan dalam *product department* dikategorikan dalam *dairy frozen, fish & seafood, drink, food, fruit, home cleaning, household, meat, non food, poultry, vegetable, saprodi* dan *others*. Selain itu dalam penjualan produk juga dikemas dalam beberapa kemasan seperti dalam Pack, KG, Sisir, PCS, dan Karton.

4.2 Pre-Processing Data

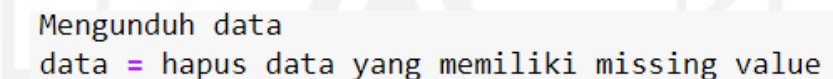
Pre-processing data bertujuan sebagai proses penemuan pengetahuan yang di dasarkan data yang berkualitas. *Pre-processing* pada data penelitian dilakukan menggunakan *Software* Python dengan beberapa library, diantaranya :

1. Numpy : *Library* yang digunakan untuk pengolahan *array* agar lebih cepat, tepat dan efisien.

2. Pandas : *Library open source* yang menyediakan struktur data untuk proses analisis data.
3. Sklearn : *Library* yang digunakan untuk membantu dalam melakukan data *slicing* guna memisahkan antara data *train* dan data *test*.
4. Statsmodels : *Library* yang digunakan untuk melakukan uji statistik dan estimasi model statistik.
5. Matplotlib : *Library* yang digunakan untuk visualisasi data berupa plot, diagram batang, histogram maupun diagram lainnya.
6. Seaborn : *Library* yang digunakan untuk visualisasi data yang berdasarkan matplotlib yang menampilkan hubungan antara dua variabel.

4.2.1 Data Cleaning

Data cleaning merupakan proses pembersihan data yang bertujuan untuk menghilangkan *noise*, data inkonsisten maupun data yang tidak relevan (Meilani,2014). Pada tahap ini dilakukan beberapa proses pembersihan data pada data yang memiliki *missing value* atau *null*. Metode yang digunakan untuk penanganan data yang memiliki *missing value* ini adalah *Listwise deletion*, yaitu sebuah metode yang digunakan untuk menghapus suatu data yang memiliki *missing value* ataupun *null* dari data sampel. Setelah melakukan penghapusan *missing value*, jumlah data yang kini tersedia sebanyak 193.511 data.



```
Mengunduh data
data = hapus data yang memiliki missing value
```

Gambar 4. 2 *Data Cleaning*

4.2.2 Data Selection

Data selection merupakan proses pengambilan data yang berhubungan dengan analisis yang dilakukan. Tahap ini akan berfokus pada pengurangan atribut yang tidak masuk dalam kategori *Farm Produce* dengan sub kategori *Vegetable*.

1. *Data farm produce*

Pada tahap ini akan dilakukan *data selection* sehingga hanya tersisa data pada kategori *farm produce* saja. *Data selection* dilakukan dengan cara membuat *dataframe* baru dengan nama `data_farm`. Setelah dilakukan *data selection*, jumlah data yang kini tersedia yaitu 158.635 data.

```
membuat dataframe data_farm
hapus data pada farm category yang termasuk non farm
```

Gambar 4. 5 *Data Selection Farm Produce*

2. *Data Kategori Vegetable*

Pada tahap ini akan dilakukan *data selection* pada entitas *product department* sehingga hanya tersisa data pada sub kategori *vegetable* saja. *Data selection* dilakukan dengan cara membuat *dataframe* baru dengan nama `data_vege`. Setelah dilakukan *data selection*, jumlah data yang kini tersedia yaitu berjumlah 87.596 data

```
membuat dataframe data_vege
hapus data_farm pada product development yang tidak termasuk vegetable
```

Gambar 4. 6 *Data Selection Vegetable*

4.2.3 *Data Transformation*

Data transformation merupakan proses transformasi dan konsolidasi data ke dalam bentuk yang disesuaikan untuk mining (Han, 2012).

1. *Tranformasi Unit Name*

Dalam data historis permintaan PT TaniHub Indonesia pada atribut *unit name* yang memuat satuan produk masih terindikasi inkonsisten dikarenakan memiliki lebih dari satu jenis satuan. Berdasarkan hal tersebut akan dilakukan tranformasi pada atribut *unit name* dengan mengubah pada satuan yang sama, yaitu gram.

```

Jika nama = KG, maka berat = Qty_ordered*1000
nama = Pack 50 GR, maka berat = Qty_ordered*50
nama = Pack 100 GR, maka berat = Qty_ordered*100
nama = Pack 250 GR, maka berat = Qty_ordered*250
nama = Pack 500 GR, maka berat = Qty_ordered*250
nama = ....., maka berat = Qty_ordered*...

```

Gambar 4. 7 *Data Transformation Unit Name*

2. Transformasi *Product Name*

Nama produk dalam data historis permintaan PT TaniHub Indonesia masih inkonsisten sehingga menyebabkan satu produk memiliki lebih dari satu nama dalam suatu *database*. Pada tahap ini dilakukan penamaan ulang agar dalam *dataframe* hanya memiliki satu nama pada setiap produk.

```

Munculkan semua nama produk
Ganti nama 'Bawang Merah KG' menjadi 'Bawang Merah'
Ganti nama 'Bawang Merah Pack 500 GR' menjadi 'Bawang Merah'
Ganti nama 'Bawang Merah Pack 250 GR' menjadi 'Bawang Merah'
Ganti nama '....' menjadi '....'

```

Gambar 4. 8 *Data Transformation Product Name*

4.2.4 Penjumlahan Harian (Berat Produk)

Pada tahap ini akan dibuat suatu *database* baru yang akan memuat hasil penjumlahan harian pada kolom 'satuan gram' yang merupakan kolom yang berisi informasi terkait jumlah berat produk dalam satuan yang sudah di transformasikan menjadi gram dengan menggunakan *Library Pandas*. Selain itu akan dilakukan penghapusan kolom pada '*farm category*' dan '*product department*' sehingga kolom yang tersisa hanyalah '*Order_dt*', '*Product_Name*', '*ID*' dan 'Satuan Gram'. Hasil akhir dari penjumlahan harian ini memuat menjadi 364 data.

```

Mengelompokkan berdasarkan tanggal sebagai index dan menjumlahkan satuan gram
Menyimpan data pada dataframe baru

```

Gambar 4. 9 Penjumlahan Harian

4.2.5 Persiapan Data

Persiapan data dilakukan untuk menyiapkan data *training* dan data *testing* sebelum dilakukan pemodelan. Pada tahap ini akan dilakukan normalisasi data menggunakan *Library Sklearn* dengan *feature scaling range* (0,1). Selanjutnya akan dilakukan pembagian data *training* dan data *testing* yang akan digunakan dengan komposisi data *training* 80% dan data *testing* 20%.

```
membuat fungsi normalisasi data pada range (0,1)
data_train = mengambil 80% dari data secara random
data_test = mengambil 20% dari data secara random
menerapkan fungsi normalisasi pada data
scaled_train = mentransform fungsi normalisasi pada data train
scaled_test = mentransform fungsi normalisasi pada data test
```

Gambar 4. 10 Persiapan Data

4.3 Pemodelan LSTM

Long Short Term Memory merupakan algoritma yang dikembangkan dari algoritma *Recurral Neural Network* (RNN) yang memiliki tingkat keakuratan yang tinggi untuk melakukan peramalan pada data *time series* (Roby, Rizki 2020). Dalam melakukan pemodelan LSTM peneliti menggunakan *Library Keras* yang merupakan *library machine learning* yang dikembangkan untuk penerapan model pembelajaran yang mencakup pembuatan prototipe berbasis *neural network*.

4.3.1 Mendefinisikan Model

Langkah pertama yang harus dilakukan adalah membuat *instance* atau mendefinisikan modelnya terlebih dahulu. Tipe model yang digunakan pada penelitian ini adalah *Class Sequential* yang berfungsi untuk membuat, menambahkan dan menghubungkan suatu *layer ke layer lainnya secara berurutan*. Pada penelitian ini terdapat penambahan 3 *layer* yang menggunakan fungsi '*add()*' dengan rincian 1 *layer* utama, 1 *dense hidden layer* dan 1 *dense layer output*. Pada *layer* utama yaitu menggunakan *default* 64 neuron dengan 7 *timesteps* dan 1 *feature*. Sedangkan untuk *layer* kedua atau lapisan tersembunyi yang

biasanya disebut *hidden layer* menggunakan 8 *neuron*. Sedangkan untuk *dense output layer* menggunakan 1 *neuron*.

Pada penelitian ini *activation function layer* yang digunakan adalah *Rectified Linear Unit* (ReLU). Selain itu untuk *compile model* menggunakan 3 parameter yaitu *Optimizer*, *Metric* dan *Loss*. Adapun *optimizer* yang digunakan adalah ‘Adam’ sedangkan untuk *metric* yang digunakan adalah ‘*root mean squared error*’ dan *loss* yang digunakan adalah ‘*mean squared error*’

```
Membuat RNN
Membuat layer utama sebanyak 64 dengan fungsi aktivasi ReLU
Menambah hidden layer sebanyak 8 dengan fungsi aktivasi ReLU
Menambahkan output layer sebanyak 1 dengan fungsi aktivasi ReLU
Melihat rancangan network
Menyusun RNN
```

Gambar 4. 11 Mendefinisikan Model

4.3.2 Training Model LSTM

Langkah selanjutnya adalah melakukan proses *training* dengan fungsi model yang telah dibuat. Proses *training* akan menggunakan fungsi ‘fit()’. Proses *training* ini menerapkan Algoritma *Backpropagation Through Time* yang diselaraskan dengan algoritma optimasi dan pengujian kesalahan yang sudah didefinisikan pada fungsi model sebelumnya. Algoritma *Backpropagation Through Time* akan melakukan *training* sejumlah dengan *epoch*. *Epoch* merupakan jumlah model akan berputar pada data. Semakin tinggi nilai *epoch* maka performa yang dihasilkan akan semakin baik (Nayoan, 2019). Jumlah *epoch* yang digunakan dalam penelitian ini adalah adalah 100 *epoch*.

```
Menjalankan RNN ke data training menggunakan perulangan sebanyak 100x
```

Gambar 4. 12 Training Model LSTM

4.3.3 Evaluasi Model LSTM

Langkah terakhir adalah melakukan evaluasi terhadap model yang digunakan dengan melakukan pengujian tingkat error. Pengukuran akurasi hasil peramalan dilakukan dengan

menggunakan metode *Root Mean Squared Error* (RMSE). Hasil yang ditampilkan merupakan besarnya perbedaan antara hasil estimasi dengan hasil yang diestimasi. Evaluasi model dilakukan menggunakan *Library Sklearn* dan *Math*.

```
memanggil library sklearn.metrics dan mengimport library mean_squared_error
memanggil library math dan mengimport sqrt
menghitung nilai mean squared error LSTM
menghitung nilai root dari nilai mean squared error LSTM
```

Gambar 4. 13 Evaluasi Model LSTM

4.4 Pemodelan ARIMA

ARIMA merupakan suatu metode peramalan yang menggunakan data masa lampau sebagai objek pengamatan terhadap suatu peristiwa ataupun kejadian yang ada pada data tersebut (Makridakis, 1998). Dalam pemodelan ARIMA, peneliti menggunakan *Library Pmdarima* yang merupakan *library* statistik yang biasanya digunakan untuk melakukan analisis *time series*. *Library Pmdarima* sangat berguna dalam membangun model ARIMA karena dapat mengoptimalkan penemuan parameter (p, d, q) yang paling cocok.

4.4.1 Identifikasi Model

Langkah pertama yang harus dilakukan dalam pemodelan ARIMA ialah melakukan identifikasi model. Hal ini bertujuan untuk mengetahui stasioneritas kondisi data *time series* yang akan dilakukan pemodelan prediksi. Identifikasi model dilakukan dengan melakukan *Augmented Dickey-Fuller (ADF) Test*. Stasioneritas data diketahui dengan melakukan perbandingan antara nilai *p-value* pada hasil *ADF test* dengan nilai α (0,05). Apabila nilai *p-value* lebih kecil dari α maka data tidak mengandung unit akar atau dapat dikatakan data sudah stasioner. Sedangkan jika nilai *p-value* lebih besar dari α maka data tersebut mengandung unit akar atau dapat dikatakan data tidak stasioner.

```
Memanggil library adfuller
datatest = menerapkan library adfuller pada data (lag otomatis = AIC)
dataoutput = test statistik, p-value, jumlah lag, jumlah observasi
untuk nilai pada data :
    dataoutput = critical value untuk 1%, 5%, 10%
cutoff = 0.05
apabila p-value < cutoff :
    tampilkan data sudah stasioner
kalau tidak
    tampilkan data belum stasioner
```

Gambar 4. 14 Identifikasi Model

4.4.2 Differencing

Langkah kedua ialah melakukan *differencing*. *Differencing* dilakukan dikarenakan hasil *ADF Test* menunjukkan bahwa data tidak stasioner dengan menunjukkan $ADF Test > p\text{-value}$. Cara untuk mengubah data yang bersifat *non* stasioner menjadi stasioner adalah dengan cara melakukan *differencing* (Marvilla, 2013). *Differencing* merupakan proses perubahan data yang tidak stasioner menjadi stasioner dengan cara menghitung nilai selisih data observasi dengan data observasi sebelumnya (Asa, 2017). Setelah dilakukan *differencing* maka diperlukan *ADF Test* lagi untuk mengetahui apakah data yang telah dilakukan satu kali *differencing* sudah stasioner atau belum, jika data masih belum stasioner maka diperlukan *differencing* kembali hingga data stasioner.

```
melakukan differensiasi orde satu
menghapus bagian yang kosong pada hasil differensiasi order satu
melakukan uji stasioner hasil differensiasi orde satu
```

Gambar 4. 17 *Differencing*

4.4.3 Estimasi Parameter Model

Langkah ketiga adalah melakukan pemilihan estimasi parameter model ARIMA. Pemilihan model ARIMA menggunakan fungsi Auto ARIMA yang secara otomatis akan menemukan model ARIMA yang tepat. Pada dasarnya Auto ARIMA menerapkan penggunaan model *selection criteria* yaitu *Akaike Information Criterion* (AIC) dalam memilih model regresi terbaik. Cara kerja Auto ARIMA adalah berdasarkan pendekatan nilai AIC dengan melihat nilai AIC yang paling terkecil untuk diambil menjadi suatu model parameter dengan *output* (p, d, q).

```
Mencari model dengan Auto ARIMA
Menampilkan ringkasan model
```

Gambar 4. 20 Mengestimasi Parameter Model

4.4.4 Training Model ARIMA

Langkah keempat adalah melakukan *training model* ARIMA. Berdasarkan hasil estimasi parameter model yang sebelumnya telah dilakukan akan dilakukan pemodelan dengan order (5, 1, 0) terhadap data *training* yang telah ditentukan.

```
model arima = melatih data train menggunakan ARIMA dengan model (5,1,0)
hasilarima = melakukan fungsi penyesuaian pada model arima
```

Gambar 4. 21 Training Model ARIMA

4.4.5 Evaluasi Model ARIMA

Langkah terakhir adalah melakukan evaluasi terhadap model yang digunakan dengan melakukan pengujian tingkat error. Pengukuran akurasi hasil peramalan dilakukan dengan menggunakan metode *Root Mean Squared Error* (RMSE). Hasil yang ditampilkan merupakan besarnya perbedaan antara hasil estimasi dengan hasil yang diestimasi

```
memanggil library sklearn.metrics dan mengimport library mean_squared_error
memanggil library math dan mengimport sqrt
menghitung nilai mean squared error ARIMA
menghitung nilai root dari nilai mean squared error ARIMA
```

Gambar 4. 22 Evaluasi Model ARIMA

BAB V

HASIL DAN PEMBAHASAN

5.1 Analisis *Pre-Processing* Data

Pada penelitian ini, *pre-processing* data dilakukan untuk menghilangkan *missing value* maupun inkonsisten data. Hal ini disebabkan oleh adanya *human error* yang menyebabkan kesalahan dalam entri data ataupun ketika proses *update* data. Selain itu, hal ini juga disebabkan oleh pengendalian sistem basis data yang masih belum baik dikarenakan masih dalam upaya perkembangan. Berikut merupakan hasil analisis dari *pre-processing* yang telah dilakukan :

5.1.1 Analisis *Data Cleaning*

Dalam suatu data tentunya akan ditemukan *missing value*. *Missing value* merupakan hilangnya informasi terkait sebuah data. Pada data historis permintaan produk PT. TaniHub Indonesia, *missing value* disebabkan oleh kesalahan pada entri data. Beberapa *commercial specialist* tidak melakukan *input* data dengan lengkap. Tentunya, hal ini akan menjadi masalah karena jumlah data yang digunakan adalah sangat besar sehingga menyebabkan keakuratan suatu data menjadi melemah dan kualitas data yang akan digunakan menjadi jelek. Oleh karena itu diperlukan *data cleaning* untuk menghilangkan *missing value* dengan metode *listwise deletion*. Penggunaan metode *listwise deletion* didukung dikarenakan merupakan metode yang cocok dengan permasalahan yang ada yaitu digunakan pada data sampel yang cukup besar dan memiliki pola *missing value* yang acak (Acock, 2005).

5.1.2 Analisis *Data Selection*

Pada *data selection* terdapat beberapa proses yang dilakukan, diantaranya adalah sebagai berikut :

1. *Data Farm Produce*

Dalam data historis permintaan PT TaniHub Indonesia memiliki beberapa entitas dan atribut di dalamnya. Adanya entitas dan atribut tersebut adalah untuk membantu *user* dalam abstraksi pada suatu sistem basis data. Selain itu, hal ini mempermudah ketika akan melakukan suatu analisis baik dari produk, *client* maupun *revenue*. Salah satu contoh yang ada pada data historis PT TaniHub Indonesia adalah pada entitas *product info* dengan atribut *farm category*. Entitas *product info* memuat segala informasi terkait spesifikasi ataupun kategori produk yang dijual. Sedangkan untuk atribut *farm category* merupakan suatu kolom yang memuat kategori produk yang diklasifikasikan menjadi *farm produce* yaitu produk yang dihasilkan oleh para petani dan *non farm* yaitu produk manufaktur. Berdasarkan batasan masalah yang telah ditentukan, penelitian ini hanya dilakukan pada *perishable product* yang berarti tergolong dalam *farm produce*. Oleh karena itu maka diperlukan *data selection* pada kategori *farm produce*.

2. Data Vegetable

Setelah dilakukan *data selection* pada atribut *farm produce* maka selanjutnya dilakukan *data selection* pada atribut *product department* yang memuat departemen produk seperti : *dairy frozen, fish & seafood, drink, food, fruit, home cleaning, household, meat, non food, poultry, vegetable, saprodi* dan *others*. Adanya atribut *product department* digunakan untuk mengklasifikasikan jenis produk yang digunakan untuk memudahkan pembagian kategori baik dalam basis data dan juga *interface website* dan aplikasi. Berdasarkan batasan masalah yang telah ditentukan, penelitian ini hanya dilakukan pada *perishable product* yang berarti tergolong dalam *farm produce*. Oleh karena itu maka diperlukan *data selection* pada kategori *farm produce*.

5.1.3 Analisis Data Transformation

Pada *data transformation* terdapat beberapa proses yang dilakukan, diantaranya adalah sebagai berikut :

1. Transformasi Unit Name

Pada data historis penjualan PT TaniHub Indonesia memiliki beberapa satuan produk diantaranya : Kilogram (KG), Pack dan PCS. Hal ini dikarenakan sebagai salah satu strategi

dalam memenuhi permintaan konsumen yang ingin membeli dalam jumlah tertentu. Pada *unit name* 'Pack' produk dikemas dengan berat tertentu dimulai dari yang paling kecil yaitu 50 gram hingga 900 gram. Sedangkan pada *unit name* 'PCS' produk dijual secara jumlah satuan. Terakhir, pada *unit name* 'Kilogram' produk dijual secara utuh dalam satuan kilogram dimulai dari 1 KG. Berdasarkan inkonsisten *unit name* tersebut tentunya akan menghambat ketika akan dilakukan pemodelan karena memiliki banyak variabel. Oleh karena itu diperlukan transformasi *unit name* menjadi agar seluruh satuan menjadi seragam. Satuan *unit name* akan ditransformasikan menjadi gram. Pemilihan satuan gram dikarenakan untuk menghindari adanya nilai desimal yang akan terjadi apabila di transformasi dalam kilogram.

2. Tranformasi Nama Produk

Inkosisten data yang terjadi pada nama produk dikarenakan penamaan produk pada PT TaniHub Indonesia disertai dengan berat produk ataupun merk produk, seperti : Bawang Putih 1 KG. Alasan penamaan produk tersebut digunakan untuk identifikasi produk yang sama namun dikemas dengan *product size* tertentu. Selain itu, terdapat inkonsisten penulisan kapital pada beberapa nama produk. Terdapat beberapa nama yang tidak ditulis kapital pada awal nama produk sedangkan beberapa nama lainnya dituliskan kapital pada awal nama produk tersebut. Hal ini disebabkan karena adanya *human error* ketika melakukan entri data produk dalam *database*. Berdasarkan kasus-kasus tersebut tentunya inkonsisten dalam penamaan nama produk akan menghambat dalam melakukan pemodelan. Oleh karena itu, diperlukan transformasi nama produk dengan hanya menyajikan 1 nama produk saja. Setelah dilakukan transformasi nama produk yang sebelumnya berjumlah 486 nama produk menjadi 84 nama produk.

Tabel 5. 1 Hasil Transformasi *Product Name*

No	Nama Produk
1	Adas
2	Arugula
3	Basil

No	Nama Produk
4	Bawah Putih
5	Bawang Bombay
6	Bawang Merah
7	Bayam
8	Beetroot
9	Brokoli
10	Buncis
11	Bunga Sedap Malam
12	Butterhead Lettuce
13	Cabai Hijau
14	Cabai Merah
15	Cabai Rawit Hijau
16	Cabai Rawit Merah
17	Caisim
18	Cengkeh
19	Daun Bawang
20	Daun Ketumbar
21	Daun Pisang
22	Daun Rosemary
23	Edamame
24	Horenzo
25	Jagung
26	Jahe
27	Jamur Kering
28	Jamur Shimeji
29	Jamur Shitake
30	Jinten
31	Kacang Kedelai
32	Kacang Panjang

No	Nama Produk
33	Kailan
34	Kale
35	Kangkung
36	Kapri
37	Kapulaga
38	Kayu Manis
39	Kelapa Parut
40	Kemangi
41	Kembang Kol
42	Kemiri
43	Kencur
44	Kentang
45	Ketumbar
46	Kol
47	Kulit Melinjo
48	Kunyit
49	Labu Kabocha
50	Labu Parang
51	Labu Siam
52	Lada Hitam
53	Lada Putih
54	Lengkuas
55	Lettuce Head
56	Lettuce Romaine
57	Lobak
58	Lolorosa
59	Mizuna
60	Naibai
61	Nangka Muda

No	Nama Produk
62	Okra
63	Oregano
64	Pakcoy
65	Pala
66	Parsley
67	Petai
68	Pumpkin Butternut
69	Radish
70	Salad
71	Sawi Putih
72	Selada Keriting
73	Seledri
74	Sereh
75	Shisho
76	Singkong
77	Siomak
78	Tauge
79	Temulawak
80	Terong
81	Timun
82	Tomat
83	Ubi
84	Wortel

5.1.4 Analisis Penjumlahan Harian

Perishable product adalah produk yang memiliki umur tahan yang sangat pendek. Untuk mengurangi adanya *food waste* yang disebabkan oleh *perishable product* maka dilakukan

prediksi secara harian. Hasil dari penjumlahan harian data jumlah *dataframe* menjadi 364 data.

5.1.5 Analisis Persiapan Data

Normalisasi data digunakan sebagai suatu teknik analisis data guna mengorganisasikan atribut-atribut data dengan mengelompokkan data sehingga terbentuk suatu entitas yang stabil, fleksibel dan *non-redundant* (Wahidin, 2021). Pada hasil analisis statistik diketahui bahwa nilai minimal dan maksimal pada data historis permintaan PT TaniHub Indonesia memiliki jangkauan yang sangat jauh yaitu 2.250 – 2.233.010. Berdasarkan hal tersebut maka diperlukan normalisasi data agar data tersebut ada pada *range* yang sama. Pada penelitian ini dilakukan *feature scaling* dengan *range* (0,1) yang berarti seluruh data akan dijadikan pada range 0 – 1. Selanjutnya akan dilakukan pembagian data *training* dan data *test*. Data *training* adalah data yang akan digunakan untuk melatih algoritma sedangkan data *test* adalah data yang digunakan untuk mengetahui performa algoritma yang sudah dilatih menggunakan data *training*. Pembagian data *training* dan data *test* terbagi dalam 80% data *training* dan 20% data *test*. Hal ini didasari oleh studi yang dilakukan oleh Gholami (2018) yang mengkaji bahwa komposisi data *training* 80% dan data *test* 20% memiliki performa yang baik untuk pemodelan prediksi dibandingkan dengan komposisi data *training* 70% dan data *test* 30%.

5.2 Analisis Pemodelan ARIMA

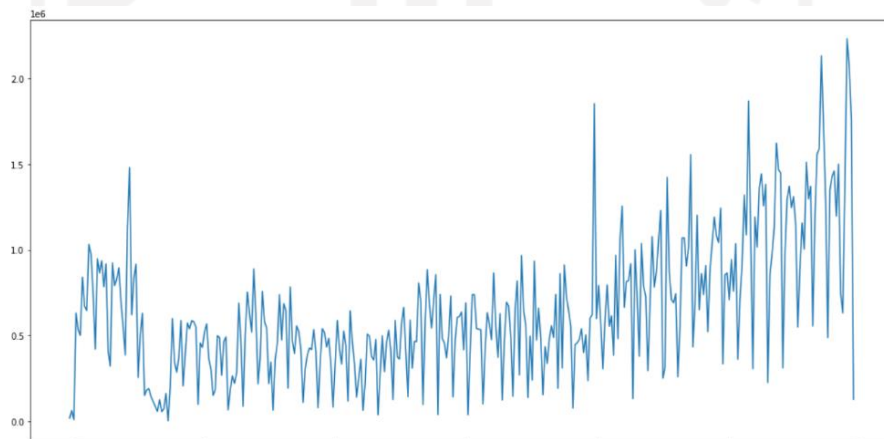
Pada penelitian ini dilakukan analisis deskriptif untuk mengetahui gambaran umum dari data permintaan produk pada kategori ‘*vegetable*’. Adapun hasil dari analisis deskriptif adalah sebagai berikut :

Tabel 5. 2 Analisis Deskriptif

Deskripsi	Nilai
Rata-rata	642.572

Standar Deviasi	399.989
Nilai Minimal	2.250
Nilai Maximum	2.233.010
Kuartil Pertama	375.187
Median	558.150
Kuartil Ketiga	857.687

Berdasarkan Tabel 5.1 dapat dilihat bahwa jumlah berat harian produk pada kategori '*vegetable*' memiliki nilai minimal 2.250 gram dan nilai maksimal 2.233.010 gram selama periode 29 Desember 2020 – 31 Desember 2021. Jumlah berat produk memiliki nilai rata-rata 642.572. Hal tersebut membuktikan terdapat jangkauan nilai yang jauh sehingga terjadi naik turun jumlah permintaan produk harian. Untuk lebih jelasnya dapat dilihat melalui plot permintaan produk '*vegetable*' berikut :



Gambar 5. 1 Plot Data

Pada Gambar 5.1 diketahui bahwa data terlihat mengalami naik turunnya penjualan (fluktuatif). Berdasarkan hal tersebut dapat diketahui bahwa data terindikasi tidak

stasioner. Untuk lebih memastikan kestasioneran data, peneliti menggunakan *Augmented Dickey-Fuller Test* (ADF Test) untuk mengetahui derajat stasioneritas dari seluruh variabel. Adapun *output* dari ADF Test adalah sebagai berikut :

Tabel 5. 3 Hasil ADF Test

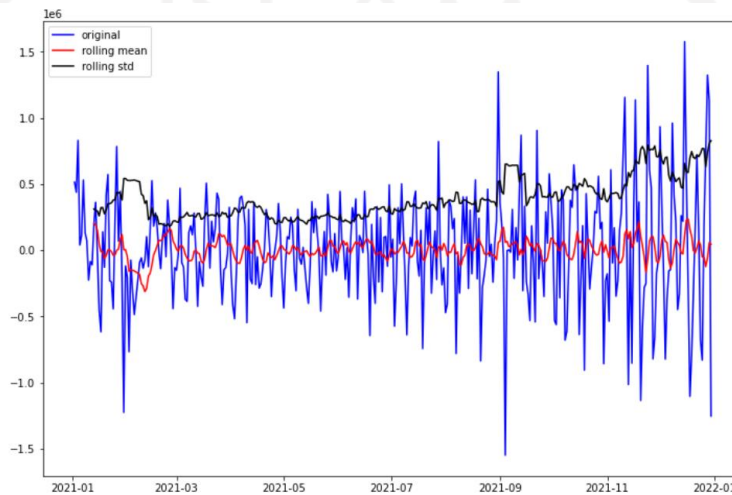
Parameter	Nilai
<i>p-value</i>	0,8833
<i>Lag used</i>	14
<i>Number of Observation</i>	364

Berdasarkan Tabel 5.3 dapat dilihat hasil hipotesis adalah sebagai berikut :

- i. Hipotesis
 H_0 : Data mengandung unit akar (data tidak stasioner)
 H_1 : Data tidak mengandung unit akar (data stasioner)
- ii. Tingkat Signifikansi
 $\alpha = 5\%$ atau 0,05
- iii. Statiska Uji
 $p\text{-value} = 0,883$
- iv. Daerah Kritis
 $p\text{-value} < \alpha$ maka H_0 ditolak
- v. Keputusan
 Karena nilai $p\text{-value} > \alpha$ maka gagal tolak H_0
- vi. Kesimpulan
 Data mengandung unit akat (data tidak stasioner)

Berdasarkan ADF Test tersebut dapat disimpulkan bahwa data mengandung unit akar atau bisa dikatakan data tidak stasioner. Menurut Widarjono (2018), data yang tidak stasioner akan mempengaruhi kinerja model yang diestimasi. Hal ini dikarenakan jika data tidak stasioner akan terjadi regresi semu yaitu regresi antara variabel independen dengan variabel dependen yang sebenarnya tidak ada hubungan korelasi namun jika di analisis

akan menghasilkan variabel yang berkorelasi (Rosadi, 2011). Berdasarkan hal tersebut maka data harus di stasionerkan terlebih dahulu. Cara yang dapat digunakan untuk menstasionerkan data adalah dengan melakukan pembedaan (*differencing*). Konsep *differencing* adalah dengan mengurangkan antara data pengamatan saat ini dengan data pengamatan sebelumnya. Berikut merupakan hasil dari data yang telah dilakukan *differencing* :



Gambar 5. 2 Hasil *Differencing*

Tentunya setelah dilakukan *differencing* diperlukan ADF Test lagi untuk mengetahui apakah data yang sudah stasioner atau belum. Berdasarkan Tabel 5.4 dapat disimpulkan bahwa hasil uji adalah sebagai berikut :

Tabel 5. 4 Hasil ADF Test *Differencing* Orde Satu

Parameter	Nilai
<i>p-value</i>	0,004
<i>Lag used</i>	14
<i>Number of Observation</i>	364

i. Hipotesis

H₀ : Data mengandung unit akar (data tidak stasioner)

H₁ : Data tidak mengandung unit akar (data stasioner)

- ii. Tingkat Signifikansi
 $\alpha = 5\%$ atau 0,05
- iii. Statiska Uji
 $p\text{-value} = 0,004$
- iv. Daerah Kritis
 $p\text{-value} < \alpha$ maka H_0 ditolak
- v. Keputusan
Karena nilai $p\text{-value} < \alpha$ maka H_0 ditolak
- vi. Kesimpulan
Data tidak mengandung unit akar (data stasioner)

Setelah dipastikan data stasioner selanjutnya dilakukan pemodelan ARIMA dengan menggunakan variabel (p,d,q). Variabel yang digunakan merupakan variabel yang dihasilkan dari estimasi parameter model menggunakan fungsi Auto ARIMA. Dalam Auto ARIMA akan memilih model berdasarkan nilai AIC yang paling kecil. Semakin kecil nilai AIC maka akan semakin baik juga performa model tersebut (Darvi, 2019). Pada Gambar 5.3 terdapat beberapa model yang dipilih. Namun, jika dilihat dari nilai AIC pada model ARIMA (5,1,0) memiliki nilai AIC yang paling kecil yaitu 433.430 dengan *time* 3,18 sec.

```

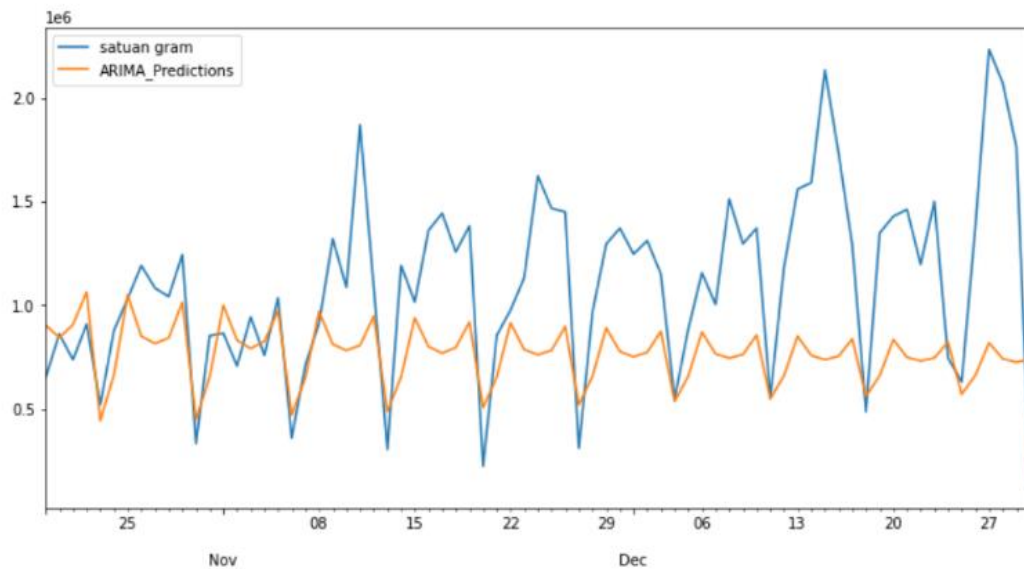
Performing stepwise search to minimize aic
ARIMA(2,1,2)(0,0,0)[0] intercept : AIC=inf, Time=17.76 sec
ARIMA(0,1,0)(0,0,0)[0] intercept : AIC=441902.752, Time=1.44 sec
ARIMA(1,1,0)(0,0,0)[0] intercept : AIC=437282.520, Time=1.11 sec
ARIMA(0,1,1)(0,0,0)[0] intercept : AIC=inf, Time=8.53 sec
ARIMA(0,1,0)(0,0,0)[0] : AIC=441900.752, Time=0.48 sec
ARIMA(2,1,0)(0,0,0)[0] intercept : AIC=435500.845, Time=2.23 sec
ARIMA(3,1,0)(0,0,0)[0] intercept : AIC=434483.461, Time=3.43 sec
ARIMA(4,1,0)(0,0,0)[0] intercept : AIC=433875.345, Time=4.49 sec
ARIMA(5,1,0)(0,0,0)[0] intercept : AIC=433432.898, Time=4.47 sec
ARIMA(5,1,1)(0,0,0)[0] intercept : AIC=inf, Time=40.25 sec
ARIMA(4,1,1)(0,0,0)[0] intercept : AIC=inf, Time=37.48 sec
ARIMA(5,1,0)(0,0,0)[0] : AIC=433430.898, Time=3.18 sec
ARIMA(4,1,0)(0,0,0)[0] : AIC=433873.345, Time=2.32 sec
ARIMA(5,1,1)(0,0,0)[0] : AIC=inf, Time=22.60 sec
ARIMA(4,1,1)(0,0,0)[0] : AIC=inf, Time=18.67 sec

Best model: ARIMA(5,1,0)(0,0,0)[0]

```

Gambar 5. 3 Hasil Auto ARIMA

Selanjutnya akan dilakukan peramalan pada model (5,1,0). Terlihat pada Gambar 5.4 plot data hasil peramalan yang memuat hasil prediksi dengan data aktual. Plot data tersebut dapat dilihat masih adanya gap antara hasil peramalan dengan data aktual. Pada hasil peramalan 25 November – 8 Desember hasil peramalan sudah hampir mendekati data aktual. Namun, setelahnya hasil peramalan menghasilkan pola yang berulang dan cenderung jauh dari data aktual. Hal ini menunjukkan hasil peramalan dengan ARIMA tidak membentuk suatu pola yang linier. Zufar (2018) telah mengkaji bahwa penyebab dari data aktual dengan hasil peramalan yang cenderung menjauh dikarenakan p -value pada lag tertentu lebih dari α sehingga masih ada yang belum stasioner. Hal ini membuktikan bahwa masih terdapat korelasi antara model yang digunakan dengan $mean$ sama dengan nol ataupun varian yang konstan. Berdasarkan hal tersebut maka diperlukan melakukan uji diagnostik yaitu *uji white noise* pada model untuk mengetahui apakah masih ada korelasi atau tidak.



Gambar 5. 4 Plot Hasil ARIMA

5.3 Analisis Pemodelan LSTM

Dalam pemodelan LSTM membutuhkan beberapa parameter yang akan digunakan untuk menghasilkan suatu model yang baik. Pada penelitian ini menggunakan model *class*

sequential dengan *optimizer* Adam. Pemilihan Adam didukung dikarenakan merupakan salah satu algoritma *optimizer* yang mudah diimplementasikan, efisien, tidak membutuhkan banyak memori serta dapat menyelesaikan berbagai macam permasalahan yang berhubungan dengan parameter maupun data (Kingma, 2014). Selain itu penggunaan *activation function layer* juga menjadi salah satu parameter yang penting. Pada penelitian ini menggunakan fungsi aktivasi ReLU. Hal ini dikarenakan fungsi aktivasi yang memberikan hasil akurasi terbaik adalah ReLU (Hakim, 2020)

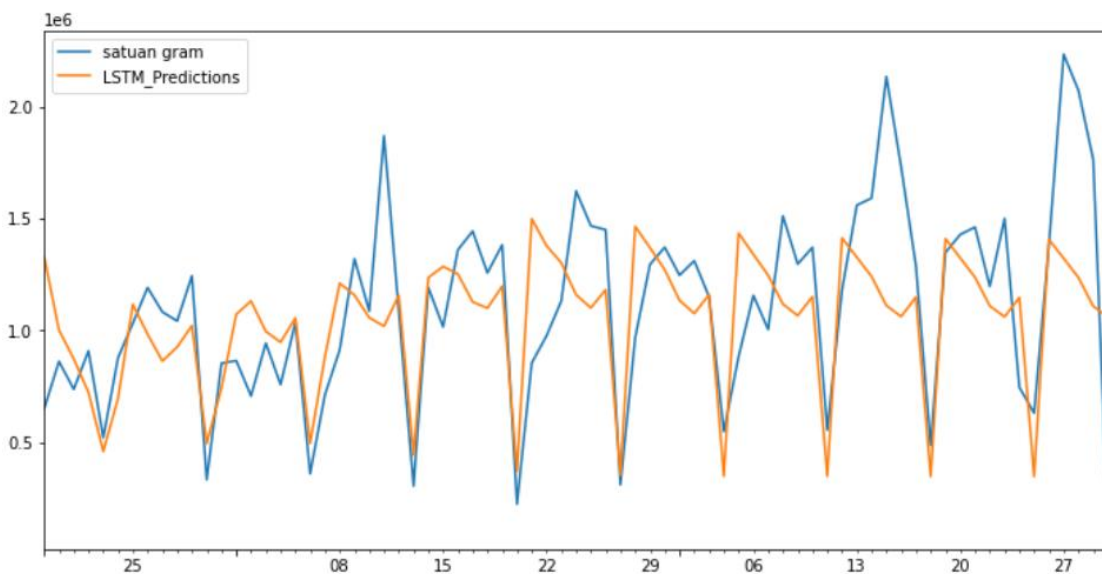
Penentuan jumlah *neuron* pada *hidden layer* dan *epoch* juga menjadi suatu parameter yang penting. Dalam suatu peramalan jumlah *neuron* pada *hidden layer* dapat mempengaruhi hasil akurasi, semakin banyak jumlah *neuron* belum tentu memiliki hasil peramalan dengan akurasi yang tinggi (Putra, 2019). Oleh karena itu pada penelitian ini dilakukan beberapa uji coba jumlah *neuron* pada *hidden layer* serta *epoch* untuk memperoleh pemodelan dengan performa yang baik.

Tabel 5. 5 Hasil Uji Coba *Hidden Layer* & *Epoch*

No	<i>Hidden Layer</i>	<i>Epoch</i>	RMSE
1	8	100	358.633
2		300	761.459
3		500	672,764
4		700	912.432
5		1000	798.539
6	32	100	562.237
7		300	687.725
8		500	823.029
9		700	533.132
10		1000	473.179
11	64	100	594.485
12		300	631.613
13		500	675.821
14		700	511.949

No	Hidden Layer	Epoch	RMSE
15		1000	522.668

Berdasarkan hasil uji coba pada Tabel 5.5 maka dapat disimpulkan bahwa dengan 8 neuron pada hidden layer dan 100 epoch akan memperoleh pemodelan dengan performa yang baik berdasarkan hasil RMSE yang paling kecil yaitu 358.633. Pada Gambar 5.5 terlihat plot data hasil peramalan dengan menggunakan metode LSTM dan parameter-parameter yang telah ditentukan. Pada plot data tersebut diketahui bahwa hasil prediksi tidak memiliki jarak yang cukup jauh dibandingkan hasil prediksi dengan ARIMA sebelumnya. Terlihat pada plot 25 November 2021 – 22 Desember 2021 model memiliki performa yang baik dalam melakukan prediksi dengan memiliki grafik yang tidak signifikansi jauh. Namun, setelah itu model memiliki pola yang berulang.



Gambar 5. 5 Plot Hasil LSTM

5.4 Analisis Evaluasi Model ARIMA & LSTM

Dalam suatu peramalan terdapat 2 kesalahan yang dapat mempengaruhi keputusan yaitu kesalahan dalam memilih metode peramalan dan kesalahan dalam mengevaluasi metode peramalan tersebut (Andini, 2016). Metode yang telah dilakukan peramalan akan dilakukan evaluasi performa model dengan menggunakan metode pengujian tingkat error, *Root Mean*

Squared Error (RMSE). Metode evaluasi yang dilakukan pada dua metode tersebut dilakukan dengan cara yang sama yaitu, dengan membandingkan antara data asli permintaan dengan hasil peramalan.

Tabel 5. 6 Hasil RMSE

No	Model	Nilai RMSE
1	ARIMA	498.337
2	LSTM	358.633

Menurut Astur (2021), apabila ingin melihat suatu nilai RMSE besar atau tidak bisa dilihat dari *range* datanya seperti contoh jika *range* data 0 hingga 1000 dan RMSE 0,7 maka tingkat erornya dapat dikategorikan kecil namun, apabila *range* data 0 hingga 1 tingkat erornya tidak dapat dikategorikan kecil lagi. Pada penelitian ini memiliki jangkauan *range* data aktual yang sangat tinggi yaitu, 2.250 hingga 2.230.010. Pada Tabel 5.6 diketahui bahwa nilai RMSE metode ARIMA adalah 498.337 dan nilai RMSE metode LSTM adalah 358.633. Hal ini dapat disimpulkan bahwa metode LSTM memiliki kesalahan prediksi yang lebih kecil dibanding metode ARIMA. Semakin kecil nilai error yang dihasilkan maka performa akurasi model tersebut semakin bagus (Raharja, 2010). Berdasarkan hal tersebut, maka LSTM merupakan metode yang lebih cocok untuk digunakan dalam melakukan pemodelan prediksi pada *perishable product*.

Selanjutnya dilakukan peramalan untuk menguji performa metode LSTM. Peramalan dilakukan pada produk Bayam yang merupakan produk dengan jumlah *demand* yang paling tinggi yaitu sebanyak 10.504 transaksi dalam jangka waktu 29 Desember 2020 – 30 Desember 2021. Peramalan dilakukan untuk Bulan Januari 2022

Tabel 5. 7 Hasil Peramalan Produk Bayam

No	Actual (Gram)	LSTM (Gram)	MAPE
1	18,000	20,924	16%
2	4,000	4,735	10%
3	28,400	22,608	20%

No	Actual (Gram)	LSTM (Gram)	MAPE
4	4,000	3,524	12%
5	5,000	3,964	21%
6	26,800	29,001	8%
7	12,000	9,905	17%
8	13,400	10,685	20%
9	6,000	5,333	11%
10	27,200	22,346	18%
11	4,000	5,923	48%
12	9,000	7,837	13%
13	26,800	33,203	24%
14	6,000	9,587	60%
15	10,000	10,557	6%
16	8,000	9,992	25%
17	26,250	26,738	2%
18	1,000	1,376	38%
19	23,000	28,473	24%
20	8,500	11,272	33%
21	10,600	9,297	12%
22	2,000	1,946	3%
23	16,800	9,161	45%
24	4,000	4,264	7%
25	11,000	15,986	45%
26	3,400	3,526	4%
27	4,000	5,061	27%
28	21,000	19,334	8%
29	21,000	16,408	22%
30	11,000	9,623	13%
31	27,100	23,489	13%
		Rata-Rata	20%

Tabel 5.7 menunjukkan hasil peramalan untuk produk *Baby Bayam* dengan metode LSTM. Pada peramalan *single product* ini menggunakan evaluasi pengujian tingkat error yaitu *Mean Absolute Percentage Error* (MAPE). Penggunaan MAPE untuk mengetahui persentase perbedaan antara data aktual dengan hasil peramalan. Berdasarkan tabel tersebut diketahui rata-rata hasil pengujian tingkat error adalah 20% dengan hanya memiliki satu hasil performa peramalan diatas 60% yaitu yang ada pada 14 Januari 2022. Suatu model memiliki hasil peramalan yang sangat akurat apabila nilai MAPE < 10%, dan jika MAPE 10% - 20% dapat dikategorikan memiliki hasil peramalan yang baik, sedangkan jika nilai MAPE 21% - 50% dapat dikategorikan memiliki hasil peramalan yang layak (cukup baik) dan jika nilai MAPE > 50% memiliki hasil peramalan yang tidak akurat (Lewis, 1982). Berdasarkan interpretasi tersebut maka dapat disimpulkan bahwa peramalan *single product* pada produk *Baby Bayam* dengan metode LSTM memiliki hasil peramalan yang baik.

5.5 Analisis Pengurangan *Food Waste*

Setelah ditentukan metode peramalan yang memiliki akurasi yang baik maka selanjutnya dilakukan perbandingan rata-rata *food waste* yang dihasilkan per harinya antara menggunakan metode intuisi setiap *commercial specialist* dengan menggunakan peramalan metode LSTM. Dalam perhitungan dibawah ini dilakukan pengurangan antara data hasil prediksi berdasarkan intuisi commercial dengan menggunakan data aktual permintaan produk bayam untuk menghitung jumlah *food waste* yang terjadi.

Tabel 5. 8 Jumlah *Food Waste* Dengan Menggunakan Metode Intuisi *Commercial*

No	Intuisi <i>Commercial</i> (Gram)	Actual (Gram)	<i>Food Waste</i> (Gram)
1	25,000	18,000	7,000
2	25,000	4,000	21,000
3	25,000	28,400	0
4	25,000	4,000	21,000
5	25,000	5,000	20,000

No	Intuisi Commercial (Gram)	Actual (Gram)	Food Waste (Gram)
6	25,000	26,800	0
7	25,000	12,000	13,000
8	25,000	13,400	11,600
9	25,000	6,000	19,000
10	25,000	27,200	0
11	25,000	4,000	21,000
12	25,000	9,000	16,000
13	25,000	26,800	0
14	25,000	6,000	19,000
15	25,000	10,000	15,000
16	25,000	8,000	17,000
17	25,000	26,250	0
18	25,000	1,000	24,000
19	25,000	23,000	2,000
20	25,000	8,500	16,500
21	25,000	10,600	14,400
22	25,000	2,000	23,000
23	25,000	16,800	8,200
24	25,000	4,000	21,000
25	25,000	11,000	14,000
26	25,000	3,400	21,600
27	25,000	4,000	21,000
28	25,000	21,000	4,000
29	25,000	21,000	4,000
30	25,000	11,000	14,000
31	25,000	27,100	0
		Rata-Rata	12,525

Berdasarkan Tabel 5.8 diketahui bahwa dalam menggunakan metode yang mengandalkan intuisi setiap *commercial specialist* dalam memprediksi suatu produk dalam satu bulan memiliki nilai yang konstan yaitu sebanyak 25.000 Gram atau 25 Kilogram. Alhasil food waste yang dihasilkan juga cenderung besar dikarenakan prediksi permintaan tidak disesuaikan dengan kondisi ataupun fenomena permintaan *customer*. Sehingga *food waste* yang dihasilkan dengan hanya menggunakan intuisi setiap *commercial specialist* dalam memprediksi jumlah permintaan produk memiliki rata-rata per harinya mencapai 12.525 Gram atau 12,5 Kilogram.

Tabel 5. 9 Jumlah *Food Waste* Dengan Menggunakan Metode LSTM

No	LSTM (Gram)	Actual (Gram)	Food Waste (Gram)
1	20,924	18,000	2,924
2	4,735	4,000	735
3	22,608	28,400	0
4	3,524	4,000	0
5	3,964	5,000	0
6	29,001	26,800	0
7	9,905	12,000	2,201
8	10,685	13,400	0
9	5,333	6,000	0
10	22,346	27,200	0
11	5,923	4,000	1,923
12	7,837	9,000	0
13	33,203	26,800	6,403
14	9,587	6,000	3,587
15	10,557	10,000	557
16	9,992	8,000	1992
17	26,738	26,250	488
18	1,376	1,000	376
19	28,473	23,000	5,473

No	LSTM (Gram)	Actual (Gram)	Food Waste (Gram)
20	11,272	8,500	2,772
21	9,297	10,600	0
22	1,946	2,000	0
23	9,161	16,800	0
24	4,264	4,000	126
25	15,986	11,000	4,986
26	3,526	3,400	126
27	5,061	4,000	1,061
28	19,334	21,000	0
29	16,408	21,000	0
30	9,623	11,000	0
31	23,489	27,100	0
		Rata-Rata	1,152

Berdasarkan Tabel 5.9 diketahui bahwa *food waste* yang dihasilkan dengan menggunakan peramalan metode LSTM dalam memprediksi jumlah permintaan produk memiliki rata-rata per harinya mencapai 1.152 Gram atau 1,1 Kilogram untuk produk bayam. Hal ini membuktikan bahwa adanya pengurangan jumlah *food waste* yang cukup signifikan antara menggunakan metode sebelumnya yaitu dengan hanya mengandalkan intuisi *commercial specialist* dan dengan menggunakan peramalan metode LSTM. Dengan demikian maka jumlah pengurangan *food waste* yang dapat diatasi dengan menggunakan metode peramalan LSTM mencapai 91%. Tentunya hal ini menandakan bahwa dengan menggunakan implementasi *machine learning* dengan metode peramalan LSTM dapat mengurangi *food waste* yang terjadi di PT TaniHub Indonesia.

BAB VI

PENUTUP

6.1 Kesimpulan

Berdasarkan penelitian yang telah dilakukan, berikut merupakan kesimpulan untuk menjawab rumusan masalah :

1. Dalam penelitian ini model prediksi permintaan yang berbasis *machine learning* untuk item *vegetable* pada produk *farm* dalam sektor *business-to-business* di PT. TaniHub Indonesia adalah dengan menggunakan komposisi data *training* 80% dan data *testing* 20%. Dalam pemodelan ARIMA menggunakan fungsi Auto ARIMA dengan model (5,1,0). Sedangkan untuk Pemodelan LSTM dilakukan dengan menggunakan *activation function layer* 'ReLU' dan *optimizer* 'Adam' dengan jumlah *layer* utama 64 *neuron*, 8 *neuron hidden layer*, 1 *neuron output layer* dan 100 *epoch*.
2. Metode yang memiliki tingkat error yang paling rendah sehingga dapat dijadikan sebagai rekomendasi metode untuk dapat mengurangi *food waste*. adalah LSTM dengan memiliki nilai RMSE 358.633 yang mampu mengurangi *food waste* mencapai 91%.

6.2 Saran

Berikut merupakan beberapa saran berdasarkan penelitian yang telah dilakukan :

1. Pada penelitian selanjutnya, pemodelan ARIMA dapat melakukan identifikasi terhadap residual yang masih memiliki nilai *p-value* pada lag tertentu yang masih memiliki hipotesis $p\text{-value} > \alpha$. Identifikasi residual dapat dilakukan dengan menggunakan uji *white noise* sehingga dapat mengetahui kestasioneran data secara merata.
2. Pada penelitian selanjutnya, pemodelan LSTM dapat melakukan uji coba menggunakan kombinasi parameter pada jumlah *layer*, *epoch*, dan *activation function layer* agar memperoleh model yang lebih baik.

DAFTAR PUSTAKA

- Acock, A. C. (2005). *Working With Missing Values*. *Journal of Marriage and family*, 67(4), 1012-1028.
- Agus Widarjono. 2018. *Ekonometrika Pengantar Dan Aplikasinya Disertai Panduan Eviews*. Edisi keli. Yogyakarta: UPP STIM YKPN Yogyakarta.
- Arunraj, N., Ahrens, D., Fernandes, M., & Müller, M. (2014). *Time series sales forecasting to reduce food waste in retail industry*. Rotterdam.
- Asa, M., Wibowo, S. S., & Sophia, E. (2017). Peramalan Jumlah Mahasiswa Baru Dengan Model *Autoregressive Integrated Moving Average* (Arima). *JIMP-Jurnal Informatika Merdeka Pasuruan*, 2(3).
- Ashfia, H. (2020). Penerapan Metode ARIMA Ensemble Pada Peramalan Inflasi di Provinsi Kalimantan (*Doctoral dissertation*, Institut Teknologi Kalimantan).
- C. D. Lewis, 1982. *Industrial and Business Forecasting Methods: A Practical Guide To Exponential Smoothing and Curve Fitting*. Butterworth-Heinemann
- Chen, F. L., & Ou, T. Y. (2008, October). *A neural-network-based forecasting method for ordering perishable food in convenience stores*. In *2008 Fourth International Conference on Natural Computation* (Vol. 2, pp. 250-254). IEEE.
- Da Veiga, C. P., Da Veiga, C. R. P., Catapan, A., Tortato, U., & Da Silva, W. V. (2014). *Demand forecasting in food retail: A comparison between the Holt-Winters and ARIMA models*. *WSEAS transactions on business and economics*, 11(1), 608-614.
- Fardani, D. P., Wuryanto, E., & Werdiningsih, I. (2015). Sistem Pendukung Keputusan Peramalan Jumlah Kunjungan Pasien Menggunakan Metode *Extreme Learning Machine* (Studi Kasus: Poli Gigi Rsu Dr. Wahidin Sudiro Husodo

- Mojokerto). *Journal of Information Systems Engineering and Business Intelligence*, 1(1), 33-40.
- Gholamy, A., Kreinovich, V., & Kosheleva, O. (2018). *Why 70/30 or 80/20 Relation Between Training and Testing Sets: A Pedagogical Explanation*.
- Hakim, A. A. A. (2020). Deteksi Defisiensi Unsur Hara Makro pada Tanaman Kopi Berdasarkan Karakteristik Gejala Visual Daun menggunakan MTCD dan JST (*Doctoral dissertation*, Universitas Muhammadiyah Malang).
- Kolker, A. (2011). *Forecasting Time Series. Healthcare Management Engineering: What Does This Fancy Term Really Mean?: The Use of Operations Management Methodology for Quantitative Decision Making in Healthcare Settings* (hal. 89). Milwaukee: Springer Science Business Media
- Kurhade, S., Kadam, A., Warekar, D., & Kamble, N. *Forecasting the Daily Sales of Perishable Food to Reduce Spoilage in Hypermarket*.
- Makridakis, S., Wheelwright, S.C. & Hyndman, R.J. (1998). *Forecasting Methods and Applications (3rd edition)*. United States of America: John Wiley & Sons, Inc.
- Marvillia, Bunga Lety . Pemodelan dan Peramalan Penutupan Harga Saham PT. TELKOM Dengan Metode ARCH-GARCH. Mathunesa: Jurnal Ilmiah Matematika, 2013, 2.1.
- Nayoan, R. A. N. (2019). Analisis Sentimen Berbasis Fitur Pada Ulasan Tempat Wisata Menggunakan Metode *Convolutional Neural Network* (CNN) (*Doctoral dissertation*, Universitas Islam Indonesia).
- Puchalsky, W., Ribeiro, G. T., da Veiga, C. P., Freire, R. Z., & dos Santos Coelho, L. (2018). *Agribusiness time series forecasting using Wavelet neural networks and metaheuristic optimization: An analysis of the soybean sack price and perishable products demand. International Journal of Production Economics*, 203, 174-189.

- Putra, B. B. W., Albar, M. A., & Irmawati, B. (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).
- R.Astur (<https://stats.stackexchange.com/users/20211/r-astur>), What are good RMSE values?, URL (version: 2016-04-26): <https://stats.stackexchange.com/q/56332>
- Raharja, A., Angraeni, W., & Vinarti, R. A. (2010). Penerapan Metode *Exponential Smoothing* Untuk Peramalan Penggunaan Waktu Telepon di PT Telkomsel Divre 3 Surabaya. *Jurnal Sistem Informasi*, 1-9.
- Rosadi, D. 2011. *Ekonometrika dan Analisis Runtun Waktu Terapan dengan Eviews*. Yogyakarta: ANDI.
- Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding machine learning: From theory to algorithms*. Cambridge university press.
- Shen, S., Jiang, H., & Zhang, T. (2012). *Stock market forecasting using machine learning algorithms*. Department of Electrical Engineering, Stanford University, Stanford, CA, 1-5.
- Sridama, P., & Siribut, C. (2018). *Decision Support System for Customer Demand Forecasting and Inventory Management of Perishable Goods*. *Journal of Advanced Management Science Vol*, 6(1).
- The Economist Intelligence Unit. (2017). *Fixing Food – Towards a More Sustainable Food System*
- Wahidin, A. J., & Maulana, R. (2021). *Classification of Super Air Jet Initial Cabin Crew Candidates Using K-Nearest Neighbor (KNN) Method: Klasifikasi Calon Awak Kabin Awal Super Air Jet Menggunakan Metode K-Nearest Neighbor (KNN)*. *SYSTEMATICS*, 3(2), 249-262.
- Levenson, A. (2005). *Harnessing the power of HR analytics*. *Strategic HR Review*.

- Wibowo, S. (2005). Peramalan Jumlah Pengunjung di UII Internet *Student Centre* dengan Model ARIMA.
- Wijaya, M. I., Hendayani, R., & Widiyanesti, S. (2014). *Demand Forecasting For Perishable Asset in Importer Company* (PT. TMM).
- Wiyanti, D. T., & Pulungan, R. (2012). Peramalan Deret Waktu Menggunakan Model Fungsi Basis Radial (RBF) dan *Auto Regressive Integrated Moving Average* (ARIMA). *Indonesian Journal of Mathematics and Natural Sciences*, 35(2).
- Wulansari, S. (2017). Optimasi Persediaan Untuk Produk *Perishable* (Studi Kasus: Rumah Sakit Sunan Kalijaga, Demak).
- Zufar, M. A. (2018). Perbandingan metode ARIMA dengan RBFNN dalam peramalan rata-rata banyaknya bilangan Sunspot (*Doctoral dissertation*, Universitas Islam Negeri Maulana Malik Ibrahim).