



Comparison of RNN and LSTM Algorithm Performance in Predicting the Number of Umrah Pilgrims at PT. Hajar Aswad

Perbandingan Performa Algoritma RNN dan LSTM dalam Prediksi Jumlah Jamaah Umrah pada PT. Hajar Aswad

**Razanul Al Kiramy^{1*}, Inggih Permana², Arif Marsal³,
Medyantiwi Rahmawita Munzir⁴, Megawati⁵**

^{1,2,3,4,5}Program Studi Sistem Informasi, Fakultas Sains dan Teknologi,
Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia

E-Mail: ¹12050310400@students.uin-suska.ac.id, ²inggihpermana@uin-suska.ac.id,
³arif.marsal@uin-suska.ac.id, ⁴medyantiwi.rahmawita@uin-suska.ac.id, ⁵megawati@uin-suska.ac.id

Received May 3rd 2024; Revised Jun 19th 2024; Accepted Jul 20th 2024
Corresponding Author: Razanul Al Kiramy

Abstract

In terms of language, umrah means pilgrimage or visiting, while technically umrah refers to a journey to the Baitullah outside of the Hajj season with the purpose of performing specific religious rituals and meeting certain requirements. PT Hajar Aswad is a umrah travel company operating in Indonesia. PT Hajar Aswad is responsible for organizing journeys, accommodations, transportation, and various other needs for umrah pilgrims. Therefore, it's necessary to have a good understanding of the patterns and trends in the number of umrah pilgrims in order to optimize operations and provide quality services. Hence, this research was conducted to predict the number of umrah pilgrims at PT Hajar Aswad using RNN and LSTM algorithms. The comparison results of the two algorithms show that LSTM is capable of providing slightly better prediction results compared to RNN with parameters: window size of 7, Adam optimizer, batch size of 8, and learning rate of 0.01. This model has an RMSE value of 0.1758, MAPE of 0.4846, and R2 of 0.5198.

Keyword: Comparison, LSTM, Prediction, RNN, Umrah

Abstrak

Secara bahasa umrah bermakna ziarah atau berkunjung, sedangkan secara istilah umrah adalah perjalanan ke Baitullah di luar waktu haji dengan tujuan melaksanakan ibadah tertentu dan memenuhi syarat-syarat khusus. PT Hajar Aswad merupakan sebuah perusahaan travel yang beroperasi di Indonesia. PT Hajar Aswad bertanggung jawab untuk mengatur perjalanan, akomodasi, transportasi, dan berbagai keperluan lainnya bagi para jemaah umrah, untuk itu perlu memiliki pemahaman yang baik mengenai pola dan tren jumlah jemaah umrah agar dapat mengoptimalkan operasional dan memberikan pelayanan yang memuaskan kepada jemaah. Oleh karena itu penelitian ini dilakukan untuk memprediksi jumlah jemaah umrah pada PT Hajar Aswad menggunakan algoritma RNN dan LSTM agar PT Hajar Aswad. Hasil perbandingan kedua algoritma menunjukkan bahwa LSTM mampu memberikan hasil prediksi yang sedikit lebih baik dibandingkan RNN dengan parameter window size 7, optimizer Adam, batch size 8, dan learning rate 0,01. Model ini memiliki nilai RMSE sebesar 0,1758, MAPE sebesar 0,4846, dan R2 sebesar 0,5198.

Kata Kunci: LSTM, Perbandingan, Prediksi, RNN, Umrah

1. PENDAHULUAN

Umrah merupakan ibadah dalam agama Islam yang dilaksanakan dengan cara berkunjung ke Kota Mekah dan Madinah di Arab Saudi [1]. Secara bahasa Umrah bermakna ziarah atau berkunjung, sedangkan secara istilah Umrah adalah perjalanan ke Baitullah selain di waktu haji dengan tujuan melaksanakan ibadah tertentu dan memenuhi syarat-syarat khusus [2]. Umrah adalah salah satu dari dua jenis ritual keagamaan dalam Islam, selain Haji. Umrah dapat dilakukan sepanjang tahun tanpa Batasan waktu, tidak seperti Haji yang hanya bisa dikerjakan pada saat tertentu [3]. Meskipun Umrah bukanlah kewajiban, banyak umat Muslim yang berkeinginan untuk melaksanakannya sebagai salah satu dari ibadah yang meningkatkan iman dan mendekatkan diri kepada Allah SWT.

Pelaksanaan Ibadah Haji dan Umrah di Indonesia diatur oleh Undang-Undang No. 13 Tahun 2008 yang berisikan Penyelenggaraan Ibadah Haji. Undang-undang ini mengatakan bahwa tujuan penyelenggaraan ibadah Haji dan Umrah adalah untuk memberikan bimbingan dan perlindungan optimal melalui manajemen pelayanan yang baik dan efisien. Hal ini bertujuan untuk memastikan pelaksanaan ibadah Haji dan Umrah berlangsung dengan lancar, aman, dan tertib, serta nyaman dan sesuai dengan ketentuan agama Islam. Tujuan akhirnya adalah meningkatkan kemungkinan jamaah Haji dan Umrah agar bisa beribadah dengan mandiri [4].

PT Hajar Aswad merupakan sebuah perusahaan travel Umrah yang beroperasi di Indonesia, hal ini bisa dilihat pada tahun 2023 berhasil ditetapkan sebagai salah satu penyedia layanan Umrah terbaik pertama di Riau. PT ini merupakan salah satu penyedia jasa perjalanan Umrah yang memiliki banyak calon jamaah, hal ini bisa dilihat dari sebanyak 3 ribu jamaah yang telah diberangkatkan terhitung dari awal tahun 2023 sampai bulan oktober 2023. Mereka bertanggung jawab untuk mengatur perjalanan, akomodasi, transportasi, dan berbagai keperluan lainnya bagi para jamaah Umrah. Dalam menjalankan bisnis ini, perlu memiliki pemahaman yang baik mengenai pola dan tren jumlah jamaah Umrah agar dapat mengoptimalkan operasional dan memberikan pelayanan yang berkualitas [5]. Berdasarkan pemahaman tersebut, PT Hajar Aswad dapat melakukan perencanaan yang lebih baik, seperti menyesuaikan kapasitas akomodasi, memprediksi kebutuhan transportasi, dan mengelola ketersediaan paket Umrah. Fokus utama dalam menjalankan bisnisnya yaitu pada kepuasan konsumen atau calon jamaah, dengan memberikan penekanan khusus pada pola pelayanan kepada jamaah. Mereka menyediakan pendampingan penuh kepada jamaah dalam pengurusan administrasi, termasuk segala persyaratan yang diperlukan untuk melaksanakan ibadah Umrah. Selain itu juga diberikan pelayanan yang baik untuk memenuhi semua kebutuhan jamaah, terutama dari segi fasilitas yang tersedia di kantor maupun fasilitas pendukung lainnya. Semua fasilitas ini ditujukan untuk memberikan kemudahan kepada jamaah yang menggunakan layanan mereka dalam melaksanakan ibadah Umrah [6].

Prediksi jumlah jamaah Umrah penting untuk dilakukan, karena dengan memiliki prediksi yang akurat penyedia jasa perjalanan Umrah dapat menghindari ketidakseimbangan antara permintaan dan penawaran serta mengurangi risiko kerugian finansial. Seperti pada saat hari libur sekolah dan libur nasional, jamaah umrah akan mengalami peningkatan cukup signifikan dibandingkan dengan hari lainnya. Selain itu, jika mengetahui prediksi terhadap jumlah jamaah maka bisa mempersiapkan barang yang diperlukan seperti tas, koper, baju batik, syal dan perlengkapan lainnya serta sumber daya yang diperlukan dengan lebih akurat dan efisien [7]. Prediksi yang akurat juga memberikan manfaat bagi pemerintah dan otoritas terkait dalam mengatur infrastruktur dan layanan Umrah, seperti penyediaan fasilitas kesehatan, keamanan, dan fasilitas publik lainnya. Melakukan prediksi jumlah jamaah Umrah juga bisa dimanfaatkan untuk pengembangan strategi berdasarkan hasil prediksi dalam meningkatkan jumlah jamaah pendaftar Umrah pada PT Hajar Aswad [8].

Sumber data yang digunakan dalam melakukan prediksi jumlah jamaah Umrah meliputi data historis jumlah jamaah Umrah dari tahun sebelumnya. Data-data ini dapat dikumpulkan dari perusahaan travel Umrah PT Hajar Aswad. Menurut data statistik yang dirilis oleh Otoritas Umum Statistik (GaStat) Arab Saudi, pada tahun 2017, jumlah jamaah Umrah mencapai 19.07 juta orang, yang merujuk pada data yang diperoleh dari Kementerian Haji dan Umrah Arab Saudi [9]. Selanjutnya, sepanjang tahun 2022, jumlah jamaah Umrah meningkat menjadi 24,71 juta orang. Dari jumlah tersebut, jamaah pria berjumlah 14,73 juta orang atau sekitar 59,6%, sedangkan jamaah perempuan berjumlah 9,98 juta orang atau sekitar 40,4%. Dalam kategori jamaah Umrah, 6,64 juta orang adalah warga negara Arab Saudi, yang menyumbang sekitar 26,9%, sementara 18,07 juta orang adalah jamaah Umrah yang berasal dari luar Arab Saudi, yang berkontribusi sekitar 73,1% [10].

Recurrent Neural Networks (RNN) dan Long Short Term Memory (LSTM) merupakan algoritma dalam machine learning yang sesuai untuk digunakan dalam melakukan prediksi berdasarkan data sekuensial, seperti data jumlah jamaah Umrah dari waktu ke waktu. RNN merupakan salah satu algoritma deep learning yang dapat melakukan prediksi pada data time series bertipe numerik, yang cara kerjanya yaitu dengan melakukan pemrosesan yang dilakukan secara berulang-ulang [11]. Selain itu deep learning untuk data deret waktu yang juga banyak digunakan adalah LSTM. Algoritma LSTM ditemukan pada tahun 1997, yang di kemukakan oleh Hochreiter dan Schmidhuber. Jaringan LSTM adalah jenis jaringan saraf berulang yang sering digunakan dalam deep learning karena kemampuannya untuk berhasil melatih arsitektur yang sangat kompleks [12] Jaringan ini tersusun dari lapisan LSTM dengan proses dilakukan berulang [13].

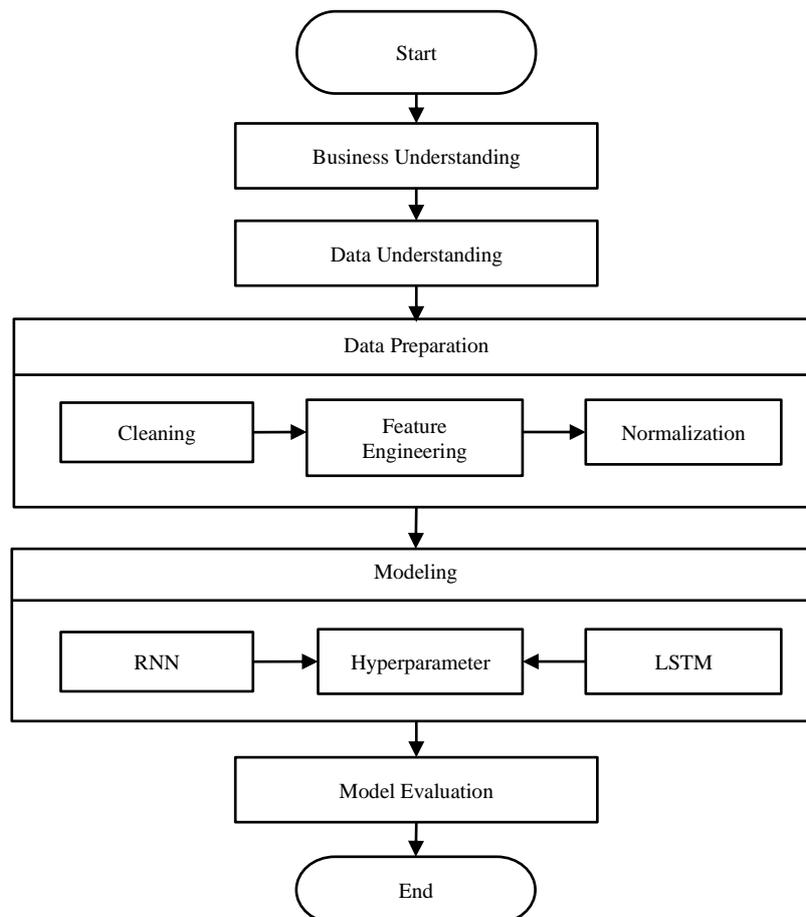
Penelitian sebelumnya yang telah dilakukan terkait algoritma yang akan diaplikasikan dalam penelitian ini salah satunya yaitu penelitian oleh Nilsen, (2022) tentang Perbandingan Model RNN dan LSTM dalam prediksi Harga Saham LQ45. Hasil dari penelitian ini didapat nilai RMSE model RNN sebesar 346.91758, sedangkan nilai RMSE model LSTM sebesar 341.45668. Jika dibandingkan dengan model RNN, model LSTM memiliki akurasi yang lebih tinggi sebanyak 11% dibandingkan model RNN [15]. Penelitian lainnya yang dilakukan oleh Nurashila (2023) mendapat kesimpulan bahwa model LSTM memberikan perbandingan yang lebih baik daripada model RNN dalam hal error dan akurasi R-Squared pada algoritma Deep Learning. Model LSTM mendapat hasil nilai MAE sebesar 0.70780 dan nilai pada RMSE sebesar 1.39910, sementara akurasi S-Squared mencapai 0.99263 atau setara dengan 99.2% [16]. Pada penelitian lainnya yang dilakukan oleh Gopali (2021) tentang Sebuah Studi Perbandingan dalam Mendeteksi Anomali pada data deret waktu menggunakan model LSTM dan TCN. Dari penelitian ini didapatkan kesimpulan bahwa model model TCN

melaporkan kinerja yang lebih baik dibandingkan dengan model LSTM, berdasarkan dari perbandingan nilai dari hasil evaluasi yang didapatkan [17]. Penelitian yang dilakukan oleh Maharani (2023) tentang analisis perbandingan kinerja model ARIMA, LSTM dan GRU, hasil yang didapat setelah melakukan evaluasi yaitu perbandingan kinerja dari tiap-tiap model menunjukkan model ARIMA mendapatkan nilai terbaik [18]. Selain itu pada penelitian yang dilakukan Selle et al., (2022) tentang perbandingan prediksi penggunaan listrik menggunakan algoritma LSTM dan RNN, diperoleh kesimpulan bahwa LSTM menunjukkan kinerja yang lebih unggul dibandingkan dengan RNN ketika mengelola deret waktu yang lebih panjang, misalnya 20 dan 30 urutan, didapat hasil nilai RMSE yang lebih rendah, yaitu 49,18 dan 49,47 untuk fitur pada data siang sedangkan 58,68 dan 57,08 untuk fitur pada data malam. Namun, ketika deret waktu yang digunakan lebih pendek, seperti 5 urutan, perbedaan kinerja antara RNN dan LSTM tidak terlalu signifikan [19]. Berdasarkan penelitian penelitian sebelumnya yang telah dilakukan, belum ada yang melakukan prediksi pada jumlah jemaah umrah dengan algoritma LSTM dan RNN. Algoritma ini dipilih karena selain dapat melakukan prediksi dalam jangka waktu yang relatif panjang juga dapat memberikan hasil yang akurat, sehingga sangat sesuai dengan kebutuhan PT Hajar Aswad. Selain itu juga dibandingkan penelitian sebelumnya, penelitian ini menggunakan evaluasi RMSE, MAPE, dan R^2 untuk mengukur seberapa akurat hasilnya.

Dalam konteks prediksi jumlah jemaah Umrah, perbandingan penggunaan algoritma RNN dan LSTM dapat memberikan potensi untuk menghasilkan prediksi yang lebih akurat berdasarkan data historis dan faktor-faktor yang relevan. Setelah memahami hubungan antara berbagai faktor dan jumlah jemaah Umrah, PT Hajar Aswad dapat melakukan prediksi jumlah jemaah Umrah di masa depan, yang akan membantu mereka dalam pengambilan keputusan bisnis yang lebih baik dan pelayanan yang lebih baik pula kepada jemaah Umrah.

2. METODOLOGI PENELITIAN

Penelitian ini mengikuti struktur metodologi *Cross-Industry Standard Process for Data Mining* (CRISP-DM) yang merupakan sebuah metodologi untuk mengembangkan solusi data mining. yang menjelaskan enam langkah kunci dalam penerapan model data mining. Langkah-langkah tersebut akan digunakan untuk membuat model prediksi jumlah jemaah Umrah pada PT. Hajar Aswad dengan algoritma RNN dan LSTM. Penelitian ini menggunakan tahapan yang bisa dilihat pada gambar 1.



Gambar 1. Metodologi Penelitian

2.1. *Business Understanding*

Tahap ini dilakukan untuk mendapatkan pemahaman mendalam tentang permasalahan yang diidentifikasi dari latar belakang dan tujuan penelitian. Tahapan ini dilakukan dengan melakukan wawancara dan observasi ke PT Hajar Aswad.

2.2. *Data Understanding*

Data yang digunakan adalah data *time series* jamaah Umrah dengan rentang waktu 8 tahun terakhir mulai dari Januari 2015 – Oktober 2023. Namun pada periode Mei 2020 – Januari 2022, pemberangkatan jamaah Umrah tidak dilakukan karena terjadi pandemi wabah COVID-19, sehingga data pada rentang tersebut tidak tersedia. Kemudian setiap tahunnya terdapat 2 bulan tidak dilakukan pemberangkatan Umrah karena adanya pelaksanaan haji, sehingga terdapat 15 baris data bernilai nol (0). Data tersebut akan tetap dipertahankan agar model yang dilatih dapat mempelajari dan memahami pola pelaksanaan haji dan tidak haji. Secara keseluruhan dalam penelitian ini data yang digunakan berjumlah 85 baris data.

2.3. *Data Preparation*

Penelitian ini melakukan data preparation atau persiapan data dengan proses sebagai berikut.

2.3.1 *Cleaning*

Proses yang dilakukan pada tahap ini yaitu melakukan deteksi dan penanganan *missing value* serta memastikan format tanggal dan waktu tetap konsisten.

2.3.2 *Feature Engineering*

Pada penelitian ini, fitur baru yang dibentuk yaitu ‘bulan masehi’, ‘tahun masehi’, ‘bulan hijriah’, ‘tahun hijriah’, ‘haji’, dan ‘tidak haji’. Hal ini dilakukan dengan tujuan untuk memberikan pola temporal berupa informasi waktu untuk meningkatkan kinerja model machine learning. Selain itu, untuk mempelajari trend dan pola historis dimasa lampau, maka perlu dilakukan tahapan *windowing*. Pada penelitian ini *window size* akan ditetapkan melalui rangkaian percobaan yaitu 3, 5 dan 7. Nilai tersebut ditetapkan untuk menguji rangkaian input panjang dan pendek untuk mendapatkan hasil yang maksimal seperti yang dilakukan oleh penelitian rahmadeyan dan mustakim (2024) [20], meskipun dengan nilai yang berbeda.

2.3.3 *Normalization*

Teknik normalisasi yang digunakan adalah min-max normalization. Min-max normalization adalah teknik yang dilakukan dengan skala ulang nilai-nilai pada data dalam rentang tertentu, umumnya 0 sampai 1 [21]. Untuk melakukan normalisasi dengan min-max normalization dapat menggunakan persamaan 1:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

2.4 *Modeling*

Pada tahapan ini dilakukan menggunakan algoritma RNN dan LSTM. Selain itu, untuk mendapatkan model dengan kemampuan yang optimal, dilakukan juga pengujian *hyperparameter* seperti *learning rate*, *optimizer*, dan lainnya. *Tools* yang digunakan yaitu pemrograman python dan google colaboratory.

2.4.1 *Recurrent Neural Network*

Recurrent Neural Network (RNN) merupakan salah satu jenis jaringan syaraf tiruan yang mampu untuk memproses data berurutan (*sequential*) dengan konteks temporal sehingga dapat digunakan pada data *time series* [22]. Peran yang signifikan dari RNN adalah dalam memori *input* dan menghasilkan *output* yang tepat sesuai dengan konteksnya. Kemampuan RNN untuk menyimpan informasi dari *input* dan mengeluarkan *output* yang sesuai dengan konteksnya adalah aspek yang krusial. Memori internal menjadi fokus penting dalam RNN karena memungkinkannya untuk memprediksi kejadian selanjutnya. Karena alasan tersebut, RNN sangat sesuai untuk aplikasi pada deret waktu, audio, video, mesin pencarian, teks dan juga mesin kasir [23].

2.4.2 *Long Short-Term Memory*

Long Short-Term Memory (LSTM) adalah pengembangan dari RNN yang dirancang khusus untuk menangani masalah ketergantungan jangka panjang [24]. Persamaan umum RNN dan LSTM yaitu arsitektur jaringan saraf yang menggunakan struktur rekuren dan fungsi aktivasi untuk memproses data berurutan. LSTM mampu mempertahankan informasi dalam jangka waktu yang panjang dan mengatasi masalah penurunan gradien yang terjadi pada RNN. Struktur LSTM terdiri dari tiga gerbang: gerbang masukan (*input gate*), gerbang lupakan (*forget gate*), dan gerbang keluaran (*output gate*). Gerbang masukan mengontrol jumlah informasi yang disimpan pada memori jangka panjang. Gerbang lupakan mengatur seberapa banyak informasi

yang akan dihapus dari memori jangka panjang. Sementara itu, gerbang keluaran mengendalikan berapa banyak informasi yang akan dikeluarkan dari memori jangka panjang.

2.5 Model Evaluation

2.5.1 Data Validation

Untuk mengetahui kemampuan model, maka perlu dilakukan validasi. Ini dilakukan dengan proses membandingkan data hasil prediksi dengan data sebenarnya. Pada penelitian ini teknik validasi yang akan digunakan adalah K-Fold Cross Validation. Teknik ini dipilih karena efektif digunakan pada jumlah data yang terbatas. Pembagian yang ditetapkan adalah 4-fold, yang artinya data akan dibagi menjadi 4 bagian, dimana 3 bagian akan dialokasikan sebagai data training, sedangkan 1 bagian lainnya akan dialokasikan sebagai data testing. Setiap bagian akan akan saling bergantian menjadi data pelatihan dan data uji kemudian akan dilatih sebanyak jumlah fold [25].

2.5.2 Evaluation Metrics

Model yang telah dilatih kemudian akan dievaluasi dengan menggunakan beberapa metrik, yaitu *Root Mean Square Error* (RMSE), *Mean Absolute Percentage Error* (MAPE), dan *R-squared* (R²).

RMSE merupakan metode pengukuran yang digunakan untuk mengevaluasi seberapa akurat hasil perhitungan dengan membandingkan tingkat kesalahan [26]. Untuk menghitung nilai RMSE dapat menggunakan persamaan 2.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (2)$$

MAPE adalah rata-rata diferensiasi absolut dari nilai prediksi dan aktual, dinyatakan dalam persentase dari nilai aktual [27]. Untuk menghitung nilai MAPE dapat menggunakan persamaan 3:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \quad (3)$$

Sedangkan R² merupakan metrik untuk menentukan besarnya variasi di dalam variabel dependen yang dijelaskan pada variabel independen, sehingga dapat digunakan untuk mengetahui seberapa baik model prediksi menjelaskan variasi data aktual [28]. Untuk menghitung nilai R² dapat menggunakan persamaan 4:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (4)$$

Pada metrik RMSE dan MAPE, semakin kecil nilai nilai yang diperoleh menunjukkan hasil prediksi dari model tersebut semakin baik. Sementara pada R² jika hasilnya mendekati 1 maka akan semakin baik dan jika mendekati 0 akan semakin buruk. Namun R² juga dapat menghasilkan nilai negatif yang menandakan bahwa hasil yang didapatkan sangat buruk.

3. HASIL DAN PEMBAHASAN

Data yang telah didapatkan akan dilakukan proses *cleaning* terlebih dahulu agar sesuai format dan siap untuk dianalisis. Hasilnya data yang akan digunakan dalam proses analisis dan pemodelan memiliki jumlah 85 baris data. Penelitian ini menggunakan data yang bisa dilihat pada tabel 1.

Tabel 1. Dataset Jamaah Umrah

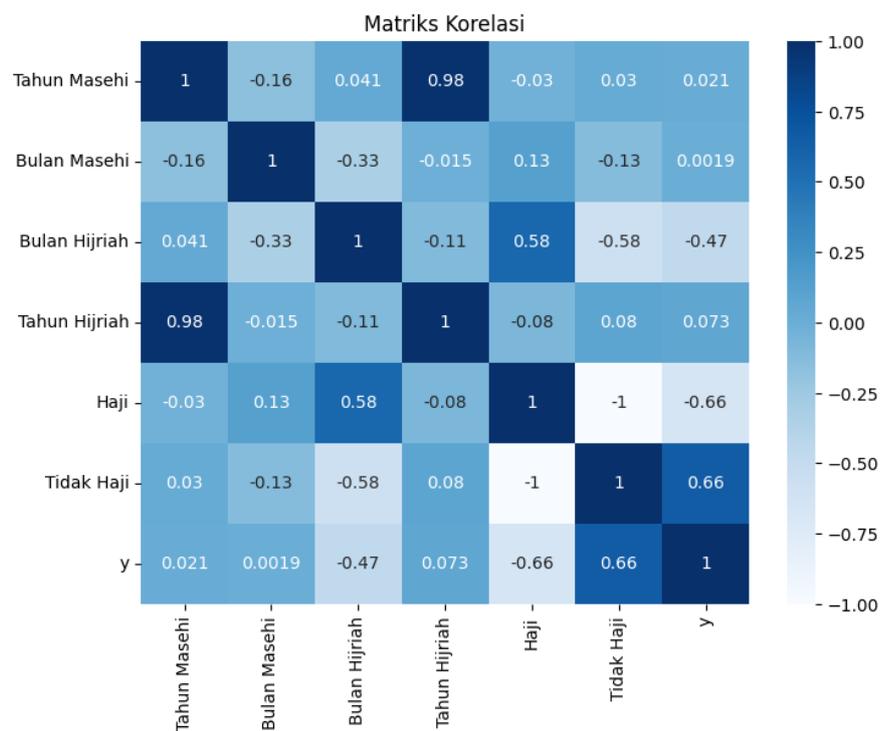
No	Bulan	Jumlah
1	Jan-15	112
2	Feb-15	623
3	Mar-15	873
4	Apr-15	516
...
82	Jul-23	420
83	Aug-23	656
84	Sep-23	209
85	Okt-23	615

Data yang sudah diberishkan kemudian dilakukan *feature engineering*. Ini merupakan teknik yang digunakan untuk membuat fitur baru berdasarkan fitur yang telah ada. Fitur baru yang akan dibentuk yaitu 'bulan hijriah', 'tahun hijriah', 'bulan masehi', 'tahun masehi', 'haji', dan 'tidak haji' berdasarkan fitur 'bulan'. Hasil implementasi dari *feature engineering* bisa dilihat pada tabel 2.

Tabel 2. Hasil Implementasi *Feature Engineering*

Tahun Masehi	Bulan Masehi	Bulan Hijriah	Tahun Hijriah	Haji	Tidak Haji
2015	7	10	1436	0	1
2015	8	11	1436	1	0
2015	9	12	1436	1	0
...
2023	8	2	1445	0	1
2023	9	3	1445	0	1
2023	10	4	1445	0	1

Saat fitur baru selesai terbentuk, fitur tersebut akan dihitung korelasinya terhadap target untuk mengetahui seberapa besar pengaruhnya. Hasilnya, dari 6 fitur baru yang dibentuk, hanya 3 fitur yang memiliki pengaruh dan korelasi yang cukup signifikan terhadap target, yaitu fitur ‘bulan hijriah’, ‘haji’, dan ‘tidak haji’. Untuk itu, fitur yang hampir tidak memiliki korelasi sama sekali terhadap target akan dihapus karena tidak memberikan banyak informasi yang berguna bagi model sehingga dapat mempercepat waktu komputasi dan mencegah *overfitting*. Matriks korelasi setiap fitur bisa dilihat pada gambar 2.



Gambar 2. Matriks Korelasi pada Setiap Fitur

Selanjutnya adalah membentuk fitur sekuensial dan temporal menggunakan teknik *windowing*. Adapun *window size* yang ditetapkan pada teknik *windowing* dilakukan melalui proses pengujian dengan nilai 3, 5, dan 7. Berikut adalah hasil dari implementasi teknik *windowing*.

Tabel 3. Hasil *Windowing* dengan 7 Window Size

X-6	X-5	X-4	X-3	X-2	X-1	y
112	623	873	516	1274	423	572
623	873	516	1274	423	572	0
...
327	0	0	420	656	209	615

Hasil *feature engineering* dan *windowing* akan disatukan menjadi satu kesatuan dataset. Dataset inilah nantinya yang akan digunakan pada tahap pemodelan. Namun, dataset ini masih memiliki nilai dengan rentang yang berbeda di setiap fiturnya. Untuk itu dataset perlu dinormalisasi terlebih dahulu dengan teknik *min-max normalization* agar memiliki rentang nilai yang sama. Hasil dari normalisasi data bisa dilihat pada tabel 4.

Tabel 4. Hasil Normalisasi Data

Bulan Hijriah	Haji	Tidak Haji	X-6	X-5	X-4	X-3	X-2	X-1	y
0,8182	0,0000	1,0000	0,0879	0,4890	0,6852	0,4050	1,0000	0,3640	0,4923
0,9091	1,0000	0,0000	0,4890	0,6852	0,4050	1,0000	0,3320	0,4923	0,0000
...
0,2727	0,0000	1,0000	0,2567	0,0000	0,0000	0,3297	0,5149	0,1799	0,5293

Saat melakukan pemodelan, data yang sudah dinormalisasi akan dibagi menjadi data pelatihan dan data validasi menggunakan teknik k-fold cross validation. Jumlah pembagian yang ditetapkan adalah 4-fold. Pemodelan algoritma RNN dan LSTM dilakukan menggunakan Bahasa pemrograman Python dengan library tensorflow dan *tools* google colaboratory. Arsitektur yang digunakan pada algoritma RNN dan LSTM terdiri dari 4 layer, yaitu 1 *input layer*, 2 *hidden layer* dengan 100 *units*, dan 1 *output layer*. Untuk mendapatkan model dengan kemampuan terbaik, pengujian parameter (*hyperparameter*) juga dilakukan pada kedua algoritma yang bisa dilihat pada tabel 5.

Tabel 5. Pengujian Parameter (*Hyperparameter*) RNN dan LSTM

Parameter	Value
Window Size	3, 5, 7
Optimizer	Adam, SGD, RMSprop
Batch Size	8, 16
Learning Rate	0.01, 0.001, 0.0001
Epoch	200

Tahap pemodelan diawali dengan mengimplementasikan algoritma RNN sesuai dengan *hyperparameter* yang telah ditentukan sebelumnya. Melalui hasil pemodelan didapatkan bahwa algoritma RNN dengan *window size* 7, *optimizer* RMSprop, batch size 8, dan *learning rate* 0,001 sebagai model yang paling baik dari seluruh percobaan yang dilakukan. Model tersebut memiliki nilai RMSE sebesar 0,1760, MAPE sebesar 0,4918, dan R2 sebesar 0,4696. Seluruh hasil evaluasi pemodelan algoritma RNN bisa dilihat pada tabel 6.

Tabel 6. Hasil Evaluasi Model RNN

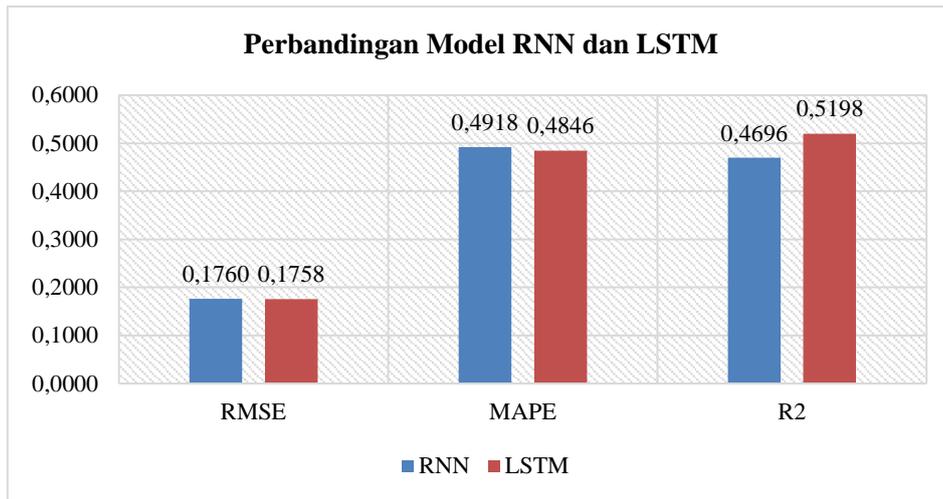
Window Size	Optimizer	Batch Size	Learning Rate	RMSE	MAPE	R ²
3	Adam	8	0,01	0,1851	0,5105	0,4252
3	Adam	8	0,001	0,1879	0,5257	0,4107
...
7	RMSprop	16	0,0001	0,1964	0,5854	0,4183

Selanjutnya adalah mengimplementasikan algoritma LSTM sesuai dengan *hyperparameter* yang telah ditentukan sebelumnya. Melalui hasil pemodelan didapatkan bahwa algoritma LSTM dengan *window size* 7, *optimizer* Adam, batch size 8, dan *learning rate* 0,01 sebagai model terbaik dari semua percobaan yang telah dilakukan. Model tersebut memiliki nilai RMSE sebesar 0,1758, MAPE sebesar 0,4846, dan R2 sebesar 0,5198. Seluruh hasil evaluasi pemodelan algoritma RNN bisa dilihat pada tabel 7.

Tabel 7. Hasil Evaluasi Model LSTM

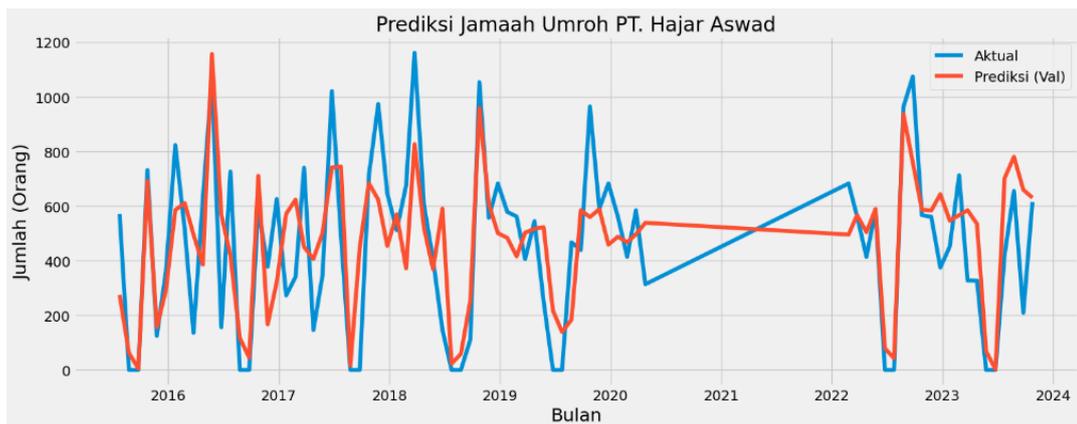
Window Size	Optimizer	Batch Size	Learning Rate	RMSE	MAPE	R ²
3	Adam	8	0,01	0,1785	0,5011	0,4684
3	Adam	8	0,001	0,1963	0,5206	0,3673
...
7	RMSprop	16	0,0001	0,2195	0,5274	0,2970

Setelah melakukan pemodelan, tahapan selanjutnya adalah membandingkan algoritma RNN dan LSTM untuk mendapatkan model dengan kemampuan yang paling terbaik. Hasil perbandingan menunjukkan bahwa model LSTM sedikit lebih unggul dibandingkan dengan RNN berdasarkan metrik RMSE, MAPE dan R2. Hasil perbandingan algoritma RNN dan LSTM dapat dilihat pada gambar 3.



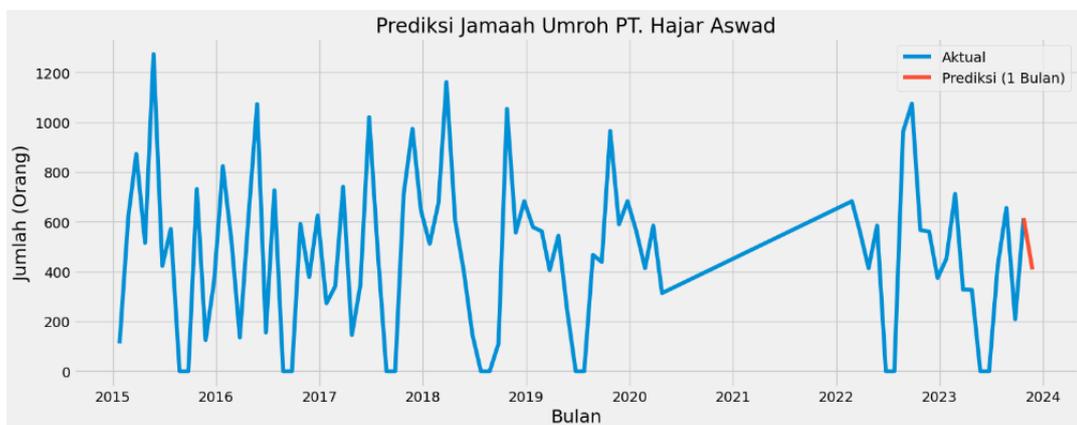
Gambar 3. Perbandingan Model RNN dan LSTM

Sebagai model terbaik, LSTM akan diimplementasikan untuk memprediksi jamaah Umrah periode berikutnya. Namun sebelum itu, model LSTM akan diuji terlebih dahulu untuk memprediksi data validasi untuk melihat dan mengetahui bagaimana kemampuan prediksinya. Berdasarkan visualisasi terlihat bahwa model memiliki kemampuan yang cukup baik dalam mengikuti fluktuatif data aktual meskipun tidak sempurna. Hasil prediksi model LSTM terhadap data validasi dapat dilihat pada gambar 4.



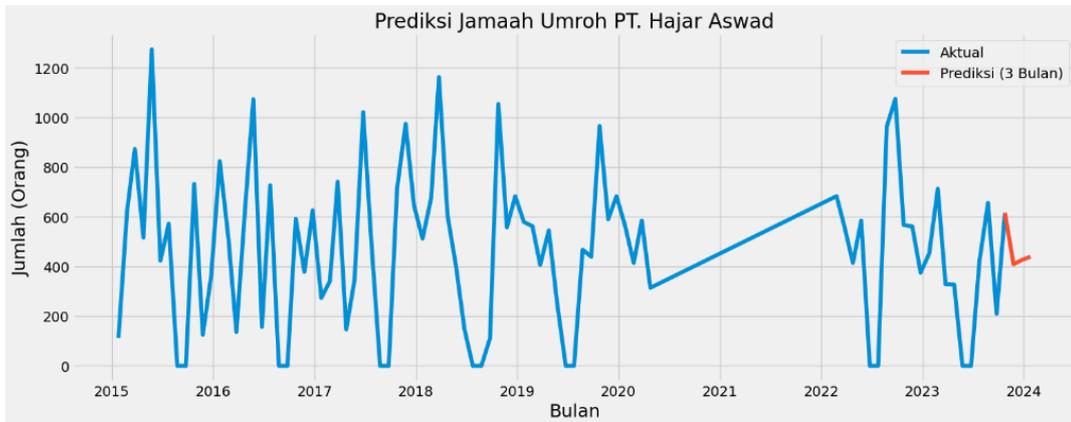
Gambar 4. Prediksi Jamaah Umrah pada Data Validasi

Tahapan terakhir adalah melakukan prediksi jamaah Umrah periode berikutnya menggunakan algoritma LSTM. Proses prediksi akan dilakukan dalam jangka waktu 1 bulan, 3 bulan, 6 bulan, hingga 1 tahun ke depan. Hasil prediksi periode berikutnya bisa dilihat pada gambar 5, 6, 7, dan 8.



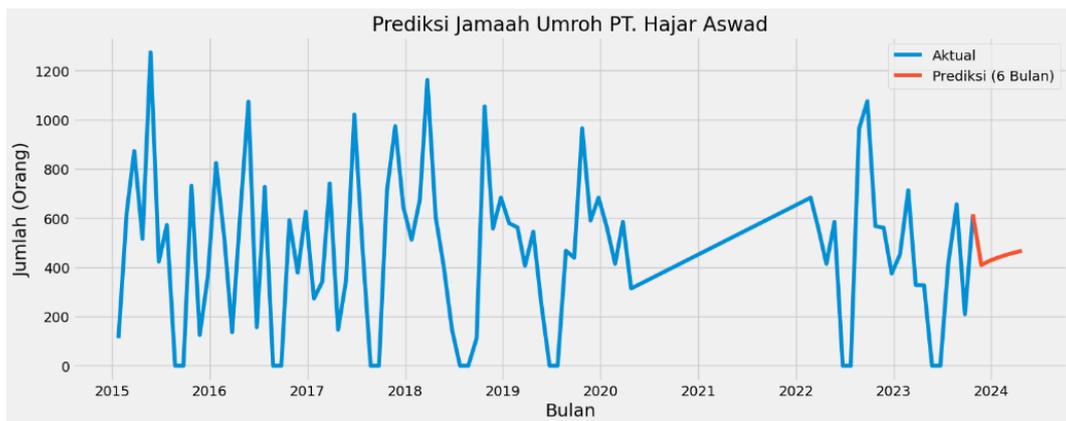
Gambar 5. Prediksi Jamaah Umroh 1 Bulan Berikutnya

Hasil peramalan model LSTM untuk 1 bulan ke depan menunjukkan bahwa jamaah umroh akan terjadi penurunan dibandingkan bulan sebelumnya, namun penurunan jumlah jamaah yang terjadi tidak terlalu signifikan.



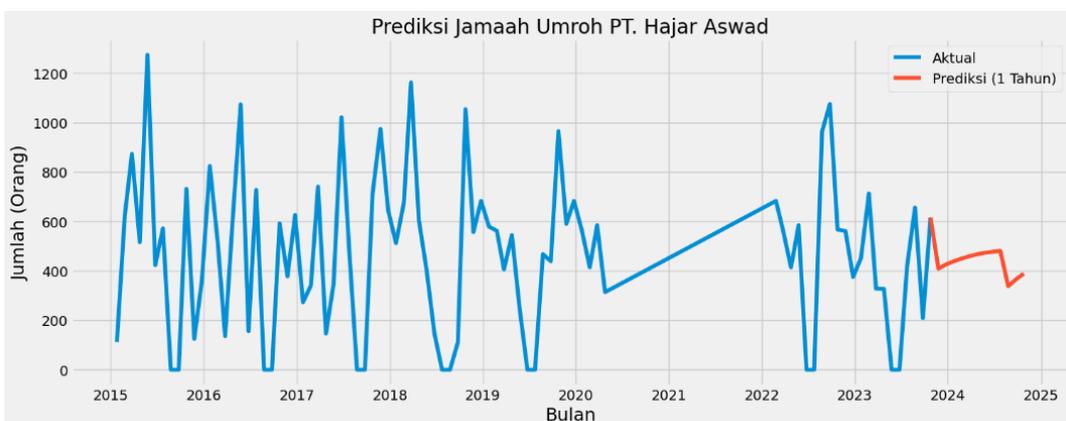
Gambar 6. Prediksi Jamaah Umroh 3 Bulan Berikutnya

Selanjutnya, peramalan model LSTM untuk 3 bulan kedepan menunjukkan bahwa jamaah umroh akan kembali meningkat setelah pada bulan sebelumnya terjadi penurunan. Namun peningkatan jumlah jamaah yang terjadi tidak terlalu signifikan.



Gambar 7. Prediksi Jamaah Umroh 6 Bulan Berikutnya

Kemudian, peramalan model LSTM untuk 6 bulan kedepan menunjukkan bahwa akan kembali terjadi peningkatan jamaah umroh dibandingkan periode sebelumnya. Peningkatan jumlah jamaah yang terjadi pada periode ini cukup stabil tetapi tidak terlalu signifikan.



Gambar 8. Prediksi Jamaah Umroh 1 Tahun Berikutnya

Terakhir, peramalan model LSTM untuk 1 tahun kedepan menunjukkan akan terjadi kenaikan dan penurunan jumlah jamaah umroh yang kemudian diakhiri dengan peningkatan kembali. Melalui hasil tersebut menunjukkan bahwa model LSTM mampu memproyeksikan bagaimana fluktuatif jamaah umrah dalam rentang waktu 1 bulan, 3 bulan, 6 bulan, dan 1 tahun kedepan. Hasil ini dapat digunakan oleh perusahaan dalam menyusun rencana strategis dan pengambilan keputusan seperti pemasaran, pengelolaan pelanggan, penentuan target penawaran dan bahkan inovasi layanan. Menggunakan langkah strategis yang tepat, perusahaan tentu dapat mencegah penurunan jumlah pelanggan dan bahkan meningkatkan pertumbuhan bisnis perusahaan.

Berdasarkan hasil pemodelan yang dilakukan menggunakan algoritma RNN dan LSTM menghasilkan algoritma LSTM sebagai yang terbaik dengan perbedaan performa yang tidak signifikan. Dalam kasus ini, *window size* yang lebih besar mampu menangkap pola historis dan *feature engineering* juga mampu menangkap pola temporal sehingga dapat memberikan informasi yang penting pada model. *Learning rate* dan *batch size* yang kecil cenderung menghasilkan model dengan tingkat kesalahan yang rendah pada setiap algoritma. Sementara penggunaan optimizer Adam dan RMSprop memberikan hasil yang bervariasi.

Meskipun model yang didapatkan sudah cukup baik, namun tingkat kesalahan model masih perlu diminimalisir. Jumlah data menjadi batasan pada penelitian ini karena adanya periode COVID-19. Untuk itu, penelitian berikutnya masih dapat menyempurnakan penelitian ini untuk menghasilkan model prediksi jamaah umroh yang lebih akurat dengan menambahkan jumlah, menggunakan data setelah periode COVID-19, menggunakan algoritma yang lebih canggih, melakukan lebih banyak percobaan *hyperparameter*, dan implementasi teknik lainnya.

4. KESIMPULAN

Penelitian ini telah berhasil menerapkan algoritma RNN dan LSTM dalam prediksi jumlah jamaah umroh di PT Hajar Aswad. Hasil perbandingan kedua algoritma memperlihatkan bahwa LSTM mampu memberikan hasil prediksi yang sedikit lebih baik dibandingkan RNN dengan parameter *window size* 7, *optimizer* Adam, *batch size* 8, dan *learning rate* 0,01. Model ini memiliki nilai RMSE sebesar 0,1758, MAPE sebesar 0,4846, dan R^2 sebesar 0,5198. *Window size* yang lebih besar mampu menangkap pola historis dan *feature engineering* juga mampu menangkap pola temporal pada data. *Learning rate* dan *batch size* yang kecil cenderung menghasilkan model dengan tingkat kesalahan yang rendah pada setiap algoritma. Sementara penggunaan *optimizer* Adam dan RMSprop memberikan hasil yang bervariasi. Sebagai model terbaik, LSTM diimplementasikan untuk memprediksi jamaah umroh dalam 1 bulan, 3 bulan, 6 bulan, dan 1 tahun mendatang. Hasilnya, LSTM memprediksi akan terjadi penurunan dan kenaikan jumlah jamaah yang fluktuatif meskipun tidak signifikan. Perusahaan dapat menggunakan hasil ini untuk menyusun rencana strategis dan pengambilan keputusan seperti pemasaran, pengelolaan pelanggan, penentuan target penawaran dan bahkan inovasi layanan sehingga dapat mencegah penurunan jumlah pelanggan dan bahkan meningkatkan pertumbuhan bisnis perusahaan.

REFERENSI

- [1] Mustafa Mat Jubri Shamsuddin, "Aplikasi Masalah Ke Atas Pencyaratan Pelaksanaan Haji Atau Umrah Melalui Agensi Pengelola," *Sains Insani*, vol. 04, 2019.
- [2] Y. C. Quomas, *Buku Tuntunan Manasik Haji dan Umrah Tahun 2024*. Jakarta: Kementerian Agama Republik Indonesia, 2024.
- [3] K. Chikhaoui, Mohammed Elrashidy, Motaz Alfarradj, Ali H. Muqaibel, Rida Sadagah, and Abdullah Sharqawi, "Automatic Hajj and Umrah Ritual Detection Using IMU Sensors," *IEEE Access*, vol. 10, 2022.
- [4] Khoir Rozzaku and Hasnun Jauhari Ritonga, "Manajemen Pelayanan Dalam Meningkatkan Minat Calon Jamaah Haji Dan Umroh Pada Pt.Ameera Mekkah Kota Medan," *Al-Mada*, vol. 6, 2023.
- [5] T. HIKMAH, "Pemanfaatan Media Sosial Dalam Meningkatkan Jumlah Calon Jamaah Umrah Di Pt. Bimalyndo Hajar Aswad Pekanbaru," UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU, PEKANBARU, 2021.
- [6] RAFIE HASAN, "Pemanfaatan Media Sosial Facebook Sebagai Strategi Pemasaran Paket Umrah Pt Hajar Aswad Mubaroq Kabupaten Siak," UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU, PEKANBARU, 2020.
- [7] Muhammad Baidhawi, "Optimalisasi Pelayanan Dalam Meningkatkan Jumlah Jamaah Umrah Pada Pt Sahid Gema Wisata Tahun 2019 Jakarta," Universitas Islam Negeri Syarif Hidayatullah Jakarta, Jakarta, 2022.
- [8] ZAHDI, "Manajemen Haji Dan Umrah (Analisis Pencapaian Kepuasan Layanan Haji Dan Umrah Di Kota Bandar Lampung)," Universitas Islam Negeri (Uin) Raden Intan Lampung, Lampung, 2021.
- [9] Aditya and Milkhatun, "Hubungan Pengetahuan Vaksinasi Meningitis dan Dukungan Keluarga terhadap Kepatuhan Vaksinasi Meningitis Jamaah Umrah di Kantor Kesehatan Pelabuhan Kelas II Samarinda," *Borneo Student Research*, vol. 1, 2020.

- [10] Humas BPKH, "Jumlah Jamaah Umroh 2022 Berkali-kali Lipat Dibandingkan Jamaah Haji Setiap Tahunnya."
- [11] B. Bramantyo, M. P. K. Putra, and H. Nirwana, "Implementasi Recurrent Neural Network Pada Multiclass Text Classification Judul Berita," *Media Borneo*, vol. 1, 2023.
- [12] P. A. Qori, D. S. Oktafani, and I. Kharisudin, "Analisis Peramalan dengan Long Short Term Memory pada Data Kasus Covid-19 di Provinsi Jawa Tengah," *PRISMA*, vol. 5, 2022.
- [13] Eko Supriyadi, "PREDIKSI PARAMETER CUACA MENGGUNAKAN DEEP LEARNING LONG-SHORT TERM MEMORY (LSTM)," *METEOROLOGI DAN GEOFISIKA*, vol. 21, 2021.
- [14] Aini Hanifa, Sugih Ahmad Fauzan, Muhammad Hikal, and Muhammad Bahrul Ashfiya, "Perbandingan Metode Lstm dan Gru (Rnn) Untuk Klasifikasi Berita Palsu Berbahasa Indonesia," *DINAMIKA REKAYASA*, vol. 17, 2021.
- [15] A. Nilsen, "Perbandingan Model RNN, Model LSTM, dan Model GRU dalam Memprediksi Harga Saham-Saham LQ45," *Statistika dan Aplikasinya*, vol. 6, 2022.
- [16] S. S. Nurashila, Faqih Hamami, and Tien Fabrianti Kusumasari, "Perbandingan Kinerja Algoritma Recurrent Neural Network (RNN) Dan Long Short-Term Memory (LSTM): Studi Kasus Prediksi Kemacetan Lalu Lintas Jaringan Pt XYZ," *Jurnal Ilmiah Penelitian dan Pembelajaran Informatika*, vol. 8, 2023.
- [17] S. Gopali, Faranak Abri, Sima Siami-Namini, and Akbar Siami Namin, "A Comparative Study of Detecting Anomalies in Time Series Data Using LSTM and TCN Models," *Research Gate*, 2021.
- [18] C. A. Maharani, Budi Warsito, and Rukun Santoso, "Analisis Sentimen Vaksin Covid-19 Pada Twitter Menggunakan Recurrent Neural Network(RNN) Dengan Algoritma Long Short-Term Memory (LSTM)," *GAUSSIAN*, vol. 12, 2023.
- [19] Nurfatima Selle, Novanto Yudistira, and Candra Dewi, "Perbandingan Prediksi Penggunaan Listrik Dengan Menggunakan Metode Long Short Term Memory (LSTM) Dan Recurrent Neural Network (RNN)," *Teknologi Informasi dan Ilmu Komputer (JTIK)*, vol. 9, 2022.
- [20] A. Rahmadeyan and Mustakim, "Long Short-Term Memory and Gated Recurrent Unit for Stock Price Prediction," *Procedia Computer Science*, vol. 234, 2024.
- [21] D. A. Nasution, Hidayah Husnul Khotimah, and Nurul Chamidah, "Perbandingan Normalisasi Data Untuk Klasifikasi Wine Menggunakan Algoritma K-NN," *Computer Engineering System and Science*, vol. 4, 2019.
- [22] N. T. Luchia, Ena Tasia, Indah Ramadhani, Akhas Rahmadeyan, and Raudiatul Zahra, "Performance Comparison Between Artificial Neural Network, Recurrent Neural Network and Long Short-Term Memory for Prediction of Extreme Climate Change," *Institute of Research and Publication Indonesia*, vol. 1, 2024.
- [23] NESMA M. REZK, MADHURA PURNAPRAJNA, TOMAS NORDSTRÖM, and ZAIN UL-ABDIN, "Recurrent Neural Networks: An Embedded Computing Perspective," *IEEE Access*, vol. 8, 2020.
- [24] N. Cahyani Amaliawati Rahmat and Y. Sabilu, "Factors Related To Traders Participation In Waste Management In Laino Market Muna District In 2018," 2019.
- [25] A. Rahmadeyan and Mustakim, "Seleksi Fitur pada Supervised Learning: Klasifikasi Prestasi Belajar Mahasiswa Saat dan Pasca Pandemi COVID-19," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 9, 2023.
- [26] F. Husaini, Inggih Permana, M. Afdal, and Febi Nur Salisah, "Penerapan Algoritma Long Short-Term Memory untuk Prediksi Produksi Kelapa Sawit," *MALCOM*, vol. 4, 2024.
- [27] R. B. R. Putra and Hendry, "Multivariate Time Series Forecasting pada Penjualan Barang Retail dengan Recurrent Neural Network," *INOVTEK POLBENG*, vol. 7, 2022.
- [28] N. Hasibuan, Yurmaini, and Erliyanti, "Pengaruh Perubahan Lingkungan Terhadap Kinerja Karyawan Pada Bank Syariah Indonesia," *Expensive*, vol. 2, 2023.