

## Peramalan Deret Waktu untuk Bisnis : Pendekatan algoritma LGBM Regressor

Fachrul Rozi Lubis<sup>1\*</sup> Eddy Rahman Syahputra<sup>2</sup>

<sup>1,2</sup> Universitas Harapan Medan, Fakultas Teknik dan Komputer, Sistem Informasi, Medan

<sup>1\*</sup> fachrul\_rozi.unhar@harapan.ac.id

Received: 15 Jan 2022

Accepted: 20 Jan 2022

Published: 25 Jan 2022



\***Fachrul Rozi Lubis**

**Keywords:** Peramalan Deret Waktu, Data Science, Analisis Bisnis, LGBM Regressor, CRISP-DM

**DSI: Jurnal Data Science Indonesia** is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).

**Abstrak :** Peramalan deret waktu adalah topik yang cukup umum di bidang *data science* (ilmu data). Perusahaan menggunakan model peramalan untuk mendapatkan pandangan yang lebih jelas tentang bisnis masa depan. Data masa lalu dikumpulkan dan dianalisis melalui model kuantitatif atau kualitatif sehingga pola dapat diidentifikasi dan dapat mengarahkan perencanaan bisnis di masa depan akan tetapi memilih algoritme yang tepat merupakan salah satu keputusan sulit ketika akan mengembangkan model peramalan deret waktu. Penelitian ini menyajikan hasil analisis data dengan mengadopsi kerangka kerja *data science* CRISP-DM dan membandingkan lima algoritma berbeda untuk memperkirakan penjualan harian selama 28 hari ke depan. Berdasarkan hasil evaluasi kinerja RMSE, algoritma LGBM Regressor menghasilkan tingkat kesalahan 7.53 %, paling rendah dibandingkan algoritma lain, akan tetapi waktu pelatihan dan pengujian paling tinggi.

### PENDAHULUAN

Analisis deret waktu merupakan salah satu jenis data yang paling umum digunakan oleh perusahaan untuk membantu mengembangkan strategi bisnis [1]. Deret waktu adalah urutan titik data berbasis waktu yang dikumpulkan pada interval tertentu dari kumpulan data tertentu yang mengalami perubahan dari waktu ke waktu yang diindeks menurut waktu [2]. Analisis deret waktu digunakan untuk menentukan model yang baik untuk diterapkan dalam meramalkan metrik bisnis [3], seperti harga pasar saham [4]–[8], penjualan [9]–[11] dan lainnya. Hal ini memungkinkan manajemen dapat memahami pola tepat waktu dalam data serta menganalisis tren dalam metrik bisnis. Dengan melacak data masa lalu, peramal berharap mendapatkan gambaran di masa dengan yang lebih baik.

Peramalan deret waktu adalah topik yang cukup umum di bidang *data science* (ilmu data). Perusahaan menggunakan model peramalan untuk mendapatkan pandangan yang lebih jelas tentang bisnis masa depan [11]. Peramalan bisnis mengacu pada alat dan teknik yang digunakan untuk memprediksi perkembangan bisnis, seperti penjualan, pengeluaran, dan keuntungan [10]. Tujuan utama dari peramalan bisnis adalah untuk mengembangkan strategi yang lebih baik berdasarkan prediksi informasi ini [9]. Data masa lalu dikumpulkan dan dianalisis melalui model kuantitatif atau kualitatif sehingga pola dapat diidentifikasi dan dapat mengarahkan perencanaan permintaan, operasi keuangan, produksi masa depan, dan operasi pemasaran. Memilih algoritme yang tepat mungkin merupakan salah satu keputusan sulit ketika akan mengembangkan model peramalan deret waktu.

Pada artikel ini, kami membandingkan lima algoritma yang berbeda, yaitu *Linear Regression* (LR), *Ridge Regressor* (RR), *XGBoost Classifier*, LGBM dan *Long short-term memory* (LSTM), pada tipe dataset deret waktu M5 Forecasting sumber terbuka dari kompetisi kaggle. Metode yang diusulkan dalam penelitian menggunakan kerangka kerja ilmu data CRISP-DM mengikuti semua proses yang diperlukan yaitu analisis bisnis, pemahaman data, analisis data eksplorasi dan pemodelan. Hasil peralaman dievaluasi dengan pendekatan RMSE (*Root Mean Square Error*)

## TINJAUAN LITERATUR

Peramalan memberikan pengetahuan tentang sifat kondisi masa depan. Peramalan penjualan membantu setiap bisnis membuat keputusan bisnis yang lebih baik secara keseluruhan, penganggaran, dan manajemen risiko. Memungkinkan perusahaan untuk secara efisien mengalokasikan sumber daya untuk pertumbuhan di masa depan dan mengelola arus kasnya. Kinerja penjualan masa lalu adalah indikator utama yang baik untuk kinerja penjualan di masa depan. Secara umum, peramalan dilakukan secara berulang secara teratur untuk berbagai tujuan perencanaan. Seperti, permintaan produk dan jasa di masa depan dapat diramalkan untuk mendukung perencanaan produksi, aktivitas pemasaran, penjadwalan sumber daya, dan perencanaan keuangan. Peramal sering mengikuti proses iteratif yang sama dengan setiap siklus peramalan dan perencanaan. Dalam beberapa tahun terakhir, metode peramalan deret waktu seperti *autoregressive moving average* (ARMA) dan *autoregressive integrated moving average* (ARIMA) merupakan salah satu metodologi yang sukses dan banyak diterapkan oleh peneliti [12], namun kemampuan prediksi model ini dibatasi pada perilaku linier, sehingga sulit diterima khususnya untuk diterapkan dalam kasus perkiraan deret waktu non-linier [13]. Kerangka kerja multi-seri untuk perkiraan permintaan diperkenalkan oleh Garnier & Belletoile, 2019 [14] dengan memanfaatkan korelasi antar rangkaian untuk meningkatkan akurasi prediksi pada penjualan e-commerce. Selanjutnya, M. Li et al., 2018 [15] melakukan perbandingan tiga model yaitu model Empiris ARIMA, Jaringan Saraf Tiruan *Autoregressive Nonlinear*, dan Model Gabungan ARIMA-NARNN untuk peramalan penjualan e-commerce dimana model kombinasi ARIMA-NARNN adalah yang terbaik dibandingkan model lainnya. Kemudian Khalil Zadeh et al., 2014 [13] mengusulkan tiga model peramalan berbeda untuk prediksi penjualan pada perusahaan distribusi obat berdasarkan pendekatan data mining. Penelitian ini membandingkan model ARIMA, *Neural Network* dan *hybrid neural network* selain itu juga mengamati penerapan analisis jaringan untuk pengelompokan obat-obatan dan menggunakannya sebagai variabel input. Hasil penelitian melaporkan bahwa penggabungan model linier dan nonlinier dapat meningkatkan kinerja peramalan. Penerapan pembelajaran mesin disajikan oleh Pavlyshenko, 2019 [16] untuk analitik prediktif penjualan. Penelitian ini melaporkan bahwa pendekatan regresi untuk peramalan penjualan dapat memberikan hasil yang lebih baik dibandingkan dengan metode deret waktu. Selain itu pendekatan algoritma LightGBM [17], [18] maupun Prophet [19] merupakan salah algoritma yang perlu di pertimbangkan dalam membangun peramalan yang optimal dalam bisnis [20]–[23].

## BAHAN DAN METODE

Metode yang diusulkan dalam penelitian menggunakan CRISP-DM untuk mengikuti semua proses yang diperlukan selama mengerjakan proyek, yaitu:

### 1. Analisis Bisnis

Peramalan penjualan adalah proses memperkirakan penjualan masa depan baik jangka pendek atau jangka panjang dengan menggunakan angka penjualan historis. Peramalan penjualan sangat penting bagi perusahaan, misalnya, berinvestasi dalam metodologi baru untuk meningkatkan pendapatan produk mereka yang dapat menyebabkan penjualan rendah di masa depan melalui promosi/diskon, dll. Oleh karena itu, untuk bisnis yang sudah mapan, penting untuk memperkirakan penjualan di masa depan akurat. Dalam penelitian ini tujuan utama adalah memperkirakan penjualan harian selama 28 hari ke depan dengan menggunakan algoritma peramalan yang berbeda. Data penjualan yang disediakan mencakup data toko di tiga negara bagian AS (California, Texas, dan Wisconsin) dan mencakup tingkat barang, departemen, kategori produk, dan detail toko. Selain itu, data memiliki variabel penjelas seperti harga, promosi, hari dalam seminggu, dan acara khusus.

### 2. Dekripsi Data

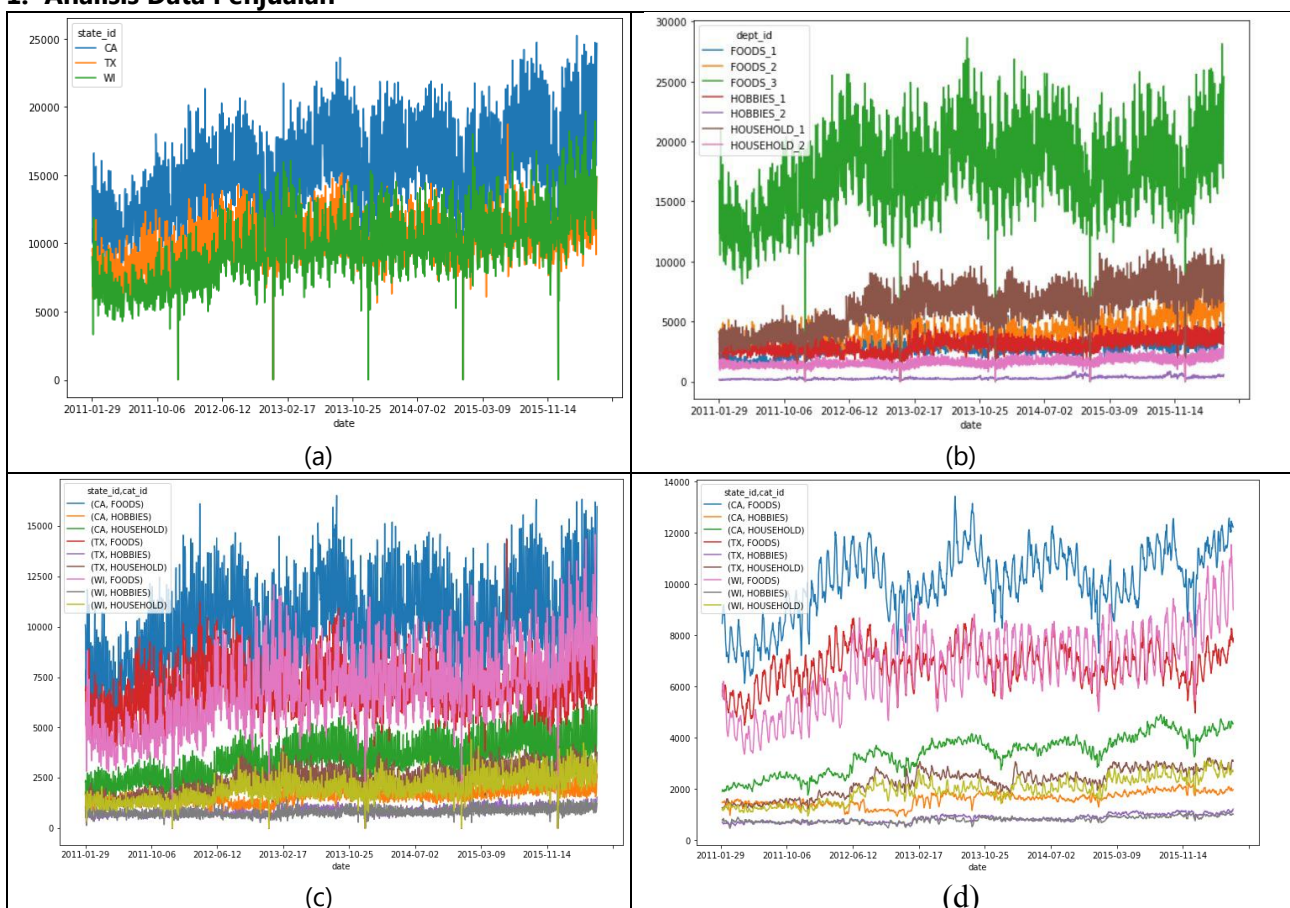
Dataset M5 Forecasting terdiri dari riwayat penjualan 3049 produk unik pada 3 kategori berbeda di tiga negara bagian AS bersama dengan detail harga setiap produk dan acara yang diadakan setiap hari. Data ini dibagi menjadi tiga file CSV, yaitu:

- Calendar.csv: Ini berisi informasi tentang tanggal penjualan produk dan acara serta program yang diadakan pada hari itu.
- Sales\_train\_evaluation: Ini berisi penjualan unit harian historis dari setiap produk di setiap toko dari hari 1 hingga hari 1941.
- Sell\_prices: Ini berisi informasi tentang harga produk pada setiap minggu per toko.

### 3. Analisis Data Eksplorasi

Tahapan pertama dalam pendekatan untuk memecahkan masalah pembelajaran mesin adalah mendapatkan wawasan dari data. Dalam langkah ini, menganalisis data jika ada nilai yang hilang untuk diproses sebelumnya, sampai pada kesimpulan yang berarti tentang bagaimana fitur tertentu memengaruhi target penjualan. langkah paling penting untuk memahami data dan membuat fitur yang akan digunakan. Berikut hasil analisis data pada langkah ini.

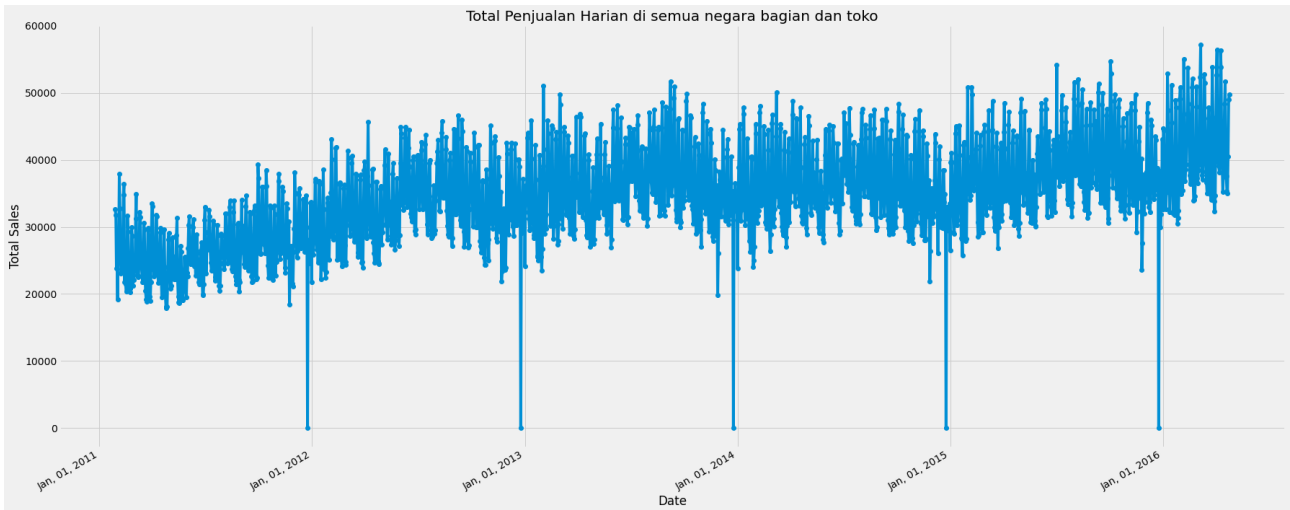
#### 1. Analisis Data Penjualan



Gambar 1 Analisis Data Penjualan berdasarkan (a) kategori, (b) departemen, (c) kota, (d) tren penjualan Berdasarkan gambar 1 dapat dilihat bahwa produk kebanyakan dijual di CA khususnya di toko CA\_3, kemudian produk departemen Food\_4 paling banyak terjual dibandingkan produk makanan lainnya dan kategori lainnya dan terdapat tren peningkatan pertumbuhan penjualan yang mendorong pendapatan perusahaan. Pada gambar 1(d) dengan rata-rata bergulir 7, tren musiman naik yang unik untuk setiap kategori dapat diamati di setiap negara bagian.

#### 2. Analisis Transaksi Penjualan Harian dan Tahunan

Analisis tren transaksi penjualan harian dan tahunan merupakan salah bagian penting dalam analisis bisnis untuk membangun model peralaman.



Gambar 2 Tren Penjualan harian dan Tahunan

Dari gambar 2 tren penjualan harian dapat diamati bahwa ada sedikit tren peningkatan total penjualan dari 2011 hingga 2016 dan pada 1 Januari setiap tahun, kami mengamati adanya penjualan nol yang kemungkinan toko tutup pada Malam Tahun Baru. Selanjutnya, pola yang sama dari keseluruhan penjualan sepanjang tahun dan memiliki musiman tahunan. Dari 2012 hingga 2015 penjualan keseluruhan kami mengamati setiap penjualan 2 bulan memiliki pola yang sama dari sini dapat mengasumsikan bahwa setiap tahun deret waktu 2012-2015 adalah deret waktu stasioner.

#### 4. Metrik Kinerja

Kami akan menggunakan RMSE (*Root Mean Square Error*) untuk mengevaluasi model. RMSE adalah ukuran yang baik ketika kita lebih peduli tentang prediksi. Ini adalah akar kuadrat dari kesalahan kuadrat rata-rata.

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (\text{Predicted}_i - \text{Actual}_i)^2}{N}} \quad (1)$$

RMSE lebih mementingkan kesalahan yang paling signifikan. Jika ada sedikit kesalahan dalam permintaan RMSE kami akan meningkat secara signifikan.

#### 5. Permodelan

Pada tahapan ini diusulkan lima algoritma berbeda yaitu Linear Regression (LR), Ridge Regressor (RR), XGBoost Classifier, LGBM dan Long short-term memory (LSTM). Semua hasil peramalan algoritma yang diusulkan akan dievaluasi berdasarkan *root mean squared error* (RMSE).

### HASIL PENELITIAN

Pada penelitian ini bertujuan untuk memperkirakan Penjualan adalah faktor kunci dalam masalah ini, jadi kami hanya akan mengambil penjualan sebagai fitur yaitu kumpulan data yang memiliki penjualan d1 hingga d1941. Jadi, kami hanya perlu mengambil 1941 kolom sebagai data. kemudian memprediksi penjualan 28 hari ke depan dengan benar. Dalam model dasar, kami hanya mengambil penjualan sebagai fitur dan itu juga variabel output (y) berisi 28 variabel, jadi yang memprediksi penjualan 28 hari ke depan ketika diberikan produk dengan penjualan sebelum 28 hari, tetapi di sini kami tidak menggunakan data yang lain seperti fitur kalender, kolom harga jual. Jadi, salah satu cara efektif untuk menampilkan data deret waktu ini adalah membingkai ulang data yang diberikan ke dalam pembelajaran terawasi sebagai kumpulan data tunggal yaitu sebagai variabel keluaran hanya berisi penjualan dan sisanya semua fitur sebagai masukan, jadi di sini variabel output (y) akan memprediksi hanya satu hari penjualan sehubungan dengan detail produk yang diberikan. Pertama membuat bingkai Data tunggal dengan menggabungkan data penjualan, data kalender, data harga. Ini dapat dilakukan dengan mengubah format data lebar menjadi format data panjang dan menggabungkan semua kumpulan data lainnya. Kemudian menangani nilai-nilai yang hilang karena berasal dari data harga jual, melalui teknik

imputasi rata-rata. Daripada mengambil rata-rata sederhana, kami menggunakan harga rata-rata barang sehubungan dengan tokonya karena harga dapat bervariasi berdasarkan toko.

Sebelum melatih model, perlu membagi data menjadi data latih, uji dan validasi yaitu data latih bersumber dari data penjualan hari (1–1913), validasi (1914–1941) penjualan, dan pengujian (1942–1969). Setelah melatih model dan mengukur metrik kinerja, kami perlu memprediksi penjualan untuk data yang tidak terlihat dan berikut hasil pengujian semua algoritma yang diusulkan.

Tabel 1. Evaluasi Hasil Peramalan

Model	RMSE
Linear Regression	11.07
Ridge Regressor	12.8
LGBM Regressor	7.53
XGboost	11.07
LSTM	79.61

Berdasarkan tabel 1 evaluasi hasil evaluasi kinerja masing-masing algoritma menggunakan RMSE, algoritma LGBM Regressor menghasilkan tingkat kesalahan paling rendah dibandingkan algoritma lain sebesar 7.53 %, kemudian algoritma Ridge Regressor sebesar 12.8 % dan yang paling tinggi nilai kesalahan adalah algoritma LSTM sebesar 79.61 %, hal ini sesuai yang wajar karena algoritma lebih baik digunakan untuk masalah peramalan dengan jumlah data yang lebih besar. Hasil evaluasi akurasi peramalan algoritma LGBM Regressor membutuhkan waktu pelatihan sebesar 20.979 per detik dan pengujian sebesar 3.043 lebih rendah dari pada algoritma Linear Regression dan Ridge Regressor sebesar 0.005.

Tabel 2. Evaluasi Waktu Peramalan

Model	Time (s)	
	Train	Test
Linear Regression	0.652	0.005
Ridge Regressor	1.082	0.005
LGBM Regressor	49.6	3.043
XGboost	20.979	0.014
LSTM	1.689	0.891

Hasil analisis data ketiga negara bagian tersebut memiliki penjualan akhir pekan yang tinggi dibandingkan dengan hari kerja. Di California dan Texas, penjualan hari Minggu lebih tinggi daripada hari Sabtu, tetapi di Wisconsin, penjualan hari Minggu lebih rendah daripada hari Sabtu. Di semua toko California dan Texas, penjualan akhir pekan tinggi dan penjualan hari Minggu lebih tinggi daripada penjualan hari Sabtu. Toko CA\_3 memiliki penjualan akhir pekan yang tinggi seperti yang kita lihat CA\_3 juga memiliki pendapatan tinggi dan proporsi penjualan tertinggi. Di semua toko Wisconsin, penjualan akhir pekan tinggi dan penjualan hari Minggu lebih sedikit daripada penjualan hari Sabtu. Sepanjang tahun penjualan akhir pekan tinggi dibandingkan dengan penjualan hari kerja. Pada tahun 2011 dan 2013 penjualan hari Sabtu lebih tinggi dibandingkan hari Minggu

## Diskusi

Berdasarkan hasil analisis yang dilakukan pada dataset di seluruh negara bagian dan toko untuk semua hari mulai dari 29 Januari 2011 hingga 24 April 2016 dan mengamati bahwa ada peningkatan pola penjualan selama bertahun-tahun yaitu, tren naik. Kami juga mengamati penjualan pada hari pertama tahun ini adalah nol (karena toko tutup pada hari tahun baru). Melakukan analisis tahun penjualan, mengamati bahwa ada musiman tahunan dan deret waktu setiap tahun adalah deret waktu stasioner.



Secara keseluruhan penjualan bulanan tertinggi pada bulan Maret dan terendah pada bulan November, penjualan minggu tertinggi di akhir pekan dibandingkan dengan hari kerja (Senin sampai Kamis) dan penjualan Jumat sedikit lebih tinggi dibandingkan dengan hari kerja lainnya. Negara bagian California memiliki penjualan tertinggi dan juga pendapatan tertinggi yang dihasilkan. Padahal penjualan dan pendapatan Texas dan Wisconsin yang dihasilkan hampir sama. Kategori makanan menyumbang 70% penjualan tetapi hanya menghasilkan 58% pendapatan. Penjualan hari SNAP sedikit lebih tinggi dibandingkan dengan penjualan hari non-SNAP di ketiga negara bagian. Pola penjualan keseluruhan dari setiap kategori serupa dan produk Makanan karena jumlahnya lebih banyak, keseluruhan penjualan dan pendapatan yang dihasilkan lebih tinggi. Ada banyak pola penjualan harian non-periodik dari masing-masing produk. Acara olahraga dan budaya mengalami peningkatan penjualan dan hari libur nasional mengalami penurunan penjualan. Penjualan tinggi di awal bulan dan menurun secara bertahap menjelang akhir bulan. Penjualan harian tertinggi tertinggi adalah untuk produk Makanan dan berasal dari negara bagian California dan Texas.

## KESIMPULAN

Penelitian ini menggunakan beberapa lagoritma berbeda untuk memperkirakan penjualan harian selama 28 hari ke depan. Pendekatan algoritma LGBM Regressor menghasilkan nilai kesalahan terendah dibandingkan semua algoritma, namun membutuhkan waktu komputasi paling besar dibandingkan algoritma lainnya. Model LSTM juga bisa dipilih jika dataset cukup besar karena bisa mendapatkan pola dengan benar tetapi tentu saja akan membutuhkan perangkat yang kuat yang dilengkapi dengan GPU dan RAM yang tinggi juga. Penelitian di masa depan Mencoba dengan teknik deep learning seperti model *Convolutional neural network* (CNN), *Recurrent Neural Network* (RNN), selain itu perlu juga mempertimbangkan menggunakan data keseluruhan dengan model berkinerja lebih baik mungkin ada peluang untuk mendapatkan hasil yang baik, lebih sedikit fitur yang digunakan dapat meningkatkan kinerja model.

## Supplementary Materials (optional)

Sumber dataset tersedia di <https://www.kaggle.com/c/m5-forecasting-accuracy/data>

## Kontribusi Penulis

Semua Penulis memiliki kontribusi yang sama dalam makalah ini Semua penulis telah membaca dan menyetujui versi manuskrip yang diterbitkan.

## Konflik kepentingan

Para penulis menyatakan tidak ada konflik kepentingan.

## REFERENCES

- [1] O. B. Sezer, M. Ugur Gudelek, and A. M. Ozbayoglu, "Financial time series forecasting with deep learning: A systematic literature review: 2005-2019," *arXiv*, pp. 2005–2019, 2019.
- [2] V. Kuznetsov and M. Mohri, "Theory and algorithms for forecasting time series," *arXiv*, no. 1994, pp. 1–41, 2018.
- [3] B. Lim and S. Zohren, "Time Series Forecasting With Deep Learning: A Survey," *arXiv*, Apr. 2020, [Online]. Available: <http://arxiv.org/abs/2004.13408>.
- [4] S. Lahmiri, "Minute-ahead stock price forecasting based on singular spectrum analysis and support vector regression," *Appl. Math. Comput.*, vol. 320, pp. 444–451, 2018, doi: 10.1016/j.amc.2017.09.049.
- [5] M. Abe and H. Nakayama, "Deep learning for forecasting stock returns in the cross-section," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 10937 LNAI, pp.

- 273–284, 2018, doi: 10.1007/978-3-319-93034-3\_22.
- [6] M. Bin Shams, M. J. Hossain, and S. R. H. Noori, "A Time Series Analysis of Trends With Twitter Hashtags Using LSTM," in *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, Jul. 2020, pp. 1–6, doi: 10.1109/ICCCNT49239.2020.9225349.
- [7] J. Cao and J. Wang, "Stock price forecasting model based on modified convolution neural network and financial time series analysis," *Int. J. Commun. Syst.*, vol. 32, no. 12, pp. 1–13, 2019, doi: 10.1002/dac.3987.
- [8] P. F. Procacci and T. Aste, "Forecasting market states," *Quant. Financ.*, vol. 19, no. 9, pp. 1491–1498, 2019, doi: 10.1080/14697688.2019.1622313.
- [9] N. Ahmad, "Sale Forecasting of Merck Pharma Company using ARMA Model," vol. 6, no. 21, pp. 30–36, 2015.
- [10] U. Ugurlu, O. Tas, and U. Gunduz, "Performance of Electricity Price Forecasting Models: Evidence from Turkey," *Emerg. Mark. Financ. Trade*, vol. 54, no. 8, pp. 1720–1739, 2018, doi: 10.1080/1540496X.2017.1419955.
- [11] M. A. Khan *et al.*, "Effective Demand Forecasting Model Using Business Intelligence Empowered with Machine Learning," *IEEE Access*, vol. 8, pp. 116013–116023, 2020, doi: 10.1109/ACCESS.2020.3003790.
- [12] M. Khashei and Z. Hajirahimi, "A comparative study of series arima/mlp hybrid models for stock price forecasting," *Commun. Stat. Simul. Comput.*, vol. 48, no. 9, pp. 2625–2640, 2019, doi: 10.1080/03610918.2018.1458138.
- [13] N. Khalil Zadeh, M. M. Sepehri, and H. Farvaresh, "Intelligent sales prediction for pharmaceutical distribution companies: A data mining based approach," *Math. Probl. Eng.*, vol. 2014, 2014, doi: 10.1155/2014/420310.
- [14] R. Garnier and A. Belletoile, "A multi-series framework for demand forecasts in E-commerce," *arXiv*, 2019.
- [15] M. Li, S. Ji, and G. Liu, "Forecasting of Chinese E-Commerce Sales: An Empirical Comparison of ARIMA, Nonlinear Autoregressive Neural Network, and a Combined ARIMA-NARNN Model," *Math. Probl. Eng.*, vol. 2018, pp. 1–12, 2018, doi: 10.1155/2018/6924960.
- [16] B. M. Pavlyshenko, "Using bayesian regression for stacking time series predictive models," *Proc. 2020 IEEE 3rd Int. Conf. Data Stream Min. Process. DSMP 2020*, pp. 305–309, 2020, doi: 10.1109/DSMP47368.2020.9204312.
- [17] T. Weng, W. Liu, and J. Xiao, "Supply chain sales forecasting based on lightGBM and LSTM combination model," *Ind. Manag. Data Syst.*, vol. 120, no. 2, pp. 265–279, 2019, doi: 10.1108/IMDS-03-2019-0170.
- [18] X. Chen, K. Zhou, H. Pan, and X. Chen, "Anode effect prediction based on Light Gradient Boosting Machine," *Proc. 2019 IEEE 8th Data Driven Control Learn. Syst. Conf. DDCLS 2019*, pp. 499–503, 2019, doi: 10.1109/DDCLS.2019.8908974.
- [19] A. M. Husein, J. P. Hutabarat, J. E. Sitorus, T. Giawa, and M. Harahap, "Predicting the Spread of the Corona Virus (COVID-19) in Indonesia: Approach Visual Data Analysis and Prophet Forecasting," *Int. J. Artif. Intell. Res.*, vol. 4, no. 2, p. 151, 2020, doi: 10.29099/ijair.v5i1.192.
- [20] I. Kaipov and A. Nedzved, "Sales forecasting of goods in shoe retail," *Cent. Eur. Res. J.*, vol. 6, no. September, pp. 10–17, 2020.
- [21] K. Bandara, P. Shi, C. Bergmeir, H. Hewamalage, Q. Tran, and B. Seaman, "Sales demand forecast in E-commerce using a long short-term memory neural network methodology," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11955 LNCS, pp. 462–474, 2019, doi: 10.1007/978-3-030-36718-3\_39.
- [22] B. M. Pavlyshenko, "Machine-learning models for sales time series forecasting," *Data*, vol. 4, no. 1, pp. 1–11, 2019, doi: 10.3390/data4010015.
- [23] H. Larsson, "Forecasting Daily Supermarkets Sales with Machine Learning," 2020.