



**PREDIKSI DATA TIME SERIES PENJUALAN KATEGORI PRODUK
MENGUNAKAN PENDEKATAN MACHINE LEARNING**

TUGAS AKHIR

Yessy Natalia Siboro
41517120088

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA**

MERCU BUANA
2020

**IN
REVIEW**



**PREDIKSI DATA TIME SERIES PENJUALAN KATEGORI PRODUK
MENGUNAKAN PENDEKATAN MACHINE LEARNING**

Tugas Akhir

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh:
Yessy Natalia Siboro
41517120088

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2020

LEMBAR PERNYATAAN ORISINALITAS

Yang bertanda tangan dibawah ini:

NIM : 41517120088

Nama : Yessy Natalia Siboro

Judul Tugas Akhir : Prediksi Data Time Series Penjualan Kategori Produk
Menggunakan Pendekatan Machine Learning

Menyatakan bahwa Laporan Tugas Akhir saya adalah hasil karya sendiri dan bukan plagiat. Apabila ternyata ditemukan didalam laporan Tugas Akhir saya terdapat unsur plagiat, maka saya siap untuk mendapatkan sanksi akademik yang terkait dengan hal tersebut.

Jakarta, 24 Januari 2020



Yessy Natalia Siboro



UNIVERSITAS
MERCU BUANA

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Yessy Natalia Siboro
NIM : 41517120088
Judul Tugas Akhir : Prediksi Data Time Series Penjualan Kategori Produk Menggunakan Pendekatan Machine Learning

Dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Noneksklusif** (*None-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul diatas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti/Noneksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan tugas akhir saya.

Selain itu, demi pengembangan ilmu pengetahuan di lingkungan Universitas Mercu Buana, saya memberikan izin kepada Peneliti di Lab Riset Fakultas Ilmu Komputer, Universitas Mercu Buana untuk menggunakan dan mengembangkan hasil riset yang ada dalam tugas akhir untuk kepentingan riset dan publikasi selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 24 Januari 2020

UNIVERSITAS
MERCU BUANA



Yessy Natalia Siboro

SURAT PERNYATAAN LUARAN TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Yessy Natalia Siboro
 NIM : 41517120088
 Judul Tugas Akhir : Prediksi Data Time Series Penjualan Kategori Produk Menggunakan Pendekatan Machine Learning

Menyatakan bahwa Luaran Tugas Akhir saya adalah sebagai berikut :

No	Luaran	Jenis		Status		
1	Publikasi Ilmiah	Jurnal Nasional Tidak Terakreditasi		Diajukan	✓	
		Jurnal Nasional Terakreditasi	✓			
		Jurnal International Tidak Bereputasi		Diterima		
		Jurnal International Bereputasi				
Disubmit/dipublikasikan di :	Nama Jurnal :					
		ISSN :				
2	Kertas Kerja, Merupakan material hasil penelitian sebagai kelengkapan Artikel Jurnal. Terdiri dari (minimal 4)	Literatur Review			[✓]	
		Hasil analisa & perancangan aplikasi			[]	
		Source code			[✓]	
		Data set			[✓]	
		Tahapan eksperimen			[✓]	
		Hasil eksperimen seluruhnya			[✓]	
					
3	HAKI Disubmit / Terdaftar	HKI		Diajukan		
		Paten		Tercatat		
		No & Tanggal Permohonan				
		No & Tanggal Pencatatan				

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 24 Januari 2020



 Yessy Natalia Siboro


LEMBAR PERSETUJUAN

Nama Mahasiswa : Yessy Natalia Siboro
NIM : 41517120088
Judul Tugas Akhir : Prediksi Penjualan Kategori Produk dari Data Time Series dengan Machine Learning

Tugas Akhir ini telah diperiksa dan disetujui

Jakarta, 24 Januari 2020

Menyetujui,



(Dr. Mujiono Sadikin, M.T.)
Dosen Pembimbing



UNIVERSITAS
MERCU BUANA

LEMBAR PERSETUJUAN PENGUJI


NIM : 41517120088
Nama : Yessy Natalia Siboro
Judul Tugas Akhir : Prediksi Data Time Series Penjualan Kategori
Produk Menggunakan Pendekatan Machine
Learning

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.


Jakarta, 13 Februari 2020



(Dr. Devi Fitriplanah)
Ketua Penguji



(Dr. Ida Nurhaida)
Anggota Penguji 1



(Desi Ramayanti, S.Kom., MT)
Anggota Penguji 2

UNIVERSITAS
MERCU BUANA

LEMBAR PENGESAHAN

NIM : 41517120088
Nama : Yessy Natalia Siboro
Judul Tugas Akhir : Prediksi Data Time Series Penjualan Kategori Produk
Menggunakan Pendekatan Machine Learning

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

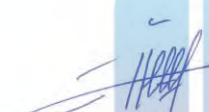
Jakarta, 13 Februari 2020

Menyetujui,



(Dr. Mujiono Sadikin, MT)
Dosen Pembimbing

Mengetahui,



(Diky Firdaus, S.Kom, MM)
Koord. Tugas Akhir Teknik Informatika



(Desi Ramayanti, S.Kom, MT)
Ka. Prodi Teknik Informatika

UNIVERSITAS
MERCU BUANA

ABSTRAK

Nama : Yessy Natalia Siboro
NIM : 41517120088
Pembimbing TA : Dr. Mujiono Sadikin, M.T
Judul : Prediksi Data Time Series Penjualan Kategori
Produk Menggunakan Pendekatan Machine
Learning

Perencanaan kebutuhan pasar diperlukan untuk mencapai kesuksesan sebuah perusahaan. Dalam perencanaan strategi pemasaran maka perlu menerapkan metode untuk meningkatkan jumlah permintaan pasar. Untuk memenuhi kebutuhan pasar dan juga jumlah data yang semakin banyak maka diperlukan penerapan metode prediksi. Dimana hal ini dilakukan untuk menghindari keterbatasan stock barang yang akan dibeli pedagang. Meskipun teknologi informasi, khususnya komputasi, telah diadopsi di berbagai bidang, namun masih banyak perusahaan tidak memanfaatkan secara optimum. Perusahaan berpotensi mendapatkan benefit dari pemanfaatan teknologi untuk komputasi misalnya berdasarkan historis transaksi penjualan dari tahun ke tahun. Sayangnya kebanyakan saat ini data tersebut, hanya digunakan sebagai bahan laporan bulanan dan tahunan. Sementara dari waktu ke waktu terjadi pertumbuhan data yang tidak dimanfaatkan dengan benar, alih - alih memudahkan tetapi yang terjadi adalah kesulitan mengakses informasi yang tepat. Pada eksperimen ini, digunakan dataset penjualan barang yang dikumpulkan selama setahun untuk memprediksi barang yang paling diminati di masa mendatang dengan perbandingan metode Long Short Term Memory dan Regresi Long Short Term Memory (LSTM). Algoritma LSTM digunakan untuk mengelola data time series. Pada penelitian ini digunakan untuk menganalisis data penjualan pada hari t guna memprediksi penjualan pada hari $t + 1$. Dimana hasil dari percobaan ini yaitu tingkat akurasi dari metode LSTM masih lebih baik dibanding dengan regresi LSTM. Dan banyaknya jumlah data training sangat mempengaruhi terhadap hasil akurasi dari prediksi.

Kata kunci:

Penjualan; Prediksi; Long Short Term Memory.

ABSTRACT

Name : Yessy Natalia Siboro
Student Number : 41517120088
Counsellor : Dr. Mujiono Sadikin, M.T
Title : Prediction Sales Data Time Series Product
Categories Using Machine Learning.

Planning market needs is needed to achieve the success of a company. In marketing strategy planning, it is necessary to apply methods to increase the amount of market demand. To meet market needs and also the amount of data that is increasingly large, it is necessary to apply predictive methods. Where this is done to avoid limited stock of goods to be purchased by traders. Although information technology, especially computing, has been adopted in various fields, there are still many companies that do not utilize it optimally. Companies have the potential to benefit from the use of technology for computing, for example, based on historical sales transactions from year to year. Unfortunately most of the time the data is only used as material for monthly and annual reports. While from time to time there is a growth of data that is not utilized properly, rather than making it easy, but what happens is the difficulty of accessing the right information. In this experiment, a dataset of sales of goods collected for a year is used to predict the most desirable items in the future by comparing the methods of Long Short Term Memory and Long Short Term Memory (LSTM) Regression. LSTM algorithm is used to manage time series data. This research is used to analyze sales data on day t to predict sales on day $t + 1$. Where the results of this experiment are the accuracy of the LSTM method is still better than the LSTM regression. And the large amount of training data greatly influences the accuracy of prediction results.

Key words:

Sales, Prediction, Long Short Term Memory

KATA PENGANTAR

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa yang telah melimpahkan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan tugas akhir yang berjudul “Prediksi Data Time Series Penjualan Kategori Produk Menggunakan Pendekatan Machine Learning” tepat pada waktunya. Tugas akhir ini disusun untuk memenuhi salah satu syarat memperoleh gelar sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Penulis menyadari bahwa tanpa bantuan dan bimbingan yang melibatkan banyak pihak, penelitian ini tidak akan terlaksana dengan baik. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Bapak Dr. Mujiono, MT, selaku dosen pembimbing tugas akhir yang telah meluangkan waktunya untuk memberikan bimbingan serta arahan dalam penyusunan tugas akhir ini hingga selesai.
2. Ibu Desi Ramayanti, S.Kom., MT, selaku Kepala Prodi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.
3. Bapak Diky Firdaus, S.Kom., MM, selaku Koordinator Tugas Akhir Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana tahun ajaran 2019/2020.
4. Kedua Orang tua yang senantiasa memberikan doa dan dukungan kepada penulis.
5. Teman-teman yang selalu menyemangati dan memberi motivasi kepada penulis selama pelaksanaan tugas akhir.
6. Semua pihak yang telah banyak membantu dalam penyusunan tugas akhir ini yang tidak dapat penulis sebutkan satu persatu.

Akhir kata, penulis berharap tugas akhir ini dapat bermanfaat bagi pembaca guna menambah pengetahuan dan wawasan.

Jakarta, 24 Januari 2020
Penulis

DAFTAR ISI

HALAMAN SAMPUL.....	i
HALAMAN JUDUL	i
LEMBAR PERNYATAAN ORISINALITAS	ii
SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR... iii	
SURAT PERNYATAAN LUARAN TUGAS AKHIR.....	iv
LEMBAR PERSETUJUAN	v
LEMBAR PERSETUJUAN PENGUJI	vi
LEMBAR PENGESAHAN	vii
ABSTRAK	viii
ABSTRACT.....	ix
KATA PENGANTAR.....	x
DAFTAR ISI.....	xi
NASKAH JURNAL	1
KERTAS KERJA.....	A
BAGIAN 1. LITERATUR REVIEW	B
BAGIAN 2 ANALISIS DAN PERANCANGAN.....	E
BAGIAN 3 SOURCE CODE	G
BAGIAN 4 DATASET	I
BAGIAN 5 TAHAPAN EKSPERIMEN	J
BAGIAN 6 HASIL SEMUA EKSPERIMEN	K

Prediksi Data Times Series Penjualan Kategori Produk Menggunakan Pendekatan Machine Learning

Yessy Natalia Siboro, Mujiono Sadikin*)

Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana
Jl. Raya Meruya Selatan, Kembangan, Jakarta, Indonesia 11650

Abstract – *Market needs planning is needed to achieve the success of a company. In marketing strategy planning it is necessary to apply methods to increase the amount of market demand. To meet market needs and also the amount of data that is increasingly large, the application of prediction methods is needed. Where this is done to avoid limited stock of goods to be purchased by traders. In this experiment, a dataset of sales of goods collected for a year is used to predict the items of most interest in the future by comparing the methods of Long Short Term Memory and Regression of Long Short Term Memory (LSTM). The results of the study used test parameters with Root Mean Square Error (RMSE) for LSTM was 5.70 and LSTM Regression was 0.43.*

Keywords – *Sales; Prediction; Long Short Term Memory*

Abstrak – *Perencanaan kebutuhan pasar diperlukan untuk mencapai kesuksesan sebuah perusahaan. Dalam perencanaan strategi pemasaran maka perlu menerapkan metode untuk meningkatkan jumlah permintaan pasar. Untuk memenuhi kebutuhan pasar dan juga jumlah data yang semakin banyak maka diperlukan penerapan metode prediksi. Dimana hal ini dilakukan untuk menghindari keterbatasan stock barang yang akan dibeli pedagang. Pada eksperimen ini, digunakan dataset penjualan barang yang dikumpulkan selama setahun untuk memprediksi barang yang paling diminati di masa mendatang dengan perbandingan metode Long Short Term Memory dan Regresi Long Short Term Memory (LSTM). Hasil penelitian menggunakan parameter uji dengan Root Mean Square Error (RMSE) untuk LSTM adalah 5.70 dan Regresi LSTM adalah 0.43.*

Kata kunci - *Penjualan; Prediksi; Long Short Term Memory.*

I. PENDAHULUAN

Platform *Point of Sales* yaitu Qasir.id merupakan penyedia aplikasi kasir yang bisa membantu pedagang mencatat penjualan, mengelola produk,

mengawasi stock, memantau laporan transaksi maupun berperan sebagai distributor untuk para pelaku UMKM. Perkembangan teknologi yang pesat di satu sisi memberikan kemudahan, namun di sisi yang lain mendatangkan tantangan. Salah satu tantangan tersebut adalah bagaimana memanfaatkan data yang terkumpul dari transaksi berbasis teknologi untuk mendukung perkembangan perusahaan. Salah satu manfaat dari data transaksi yang terkumpul adalah untuk keperluan prediksi yang mampu membantu penentuan strategi perusahaan.

Meskipun teknologi informasi, khususnya komputasi, telah diadopsi di berbagai bidang, namun masih banyak perusahaan tidak memanfaatkan secara optimum. Perusahaan berpotensi mendapatkan benefit dari pemanfaatan teknologi untuk komputasi misalnya berdasarkan historis transaksi penjualan dari tahun ke tahun. Sayangnya kebanyakan saat ini data tersebut, hanya digunakan sebagai bahan laporan bulanan dan tahunan. Sementara dari waktu ke waktu terjadi pertumbuhan data yang tidak dimanfaatkan dengan benar, alih - alih memudahkan tetapi yang terjadi adalah kesulitan mengakses informasi yang tepat. Data yang tidak diolah ini akan mengakibatkan penumpukan data yang tidak bermanfaat [1],[2].

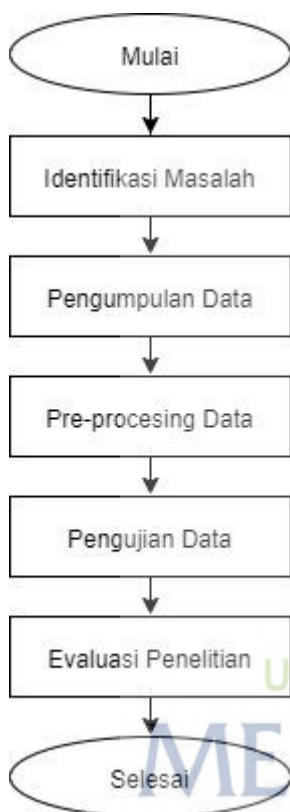
Salah satu akibat kurang dimanfaatkannya data yang dimiliki adalah promosi yang tidak tepat sasaran sehingga biaya promosi terbuang sia-sia [3]. Salah satu cara untuk membantu strategi pemasaran adalah teknologi informasi. Peramalan merupakan salah satu teknik untuk memprediksi permintaan pasar di masa mendatang berdasarkan data histori. Terdapat banyak teknik atau algoritma yang bisa digunakan untuk melakukan prediksi di masa yang akan datang. Salah satunya adalah teknik Machine Learning model prediksi Long Short Term Memory (LSTM). Pada penelitian ini digunakan teknik LSTM untuk melakukan prediksi data time series transaksi penjualan. Kinerja prediksi diukur menggunakan parameter Root Mean Square Error (RMSE).

Data transaksi penjualan yang tersimpan dapat memberikan manfaat untuk manajemen perusahaan [4]. Terdapat beberapa studi terkait yang telah dilakukan oleh para peneliti diantaranya yaitu prediksi pergerakan saham [5], prediksi kunjungan wisatawan [6], prediksi arah indeks pasar saham

harian [7], dan perbandingan metode klasifikasi untuk memprediksi potensi resiko [8]. Long Short Term Memory yang digunakan oleh Tingwei Gao, dkk [5] bertujuan untuk memprediksi pergerakan saham pada hari berikutnya.

II. METODOLOGI PENELITIAN

Pada penelitian ini dilakukan eksperimen, dengan menguji coba pada model LSTM pada data histori transaksi. Adapun tahapan-tahapan penelitian seperti Gambar 1.



Gambar 1. Tahapan Penelitian

Eksperimen dilakukan dengan memanfaatkan model prediksi algoritma Long Short Term Memory (LSTM) dan Regresi Long Short Term Memory. Algoritma LSTM digunakan untuk mengelola data time series [9]. Pada penelitian ini digunakan untuk menganalisis data penjualan pada hari t guna memprediksi penjualan pada hari $t + 1$.

A. Pengumpulan Data

Langkah awal dalam penelitian adalah pengumpulan data. Dimana data yang digunakan merupakan transaksi penjualan selama periode September 2018 – April 2019. Dari data tersebut diperoleh instan sebanyak 80476 row data dan dikonversi menjadi file .csv (comma separated values). Hasil konversi ditunjukkan pada Gambar 2.

A	B	C	D	E	F	G	H	I	J	K	
1	Date_order	Id_order	no_order	product_id	product_name	variant_product	category	brand	quantity	total_price	status
2	03/09/2018 19:40	395	18994395	10592	S-tee ctn 350 ml x 12 pcs	S-tee ctn 350 ml x 12 pcs	Minuman Kemasan	S-tee	1	24000	delivered
3	04/09/2018 05:58	396	18994396	10510	Nu Green Tea Honey ctn 450ml x 24pcs	Nu Green Tea Honey 450 ml	Minuman Kemasan	NU GREEN TEA	1	100000	delivered
4	04/09/2018 05:58	396	18994396	10583	Teh Pucuk ctn 350 ml x 24 pcs	Teh Pucuk ctn 350 ml x 24 pcs	Minuman Kemasan	Teh Pucuk Harum	2	106000	delivered
5	04/09/2018 05:58	396	18994396	10585	ABC Kopi Susu ctn 33gr x 120 pcs	ABC Kopi Susu ctn 33gr x 120 pcs	Minuman Instan	Abckopi	1	114000	delivered
6	04/09/2018 05:58	396	18994396	10579	Kapal Api Special Mix 25 gr x 120 pcs	Kapal Api Special Mix 25 gr x 120 pcs	Minuman Instan	Kapal Api	1	118000	delivered
7	04/09/2018 05:58	396	18994396	10609	Minutes Maid Gelas 296 ml x 24 pcs	Minutes Maid Gelas 296 ml x 24 pcs	Minuman Kemasan	Minutes Maid	4	140000	delivered
8	04/09/2018 05:58	396	18994396	10594	Frestea gelas ctn 300 ml x 24 pcs	Frestea gelas ctn 300 ml x 24 pcs	Minuman Kemasan	Frestea	4	144000	delivered
9	04/09/2018 05:58	396	18994396	10589	Aqua Botol ctn 600 ml x 24 pcs	Aqua Botol ctn 600 ml x 24 pcs	Minuman Kemasan	Aqua	4	174000	delivered
10	04/09/2018 05:58	396	18994396	10586	Bear Brand Kaleng ctn 189 ml x 30 pcs	Bear Brand Kaleng ctn 189 ml x 30 pcs	Minuman Kemasan	Nestle Bear Brand	1	216000	delivered
11	04/09/2018 11:50	397	18994397	10597	Beng-beng Wafer box 22 gr x 20 pcs	Beng-beng Wafer box 22 gr x 20 pcs	Makanan Instan	Beng-beng	4	106000	delivered
12	04/09/2018 11:50	397	18994397	10585	ABC Kopi Susu ctn 33gr x 120 pcs	ABC Kopi Susu ctn 33gr x 120 pcs	Minuman Instan	Abckopi	1	114000	delivered
13	04/09/2018 11:50	397	18994397	10606	Luwak White Coffee ctn 20 gr x 200 pcs	Luwak White Coffee Carton	Minuman Instan	LUNAK COFFEE	1	190000	delivered
14	04/09/2018 12:42	398	18994398	10579	Kapal Api Special Mix 25 gr x 120 pcs	Kapal Api Special Mix 25 gr x 120 pcs	Minuman Instan	Kapal Api	10	1180000	delivered
15	04/09/2018 12:42	398	18994398	10579	Kapal Api Special Mix 25 gr x 120 pcs	Kapal Api Special Mix 25 gr x 120 pcs	Minuman Instan	Kapal Api	10	1180000	delivered
16	04/09/2018 12:49	399	18994399	10589	Aqua Botol ctn 600 ml x 24 pcs	Aqua Botol ctn 600 ml x 24 pcs	Minuman Kemasan	Aqua	1	45500	delivered
17	04/09/2018 12:49	399	18994399	10608	Aqua Botol ctn 1500 ml x 12 pcs	Aqua Botol ctn 1500 ml x 12 pcs	Minuman Kemasan	Aqua	1	45500	delivered
18	04/09/2018 12:57	400	18994400	10589	Aqua Botol ctn 600 ml x 24 pcs	Aqua Botol ctn 600 ml x 24 pcs	Minuman Kemasan	Aqua	1	45500	cancel
19	04/09/2018 12:57	400	18994400	10609	Minutes Maid Gelas 296 ml x 24 pcs	Minutes Maid Gelas 296 ml x 24 pcs	Minuman Kemasan	Minutes Maid	2	70000	cancel
20	04/09/2018 13:28	401	18994401	10599	Minutes Maid Gelas 296 ml x 24 pcs	Minutes Maid Gelas 296 ml x 24 pcs	Minuman Kemasan	Minutes Maid	1	30000	cancel
21	04/09/2018 13:28	401	18994401	10594	Frestea gelas ctn 300 ml x 24 pcs	Frestea gelas ctn 300 ml x 24 pcs	Minuman Kemasan	Frestea	1	30000	cancel
22	04/09/2018 13:28	401	18994401	10589	Aqua Botol ctn 600 ml x 24 pcs	Aqua Botol ctn 600 ml x 24 pcs	Minuman Kemasan	Aqua	1	45500	cancel
23	04/09/2018 13:28	401	18994401	10608	Aqua Botol ctn 1500 ml x 12 pcs	Aqua Botol ctn 1500 ml x 12 pcs	Minuman Kemasan	Aqua	1	45500	cancel

Gambar 2. Hasil Konversi Transaksi Penjualan

Keterangan :

- *Date_order* merupakan tanggal pembelian produk.
- *Id_order* merupakan id dari pesanan yang di order oleh pelanggan.
- *Product_id* merupakan id dari produk/barang.
- *Category* merupakan kategori produk berdasarkan jenis produknya.
- *Quantity* merupakan jumlah produk yang dibeli oleh pelanggan
- *Total_price* merupakan jumlah harga yang dibayar dari total pesanan pelanggan.

B. Pra Pemrosesan Data

Tahap selanjutnya yaitu pra pemrosesan data. Pada tahap ini dilakukan proses ekstraksi dan *cleaning* data [10]. Dimana pada tahap ini dilakukan pemilihan atribut yang digunakan pada penelitian dan hasil data yang sudah di preprocessing ditunjukkan pada tabel 1.

Tabel 1. Tabel Atribut

No.	Atribut	Tipe	Keterangan
1	Date_order	Integer	
2	Id_order	Integer	
3	Product_id	Integer	
4	Category	Integer	1 = Adult Diapers 2 = Baby Diapers 3 = Bahan Kue 4 = Bahan Makanan 5 = Beras 6 = Buah Segar 7 = Bumbu Masak 8 = Energy Drink 9 = Gummy Candy 10 = Isotonic Drink 11 = Lotion Mosquito Repellen 12 = Makanan 13 = Makanan Bayi 14 = Makanan Instan 15 = Makanan

Kemasan
....
30 = Sayuran

5	Quantity	Integer	
6	Total price	Integer	Rupiah

Setelah proses ekstraksi, dilakukan proses *cleaning* data dengan menghilangkan atribut yang tidak digunakan, membersihkan data yang noisy dan yang tidak sesuai dengan rule. Pada tahap ini atribut *category* yang bertipe text akan dikonversi menjadi tipe numerik dengan pemberian label pada setiap kategorinya seperti Gambar 3.

	A	B	C	D	E	F
1	date_order	id_order	product_id	category	quantity	total_price
2	04/09/2018	396	10585	19	1	114000
3	04/09/2018	396	10589	20	4	174000
4	04/09/2018	396	10594	20	4	144000
5	04/09/2018	396	10579	19	1	118000
6	04/09/2018	396	10609	20	4	140000
7	04/09/2018	396	10586	20	1	216000
8	04/09/2018	396	10610	20	1	100000
9	04/09/2018	396	10583	20	2	106000
10	04/09/2018	397	10606	19	1	190000
11	04/09/2018	397	10585	19	1	114000
12	04/09/2018	397	10597	15	4	106000
13	04/09/2018	398	10579	19	10	1180000
14	04/09/2018	398	10579	19	10	1180000
15	04/09/2018	399	10589	20	1	43500
16	04/09/2018	399	10608	20	1	43500
17	04/09/2018	400	10609	20	2	70000
18	04/09/2018	400	10589	20	1	43500
19	04/09/2018	401	10589	20	1	43500
20	04/09/2018	401	10608	20	1	43500
21	04/09/2018	401	10594	20	1	36000
22	04/09/2018	401	10609	20	1	35000
23	04/09/2018	402	10581	15	1	86000

Gambar 3. Data setelah Preprocessing

C. Desain Eksperimen

Pada tahap ini, digunakan Long Short Term Memory (LSTM) dan Regresi Long Short Term Memory (LSTM) yang dievaluasi dengan Root Mean Square Error dan (RMSE) dan Mean Absolute Percentage Error (MAPE).

1. Long Short Term Memory Process

Pada tahap Processing data dalam analisis time series dengan model prediksi LSTM meliputi normalisasi data. Dimana nilai data dibagi dengan nilai range data (nilai data maksimum – nilai data minimum) [11]. Dan data tersebut dinormalisasi ke dalam 2 interval [-1,1] dan [0,1] dengan tujuan memastikan interval mana yang menghasilkan nilai terbaik. Tujuan dari Normalisasi yaitu:

1. Untuk menghilang kerangkapan data.
2. Untuk mengurangi kompleksitas.
3. Untuk mempermudah pemodifikasian data.

$$X_n = \frac{X_0 - X_{min}}{X_{max} - X_{min}}$$

Keterangan:

X_n = Nilai hasil normalisasi

X_0 = Nilai data aktual

X_{min} = Nilai minimum dari data aktual

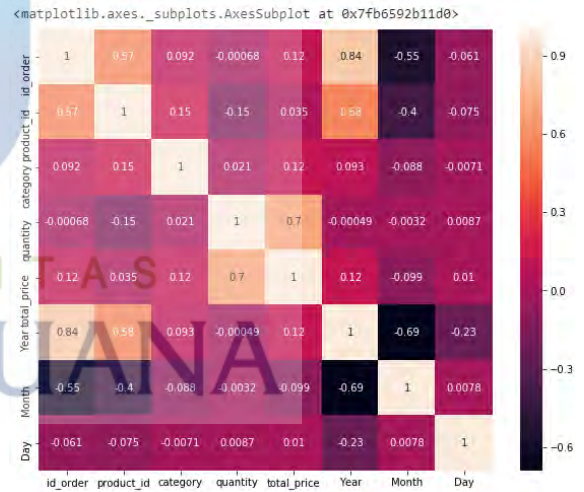
X_{max} = Nilai maximum dari data aktual

2. Long Short Term Memory Regression Prediction

Long Short Term Memory (LSTM) disebutkan untuk pertama kali pada tahun 1997 dijelaskan oleh Hochreiter dan Schmidhuber[13]. Metode ini merupakan salah satu jenis Recurrent Neural Network (RNN) dimana memiliki kemampuan untuk mengolah informasi yang lebih akurat dan menyimpan informasi untuk jangka waktu yang lama. Prediksi dengan regresi akan menganalisa jumlah penjualan pada hari t dalam kolom pertama dan untuk memprediksi penjualan pada hari ke t + 1 pada kolom kedua. Metode ini memiliki lapisan tersembunyi yang terdiri dari sel memori, satu sel memori memiliki tiga gate yaitu input gate, forget gate, output gate [12].

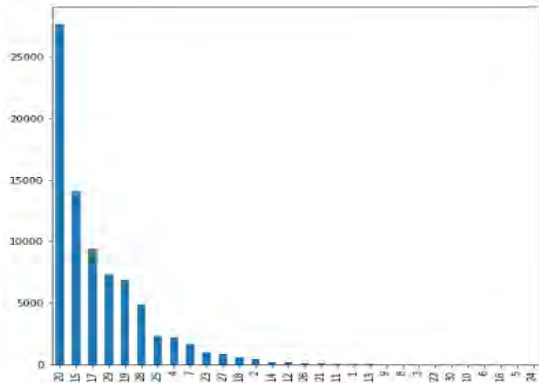
D. Eksperimen dan Pengujian

Pada penelitian ini, diperoleh nilai akurasi dari algoritma prediksi yang akan digunakan dalam pengolahan data. Beberapa atribut yang dekat kolerasinya akan dipilih seperti atribut *category* dan *quantity* sehingga menghasilkan bobot LSTM yang optimal. Dimana parameternya berupa komposisi data dari time series yang terdapat korelasi antar atributnya seperti Gambar 4.



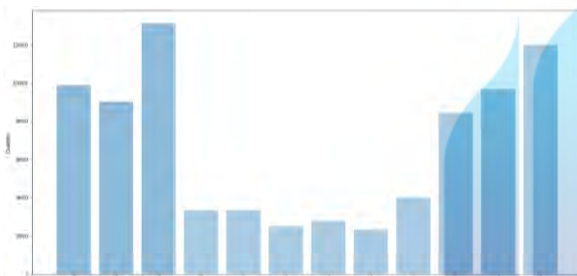
Gambar 4. Korelasi antar atribut

Setelah melihat korelasi tersebut, maka data kemudian di visualisasikan untuk membantu melihat hubungan terstruktur antara atribut *id_order*, *category* dan *quantity*. Data yang sudah divisualisasi menghasilkan grafik data penjualan kategori produk yang paling laris terjual pada Gambar 5.



Gambar 5. Korelasi *category* dengan *quantity*

Data penjualan yang paling laris kemudian divisualisasikan kembali dan menghasilkan grafik data penjualan harian berdasarkan quantity setiap bulannya pada Gambar 6 dan grafik data penjualan per hari pada Gambar 7.



Gambar 6. Data penjualan per bulan



Gambar 7. Grafik data penjualan per hari

Sebelum memulai prediksi, dilakukan ekstraksi array dari dataframe dan mengkonversi nilai integer ke nilai floating. Kemudian membangun model prediksi LSTM dan Regresi LSTM dari normalisasi dataset yang menerapkan MinMaxScaler dan scikit-learn.. Dari parameter yang disediakan dapat dibentuk skenario untuk beberapa pengujian ulang dalam penelitian ini. Sebelum melakukan pengujian, terdapat parameter uji yang akan mempengaruhi hasil prediksi disajikan pada Tabel 2.

Tabel 2. Skenario Percobaan

	1	2	3
Dataset	70/30	80/20	90/10
Range Interval	[-1,1]	[0,1]	
Look Back	1	2	3

Pengguna parameter untuk LSTM hanya menggunakan dua parameter yaitu komposisi dataset dan nilai range interval sedangkan untuk Regresi LSTM menggunakan ketiga parameter tersebut. Setelah komposisi dataset yang dijadikan data training dan testing telah ditentukan, selanjutnya data dilatih dengan *epoch* sebanyak 100 kali dengan nilai *loss* yang hampir sama seperti Gambar 8.

```
Epoch 33/100
- 1055 - loss: 6.8292e-04
Epoch 34/100
- 1045 - loss: 6.8368e-04
Epoch 35/100
- 1055 - loss: 6.8382e-04
Epoch 36/100
- 1055 - loss: 6.8332e-04
Epoch 37/100
- 1055 - loss: 6.8209e-04
Epoch 38/100
- 1075 - loss: 6.8204e-04
Epoch 39/100
- 1085 - loss: 6.8274e-04
Epoch 40/100
- 1085 - loss: 6.8237e-04
Epoch 41/100
- 1135 - loss: 6.8222e-04
Epoch 42/100
- 1115 - loss: 6.8325e-04
Epoch 43/100
- 1105 - loss: 6.8241e-04
Epoch 44/100
- 1115 - loss: 6.8293e-04
Epoch 45/100
- 1075 - loss: 6.8240e-04
Epoch 46/100
- 1085 - loss: 6.8218e-04
Epoch 47/100
- 1085 - loss: 6.8228e-04
Epoch 48/100
- 1065 - loss: 6.8232e-04
Epoch 49/100
- 1075 - loss: 6.8220e-04
Epoch 50/100
- 1055 - loss: 6.8164e-04
```

Gambar 8. Training data dengan *epoch*

F. Evaluasi Penelitian

Tahapan ini akan membahas hasil evaluasi penelitian yang sudah dilakukan. Tujuannya untuk menemukan hasil yang terbaik agar menentukan pola yang menarik atau bernilai dalam strategi pemasaran perusahaan. Hasil penelitian berupa pola-pola yang khas berdasarkan data yang dimiliki.

III. HASIL DAN PEMBAHASAN

Pada penelitian ini, diperoleh sebuah model prediksi LSTM dan Regresi LSTM dengan menggunakan framework Keras dari Python. Peramalan terbaik didasarkan pada tingkat kesalahan prediksi, di mana semakin kecil tingkat kesalahan yang dihasilkan, semakin tepat sebuah metode dalam memprediksi. Tabel 3 dan Tabel 4 menunjukkan hasil pembagian dalam pembentukan skenario eksperimen untuk menemukan metode terbaik dalam pengujian yang dilakukan oleh LSTM dan Regresi LSTM.

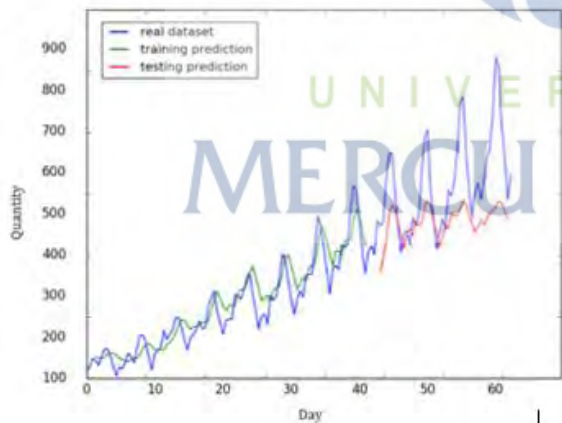
Tabel 3. Hasil RMSE Percobaan LSTM

Percobaan	Parameter	RMSE
1	D1, R1	6.44
2	D1, R2	5.83
3	D2, R1	5.21
4	D2, R2	4.45
5	D3, R1	6.62
6	D3, R2	8.04
Average		6.09

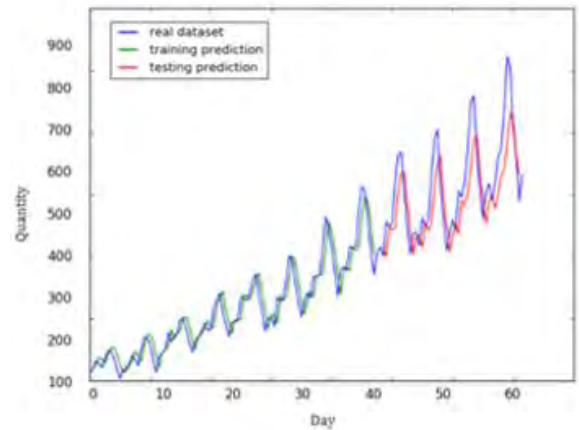
Tabel 4. Hasil RMSE Percobaan Regresi LSTM

Percobaan	Parameter	RMSE
1	D1, R1, L1	0.03
2	D1, R1, L2	0.05
3	D1, R1, L3	0.04
4	D1, R2, L1	0.06
5	D1, R2, L2	0.11
6	D1, R2, L3	0.08
7	D2, R1, L1	0.02
8	D2, R1, L2	0.05
9	D2, R1, L3	0.04
10	D2, R2, L1	0.09
11	D2, R2, L2	0.13
12	D2, R2, L3	0.10
13	D3, R1, L1	0.11
14	D3, R1, L2	0.08
15	D3, R1, L3	0.07
16	D3, R2, L1	0.09
17	D3, R2, L2	0.05
18	D3, R2, L3	0.10
Average		0.07

Data yang telah diolah melalui beberapa kali percobaan dan menghasilkan nilai tertentu dari setiap eksperimen tidak ditampilkan begitu saja dengan tabel. Data disajikan dengan grafik berwarna yang disebut plot. Contohnya melalui plot yang disajikan pada Gambar 9 dan 10 dapat dilihat perbandingan antara nilai hasil prediksi dengan nilai aktual dari kedua metode penelitian yang sedang diuji.



Gambar 9. Plot Hasil Eksperimen LSTM



Gambar 10. Plot Hasil Eksperimen Regresi LSTM

Data disajikan juga melalui tabel yang berisikan informasi tentang perbandingan antara nilai prediksi dan nilai aktual pada data penjualan perhari pada Tabel 5 dan 6.

Tabel 5. Perbandingan Nilai Prediksi dan Aktual Eksperimen LSTM

Day	Predicted	Expected
1	103	58
2	95	33
3	150	100
4	114	50
5	120	82
...
11	80	35
12	106	70
13	130	96
14	169	53
15	97	40

Tabel 6. Perbandingan Nilai Prediksi dan Aktual Eksperimen Regresi LSTM

Day	Predicted	Expected
1	50	52
2	31	35
3	69	70
4	34	38
5	48	50
...
12	70	76
14	57	58
14	82	91
15	66	70

IV. KESIMPULAN

Dalam penelitian ini telah dilakukan perbandingan antara metode LSTM dan regresi LSTM yang diimplementasikan pada dataset penjualan barang untuk memprediksi penjualan di masa yang akan datang. Dimana hasil dari percobaan ini yaitu tingkat akurasi dari metode LSTM masih lebih baik dibanding dengan regresi LSTM. Dan banyaknya jumlah data training sangat mempengaruhi terhadap hasil akurasi dari prediksi. Dimana nilai RMSE yang didapat dari percobaan ini yaitu sebesar 6.09 untuk LSTM dan 0.07 Regresi LSTM.

Adapun saran yang dapat diberikan untuk penelitian selanjutnya yaitu menggunakan metode prediksi dan variable yang berbeda sehingga dimungkinkan kekurangan dari metode yang diterapkan pada penelitian ini bisa menjadi pengembangan pada penelitian selanjutnya untuk mendapatkan hasil yang lebih baik.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada platform Qasir.id atas partisipasinya untuk memberikan data transaksi penjualan dalam pelaksanaan penelitian.

DAFTAR PUSTAKA

- [1] M. Fauzy, K. R. S. W, and I. Asror, "Penerapan Metode Association Rule Menggunakan Algoritma Apriori Pada Simulasi Prediksi Hujan Wilayah Kota Bandung," *J. Ilm. Teknol. Inf. Terap.*, vol. 13, no. 2, pp. 115–124, 2016.
- [2] Fitriyani, "Implementasi Algoritma Fp-Growth Menggunakan Association Rule Pada Market Basket Analysis," *Informatika*, vol. II, no. 1, pp. 296–305, 2015.
- [3] M. I. Ghazali, R. Z. Ehwan, and W. H. Sugiharto, "Analisa Pola Belanja Menggunakan Algoritma Fp Growth, Self Organizing Map (Som) Dan K Medoids," vol. 8, no. 1, pp. 317–326, 2017.
- [4] G. Mathisen, "Forecasting Multivariate Time Series Data Using Neural Networks Sigurd Øyen," no. June, 2018.
- [5] T. Gao, Y. Chai, and Y. Liu, "Applying long short term memory neural networks for predicting stock closing price," *Proc. IEEE Int. Conf. Softw. Eng. Serv. Sci. ICSESS*, vol. 2017-Novem, pp. 575–578, 2018.
- [6] A. A. Rizal and S. Soraya, "Multi Time Steps Prediction dengan Recurrent Neural Network Long Short Term Memory," *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 18, no. 1, pp. 115–124, 2018.
- [7] Q. Mingyue, L. Cheng, and S. Yu, "Application of the Artificial Neural Network in Predicting the Direction of Stock Market Index," *Proc. - 2016 10th Int. Conf. Complex, Intelligent, Softw. Intensive Syst. CISIS 2016*, pp. 219–223, 2016.
- [8] M. Sadikin and F. Alfiandi, "Comparative study of classification method on customer candidate data to predict its potential risk," *Int. J. Electr. Comput. Eng.*, vol. 8, no. 6, pp. 4763–4771, 2018.
- [9] D. I. Komputer, F. Matematika, D. A. N. Ilmu, and P. Alam, "Prediksi temporal untuk kemunculan titik panas di kabupaten rokan hilir riau menggunakan long short term memory rnn hafshah luthfiah," 2018.
- [10] X. Qu, X. Kang, Z. Chao, J. Shuai, and X. Ma, "Short-term prediction of wind power based on deep Long Short-Term Memory," *Asia-Pacific Power Energy Eng. Conf. APPEEC*, vol. 2016-Decem, pp. 1148–1152, 2016.
- [11] R. Hidayat, "Meminimalisasi nilai error peramalandengan algoritma," pp. 187–192, 2012.
- [12] R. Vinayakumar, K. P. Soman, and P. Poornachandran, "Long short-term memory based operation log anomaly detection," *2017 Int. Conf. Adv. Comput. Commun. Informatics, ICACCI 2017*, vol. 2017-Janua, pp. 236–242, 2017.
- [13] L. Skovajsová, "Long Short-Term Memory Description and its Application in Text Processing."

KERTAS KERJA

Ringkasan

Pada eksperimen ini, digunakan dataset penjualan barang yang dikumpulkan selama setahun untuk memprediksi barang yang paling diminati di masa mendatang dengan perbandingan metode Long Short Term Memory dan Regresi Long Short Term Memory (LSTM). Algoritma LSTM digunakan untuk mengelola data time series. Pada penelitian ini digunakan untuk menganalisis data penjualan pada hari t guna memprediksi penjualan pada hari $t + 1$. Dimana hasil dari percobaan ini yaitu tingkat akurasi dari metode LSTM masih lebih baik dibanding dengan regresi LSTM. Dan banyaknya jumlah data training sangat mempengaruhi terhadap hasil akurasi dari prediksi.

