

Time Series Forecasting Menggunakan Deep Gated Recurrent Units

Arwansyah¹, Suryani^{2*}, Hasyrif SY³, Usman⁴, Ahyuna⁵, Samsu Alam⁶

^{1,2,3,4,5,6}Universitas Dipa Makassar, Indonesia

¹arwansyah@undipa.ac.id, ²suryani187@undipa.ac.id, ³hasyrif@gmail.com, ⁴usmanstmikdp@gmail.com,

⁵Ahyuna@undipa.ac.id, ⁶alam@undipa.ac.id



Histori Artikel:

Diajukan: 22 Juni 2024

Disetujui: 27 Juni 2024

Dipublikasi: 30 Juni 2024

Kata Kunci:

Time series forecasting;
Gated Recurrent Unit
(GRU); Deep Learning;
Recurrent Neural Networks
(RNN); Mean Absolute
Error (MAE); Root Mean
Squared Error (RMSE)

Digital Transformation

*Technology (Digitech) is an
Creative Commons License This
work is licensed under a
Creative Commons Attribution-
NonCommercial 4.0
International (CC BY-NC 4.0).*

Abstrak

Time series forecasting memiliki aplikasi yang luas dalam berbagai bidang seperti keuangan, ekonomi, meteorologi, dan pengelolaan rantai pasokan. Penelitian ini bertujuan untuk mengevaluasi kinerja model Gated Recurrent Unit (GRU) yang mendalam dalam konteks prediksi data time series. Akurasi prediksi time series sangat penting untuk pengambilan keputusan yang lebih baik dan efisien. GRU diperkenalkan sebagai varian Recurrent Neural Networks (RNN) yang lebih efisien dan efektif dalam menangani data time series. RNN sering digunakan untuk tugas ini, namun memiliki kelemahan dalam mempelajari ketergantungan jangka panjang akibat masalah vanishing gradient. Penelitian ini mengusulkan penggunaan model GRU mendalam untuk meningkatkan akurasi prediksi data time series. Arsitektur model yang diusulkan terdiri dari tiga lapisan GRU dengan 64 unit masing-masing, diikuti oleh satu lapisan output dense. Dataset yang digunakan dalam penelitian ini adalah dataset time series yang terdiri dari AQI, Weather, Exchange Rate, dan ETT. Model dilatih menggunakan optimizer Adam dan loss function mean squared error, dengan jumlah epoch sebanyak 100 dan batch size 32. Data dibagi menjadi set training dan testing dengan rasio 80:20. Kinerja model dievaluasi menggunakan metrik Mean Absolute Error (MAE) dan Root Mean Squared Error (RMSE). Hasil eksperimen menunjukkan bahwa model GRU mendalam menghasilkan prediksi yang lebih akurat dibandingkan dengan model ARIMA, RNN, dan LSTM. Model GRU mendalam menunjukkan kemampuan yang lebih baik dalam mempelajari ketergantungan jangka panjang dalam data time series, sehingga meningkatkan akurasi prediksi. Hasil penelitian ini memberikan kontribusi signifikan dalam bidang time series forecasting dan dapat menjadi alternatif yang lebih efektif dibandingkan model tradisional dan model machine learning lainnya.

PENDAHULUAN

Time series forecasting adalah teknik yang digunakan untuk memprediksi nilai masa depan berdasarkan data historis yang berurutan dalam waktu. Metode ini sangat penting dalam berbagai bidang, seperti keuangan, ekonomi, meteorologi, kesehatan, dan manajemen rantai pasokan. Misalnya, dalam bidang keuangan, peramalan harga saham atau nilai tukar mata uang dapat membantu investor dalam membuat keputusan yang lebih baik. Di bidang meteorologi, prediksi cuaca dapat membantu dalam perencanaan aktivitas sehari-hari dan mitigasi bencana alam. Secara tradisional, metode statistik seperti ARIMA (AutoRegressive Integrated Moving Average) telah banyak digunakan untuk time series forecasting. Namun, metode ini memiliki keterbatasan dalam menangani data yang sangat kompleks dan tidak linear. Dalam beberapa dekade terakhir, metode berbasis machine learning dan deep learning telah mendapatkan perhatian yang lebih besar karena kemampuannya dalam menangani kompleksitas dan non-linearitas dalam data. Recurrent Neural Networks (RNN) adalah salah satu metode deep learning yang dirancang khusus untuk data berurutan seperti time series. RNN memiliki koneksi berulang yang memungkinkan mereka untuk mempertimbangkan informasi dari langkah waktu sebelumnya saat membuat prediksi. Namun, RNN tradisional menghadapi masalah vanishing gradient, yang menyebabkan kesulitan dalam mempelajari ketergantungan jangka panjang dalam data time series. Beberapa penelitian menggunakan sekunsial model seperti RNN telah dilakukan diantaranya Forecasting across time series databases using recurrent neural networks on groups of similar series: A clustering approach (Bandara, Bergmeir, & Smyl, 2020), penelitian ini Menggunakan pendekatan clustering untuk peramalan dengan RNN pada kelompok data time series yang serupa, meningkatkan akurasi prediksi. Penelitian lain yang menggabungkan 2 metode dengan judul penelitian A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting. Penelitian ini Menggabungkan smoothing eksponensial dengan RNN untuk peramalan time series, menghasilkan peningkatan akurasi (Smyl, 2020). Sekuensial model adalah salah satu metode yang efektif dalam memprediksi data time series

terlebih setelah LSTM banyak digunakan oleh para peneliti yang memprediksi konsumsi energi dimana penelitian ini menggunakan model LSTM hibrida untuk memprediksi konsumsi energi rumah tangga, menghasilkan prediksi yang lebih akurat (Yan et al., 2019). Selain itu terdapat juga literature yang melakukan survey terhadap penggunaan deep learning pada data time series seperti Time-series forecasting with deep learning: A survey (Lim & Zohren, 2021), survey ini adalah survei komprehensif tentang teknik deep learning untuk peramalan time series, mencakup model dan aplikasi terkini.

Untuk mengatasi masalah ini, arsitektur RNN yang lebih canggih, seperti Long Short-Term Memory (LSTM) dan Gated Recurrent Unit (GRU), telah dikembangkan. GRU, khususnya, menawarkan keuntungan dalam hal kesederhanaan dan efisiensi komputasi dibandingkan LSTM, sambil tetap efektif dalam menangani masalah vanishing gradient. Meskipun GRU telah terbukti efektif dalam berbagai aplikasi time series forecasting, sebagian besar penelitian yang ada berfokus pada penggunaan GRU dengan satu atau dua lapisan. Penelitian ini berangkat dari premis bahwa model GRU yang mendalam, yang terdiri dari beberapa lapisan GRU, dapat lebih baik dalam menangkap pola yang kompleks dan ketergantungan jangka panjang dalam data time series. Oleh karena itu, penelitian ini mengusulkan penggunaan model GRU mendalam dan mengevaluasi kinerjanya dalam prediksi data time series.

STUDI LITERATUR

Penelitian terkait berikut memberikan dasar yang kuat untuk memahami dan mengimplementasikan model LSTM serta variannya dalam berbagai aplikasi time series, serta memberikan panduan untuk fine tuning parameter guna mencapai performa optimal.

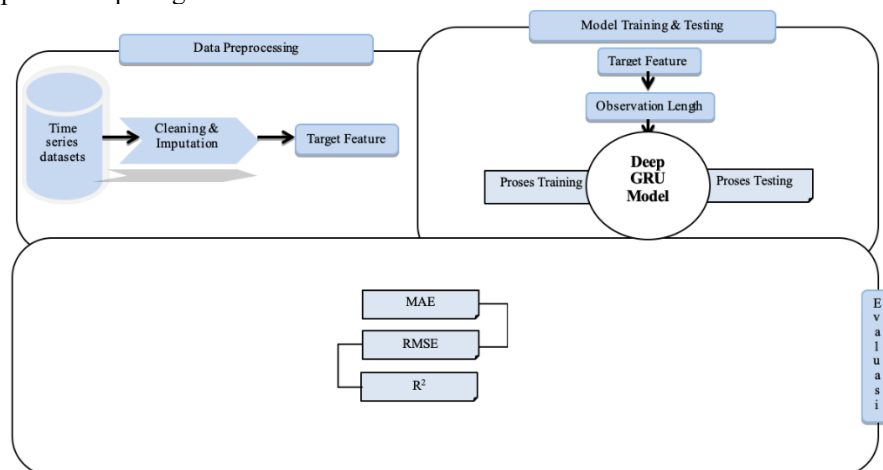
1. Penelitian terkait yang dilakukan oleh (Bandara et al., 2020) dengan judul "Forecasting across time series databases using recurrent neural networks of groups similar series : A clustering approach", menggunakan pendekatan clustering untuk peramalan dengan RNN pada kelompok data time series yang serupa untuk meningkatkan akurasi prediksi.
2. Penelitian yang dilakukan oleh (Smyl, S., 2020) dengan judul "A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting", menggabungkan smoothing eksponensial dan RNN untuk peramalan time series untuk meningkatkan akurasi.
3. Penelitian yang dilakukan oleh (Fan, J., et al., 2019) dengan judul "A hybrid LSTM Neural Network for energy consumption forecasting of individual households". Menggunakan hibrid LSTM untuk memprediksi konsumsi energi rumah tangga.
4. Penelitian berupa survey yang dilakukan oleh (Lim, B., and Zohren, S., 2020) dengan judul "Time series forecasting with deep learning : A survey", survey komprehensif tentang teknik deep learning untuk peramalan time series mencakup model terkini dan aplikasi.
5. Penelitian yang dilakukan oleh (Huang, Z. et al, 2019) dengan judul "T-GCN : A temporal group convolutional network for traffic prediction", memperkenalkan T-GCN, sebuah jaringan konvolusi graf temporal untuk prediksi lalu lintas dengan menggabungkan graf dan data time series.
6. Penelitian yang dilakukan oleh (Bandara, et al, 2020) dengan judul "LSTM-msnet: Leveraging forecasts on sets of related time series with multiple seasonal patterns", menggunakan LSTM untuk memanfaatkan pola musiman dalam data time series yang terkait, meningkatkan akurasi prediksi.
7. Penelitian yang dilakukan oleh (Liu, C et al, 2023) dengan judul "Multi-Scale GRU-CNN", mengembangkan multi scale GRU CNN untuk memprediksi suhu dan menunjukkan peningkatan akurasi prediksi.
8. Penelitian yang dilakukan oleh (Wen et al., 2020) dengan judul "A multi horizon quantile recurrent forecaster with mixed quantile CNN", menggunakan forecaster berulang dengan jaringan saraf konvolusi kuantile campuran untuk peramalan multi horizon.
9. Penelitian yang dilakukan oleh (Salinas et al, 2020) dengan judul "DeepAR: Probabilistic forecasting with autoregressive recurrent networks", memperkenalkan DeepAR, sebuah model peramalan probabilistik yang menggunakan jaringan autoregressive recurrent untuk data time series.
10. Penelitian yang dilakukan oleh (Yan, R & Ouyang, Y., 2022) dengan judul "A hybrid based on GRU and XGBoost for stock price prediction", menggabungkan GRU dan XGBoost untuk memprediksi harga saham dengan hasil prediksi yang lebih akurat dan stabil.
11. Penelitian yang dilakukan oleh (Sun et al., 2020) dengan judul "Forecasting tourist arrivals with machine learning and internet search index.". menggunakan GRU dan indeks pencarian internet untuk memprediksi kedatangan wisatawan.
12. Penelitian yang dilakukan oleh (Kong et al., 2021) dengan judul "Short-Term residential load forecasting based on LSTM and GRU". menggunakan LSTM dan GRU untuk peramalan beban residential jangka pendek dimana hasilnya menunjukkan kinerja yang superior dibandingkan

model tradisional.

13. Penelitian yang dilakukan oleh (T. Kim & Kim, 2019) dengan judul "Predicting stock prices with a feature fusion LSTM-CNN model using different representations of the same data", menggabungkan LSTM dan CNN untuk memprediksi harga saham, menggunakan representasi data yang berbeda untuk meningkatkan akurasi.
14. Penelitian yang dilakukan oleh (Guo et al, 2021) dengan judul "Multi-Step ahead wind speed forecasting using time series decomposition and GRU Networks", menggabungkan dekomposisi time series dengan jaringan GRU untuk peramalan kecepatan angin multi langkah dimana hasil yang diperoleh lebih akurat.
15. Penelitian yang dilakukan oleh (Cai et al. 2020) dengan judul "A novel combined model based on CNN and Bi-directional GRU for urban traffic flow prediction", menggabungkan CNN dan BiGRU untuk memprediksi arus lalu lintas perkotaan dimana hasil penelitian ini menunjukkan kinerja peningkatan akurasi prediksi.
16. Penelitian yang dilakukan oleh (Ma et al., 2020) dengan judul "Daily PM 2.5 concentration forecasting using hibryd models based on GRU and LSTM", menggunakan model GRU dan LSTM untuk memprediksi konsentrasi kualitas udara berupa PM2.5
17. Penelitian yang dilakukan oleh (Zheng et al., 2020) dengan judul "Short-Term load forecasting using EMD-LSTM neural networks with a GRU attention mechanism", menggabungkan EMD, LSTM dan GRU untuk prediksi beban jangka pendek.
18. Penelitian yang dilakukan oleh (Raza,A., & Ding, C., 2021) dengan judul "An improved deep learning model for load forecasting in Smart Grid", meningkatkan model GRU untuk digunakan dalam memprediksi beban di smart grid.
19. Penelitian yang dilakukan oleh (Xu et al., 2020) dengan judul "Forecasting container throughput using deep learning with GRU network", menggunakan GRU model dalam memprediksi throughput container.

METODE

Metode penelitian yang dilakukan melalui beberapa tahapan antara lain : Pengumpulan Data, yaitu langkah awal dalam metodologi ini. Penelitian ini menggunakan kaggle public dataset yakni ETT, AQI, dan Energy dimana untuk dataset ETT kami menggunakan dua jenis yakni ETTh1,dan ETTm2. Workflow dari deep GRU model pada penelitian ini dapat dilihat pada gambar 1.



Gambar 1. Model Workflow

Tahap selanjutnya adalah Pra-Pemrosesan Data. Setelah tahap pengumpulan data dilakukan, tahap berikutnya adalah melakukan pengecekan dari dataset yang akan digunakan, hal ini mencakup, normalisasi atau mengubah nilai dari suatu kolom ke dalam range antara -1 dan 1, selain itu kami juga melakukan pembersihan data dan inputasi hal ini dilakukan untuk memastikan bahwa dataset yang akan di training dan di testing sesuai dengan standar pemrosesan data pada deep learning. Metode normalisasi yang digunakan adalah min-max scaler menggunakan sklearn library. Untuk inputasi kami menggunakan teknik mean yakni mengambil nilai rata-rata dari dua data yang terdekat.

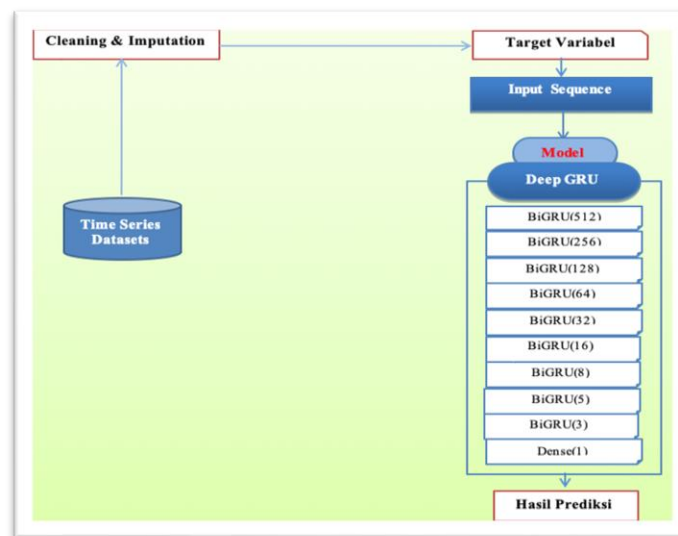
Setelah tahap Pra-Pemrosesan Data dilakukan Pembagian Data: Pada penelitian ini kami data akan dibagi menjadi set pelatihan dan set pengujian untuk evaluasi model. Untuk itu, dataset akan kami gunakan 80% sebagai data training dan 20% sebagai data testing. Hal ini berlaku untuk semua datasets yang digunakan.

Selanjutnya tahap pembangunan deep GRU model. Deep GRU model dibangun menggunakan beberapa layer Bidirectional dimana masing- masing layer memiliki jumlah neuron yang berbeda. Terdapat 3 layer yaitu input layer, hidden layer dan output layer. Untuk hidden layer memiliki beberapa layer diantaranya layer pertama dengan 512 neuron, layer kedua 256 neuron, layer ketiga 128 neuron, layer ke empat 64 neuron, layer kelima 32 neuron, layer keenam 16 neuron, layer ketujuh 8 neuron, layer kedelapan 5 neuron, layer kesembilan 3 neuron, dan terakhir 1 dense layer. Deep GRU model dibangun menggunakan library TensorFlow. Struktur model mencakup beberapa layer yang memiliki jumlah neuron yang berbeda. Selain tensorflow kami juga menggunakan library lain seperti pandas, numpy dan matplotlib. Untuk mengukur akurasi dari model yang dibuat, maka penulis menggunakan 3 matriks evaluasi yaitu RMSE, MAE dan R². Parameter yang digunakan dalam membangun model Deep GRU dapat dilihat pada tabel 1.

Tabel 1. Parameter

No	Parameter
1	Activation function (tanh, sigmoid)
2	Learning rate (0.0001)
3	Beta 1 (0.9)
4	Beta 1 (0.99)
5	Batch Size (8)
6	Optimizer(Adam)
7	Loss function(MSE)
8	Epoch (30)

Model arsitektur yang dibangun dapat dilihat pada gambar 2 berikut:



Gambar 2. Model Arsitektur

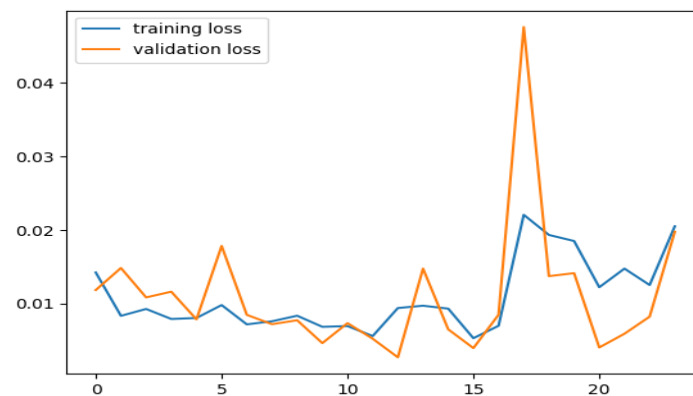
HASIL

Hasil akurasi dari model yang dibuat menggunakan 3 matrik evaluasi pada masing-masing datasets dapat dilihat pada tabel 2. Angka pada masing-masing baris menunjukkan nilai error berdasarkan hasil testing pada masing-masing dataset.

Tabel 2. Tabel Matrik Evaluasi Model

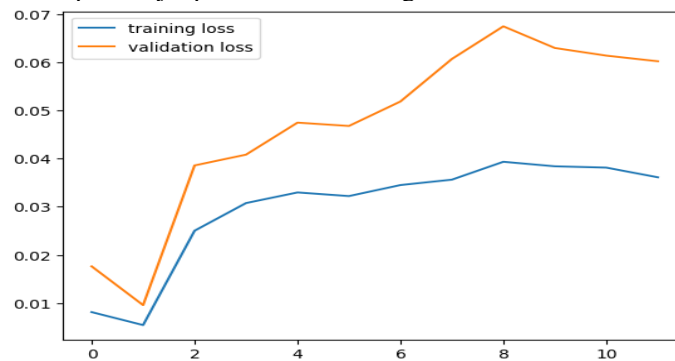
Dataset	MSE	MAE	R ²
ETTh	0.0188	0.1062	0.0530
ETTm	0.0103	0.0726	0.6170
Shanghai AQI	0.1464	0.3267	0.1100
Energy dataset	0.0001	0.0105	0.9985

Sementara training history dapat dilihat pada gambar 3 sampai gambar 6.



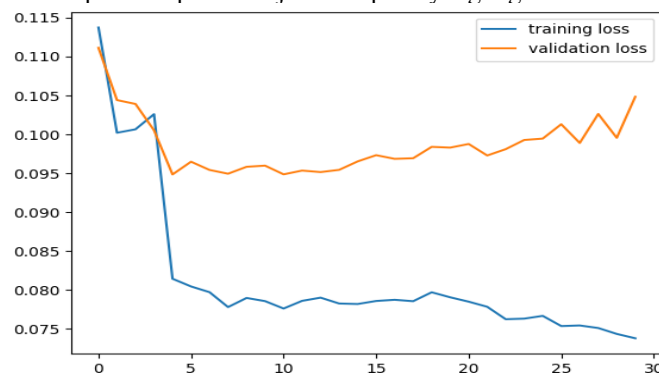
Gambar 3. Training History ETTh Dataset

Gambar 3 merupakan proses history dimana sumbu x mendefinisikan jumlah epoch dan sumbu y mendefinisikan nilai error yang didapatkan selama proses training. Proses training pada dataset ETTh pada awal epoch mengalami fluktuasi sampai pada epoch ke 15. Setelah epoch 15 training proses mengalami peningkatan error namun diakhir training model berhasil menurunkan nilai error ke angka 0.02 sehingga hal tersebut membuktikan bahwa model dapat belajar pola sekuensial dengan baik.



Gambar 4. Training History ETTm Dataset

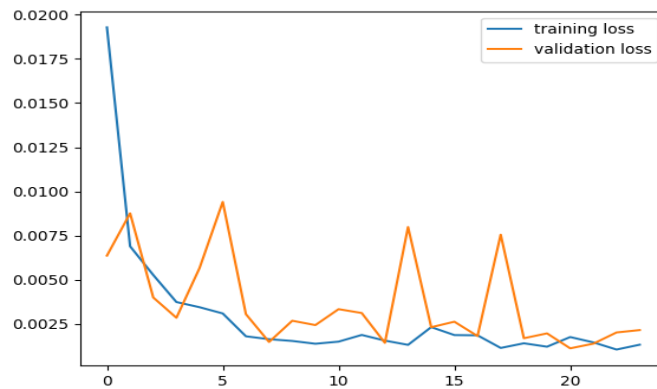
Gambar 4 merupakan training histori untuk dataset ETTm. Berdasarkan grafik pada gambar tersebut terlihat bahwa dari awal epoch terjadi peningkatan nilai error sampai pada akhir epoch. Namun hal tersebut tidak berdampak signifikan sebab model masih dapat memprediksi dengan baik yang dibuktikan diakhir epoch nilai dari loss berkisar 0.06. Nilai tersebut cukup rendah untuk model prediksi menggunakan data time series. Peningkatan nilai error disebabkan ukuran dataset yang cukup besar sehingga model membutuhkan epoch yang seharusnya lebih banyak namun pada eksperimen jumlah epoch yang digunakan adalah 12.



Gambar 5. Training History Shanghai AQI Dataset

Gambar 5 merupakan grafik dari proses training menggunakan dataset AQI. Berdasarkan grafik tersebut dapat disimpulkan bahwa proses training berjalan cukup efektif sebab terjadi penurunan nilai error di awal epoch sampai pada akhir epoch. Walaupun nilai validation mengalami peningkatan namun hal tersebut tidak berdampak pada performa model dalam memprediksi data time series. Nilai error yang berkisar 0.07 pada akhir epoch

menunjukkan model yang dibuat dapat mengekstrak fitur dan melakukan prediksi dengan baik.



Gambar 6. Training History Energy Dataset

Gambar 6 merupakan grafik training histori menggunakan dataset energy. Pada grafik tersebut terjadi fluktuasi nilai loss yang didapatkan untuk validasi namun tetap mengalami penurunan sampai pada akhir epoch. Berbeda halnya dengan training loss yang mengalami penurunan nilai loss yang cukup signifikan sampai pada akhir epoch dengan nilai yang diperoleh sekitar 0.0020. Training untuk dataset ini adalah training yang lebih baik dibandingkan training pada dataset yang lain. Dengan demikian dapat disimpulkan bahwa dataset energy adalah dataset dengan proses training yang lebih baik dan hasil prediksi yang lebih akurat.

PEMBAHASAN

Berdasarkan hasil yang diperoleh dari matrik evaluasi menggunakan MSE, MAE, R2 dapat diketahui bahwa model bekerja dengan baik pada 4 dataset yang berbeda dengan memperoleh nilai error yang cukup rendah. Dari 4 dataset yang digunakan, 3 diantaranya memiliki tingkat error yang rendah yaitu ETTh, ETTm, dan energy dataset. Sehingga perhitungan rata-rata dari tingkat error yang diperoleh cukup signifikan dimana hal tersebut membuktikan bahwa model yang dibangun dapat memprediksi secara akurat data time series yang berbeda. Sementara pada training history untuk semua dataset juga memperlihatkan bahwa terjadi penurunan nilai error selama proses training. Selain itu, nilai error masih dapat diturunkan dengan melakukan beberapa pengaturan yang meliputi jumlah epoch, sequence length, prediction length, learning rate dan pemilihan algoritma optimisasi. Pengaturan parameter dapat dilakukan apabila kriteria sumber daya berupa komputer untuk proses training memiliki spesifikasi yang baik. Sebagai contoh jika jumlah epoch ditingkatkan maka secara otomatis dapat meningkatkan akurasi atau memperkecil nilai error namun disisi lain hal tersebut membutuhkan waktu komputasi yang cukup tinggi sebab dataset yang digunakan cukup besar sehingga jika menggunakan spesifikasi komputer yang rendah maka menimbulkan kendala dalam berupa waktu training yang cukup lama. Pada gambar 3. diketahui bahwa pada proses training untuk dataset ETTh dengan jumlah epoch sebanyak 24 dimana pada epoch 1 - 15 model belajar agak stabil dan mengalami penurunan tingkat error yang tidak begitu signifikan. Namun setelah epoch 15 training dan validation loss menjadi tidak stabil dan tingkat error menjadi lebih tinggi. Hal ini bisa disebabkan oleh ekstraksi fitur pada dataset yang jumlahnya cukup banyak sulit untuk dianalisa oleh model. Namun pada akhir epoch model dapat menekan tingkat error ke angka 0.02. Hal ini membuktikan bahwa proses pembelajaran pola oleh model berjalan cukup baik dimana hasil prediksi untuk data testing memiliki tingkat error yang rendah. Dilain sisi, terjadi peningkatan error selama proses training yang dapat dilihat pada gambar 4. Hal ini dikarenakan dataset ETTm sangat besar dibandingkan dengan dataset yang lain sehingga menurut analisa kami, Dataset yang besar akan mengakibatkan model belajar mengekstraksi pola lebih sulit dibandingkan jika menggunakan dataset yang berukuran kecil. Namun hal tersebut tidak berdampak signifikan sebab pada hasil akhir model dapat memprediksi data testing dengan hasil yang akurat dimana tingkat error yang didapat sekitar 0.01 - 0.010. Lain halnya dengan dataset AQI pada gambar 5 dimana training loss dan validation loss berbanding terbalik. Training loss mengalami penurunan dan validation loss mengalami peningkatan error. Hal ini juga tidak berdampak signifikan pada model sebab model yang dibuat tetap dapat memprediksi data testing dengan akurat. Berbeda dengan dataset energy pada gambar 6 dimana training loss dan validation loss menunjukkan trend yang baik pada saat proses training walaupun validation loss tidak stabil sampai pada epoch terakhir namun training loss menunjukkan perubahan tingkat loss atau error yang signifikan yang membuktikan bahwa model dapat mengekstraksi dan menganalisa pola dengan baik dan akurat. Dapat kami simpulkan bahwa beberapa hal yang mempengaruhi perbedaan dari setiap dataset untuk validation dan training proses adalah 1). ukuran dataset, 2) jenis dataset, dan 3) jumlah epoch.

KESIMPULAN

Dalam penelitian ini, kami mengembangkan model GRU secara mendalam dimana terdapat beberapa layer yang memiliki jumlah neuron yang berbeda dengan total hidden layer sebanyak 9. kedalaman dari layer yang dibuat memiliki beberapa alasan diantaranya semakin dalam layer maka proses pembelajaran pola dapat semakin maksimal, semakin banyak neuron yang terkoneksi maka ekstraksi fitur akan semakin baik sehingga akan menghasilkan model dengan akurasi prediksi yang baik. Berdasarkan hasil experiment menggunakan 4 dataset yang berbeda, model yang dibuat dapat menghasilkan akurasi yang baik dan hal tersebut dibuktikan dengan nilai dari maktriks evaluasi RMSE dan MAE yang rendah untuk setiap dataset.

REFERENSI

- Bandara, K., Bergmeir, C., & Hewamalage, H. (2020). LSTM-MSNet: Leveraging forecasts on sets of related time series with multiple seasonal patterns. *IEEE Transactions on Neural Networks and Learning Systems*, 32(4), 1586–1599.
- Bandara, K., Bergmeir, C., & Smyl, S. (2020). Forecasting across time series databases using recurrent neural networks on groups of similar series: A clustering approach. *Expert Systems with Applications*, 140, 112896.
- Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *ArXiv Preprint ArXiv:1406.1078*.
- Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *ArXiv Preprint ArXiv:1412.3555*.
- Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to forget: Continual prediction with LSTM. *Neural Computation*, 12(10), 2451–2471.
- Graves, A., Mohamed, A., & Hinton, G. (2013). Speech recognition with deep recurrent neural networks. *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 6645–6649.
- Kim, H. Y., & Won, C. H. (2018). Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. *Expert Systems with Applications*, 103, 25–37.
- Kim, T., & Kim, H. Y. (2019). Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data. *PloS One*, 14(2), e0212320.
- Lim, B., & Zohren, S. (2021). Time-series forecasting with deep learning: a survey. *Philosophical Transactions of the Royal Society A*, 379(2194), 20200209.
- Ma, X., Tao, Z., Wang, Y., Yu, H., & Wang, Y. (2015). Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transportation Research Part C: Emerging Technologies*, 54, 187–197.
- Pascanu, R., Mikolov, T., & Bengio, Y. (2013). On the difficulty of training recurrent neural networks. *International Conference on Machine Learning*, 1310–1318.
- Rangapuram, S. S., Seeger, M. W., Gasthaus, J., Stella, L., Wang, Y., & Januschowski, T. (2018). Deep state space models for time series forecasting. *Advances in Neural Information Processing Systems*, 31.
- Salinas, D., Flunkert, V., Gasthaus, J., & Januschowski, T. (2020). DeepAR: Probabilistic forecasting with autoregressive recurrent networks. *International Journal of Forecasting*, 36(3), 1181–1191.
- Smyl, S. (2020). A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting. *International Journal of Forecasting*, 36(1), 75–85.
- Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. *Advances in Neural Information Processing Systems*, 27.
- Yan, K., Li, W., Ji, Z., Qi, M., & Du, Y. (2019). A hybrid LSTM neural network for energy consumption forecasting of individual households. *Ieee Access*, 7, 157633–157642.
- Yu, H.-F., Rao, N., & Dhillon, I. S. (2016). Temporal regularized matrix factorization for high-dimensional time series prediction. *Advances in Neural Information Processing Systems*, 29.
- Zaremba, W., Sutskever, I., & Vinyals, O. (2014). Recurrent neural network regularization. *ArXiv Preprint ArXiv:1409.2329*.
- Zhao, L., Song, Y., Zhang, C., Liu, Y., Wang, P., Lin, T., Deng, M., & Li, H. (2019). T-gcn: A temporal graph convolutional network for traffic prediction. *IEEE Transactions on Intelligent Transportation Systems*, 21(9), 3848–3858.