



Performance Comparison of RNN-Based Models in Forecasting Indonesian Economic and Financial Data

Perbandingan Kinerja Model Berbasis RNN pada Peramalan Data Ekonomi dan Keuangan Indonesia

Cahya Alkahfi^{1*}, Anang Kurnia², Asep Saefuddin³

¹Direktorat Neraca Produksi, Badan Pusat Statistik, Indonesia

^{2,3}Departemen Statistika, Institut Pertanian Bogor, Indonesia

E-Mail: ¹cahyaalkahfi@gmail.com, ²anangk@apps.ipb.ac.id, ³asaefuddin@gmail.com

Received May 27th 2024; Revised Jul 7th 2024; Accepted Jul 9th 2024
Corresponding Author: Cahya Alkahfi

Abstract

Time series forecasting stands as a pivotal element within economic and financial analysis, enabling stakeholders to anticipate various economic indicators prior to the official release of data. In this context, machine learning models such as Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU) exhibit promising potential in predicting time series data. Several studies affirm that LSTM and GRU outperform RNN in performance, attributed to their mechanisms for addressing the vanishing gradient issue commonly encountered in conventional RNN models. In Indonesia, many studies have been conducted on forecasting economic and financial data using RNN-based models including LSTM and GRU. However, very few have compared the performance of RNN, LSTM, and GRU models comprehensively to identify which model better captures the patterns of economic and financial movements in Indonesia. This research focuses on evaluating the performance of these three models comprehensively. To ensure a comprehensive outcome, the study will assess the models across three different types of data, including IHSG, export values, and GDP. The findings of this research indicate that overall, the GRU model demonstrates the best performance, followed by LSTM, which also competes favorably with RNN. Beyond accuracy, both GRU and LSTM models excel in performance stability, as evidenced by their relatively low standard deviations compared to RNN. This superiority becomes even more pronounced, particularly when applied to GDP models, which are available for short time periods only.

Keyword: Economy, Financial, GRU, LSTM, RNN

Abstrak

Peramalan deret waktu merupakan salah satu elemen kunci dalam analisis ekonomi dan keuangan. memungkinkan pemangku kepentingan untuk membuat perkiraan terhadap berbagai indikator ekonomi sebelum data resmi dirilis. Dalam konteks ini, model pembelajaran mesin seperti Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), dan Gated Recurrent Unit (GRU) menunjukkan potensi yang menjanjikan dalam memprediksi data deret waktu. Sejumlah penelitian juga menegaskan bahwa LSTM dan GRU mampu mengungguli kinerja RNN. Kedua model tersebut memiliki mekanisme untuk mengatasi masalah vanishing gradient yang sering ditemui pada model RNN konvensional. Di Indonesia cukup banyak penelitian peramalan data ekonomi dan keuangan menggunakan model-model berbasis RNN termasuk LSTM dan GRU. Namun, sangat sedikit yang membandingkan kinerja model RNN, LSTM dan GRU secara komprehensif untuk melihat model yang lebih baik dalam menangkap pola-pola pergerakan ekonomi dan keuangan di Indonesia. Penelitian ini menitikberatkan untuk menguji kinerja ketiga model tersebut pada data-data yang ada di Indonesia. Agar hasil lebih komprehensif, penelitian ini akan menguji model pada tiga jenis data yang berbeda meliputi IHSG, nilai ekspor dan PDB. Hasil penelitian ini mengindikasikan bahwa secara keseluruhan, model GRU menunjukkan kinerja terbaik, diikuti oleh model LSTM yang juga kompetitif dibandingkan RNN. Selain akurasi, model GRU dan LSTM juga menonjol dalam hal stabilitas kinerja, ditandai dengan simpangan baku yang relatif kecil jika dibandingkan dengan RNN. Keunggulan ini menjadi semakin signifikan terutama saat diterapkan pada model PDB dimana hanya tersedia untuk periode waktu yang pendek.

Kata Kunci: Ekonomi, GRU, Keuangan, LSTM, RNN

1. PENDAHULUAN

Peramalan deret waktu telah menjadi kajian penting di dalam analisis bidang ekonomi dan keuangan. Topik ini berkembang menjadi semakin berkembang bersama dengan penelitian Box dan Jenkins yang mengintegrasikan model Autoregressive (AR) dan Moving Average (MA) [1]. Model Autoregressive Integrated Moving Average (ARIMA) semakin luas digunakan dalam analisis deret waktu dimana O'Donovan menunjukkan penggunaan model ARIMA memberi hasil lebih baik dibandingkan model lainnya pada saat itu [2]. Selain pendekatan model-model parametrik, terdapat perkembangan yang menjanjikan dalam penggunaan model-model berbasis pembelajaran mesin untuk menangani data deret waktu. Salah satu model yang banyak diaplikasikan adalah Recurrent Neural network (RNN), yang memiliki dua variasi utama yaitu Long Short-Term Memory (LSTM) dan Gated Recurrent Unit (GRU). LSTM dan GRU merupakan jenis-jenis arsitektur RNN yang memiliki keunggulan dirancang khusus untuk mengatasi masalah klasik pada RNN yaitu *vanishing gradient* [3]. Penelitian peramalan data deret waktu dengan LSTM dan GRU juga menunjukkan hasil yang lebih baik dibandingkan model RNN. Peramalan konsumsi energi listrik menggunakan model LSTM memberikan performa yang lebih unggul dibandingkan model RNN [4]. Perbandingan empiris menunjukkan kinerja model GRU mengungguli RNN biasa serta hasilnya sebanding dengan model LSTM walaupun dengan jumlah gate yang lebih sedikit [5].

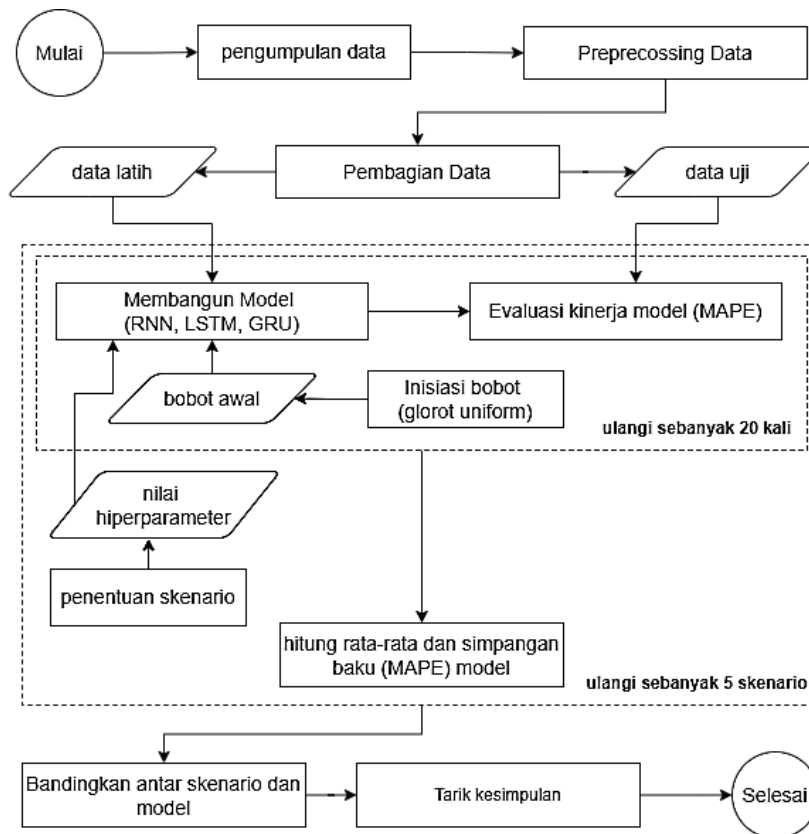
Dalam bidang ekonomi, penelitian [6] menyimpulkan bahwa model LSTM memberikan performa yang lebih baik dibandingkan model prediksi tradisional dalam meramalkan pertumbuhan ekonomi negara G20 khususnya pada negara berkembang. [7] membangun model berbasis LSTM dengan peubah penjelas untuk peramalan data perdagangan internasional pada 10 negara. Hasil empiris juga menunjukkan model LSTM hampir selalu memberikan kinerja terbaik dibandingkan model ARIMA, VAR/VECM, XGB dan beberapa model lainnya. Penelitian lainnya oleh [8] untuk peramalan indeks berjangka produk pertanian di dua pasar berjangka (Dalian dan Zhengzhou) dimana kinerja model LSTM mengungguli model ARIMA. [9] menunjukkan bahwa peramalan menggunakan model LSTM dan GRU efektif dalam menangkap pergerakan harga saham. Dalam konteks Indonesia, cukup banyak penelitian yang mengkaji peramalan data ekonomi dan keuangan Indonesia. Namun tidak banyak yang mencoba membandingkan model-model berbasis RNN secara komprehensif untuk memperoleh kesimpulan yang lebih lengkap. Penelitian [10] yang membandingkan model RNN, LSTM dan GRU pada peramalan harga saham-saham LQ45 juga menunjukkan bahwa model GRU menghasilkan kinerja terbaik. Pada penelitian ini belum mengeksplorasi berbagai kombinasi *hyperparameter* model yang bisa saja memberikan hasil berbeda. Selain itu, faktor penting lainnya yaitu penentuan bobot inisial yang bersifat acak dapat mempengaruhi kinerja akhir model [11]. Sehingga, dengan pengaturan hiperparameter model yang sama persis sangat mungkin memberikan kinerja model yang berbeda.

Penelitian ini menitikberatkan untuk membandingkan kinerja model-model berbasis RNN pada data-data ekonomi dan keuangan Indonesia, khususnya untuk memprediksi indikator ekonomi meliputi Indeks harga Saham Gabungan (IHSG), nilai ekspor serta Produk Domestik Bruto (PDB). Studi ini bertujuan untuk mengevaluasi keefektifan dan keakuratan model RNN, LSTM dan GRU dalam konteks ekonomi dan keuangan Indonesia secara komprehensif. Pengujian model akan dilakukan secara komprehensif dengan mempertimbangkan berbagai kriteria. Kriteria yang digunakan meliputi data-data dengan karakteristik dan pola yang sangat berbeda; pengaturan model untuk berbagai nilai hiperparameter; serta pengujian untuk berbagai ukuran jendela (*window*). Tidak sampai di sana, pengujian juga akan dilakukan secara repetitif. Setiap satu kombinasi karakteristik akan dilakukan beberapa kali pengulangan dengan bobot inisial yang berbeda-beda sehingga menghasilkan nilai kinerja lebih lengkap. Dengan analisis yang komprehensif, diharapkan dapat ditemukan model yang paling efektif untuk peramalan indikator perekonomian di Indonesia. Hasil penelitian ini akan bermanfaat bagi para ekonom, pembuat kebijakan, dan praktisi keuangan dalam memilih dan menerapkan model prediksi yang tepat untuk analisis ekonomi. Selain itu, penelitian ini juga diharapkan dapat membuka jalan bagi pengembangan lebih lanjut dari model-model pembelajaran mesin dalam bidang ekonomi, serta memperkuat pemahaman tentang dinamika ekonomi Indonesia yang kompleks. Dengan demikian, penelitian ini tidak hanya akan memberikan kontribusi signifikan dalam pemilihan model prediksi yang lebih baik, tetapi juga dalam pengambilan keputusan yang lebih informasional dan tepat sasaran di bidang ekonomi dan keuangan.

2. METODOLOGI PENELITIAN

Penelitian ini menitik beratkan pada perbandingan dari tiga jenis model berbasis RNN yaitu RNN, LSTM dan GRU. Adapun tahapan-tahapannya secara garis besar disajikan pada Gambar 1. Langkah utama yang menjadi bagian penting pada penelitian ini adalah penggunaan berbagai data dan skenario dalam membandingkan model. Pada setiap arsitektur model, akan dibangun 5 skenario yang meliputi mengatur hiperparameter yang berbeda-beda. Pada masing-masing skenario, juga akan dilakukan pengulangan sebanyak 20 kali dengan nilai bobot awal yang berbeda-beda sehingga kinerja model dapat diukur secara lebih lengkap. Dengan kata lain, pada satu model dengan pengaturan tertentu, kita dapat menilai berapa kinerja rata-rata,

berapa kinerja terbaik dan terburuknya. Langkah ini juga akan menghasilkan informasi bagaimana stabilitas kinerja suatu arsitektur model berdasarkan 20 percobaan tersebut melalui nilai simpangan bakunya.



Gambar 1. Tahapan Penelitian

2.1. Data

Untuk memberikan hasil yang lebih komprehensif maka penelitian akan menggunakan tiga sumber data denganyang berbeda. Data pertama yaitu data Indeks Harga Saham gabungan (IHSG) dengan periode waktu harian dari Januari 2010 hingga April 2024. Data kedua menggunakan data nilai ekspor Indonesia dalam satuan juta US\$ dengan periode bulanan Januari 2009 sampai April 2024 (184 bulan). Data ekspor diperoleh dari Badan Pusat Statistik (BPS). Data ketiga yaitu data Produk Domestik Bruto (PDB) atas dasar harga berlaku dalam satuan miliar rupiah untuk periode kuartalan (3 bulanan) dari kuartal 1 2010 sampai dengan kuartal 1 2024 (57 kuartal). Data ini juga bersumber dari BPS dan dengan periode waktu sesuai ketersediaan data yang tersedia. PDB menunjukkan besaran nilai tambah yang dihasilkan suatu negara selama periode tertentu dari seluruh sektor ekonomi [12].

Selanjutnya, masing-masing data dibagi menjadi data latih dan data uji. Untuk data IHSG data uji akan menggunakan data periode Januari-April 2024. Periode data uji untuk ekspor akan diambil data dari periode bulan Januari 2021 sampai dengan April 2024. Adapun untuk data PDB akan digunakan data kuartal ke-1 2021 sampai dengan kuartal ke-1 2024. Periode lainnya yang lebih awal pada masing-masing data dijadikan sebagai data latih.

2.2. Pra-pemrosesan Data

Terdapat dua tahapan utama yang dilakukan pada bagian pra-pemrosesan data meliputi transformasi skala data dan pembuatan *windowing* (jendela) data. Transformasi data bertujuan untuk mengubah skala data menjadi seragam. Dalam pemodelan *neural network* proses ini merupakan bagian penting karena dapat mempercepat konvergensi selama proses pembelajaran serta mampu meningkatkan akurasi model [13]. Transformasi yang akan digunakan pada penelitian ini yaitu *Min-Max Normalization* dengan rentang nilai -1 sampai 1. Penggunaan rentang nilai ini dikarenakan sejalan dengan fungsi aktivasi pada model-model *recurrent* yang berupa fungsi tanh dengan domain nilai -1 sampai +1. Adapun formula untuk melakukan transformasi *Min-Max Normalization* dengan rentang nilai tersebut diberikan pada Persamaan (1).

$$x' = 2 \cdot \frac{x - x_{\min}}{x_{\max} - x_{\min}} - 1 \tag{1}$$

Windowing, atau yang dikenal juga sebagai teknik jendela, adalah sebuah metode dalam analisis data time series yang membagi data menjadi sub-sekuen atau jendela yang lebih kecil. Setiap jendela dianalisis secara terpisah untuk mengidentifikasi pola dan tren lokal. Teknik ini bermanfaat untuk menangkap pola data, mengurangi keberadaan *noise* serta dapat meningkatkan akurasi prediksi dengan pengaturan yang sesuai [14].

2.3. Model

2.3.1. Recurrent Neural Network (RNN)

RNN adalah salah satu jenis ANN yang dapat memproses data berurutan atau sekuensial. RNN memiliki memori internal yang memungkinkannya untuk mengingat informasi dari masa lalu dan menggunakan informasi tersebut untuk memprediksi masa depan. RNN bekerja dengan cara melewatkan informasi dari setiap sampel data ke sampel berikutnya. Informasi ini kemudian disimpan di memori internal RNN dan digunakan untuk memproses sampel data berikutnya. Proses ini berulang terus menerus hingga seluruh data sekuensial telah diproses [15].

Secara umum, pada RNN keluaran dari unit-unit di dalam *hidden layer* akan dikirimkan kembali ke dalam jaringan (*feedback*). Proses ini memungkinkan RNN untuk menggunakan informasi pada proses sebelumnya sebagai bagian dalam pemrosesan data saat ini. Mekanisme pada model RNN menjadikannya cocok digunakan untuk pemodelan yang melibatkan data sekuensial seperti peramalan data deret waktu. Namun demikian, pada model RNN terdapat masalah *vanishing gradient* yang menyebabkan RNN hanya memiliki memori jangka pendek sehingga kemampuannya dalam memprediksi jangka panjang menjadi terbatas [16].

Langkah pertama pada model RNN disajikan pada Persamaan (2) yaitu menghitung *hidden state* baru h_t melalui fungsi tanh dari input saat ini x_t dan *hidden state* sebelumnya h_{t-1} . Selanjutnya, untuk *output* numerik maka nilai keluaran dari model RNN diperoleh langsung melalui Persamaan (3) yaitu total nilai *hidden state* baru berdasarkan bobot masing-masing unitnya.

$$h_t = \tau(W_x x_t + W_h h_{t-1} + b_h) \quad (2)$$

$$o_t = W_o h_t + b_o \quad (3)$$

2.3.2. Long Short-term Memory (LSTM)

LSTM merupakan jenis RNN dimana terdapat penambahan *memory cell* yang dapat menyimpan informasi untuk jangka waktu yang lebih panjang [17]. LSTM diusulkan sebagai solusi untuk menangani masalah *vanishing gradient* pada model RNN standar saat memproses data deret waktu yang panjang. Permasalahan *vanishing gradient* mengakibatkan RNN tidak dapat menangkap hubungan jangka panjang pada data deret waktu [18]. Hal ini berdampak pada berkurangnya akurasi prediksi pada model RNN [19]. Penanganan masalah *vanishing gradient* pada RNN dilakukan melalui skema memori yang terdiri dari tiga bagian yaitu *forget gate*, *input gate*, dan *output gate* [20].

$$f_t = \sigma(W_f [h_{t-1}, x_t] + b_f) \quad (4)$$

$$i_t = \sigma(W_i [h_{t-1}, x_t] + b_i) \quad (5)$$

$$\tilde{c}_t = \tau(W_c [h_{t-1}, x_t] + b_c) \quad (6)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (7)$$

$$o_t = \sigma(W_o [h_{t-1}, x_t] + b_o) \quad (8)$$

$$h_t = o_t \odot \tau(c_t) \quad (9)$$

Proses pertama dimulai dari komponen *forget gate* (f_t) dimana informasi yang tidak diperlukan hasil proses sebelumnya akan dihilangkan melalui fungsi aktivasi sigmoid pada Persamaan (4). Proses kedua adalah melalui komponen *input gate* (i_t). Berdasarkan Persamaan (5), fungsi aktivasi sigmoid akan menentukan informasi mana yang akan diperbarui ke bagian *cell state*. Selain itu, dibentuk kandidat vektor baru melalui fungsi aktivasi tanh berdasarkan Persamaan (6) yang akan ditambahkan pada *cell state*. Selanjutnya nilai *cell state* diperoleh melalui Persamaan (7) yaitu dengan perkalian titik antara *forget gate* waktu ke- t dan *cell state* pada waktu $t-1$ serta ditambah dengan perkalian titik antara *input gate* dengan kandidat *cell state* \tilde{c}_t hasil dari Persamaan (6).

Langkah terakhir yaitu pada komponen *output gate*. Nilai keluaran dari *output gate* yaitu i_t diperoleh berdasarkan fungsi aktivasi sigmoid pada Persamaan (8). Selanjutnya melalui Persamaan (9) dilakukan perkalian titik antara o_t dengan fungsi tanh pada *cell state* c_t untuk menghasilkan *output/hidden state* ke- t . Adapun vektor c_t dan vektor h_t akan menjadi nilai keluaran dari model LSTM untuk digunakan pada langkah selanjutnya.

2.3.3. Gated Recurrent Unit (GRU)

GRU pertama kali diperkenalkan oleh [5], dan telah menjadi salah satu algoritma paling menjanjikan dari berbagai arsitektur RNN. Sama halnya dengan LSTM, GRU mampu menangani masalah *vanishing gradient* yang terjadi dalam jaringan RNN standar. GRU dianggap sebagai variasi dari LSTM karena kedua algoritma ini dapat memberikan hasil yang sangat baik untuk banyak kasus empiris. GRU memiliki jumlah gate lebih sedikit dibandingkan LSTM yaitu *update gate* dan *reset gate* [21].

Tahapan pembelajaran pada model GRU sedikit lebih sederhana dibandingkan LSTM karena hanya memiliki dua *gate* yaitu *update gate* dan *reset gate*.

$$z_t = \sigma(W_z[h_{t-1}, x_t] + b_z) \tag{10}$$

$$r_t = \sigma(W_r[h_{t-1}, x_t] + b_r) \tag{11}$$

$$\tilde{h}_t = \tau(W_h[r_t \odot h_{t-1}, x_t] + b_h) \tag{12}$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \tag{13}$$

$$o_t = \tau(W_o h_t + b_o) \tag{14}$$

Langkah pertama adalah melalui Persamaan (10) menentukan bagian mana dari *hidden state* sebelumnya h_{t-1} yang akan diperbarui berdasarkan input periode saat ini x_t . Selanjutnya Persamaan (11) yaitu *reset gate* digunakan untuk mengontrol berapa banyak informasi periode sebelumnya h_{t-1} yang akan dibuang sebelum digabungkan dengan kandidat *hidden state*. Pembuatan kandidat *hidden state* \tilde{h}_t melalui Persamaan (12) dengan mempertimbangkan input saat ini x_t dan informasi masa lalu yang disesuaikan melalui perkalian titik $r_t \odot h_{t-1}$. Berikutnya dengan Persamaan (13) dilakukan penggabungan *hidden state* sebelumnya h_{t-1} dengan kandidat *hidden state* \tilde{h}_t , dan diboboti dengan *update gate* z_t . Terakhir melalui Persamaan (14) output o_t diperoleh berdasarkan *hidden state* baru h_t , menggunakan aktivasi tanah.

2.4. Perbandingan Model

Model yang akan diterapkan terdiri dari 3 jenis model *recurrent* yaitu RNN, LSTM dan GRU. Selain itu pada masing-masing model dan data juga akan diterapkan beberapa skenario seperti yang disajikan pada Tabel 1. Setiap jenis model akan menggunakan 5 skenario. Setiap skenario akan diatur 3 hiperparameter meliputi jumlah unit pada lapisan *recurrent*, jumlah unit pada lapisan *perceptron* dan jumlah *window* yang digunakan. Adapun hiperparameter lainnya akan ditetapkan secara seragam yaitu *optimizer* menggunakan *adam* dan *learning rate* sebesar 0,01 serta *batch size* 32.

Hasil akhir kinerja model memiliki ketergantungan yang sangat besar terhadap bobot inisial yang diberikan. Bobot awal ini umumnya dibangkitkan secara acak setiap kali model dilatih ulang. Hal ini menyebabkan pengaturan model yang sama persis, bisa jadi menghasilkan kinerja yang berbeda sebagai akibat dari bobot inisial acak tersebut. Agar hasil perbandingan antar model lebih komprehensif serta untuk meningkatkan peluang memperoleh bobot acak yang lebih optimal maka untuk setiap skenario akan dilakukan pengulangan sebanyak 20 kali. Setiap ulangan menggunakan bobot inisial acak yang berbeda-beda menggunakan fungsi Glorot Uniform. Selanjutnya hasil akhir kinerja model akan diputuskan berdasarkan kinerja rata-rata dari 20 ulangan tersebut.

Tabel 1. Skenario Model

Data	Skenario	Jumlah Unit recurrent	Jumlah Unit Perceptron	Window
IHSG	128-32-w1	128	32	1
	128-64-w1	128	64	1
	128-64-w2	128	64	2
	128-64-w5	128	64	5
	256-128-w5	256	128	5
EKSPOR	48-24-w1	128	32	1
	64-32-w1	128	64	1

Data	Skenario	Jumlah Unit <i>recurrent</i>	Jumlah Unit Perceptron	Window
PDB	64-32-w2	128	64	2
	128-64-w4	128	64	4
	128-64-w6	256	128	6
	32-16-w1	128	32	1
	32-16-w2	128	64	2
	32-16-w4	128	64	4
	64-32-w2	128	64	2
	64-32-w4	256	128	4

Evaluasi perbandingan model akan menggunakan metrik *Mean Absolute Percentage Error* (MAPE). Nilai MAPE menunjukkan persentase rata-rata penyimpangan absolut antara nilai prediksi terhadap nilai aktualnya. Nilai MAPE lebih mudah diinterpretasikan karena tidak tergantung dengan unit satuan data yang diukur [22]. MAPE dihitung menggunakan formula pada Persamaan (15).

$$MAPE = 100 \times \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{x}_i - x_i}{x_i} \right| \quad (15)$$

3. HASIL DAN ANALISIS

Kinerja rata-rata setiap model untuk data IHSG disajikan pada Tabel 2. Berdasarkan hasil tersebut diperoleh bahwa dari 5 skenario, model GRU memiliki kinerja yang lebih unggul pada 4 skenario. Sementara 1 skenario lainnya model LSTM menghasilkan kinerja rata-rata yang lebih baik. Secara keseluruhan rata-rata MAPE terkecil diperoleh pada model GRU untuk skenario 128-64-w5 dengan nilai 0,3695 persen. Hal ini berarti secara-rata-rata prediksi model sangat baik dengan rata-rata kesalahan sekitar 0,3695 persen saja.

Tabel 2. Rata-rata Nilai MAPE Model untuk Data IHSG (Persen)

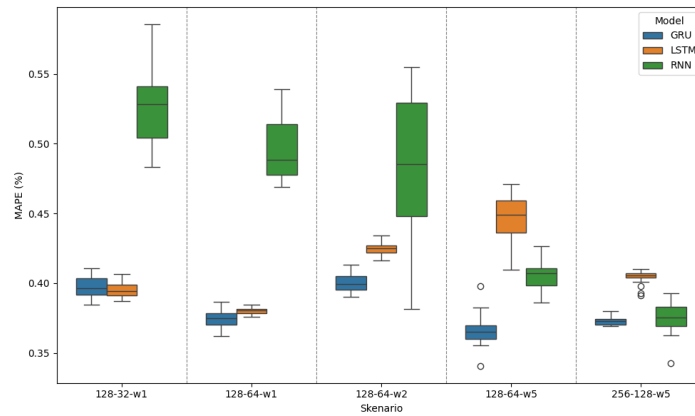
Skenario	RNN	LSTM	GRU
128-32-w1	0,5240	0,3937	0,3971
128-64-w1	0,4897	0,3792	0,3732
128-64-w2	0,4850	0,4249	0,3991
128-64-w5	0,4053	0,4487	<u>0,3695</u>
256-128-w5	0,3790	0,4054	0,3720

Tabel 3. Simpangan Baku Nilai MAPE Model untuk Data IHSG (Persen)

Skenario	RNN	LSTM	GRU
128-32-w1	0,0256	0,0058	0,0075
128-64-w1	0,0223	<u>0,0024</u>	0,0060
128-64-w2	0,0490	0,0049	0,0063
128-64-w5	0,0101	0,0162	0,0114
256-128-w5	0,0112	0,0049	0,0029

Secara lebih rinci, pada Gambar 2 dapat dilihat bagaimana sebaran kinerja setiap model dan skenario untuk data IHSG berdasarkan 20 ulangan yang dilakukan. Dari gambar ini dapat dilihat bahwa model GRU hampir selalu konsisten menghasilkan nilai MAPE yang lebih rendah dibandingkan model-model lainnya. Pada setiap skenario, nilai MAPE untuk model GRU dan LSTM menunjukkan hasil yang stabil. Hal ini berdasarkan ukuran boxplot yang cenderung kecil. Hal ini menunjukkan bahwa model mampu menghasilkan kinerja yang relatif sama meskipun dengan bobot inisial berbeda-beda. Sementara pada model RNN, khususnya pada skenario 128-32-w1, 128-64-w1 dan 128-64-w2 menunjukkan kinerja yang sangat tidak stabil atau bervariasi. Misalkan pada skenario 128-64-w2, dari 20 ulangan terdapat kinerja model dengan nilai MAPE kurang dari 0,4 namun pada ulangan yang lain terdapat kinerja model yang sangat buruk dengan nilai MAPE melebihi 0,55 persen.

Hasil tersebut dapat pula dilihat berdasarkan nilai simpangan baku kinerja model yang disajikan pada Tabel 3. Nilai simpangan baku model LSTM dan GRU cenderung lebih kecil dibandingkan model RNN. Satu pengecualian yaitu pada skenario 128-64-w5, namun pada skenario ini nilai ketiganya tidak jauh berbeda pada kisaran angka 0,01.



Gambar 2. Plot Sebaran Nilai MAPE untuk Model Data IHSG

Perbandingan kinerja model untuk data Ekspor ditampilkan pada Tabel 4, dimana menunjukkan hasil yang cenderung beragam. Dari 5 skenario, model GRU dan LSTM unggul pada masing-masing 2 skenario serta model RNN unggul pada 1 skenario. Adapun secara keseluruhan dapat dilihat bahwa model GRU menghasilkan kinerja terbaik pada skenario 128-64-w6, dengan nilai MAPE sebesar 7,36 persen.

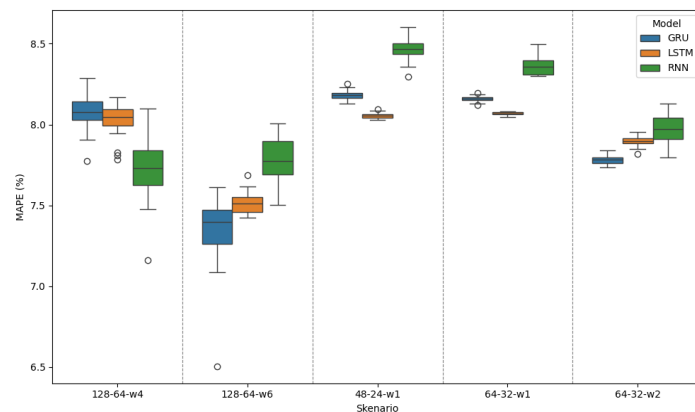
Tabel 4. Rata-rata Nilai MAPE Model untuk Data Ekspor (Persen)

Skenario	RNN	LSTM	GRU
48-24-w1	8,46	8,05	8,18
64-32-w1	8,35	8,07	8,16
64-32-w2	7,96	7,90	7,78
128-64-w4	7,87	8,06	8,11
128-64-w6	7,74	7,52	<u>7,36</u>

Tabel 5. Simpangan Baku Nilai MAPE Model untuk Data Ekspor (Persen)

Skenario	RNN	LSTM	GRU
48-24-w1	0,0697	0,0163	0,0293
64-32-w1	0,0625	<u>0,0090</u>	0,0186
64-32-w2	0,0932	0,0317	0,0273
128-64-w4	0,2080	0,1070	0,1160
128-64-w6	0,1473	0,0647	0,2320

Secara detail, melalui Gambar 3, dapat dilihat bahwa pada sebaran kinerja model GRU dan LSTM untuk data ekspor juga cenderung lebih stabil dibandingkan model RNN. Pada skenario 48-24-w1, 64-32-w1 serta 64-32-w2 model GRU dan LSTM menunjukkan kinerja yang sangat stabil dibandingkan model RNN. Sementara pada dua skenario lainnya ketiga model cenderung memiliki variasi hasil kinerja yang relatif sama. Hasil ini didukung berdasarkan tabulasi nilai simpangan baku kinerja model pada Tabel 5. Dimana terlihat bahwa model LSTM dan GRU cenderung memiliki nilai simpangan baku yang kecil.



Gambar 3. Plot Sebaran Nilai MAPE untuk Model Data Ekspor

Selanjutnya, pada Tabel 5 dapat dilihat perbandingan kinerja model untuk data PDB. Dari tabel tersebut menunjukkan bahwa kinerja model GRU dan LSTM jauh mengungguli model RNN. Nilai MAPE pada model GRU dan LSTM seluruhnya sekitar 2 kali lipat lebih baik dibandingkan model RNN. Dari 5 skenario tersebut model GRU unggul pada 3 skenario dengan nilai MAPE terbaik yaitu 1,77 persen pada skenario 64-32-w2.

Tabel 5. Rata-rata Nilai MAPE Model untuk Data PDB (Persen)

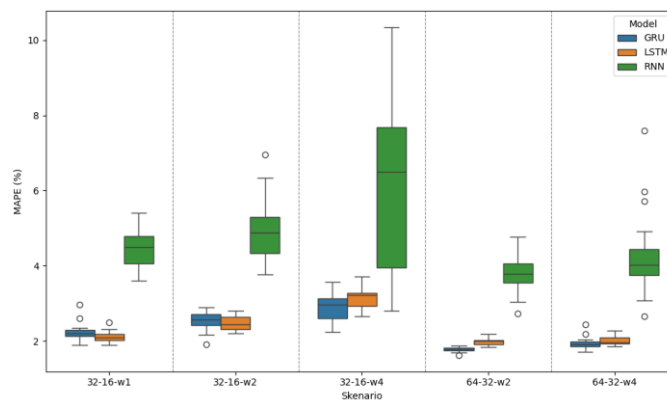
Skenario	RNN	LSTM	GRU
32-16-w1	4,47	2,09	2,20
32-16-w2	4,92	2,46	2,58
32-16-w4	5,67	3,13	2,90
64-32-w2	3,82	1,99	<u>1,77</u>
64-32-w4	4,00	1,98	1,88

Tabel 6. Simpangan Baku Nilai MAPE Model untuk Data PDB (Persen)

Skenario	RNN	LSTM	GRU
32-16-w1	0,4980	0,1354	0,2299
32-16-w2	0,8428	0,1897	0,2428
32-16-w4	2,2465	0,2665	0,3816
64-32-w2	0,4696	0,0970	0,0595
64-32-w4	1,0986	0,1185	0,1548

Melalui Gambar 4, terlihat bahwa kinerja model GRU dan LSTM unggul jauh dibandingkan model RNN. Pada setiap skenario, Gambar 4 juga menampilkan secara jelas bahwa kedua model memiliki kinerja yang sangat stabil dan berbanding terbalik dengan model RNN. Hal ini mengindikasikan bahwa pada model GRU dan LSTM, meskipun bobot inisial berbeda-beda model tetap mampu menghasilkan kinerja yang relatif sama. Sementara sebaliknya pada model RNN, kinerja yang dihasilkan sangat terpengaruh pada nilai bobot inisialnya.

Berdasarkan Tabel 6 terlihat bagaimana besaran nilai simpangan baku model RNN sangat tinggi dibandingkan model LSTM dan GRU. Khususnya pada skenario 32-16-w4 dan 64-32-w4 dimana simpangan baku model RNN hampir 10 kali lipat dibandingkan model LSTM.



Gambar 4. Plot Sebaran Nilai MAPE untuk Model Data PDB

4. KESIMPULAN

Hasil penelitian ini menunjukkan bahwa secara umum model GRU mampu menghasilkan kinerja terbaik diikuti oleh model LSTM dan RNN. Pada data dengan series yang pendek yaitu data PDB terlihat bahwa model RNN memiliki kinerja yang buruk dengan nilai MAPE lebih dari dua kali lipat dibandingkan model LSTM dan GRU. Berdasarkan nilai simpangan baku, model GRU dan LSTM juga cenderung menghasilkan kinerja yang lebih stabil dibandingkan model RNN. Nilai simpangan baku kedua model tersebut hampir selalu lebih unggul dibandingkan model RNN untuk setiap skenario. Pada data PDB yang memiliki *series* pendek perbedaan ini semakin jauh dimana simpangan baku model-model RNN hampir 10 kali lipat lebih tinggi dibandingkan model LSTM. Berdasarkan hasil ini, peneliti sangat menyarankan penggunaan GRU sebagai model peramalan data ekonomi dan keuangan di Indonesia karena memiliki kinerja yang lebih unggul serta kestabilan hasil yang juga lebih baik dibandingkan LSTM dan RNN. Saran untuk penelitian lanjutan yaitu dengan memperbanyak skenario dengan kombinasi hiperparameter lainnya seperti *optimizer*, *learning rate*, *dropout* dan sebagainya. Pencarian model terbaik dapat pula dikembangkan lebih lanjut dengan penggunaan teknik seperti *Grid Search*, *Random Search* ataupun *Bayesian Optimization*.

REFERENSI

- [1] G. E. P. Box dan G. M. Jenkins, "Time series analysis : forecasting and control," hlm. 575, 1976.
- [2] T. M. O'Donovan, "Short term forecasting : an introduction to the Box-Jenkins approach," hlm. 282, 1983.
- [3] V. Polepally, N. S. Nandini Reddy, M. Sindhuja, N. Anjali, dan K. J. Reddy, "A Deep Learning Approach for Prediction of Stock Price Based on Neural Network Models: LSTM and GRU," dalam *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, IEEE, Jul 2021, hlm. 1–4. doi: 10.1109/ICCCNT51525.2021.9579782.
- [4] N. Selle, N. Yudistira, dan C. Dewi, "Perbandingan Prediksi Penggunaan Listrik dengan Menggunakan Metode Long Short Term Memory (LSTM) dan Recurrent Neural Network (RNN)," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. Volume 9, hlm. 155–162, Feb 2022, doi: 10.25126/JTIK.2022915585.
- [5] J. Chung, C. Gulcehre, K. Cho, dan Y. Bengio, "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling," Des 2014.
- [6] S. Park dan J. S. Yang, "Interpretable deep learning LSTM model for intelligent economic decision-making," *Knowl Based Syst*, vol. 248, hlm. 108907, Jul 2022, doi: 10.1016/J.KNOSYS.2022.108907.
- [7] M. L. Shen, C. F. Lee, H. H. Liu, P. Y. Chang, dan C. H. Yang, "Effective multinational trade forecasting using LSTM recurrent neural network," *Expert Syst Appl*, vol. 182, hlm. 115199, Nov 2021, doi: 10.1016/J.ESWA.2021.115199.
- [8] S. Chen, X. Han, Y. Shen, dan C. Ye, "Application of Improved LSTM Algorithm in Macroeconomic Forecasting," *Comput Intell Neurosci*, vol. 2021, 2021, doi: 10.1155/2021/4471044.
- [9] Y. Gao, R. Wang, dan E. Zhou, "Stock Prediction Based on Optimized LSTM and GRU Models," *Sci Program*, vol. 2021, hlm. 1–8, Sep 2021, doi: 10.1155/2021/405281.
- [10] A. Nilsen, "Perbandingan Model RNN, Model LSTM, dan Model GRU dalam Memprediksi Harga Saham-Saham LQ45," *Jurnal Statistika dan Aplikasinya*, vol. 6, no. 1, hlm. 137–147, Jun 2022, doi: 10.21009/JSA.06113.
- [11] M. V. Narkhede, P. P. Bartakke, dan M. S. Sutaone, "A review on weight initialization strategies for neural networks," *Artif Intell Rev*, vol. 55, no. 1, hlm. 291–322, Jan 2022, doi: 10.1007/s10462-021-10033-z.
- [12] N. G. Mankiw, Harvard University. Department of Economics., dan Harvard University. Faculty of Arts and Sciences., "Macroeconomics," hlm. 642, Diakses: 7 Oktober 2022. [Daring]. Tersedia pada: <https://www.readallbooks.org/book/macroeconomics-9-edition/>
- [13] J. Ha, M. Kambe, dan J. Pe, "Data Mining: Concepts and Techniques," *Data Mining: Concepts and Techniques*, hlm. 1–703, Jan 2011, doi: 10.1016/C2009-0-61819-5.
- [14] W. Zhou, D. Wang, H. Li, dan W. Song, "Long-Term Forecasting of Time Series Based on Sliding Window Information Granules and Fuzzy Inference System," *2018 International Conference on Information, Cybernetics, and Computational Social Systems, ICCSS 2018*, hlm. 375–380, Des 2018, doi: 10.1109/ICSS.2018.8572415.
- [15] J. Schmidhuber, "Deep Learning in Neural Networks: An Overview," *Neural Networks*, vol. 61, hlm. 85–117, Apr 2014, doi: 10.1016/j.neunet.2014.09.003.
- [16] F. M. Bianchi, E. Maiorino, M. C. Kampffmeyer, A. Rizzi, dan R. Jenssen, "An overview and comparative analysis of Recurrent Neural Networks for Short Term Load Forecasting," Mei 2017, doi: 10.1007/978-3-319-70338-1.
- [17] N. K. Manaswi, "Deep Learning with Applications Using Python," *Deep Learning with Applications Using Python*, 2018, doi: 10.1007/978-1-4842-3516-4.
- [18] A. Saxena dan T. Sukumar, "Predicting bitcoin price using lstm And Compare its predictability with arima model," 2018.
- [19] Z. Zhao, W. Chen, X. Wu, P. C. Y. Chen, dan J. Liu, "LSTM network: a deep learning approach for short-term traffic forecast," *Iet Intelligent Transport Systems*, vol. 11, hlm. 68–75, 2017.
- [20] S. Hochreiter dan J. Schmidhuber, "Long Short-Term Memory," *Neural Comput*, vol. 9, no. 8, hlm. 1735–1780, Nov 1997, doi: 10.1162/NECO.1997.9.8.1735.
- [21] M. Pirani, P. Thakkar, P. Jivrani, M. H. Bohara, dan D. Garg, "A Comparative Analysis of ARIMA, GRU, LSTM and BiLSTM on Financial Time Series Forecasting," *IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics, ICDCECE 2022*, 2022, doi: 10.1109/ICDCECE53908.2022.9793213.
- [22] H. Hewamalage, K. Ackermann, dan C. Bergmeir, "Forecast Evaluation for Data Scientists: Common Pitfalls and Best Practices," Mar 2022.