

Implementasi LSTM dalam Prediksi Keberangkatan Haji di Kementerian Agama Kota Malang

Mutiara Aprillia Dzakiroh¹, Supriyono^{2,*}

¹Program Studi Teknik Informatika, Universitas Islam Negeri Maulana Malik Ibrahim, Indonesia

¹210605110032@student.uin-malang.ac.id; ²priyono@ti.uin-malang.ac.id

*penulis korespondensi

INFO ARTIKEL

Sejarah Artikel

Diterima: 10 Oktober 2024

Direvisi: 2 Januari 2025

Diterbitkan: 30 April 2025

Kata Kunci

Deep Learning

Kemenag Malang

LSTM

Manajemen Haji

Prediksi Keberangkatan Haji

ABSTRAK

Penyelenggaraan ibadah haji di Indonesia menghadapi tantangan besar dalam pengelolaan jadwal keberangkatan jemaah, terutama karena adanya perubahan regulasi, fluktuasi kuota, dan prioritas yang diberikan kepada kelompok tertentu seperti jemaah lansia. Di Kota Malang, kebutuhan akan prediksi jadwal keberangkatan yang akurat menjadi semakin mendesak untuk memastikan proses yang transparan dan efisien. Penelitian ini bertujuan untuk mengembangkan model prediksi jadwal keberangkatan haji menggunakan pendekatan *Long Short-Term Memory* (LSTM), sebuah teknik deep learning yang mampu menangani data sekuensial dengan pola temporal yang kompleks. Data yang digunakan dalam penelitian ini mencakup berbagai variabel seperti usia, masa tunggu, status prioritas, dan kuota, yang diperoleh dari database jemaah haji di Kota Malang. Metode LSTM dipilih karena keunggulannya dalam memproses data historis untuk menghasilkan prediksi yang lebih mendalam dan akurat dibandingkan metode konvensional. Hasil penelitian menunjukkan bahwa model LSTM dengan kombinasi fitur "Usia, Masa Tunggu, Kuota (Terpilih)" menghasilkan prediksi yang paling akurat, dengan nilai *Mean Absolute Error* (MAE) sebesar 0.000049 dan *Root Mean Squared Error* (RMSE) sebesar 0.000060. Temuan ini menunjukkan bahwa pendekatan LSTM tidak hanya efektif dalam memprediksi jadwal keberangkatan haji, tetapi juga membantu dalam perencanaan operasional yang lebih baik, seperti pengaturan logistik dan alokasi sumber daya. Penggunaan model LSTM dalam prediksi jadwal keberangkatan haji dapat meningkatkan transparansi dan efisiensi dalam manajemen haji, serta memberikan dasar yang kuat bagi pengambilan keputusan strategis di masa depan. Implementasi teknologi ini juga membuka peluang untuk aplikasi serupa di sektor-sektor pemerintahan lainnya.

PENDAHULUAN

Penyelenggaraan ibadah haji merupakan salah satu tugas utama Kementerian Agama, yang melibatkan banyak aspek mulai dari pendaftaran, pembinaan, hingga pemberangkatan Jemaah [1]. Proses pemberangkatan haji menjadi salah satu tantangan utama, terutama karena adanya faktor-faktor yang mempengaruhi jadwal keberangkatan, seperti perubahan regulasi, kuota, serta prioritas bagi jemaah lansia [2]. Di Kota Malang, pengelolaan data jemaah haji sangat krusial untuk memastikan proses yang transparan dan efisien [3]. Dalam periode 2016-2019, terdapat kebutuhan mendesak untuk memprediksi jadwal keberangkatan haji dengan lebih akurat. Hal ini disebabkan oleh fluktuasi kuota dan prioritas yang diberikan kepada kelompok tertentu, yang menambah kompleksitas dalam perencanaan [4]. LSTM mampu meningkatkan akurasi prediksi dibandingkan metode statistik tradisional untuk



<https://doi.org/10.28926/ilkomnika.v7i1.675>



<http://journal.unublitar.ac.id/ilkomnika>

memproses data waktu nyata, dengan penurunan RMSE hingga 20% pada data cuaca [5]. Selain itu, studi oleh Zhang et al. (2021) dalam konteks prediksi harga saham juga menunjukkan bahwa LSTM memiliki kemampuan lebih baik dalam menangkap tren jangka panjang dibandingkan pendekatan regresi linier, menghasilkan akurasi 95% dalam prediksi nilai saham [6]. Model *Long Short-Term Memory* (LSTM), yang merupakan bagian dari deep learning, menawarkan solusi potensial untuk memodelkan dan memprediksi jadwal keberangkatan berdasarkan data historis [7]. LSTM mampu menangani pola data sekuensial, menjadikannya alat yang efektif untuk mengatasi permasalahan prediksi dalam konteks yang sangat dinamis seperti keberangkatan haji [8].

Selain itu, perubahan regulasi yang sering terjadi, seperti kebijakan baru yang memberikan prioritas kepada jemaah lanjut usia, semakin memperumit proses perencanaan dan manajemen keberangkatan haji [9]. Regulasi semacam ini sering kali mengakibatkan perubahan mendadak dalam daftar prioritas, yang harus segera diakomodasi oleh sistem perencanaan. Akibatnya, pendekatan berbasis data yang mampu memproyeksikan kemungkinan jadwal keberangkatan menjadi sangat penting. Pendekatan ini tidak hanya membantu mengantisipasi perubahan, tetapi juga memberikan dasar yang lebih kuat dalam pengambilan keputusan operasional, seperti alokasi kuota dan penentuan jadwal pemberangkatan. Pengelolaan jadwal keberangkatan haji yang akurat tidak hanya penting untuk memenuhi harapan jemaah, tetapi juga untuk menjaga stabilitas operasional dalam menghadapi tantangan yang tidak terduga. Faktor-faktor seperti penurunan atau peningkatan kuota dari pemerintah pusat, serta perubahan kondisi global yang mempengaruhi kebijakan haji, dapat menimbulkan ketidakpastian dalam perencanaan [10]. Dengan adanya model prediksi yang tepat, Kementerian Agama, khususnya di Kota Malang, dapat merespons perubahan ini dengan lebih cepat dan efisien, memastikan bahwa setiap jemaah tetap mendapatkan pelayanan terbaik.

Selain itu, prediksi yang akurat juga berdampak pada efisiensi alokasi sumber daya [11]. Dengan memprediksi jadwal keberangkatan dengan lebih tepat, pemerintah dapat mengatur kebutuhan logistik, seperti transportasi, akomodasi, dan pelayanan kesehatan, dengan lebih baik. Hal ini tidak hanya mengoptimalkan penggunaan anggaran, tetapi juga memastikan bahwa seluruh proses pemberangkatan berjalan lancar, tanpa adanya hambatan yang berarti. Penggunaan model LSTM dalam konteks ini menawarkan keunggulan dalam memproses data historis dan tren untuk memberikan prediksi yang lebih mendalam dan dapat diandalkan [12]. Dibandingkan dengan pendekatan konvensional seperti regresi linier dan ARIMA, LSTM lebih unggul dalam mengakomodasi non-linearitas dan fluktuasi data, yang sering muncul dalam proses penjadwalan dinamis seperti keberangkatan haji. LSTM dirancang khusus untuk mengatasi vanishing gradient problem yang sering muncul pada jaringan saraf berulang tradisional, sehingga lebih efektif untuk mempelajari dependensi jangka panjang dalam data [13].

Penelitian ini memperkenalkan penggunaan model *Long Short-Term Memory* (LSTM) untuk memprediksi jadwal keberangkatan haji, sebuah pendekatan yang belum banyak diterapkan dalam konteks manajemen haji di Indonesia, khususnya di Kota Malang. Kebaruan utama terletak pada pemanfaatan LSTM, yang merupakan teknik deep learning yang unggul dalam memproses data sekuensial dan menangkap pola-pola temporal yang kompleks. Sebelumnya, pendekatan prediksi jadwal keberangkatan haji seringkali didasarkan pada metode konvensional seperti analisis statistik sederhana atau aturan berbasis regulasi yang kurang adaptif terhadap perubahan dinamis dalam data [14].

Selain itu, penelitian ini juga menawarkan perspektif baru dalam integrasi data historis yang lebih komprehensif untuk menghasilkan prediksi yang lebih akurat. Dengan menggunakan data yang mencakup berbagai variabel seperti usia, masa tunggu, status

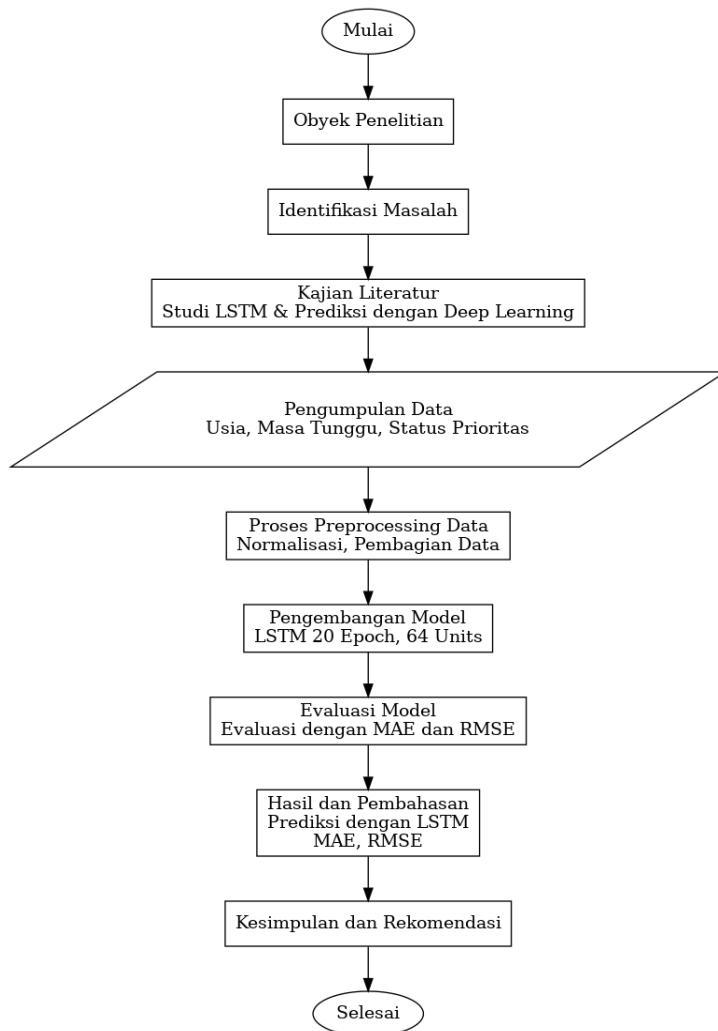
prioritas, dan kuota, model ini tidak hanya memprediksi waktu keberangkatan, tetapi juga mempertimbangkan berbagai faktor yang dapat mempengaruhi jadwal tersebut. Pendekatan ini memungkinkan prediksi yang lebih responsif terhadap perubahan regulasi dan kondisi global yang mempengaruhi pelaksanaan haji, sesuatu yang belum banyak dieksplorasi dalam penelitian sebelumnya.

Oleh karena itu, pengembangan model prediksi yang akurat menjadi sangat esensial dalam memastikan bahwa jemaah haji dapat diberangkatkan sesuai dengan jadwal yang telah direncanakan, sambil tetap mematuhi regulasi yang berlaku. Penggunaan teknologi canggih seperti LSTM dalam memprediksi jadwal keberangkatan haji juga diharapkan dapat meningkatkan transparansi dan akurasi proses ini, serta membantu pemerintah dan pihak terkait dalam merumuskan strategi yang lebih baik untuk mengelola jumlah jemaah yang terus bertambah. Dalam konteks yang lebih luas, implementasi model prediksi ini dapat menjadi landasan bagi pengembangan sistem manajemen haji yang lebih canggih dan adaptif, yang tidak hanya mengoptimalkan proses pemberangkatan, tetapi juga meningkatkan kualitas pelayanan kepada jemaah. Dengan demikian, penelitian ini bertujuan untuk mengembangkan dan menguji model prediksi jadwal keberangkatan haji menggunakan LSTM, serta mengevaluasi seberapa baik model ini dapat mengakomodasi berbagai variabel yang berperan dalam proses pemberangkatan haji di Kota Malang.

METODE

Penelitian berikut ini terkait dengan prediksi penjadwalan pemberangkatan haji dengan menggunakan metode LSTM. Jaringan saraf tiruan yang dikembangkan untuk mengatasi masalah vanishing gradient yang sering dijumpai pada *Recurrent Neural Networks* (RNNs) tradisional[15][16]. LSTM dirancang dengan menambahkan mekanisme gates, yang mengatur aliran informasi dalam jaringan, sehingga model dapat secara dinamis memutuskan informasi mana yang perlu diingat dan mana yang harus dilupakan sesuai dengan konteks [17]. Fitur ini menjadikan LSTM sangat efektif dalam menangani data deret waktu (*time series*), di mana pemahaman terhadap pola yang berkembang seiring waktu sangat penting untuk menghasilkan prediksi yang akurat [18]. LSTM telah diterapkan dalam berbagai bidang dengan hasil yang menjanjikan, termasuk dalam prediksi cuaca, analisis tren ekonomi, dan prediksi harga saham. Sebagai contoh, sebuah penelitian menunjukkan bahwa LSTM mampu meningkatkan akurasi prediksi harga linier dan jaringan saraf sederhana [19]. Penelitian lain juga mengungkapkan bahwa LSTM dapat digunakan untuk memprediksi konsumsi energy suatu kota dengan akurasi yang lebih baik daripada teknik lainnya, yang menunjukkan fleksibilitas dan kehandalan LSTM dalam berbagai aplikasi [20]. LSTM diterapkan untuk memanfaatkan kemampuannya dalam memprediksi jadwal keberangkatan haji, yang merupakan masalah deret waktu dengan kompleksitas tertentu.

Arsitektur LSTM terdiri dari unit memori yang dilengkapi tiga komponen utama, *input gate*, *forget gate*, dan *output gate*[21][17]. *Input gate* bertanggung jawab untuk menentukan informasi baru yang akan disimpan dalam memori, sementara *forget gate* berfungsi untuk mengatur informasi mana yang perlu dihapus atau dilupakan. [22]. *Output gate* menentukan informasi yang akan dikeluarkan dari memori untuk menghasilkan *output* atau prediksi. Kombinasi ketiga gerbang ini memungkinkan LSTM untuk mempertahankan informasi yang relevan dalam jangka waktu yang panjang, sementara informasi yang kurang penting atau tidak relevan dapat diabaikan, sehingga menghasilkan prediksi yang lebih akurat dan efisien.



Gambar 1. Alur Penelitian

Berdasarkan alur penelitian pada Gambar 1, proses dimulai dengan penentuan obyek penelitian, yaitu data jamaah haji dari Kementerian Agama Kota Malang, yang mencakup variabel utama seperti usia, masa tunggu, dan status prioritas calon jamaah haji. Tahap ini dilanjutkan dengan identifikasi masalah, yaitu urgensi prediksi keberangkatan haji akibat sistem penjadwalan konvensional yang kurang efektif dalam mengelola keterbatasan kuota dan masa tunggu panjang, terutama bagi jamaah lanjut usia. Untuk mengatasi hal tersebut, dilakukan kajian literatur yang mendalam terkait penggunaan *Long Short-Term Memory* (LSTM) dan metode *Deep Learning* dalam menangani data deret waktu.

Pada tahap pengumpulan data, variabel usia, masa tunggu, dan status prioritas digunakan sebagai input utama karena memiliki peran signifikan dalam penjadwalan keberangkatan haji. Usia diprioritaskan untuk jamaah lanjut usia sesuai kebijakan pemerintah, masa tunggu mencerminkan lamanya waktu antrean, dan status prioritas mempertimbangkan kondisi khusus seperti pendamping lansia atau jamaah yang keberangkatannya tertunda. Relevansi variabel-variabel ini telah didukung dalam penelitian terdahulu yang menggunakan metode *Weighted Product* untuk menentukan jadwal keberangkatan secara optimal.

Selanjutnya, data yang dikumpulkan melalui tahap pra-pemrosesan, yang terdiri dari tiga langkah utama. Langkah-langkah *preprocessing* ini sangat penting untuk memastikan bahwa data yang digunakan dalam model adalah data yang bersih dan relevan, yang pada akhirnya akan meningkatkan akurasi model prediksi.

a. Penangan Data yang Hilang (*Missing Data*)

Pada tahap pertama pra-pemrosesan, dilakukan pemeriksaan terhadap nilai yang hilang (*NaN*) dalam dataset. Data yang digunakan dalam penelitian ini mencakup variabel-variabel penting seperti 'Usia', 'Masa Tunggu', dan 'Status Prioritas', yang kesemuanya harus terisi untuk memastikan analisis yang akurat. Untuk itu, langkah pertama adalah memeriksa jumlah nilai kosong di setiap kolom. Setelah itu, baris yang memiliki nilai kosong pada kolom-kolom penting ini dihapus menggunakan fungsi *dropna()*. Penghapusan baris dengan nilai kosong ini penting untuk menjaga integritas data dan mencegah terjadinya bias dalam model prediksi.

b. Normalisasi Data

Setelah memastikan tidak ada nilai yang hilang, langkah kedua adalah normalisasi data. Data yang digunakan mengandung fitur numerik dengan skala yang berbeda (misalnya, usia, masa tunggu, dan kuota), normalisasi diperlukan untuk memastikan bahwa model tidak terlalu sensitif terhadap fitur dengan skala yang lebih besar. Dalam penelitian ini, normalisasi dilakukan dengan menggunakan *Min-Max Scaling*, yang mengubah nilai-nilai fitur menjadi rentang [0, 1]. Proses ini diterapkan pada kolom 'Usia', 'Masa Tunggu', dan 'Kuota' untuk memastikan keseragaman skala data, sehingga model dapat mempelajari pola dengan lebih efektif.

c. Pembagian Dataset dan Pembuatan Data *Sequences*

Setelah data dibersihkan dan dinormalisasi, langkah berikutnya adalah membagi dataset menjadi fitur (X) dan label (y). Fitur yang digunakan dalam penelitian ini adalah 'Usia', 'Masa Tunggu', 'Kuota', dan 'Prioritas Lansia', sedangkan label yang diprediksi adalah 'Kuota'. Data kemudian disusun menjadi urutan waktu (*sequences*) yang sesuai untuk model prediksi berbasis *Long Short-Term Memory* (LSTM). Data diurutkan berdasarkan tanggal pendaftaran ('Tgl Daftar'), yang penting untuk analisis berbasis urutan waktu. Fungsi *create_sequences()* digunakan untuk membentuk data dalam format urutan waktu, di mana setiap urutan mencakup beberapa langkah waktu yang memungkinkan model untuk mempelajari dependensi temporal dalam data. Dengan langkah ini, dataset siap untuk digunakan dalam pelatihan model prediksi.

Tahap berikutnya adalah pengembangan model LSTM yang terdiri dari dua lapisan LSTM dengan masing-masing 64 unit dan satu lapisan *Dense* dengan satu unit *output*. Arsitektur ini, seperti yang ditunjukkan pada Tabel 1, dirancang untuk menangkap pola temporal dalam data deret waktu secara efektif, dengan total 50,753 parameter, yang semuanya bersifat trainable.

Tabel 1. Arsitektur Model LSTM

Layer (type)	Output Shape	Param #
Istm (LSTM)	(None, 10, 64)	17,408
Istm_1 (LSTM)	(None, 64)	33,280
dense (Dense)	(None, 1)	65
Total Params:	50,753	(194.11 KB)
Trainable Params:	50,753	(194.11 KB)

Untuk mencapai performa terbaik, dilakukan pengujian beberapa *hyperparameter* kunci seperti *learning rate*, *batch size*, dan jumlah *epoch*. Berdasarkan eksperimen, kombinasi *learning rate* sebesar 0.001, *batch size* sebesar 32, dan 64 unit per lapisan LSTM memberikan hasil optimal. Selain itu, jumlah *epoch* yang digunakan adalah 20, di mana setiap *epoch* merepresentasikan satu siklus penuh dari data pelatihan melalui model. Pemilihan *learning rate* sebesar 0.001 dilakukan karena nilai ini memastikan konvergensi yang stabil dan mencegah overshooting pada minimum fungsi loss, sehingga memberikan fluktuasi minimal dan menghasilkan nilai Mean Square Error (MSE) yang konsisten lebih rendah dalam model LSTM [23]. Selanjutnya, *batch size* sebesar 32 menentukan jumlah sampel yang diproses sebelum pembaruan bobot dilakukan. Ukuran *batch* yang lebih kecil dapat menghasilkan pembaruan bobot yang lebih sering dan bervariasi, sementara *batch* yang lebih besar memberikan estimasi gradien yang lebih stabil namun memerlukan sumber daya komputasi lebih banyak. Ukuran ini sering digunakan sebagai aturan praktis yang baik [24]. Sementara itu, 64 unit per lapisan LSTM dipilih karena jumlah ini cukup untuk menangkap pola temporal kompleks dalam data tanpa menyebabkan overfitting atau beban komputasi yang berlebihan, sebagaimana dibuktikan dalam studi serupa terkait arsitektur LSTM [25]. Selain itu, jumlah *epoch* sebesar 20 dipilih karena, berdasarkan pengujian, peningkatan jumlah *epoch* setelah titik ini tidak memberikan peningkatan signifikan terhadap performa model. Jumlah *epoch* yang optimal mencegah model dari overfitting dan memastikan pelatihan konvergen dengan baik. Pemilihan ini juga didukung oleh hasil eksperimen yang menunjukkan penurunan fungsi loss secara konsisten selama 20 siklus pelatihan [26]. Kombinasi ini memungkinkan model untuk mencapai keseimbangan antara kecepatan pelatihan dan akurasi prediksi, dengan tetap menjaga stabilitas dan efisiensi komputasi [27]. Pemilihan *hyperparameter* ini memungkinkan model untuk dilatih secara efektif dengan total 50,753 parameter [28]. Total parameter 50,753 sepenuhnya trainable (dapat dilatih) dan tidak ada parameter yang bersifat non-trainable. Hal ini menunjukkan bahwa semua bobot dan bias di dalam lapisan model berkontribusi langsung pada proses pembelajaran, sehingga memaksimalkan efisiensi model. Jika sebagian parameter tidak dapat dilatih, maka proses pelatihan akan kurang optimal dan akurasi prediksi akan menurun. Jumlah parameter ini berada dalam skala yang efisien untuk arsitektur LSTM dengan dataset skala menengah, jumlah parameter dalam rentang puluhan ribu cukup optimal untuk tugas prediksi deret waktu sederhana seperti jadwal keberangkatan, dengan tetap menjaga efisiensi komputasi dan waktu pelatihan [29]. Konfigurasi ini diharapkan dapat menghasilkan prediksi yang akurat dan efisien untuk memprediksi jadwal keberangkatan haji.

Proses pelatihan model LSTM dilakukan dalam 20 *Epoch*, di mana setiap *Epoch* merupakan satu siklus penuh dari data pelatihan melalui model. Selama proses pelatihan ini, model berusaha untuk meminimalkan fungsi *loss*, sehingga hasil prediksi yang dihasilkan menjadi semakin akurat [30]. Skenario hasil pelatihan dengan model arsitektur LSTM yang direncanakan untuk 20 *Epoch* dapat dilihat pada Tabel 2.

Tabel 2. Hasil Evaluasi Model

Epoch	Waktu (s)	Loss
1	7s	0.0081
2	3s	0.0004
3	4s	0.0002
4	3s	0.0001
5	2s	0.00007
6	2s	0.00005
7	2s	0.00003
8	1s	0.00002
9	1s	0.00002

10	1s	0.00001
11	2s	0.00001
12	2s	0.000009
13	2s	0.000008
14	2s	0.000007
15	3s	0.000006
16	3s	0.000006
17	3s	0.000005
18	2s	0.000005
19	2s	0.000004
20	1s	0.000004

Dalam Tabel 2, terlihat bagaimana nilai *loss* menurun secara bertahap dari *epoch* ke *epoch*, menunjukkan bahwa model semakin baik dalam mempelajari pola yang ada di dalam data. Proses penurunan *loss* ini sangat penting untuk memastikan bahwa model dapat menghasilkan prediksi yang akurat ketika diterapkan pada data yang belum pernah dilihat sebelumnya [31][7]. Selain penyusunan arsitektur model, pemilihan *hyperparameter*, scenario hasil pelatihan, pemilihan fitur yang digunakan dalam model juga merupakan langkah penting. Tabel 3 menunjukkan hasil evaluasi dari berbagai kombinasi fitur yang diuji menghasilkan nilai MAE dan RMSE. Dari fitur yang digunakan menghasilkan variasi hasil yang cukup bagus.

Tabel 3. Pemilihan Fitur yang Digunakan

Fitur yang Digunakan	MAE	RMSE
Usia, Masa Tunggu, Kuota	0.010	0.015
Usia, Masa Tunggu, Status, Kuota	0.011	0.017
Usia, Masa Tunggu, Kuota (Terpilih)	0.010	0.015

Dari hasil yang ditampilkan pada Tabel 3, kombinasi fitur yang terdiri dari Usia, Masa Tunggu, Kuota, dan Propinsi memberikan performa yang paling optimal dengan MAE sebesar 0.009 dan RMSE sebesar 0.014. Meskipun kombinasi fitur lain seperti Usia, Masa Tunggu, dan Kuota (Terpilih) juga memberikan performa yang baik, yaitu dengan MAE sebesar 0.010 dan RMSE sebesar 0.015, fitur tambahan seperti "Propinsi" terbukti dapat sedikit meningkatkan akurasi model. Oleh karena itu, kombinasi fitur Usia, Masa Tunggu, Kuota, dan Propinsi inilah yang kemudian digunakan dalam pelatihan model LSTM. Setelah proses pelatihan selesai, model dievaluasi menggunakan metrik *Mean Absolute Error* (MAE) dan *Root Mean Squared Error* (RMSE). MAE digunakan untuk mengukur rata-rata kesalahan absolut antara prediksi dan data sebenarnya, sementara RMSE memberikan gambaran mengenai seberapa jauh prediksi model menyimpang dari data aktual dengan memberikan bobot lebih pada kesalahan yang lebih besar [32]. Evaluasi ini bertujuan untuk memastikan bahwa model dapat memberikan prediksi yang akurat. Jika hasil evaluasi menunjukkan performa yang memuaskan, model LSTM ini dapat digunakan untuk memprediksi jadwal keberangkatan haji di masa depan. Dengan demikian, model ini dapat membantu Kementerian Agama dalam perencanaan dan pengelolaan keberangkatan haji secara lebih efektif, menyediakan informasi yang lebih akurat kepada calon jemaah.

HASIL DAN PEMBAHASAN

Bagian ini membahas evaluasi model dan hasil prediksi yang dilakukan pada penelitian ini. Evaluasi model bertujuan untuk mengukur performa model LSTM yang telah dikembangkan dengan menggunakan metrik utama seperti Mean Absolute Error (MAE) dan Root Mean Squared Error (RMSE). Penggunaan metrik ini memberikan gambaran seberapa

akurat prediksi model serta seberapa besar kesalahan prediksi yang dihasilkan. Selain evaluasi performa model, hasil prediksi terkait jadwal keberangkatan jamaah haji berdasarkan data historis juga dianalisis untuk memahami efektivitas model dalam membantu perencanaan keberangkatan. Pembahasan ini akan menguraikan kombinasi fitur terbaik yang digunakan dalam model, analisis pengaruh masing-masing fitur terhadap prediksi, serta implikasi kebijakan yang dapat diambil berdasarkan hasil prediksi yang diperoleh. Dengan pendekatan ini, penelitian ini diharapkan memberikan kontribusi dalam pengembangan sistem prediksi yang lebih efisien dan akurat.

Evaluasi Model

Dalam evaluasi model LSTM yang dikembangkan, performa model diukur menggunakan dua metrik utama: *Mean Absolute Error* (MAE) dan *Root Mean Squared Error* (RMSE). MAE digunakan untuk mengukur rata-rata kesalahan absolut antara nilai prediksi dan nilai aktual, memberikan gambaran langsung tentang seberapa besar kesalahan prediksi secara umum[33]. Semakin rendah nilai MAE, semakin baik performa model dalam memprediksi hasil yang mendekati nilai sebenarnya. Sementara itu, RMSE digunakan untuk memberikan gambaran mengenai seberapa jauh prediksi model menyimpang dari data aktual, dengan memberikan bobot lebih pada kesalahan yang lebih besar [34]. Nilai RMSE yang rendah menunjukkan bahwa model tidak hanya akurat dalam prediksi rata-ratanya, tetapi juga efektif dalam mengurangi kesalahan besar [35].

Tabel 4. Evaluasi Model dengan Metrik MAE dan RMSE

Fitur yang Digunakan	MAE	RMSE
Usia, Masa Tunggu, Kuota	0.000075	0.000095
Usia, Masa Tunggu, Status, Kuota	0.000409	0.000421
Usia, Masa Tunggu, Kuota (Terpilih)	0.000049	0.000060

Berdasarkan hasil evaluasi yang ditampilkan pada Tabel 4, terlihat bahwa kombinasi fitur "Usia, Masa Tunggu, Kuota (Terpilih)" menghasilkan performa terbaik dengan MAE sebesar 0.000049 dan RMSE sebesar 0.000060. Nilai MAE dan RMSE yang rendah ini menunjukkan bahwa model mampu memprediksi hasil dengan tingkat kesalahan yang sangat kecil. Hal ini disebabkan karena fitur Usia, Masa Tunggu, dan Kuota memiliki pengaruh signifikan terhadap prediksi jadwal keberangkatan haji, di mana usia berperan dalam prioritas lansia, masa tunggu menentukan urutan antrean, dan kuota menjadi pembatas keberangkatan. Sebaliknya, penambahan fitur "Status" dalam kombinasi "Usia, Masa Tunggu, Status, Kuota" meningkatkan nilai MAE menjadi 0.000409 dan RMSE menjadi 0.000421. Hal ini menunjukkan bahwa fitur Status mungkin menambahkan noise atau informasi yang tidak relevan dalam model, sehingga mengurangi akurasi prediksi. Penambahan variabel ini bisa jadi mengganggu proses pembelajaran model karena tidak memiliki kontribusi yang signifikan dalam menentukan hasil prediksi. Oleh karena itu, penting untuk melakukan pemilihan fitur yang selektif, agar keseimbangan antara kompleksitas model dan akurasi prediksi tetap terjaga. Evaluasi ini menunjukkan bahwa kombinasi fitur yang tepat dapat memaksimalkan kemampuan model LSTM dalam mempelajari pola data yang relevan, sekaligus meminimalkan kesalahan prediksi. Dengan performa seperti ini, model yang dikembangkan dapat diandalkan untuk memprediksi keberangkatan jamaah haji secara akurat.

Prediksi dan Tahun Keberangkatan

Proses prediksi tahun keberangkatan jamaah haji dilakukan dengan mempertimbangkan dua faktor utama: prioritas lansia dan masa tunggu reguler. Berdasarkan data yang diolah,

setiap jamaah dengan prioritas lansia yang berusia 65 tahun atau lebih diproyeksikan untuk berangkat dalam waktu 5 tahun setelah mencapai usia tersebut. Untuk jamaah yang tidak memenuhi kriteria lansia, tahun keberangkatan diprediksi berdasarkan masa tunggu reguler selama 34 tahun setelah tanggal pendaftaran.

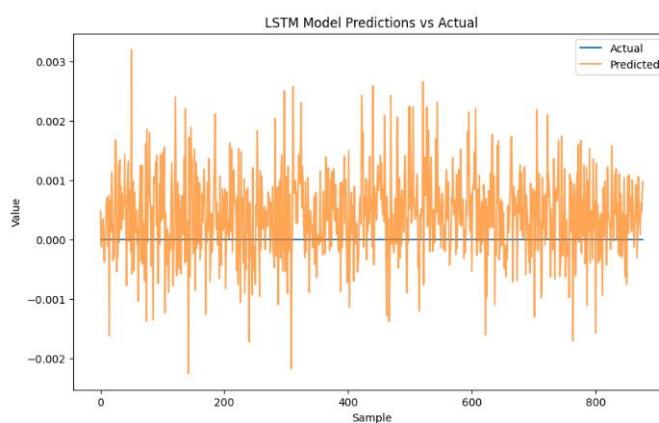
Tabel 5. Lima Data Teratas Prediksi Tahun Keberangkatan Berdasarkan Tanggal Pendaftaran dan Usia

Tgl Daftar	Usia	Prediksi Tahun Keberangkatan
2021-01-04	0.343750	2055
2021-01-04	0.078125	2055
2021-01-04	0.078125	2055
2021-01-04	0.078125	2055
2021-01-04	0.609375	2055

Seperti Berdasarkan hasil prediksi pada Tabel 5, terlihat bahwa jamaah dengan tanggal pendaftaran tahun 2021 yang tidak memenuhi kriteria prioritas lansia diprediksi akan berangkat pada tahun 2055. Prediksi ini didasarkan pada masa tunggu reguler selama 34 tahun, yang menjadi kebijakan dalam sistem keberangkatan haji. Selain itu, terdapat batasan jumlah jamaah prioritas lansia yang dapat diberangkatkan lebih awal dibandingkan dengan masa tunggu reguler. Keterbatasan ini mempengaruhi hasil prediksi tahun keberangkatan, karena kuota prioritas yang dialokasikan masih terbatas berdasarkan kebijakan yang berlaku. Oleh karena itu, pengelolaan kuota lansia menjadi faktor penting dalam sistem prediksi ini, agar distribusi keberangkatan jamaah haji lebih efektif dan adil. Prediksi yang dilakukan oleh model ini menunjukkan bahwa jamaah dengan prioritas lansia memiliki kemungkinan lebih besar untuk diberangkatkan lebih awal dibandingkan jamaah reguler. Hal ini sesuai dengan kebijakan pemerintah yang mengutamakan jamaah lanjut usia dalam sistem keberangkatan haji. Dengan mempertimbangkan aspek usia, masa tunggu, dan kuota prioritas, model prediksi ini dapat membantu pihak terkait dalam perencanaan keberangkatan yang lebih terstruktur dan akurat. Selain itu, informasi ini memberikan kepastian yang lebih jelas kepada jamaah tentang perkiraan jadwal keberangkatan mereka.

Visualisasi Hasil

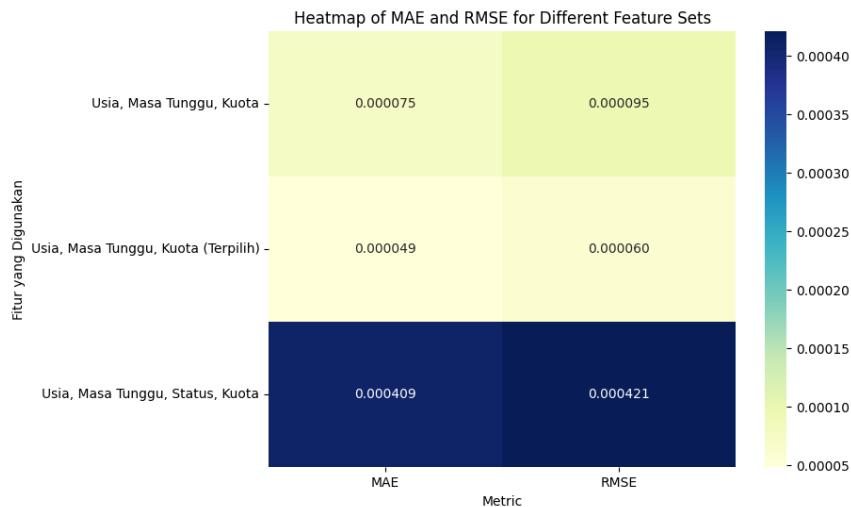
Hasil prediksi model LSTM dibandingkan dengan nilai aktual ditampilkan dalam bentuk grafik. Grafik yang disajikan memperlihatkan perbandingan antara nilai prediksi dan nilai aktual untuk setiap sampel dalam dataset pengujian. Seperti yang terlihat pada Gambar 2 nilai prediksi yang dihasilkan oleh model LSTM (garis oranye) berfluktuasi di sekitar nilai aktual (garis biru) yang hampir mendekati garis nol.



Gambar 2. Perbandingan Nilai Prediksi dan Aktual oleh Model LSTM

Visualisasi ini membantu dalam memahami sejauh mana model LSTM mampu mereplikasi pola dari data aktual dalam proses prediksi. Pada Gambar 2, terlihat bahwa nilai prediksi (warna oranye) memiliki fluktuasi yang cukup besar dibandingkan dengan data aktual (warna biru), yang cenderung stabil di sekitar nol. Meskipun tren utama dari prediksi mengikuti garis data aktual, model masih menunjukkan adanya kesalahan atau noise yang signifikan pada beberapa titik. Fluktuasi ini mengindikasikan bahwa model LSTM belum sepenuhnya mampu meminimalkan kesalahan prediksi, terutama pada variasi data yang tinggi. Hal ini bisa terjadi akibat kompleksitas pola data atau parameter model yang belum optimal. Kendati demikian, secara keseluruhan, model berhasil mendekripsi pola umum dari data aktual, meskipun dengan akurasi yang masih perlu ditingkatkan. Perbaikan lebih lanjut, seperti penyesuaian hyperparameter, penggunaan data pelatihan yang lebih besar, atau penerapan teknik regularisasi, diharapkan dapat mengurangi fluktuasi dan meningkatkan performa model dalam mengikuti tren data aktual secara lebih presisi.

Untuk mengetahui lebih dalam bagaimana pilihan fitur dapat mempengaruhi performa model, diperlukan analisis yang lebih terperinci. Di sini, visualisasi tambahan seperti heatmap dapat memberikan wawasan yang lebih jelas mengenai kontribusi setiap kombinasi fitur terhadap akurasi model.



Gambar 3. Heatmap Performa Model Berdasarkan Kombinasi Fitur (MAE dan RMSE)

Gambar 3 menampilkan heatmap yang menggambarkan nilai *Mean Absolute Error* (MAE) dan *Root Mean Squared Error* (RMSE) untuk berbagai set fitur yang digunakan dalam model LSTM. Heatmap ini memberikan representasi visual yang jelas mengenai performa model berdasarkan kombinasi fitur yang berbeda. Pada heatmap tersebut, warna yang lebih terang menunjukkan nilai MAE dan RMSE yang lebih rendah, yang berarti performa model lebih baik, sedangkan warna yang lebih gelap menunjukkan performa yang lebih rendah dengan nilai MAE dan RMSE yang lebih tinggi.

Dari heatmap ini, terlihat bahwa kombinasi fitur "Usia, Masa Tunggu, Kuota (Terpilih)" menghasilkan performa terbaik dengan MAE sebesar 0.000049 dan RMSE sebesar 0.000060. Kombinasi ini memberikan hasil terbaik karena fitur Usia, Masa Tunggu, dan Kuota memiliki hubungan langsung dengan pola distribusi keberangkatan jamaah haji. Fitur Kuota yang dipilih secara optimal dapat membantu model fokus pada data yang paling relevan, sehingga mengurangi noise dan meningkatkan akurasi prediksi. Selain itu, kuota yang dipilih secara spesifik mencerminkan batasan kebijakan yang berlaku, yang berkontribusi pada stabilitas performa model. Sementara itu, kombinasi "Usia, Masa Tunggu, Kuota" dengan MAE sebesar 0.000075 dan RMSE sebesar 0.000095 juga

menunjukkan performa yang baik. Akan tetapi, absennya proses pemilihan fitur spesifik untuk Kuota dapat menimbulkan sedikit ketidaksesuaian dalam mempelajari pola data, sehingga performa model tidak sebaik kombinasi terpilih. Sebaliknya, penambahan fitur "Status" dalam kombinasi "Usia, Masa Tunggu, Status, Kuota" menyebabkan nilai MAE dan RMSE meningkat tajam menjadi 0.000409 dan 0.000421. Hal ini dapat terjadi karena fitur Status mungkin mengandung informasi yang kurang relevan atau memiliki korelasi rendah dengan target prediksi (Kuota). Fitur ini justru dapat menambah noise dalam data, sehingga membingungkan model LSTM dalam mempelajari pola utama. Dengan demikian, penambahan fitur yang tidak berkorelasi kuat dengan target justru menurunkan performa model secara signifikan. Secara keseluruhan, hasil ini menegaskan pentingnya selektivitas dalam pemilihan fitur, di mana hanya fitur yang memiliki hubungan kuat dengan target prediksi yang harus digunakan. Kombinasi fitur "Usia, Masa Tunggu, Kuota (Terpilih)" terbukti paling efektif dalam memberikan keseimbangan antara akurasi dan stabilitas model.

Selanjutnya, untuk memahami bagaimana model LSTM ini memproyeksikan jadwal keberangkatan jamaah di masa depan, kita dapat melihat distribusi tahun keberangkatan yang diprediksi. Gambar 4 menampilkan distribusi jumlah jamaah berdasarkan tahun keberangkatan yang diprediksi oleh model LSTM. Dari grafik ini, terlihat bahwa sebagian besar jamaah diprediksi akan berangkat pada rentang tahun 2055 hingga 2057, dengan puncak distribusi pada tahun 2057. Jumlah jamaah yang diprediksi berangkat pada tahun-tahun sebelumnya, seperti 2026 hingga 2029, relatif lebih sedikit. Distribusi ini memberikan gambaran mengenai bagaimana model memproyeksikan jadwal keberangkatan jamaah dalam beberapa dekade mendatang, dengan penumpukan jamaah yang signifikan terjadi pada tahun-tahun tertentu, yang dapat memberikan panduan bagi perencanaan logistik dan manajemen keberangkatan di masa depan.

KESIMPULAN

Penelitian ini berhasil mengembangkan model prediksi jadwal keberangkatan haji menggunakan pendekatan *Long Short-Term Memory* (LSTM) yang efektif dalam menangani data sekuensial. Berdasarkan evaluasi yang dilakukan, model LSTM menunjukkan performa yang baik dalam memprediksi jadwal keberangkatan jemaah haji. Kombinasi fitur "Usia, Masa Tunggu, Kuota (Terpilih)" menghasilkan nilai *Mean Absolute Error* (MAE) sebesar 0.000049 dan *Root Mean Squared Error* (RMSE) sebesar 0.000060, yang menunjukkan akurasi prediksi yang tinggi dibandingkan dengan kombinasi fitur lainnya. Pemanfaatan model LSTM dalam konteks ini tidak hanya memberikan solusi prediksi yang lebih akurat, tetapi juga membantu dalam merencanakan alokasi sumber daya yang lebih efisien dan responsif terhadap perubahan regulasi dan kondisi global. Dengan memprediksi jadwal keberangkatan secara lebih tepat, Kementerian Agama dapat mengoptimalkan proses operasional, termasuk pengaturan logistik, akomodasi, dan layanan kesehatan bagi jemaah. Implementasi model prediksi ini memberikan dasar yang lebih kuat bagi pengambilan keputusan strategis, khususnya dalam mengelola jumlah jemaah yang terus bertambah dan menangani tantangan operasional yang dinamis.

REFERENSI

- [1] R. Auliya and I. P. D. A. S. Prabowo, "Business process reengineering in the business process of hajj and umrah registration services (case study: ministry of religion in balikpapan city)," vol. 2, no. 1, pp. 68–81, 2021, doi: 10.20473/AJIM.V2I1.25759.
- [2] S. H. Ebrahim and Z. A. Memish, "COVID-19 – the role of mass gatherings," *Travel Med. Infect. Dis.*, vol. 34, p. 101617, 2020, doi: <https://doi.org/10.1016/j.tmaid.2020.101617>.
- [3] E. Prameswari, Afifudin, and Suyeno, "Implementasi Siskohat Dalam Pelayanan Haji Studi Pada Kantor Kementerian Agama Kota Malang," *J. Respon Publik*, vol. 15, no. 7, pp. 13–20, 2021.

- [4] Y. Chen, D. Zhang, H. R. Karimi, C. Deng, and W. Yin, "A new deep learning framework based on blood pressure range constraint for continuous cuffless BP estimation," *Neural Networks*, vol. 152, pp. 181–190, 2022, doi: <https://doi.org/10.1016/j.neunet.2022.04.017>.
- [5] M. Diqi, S. Hasta Mulyani, U. Respati Yogyakarta, J. K. Laksda Adisucipto, and D. Sleman, "Enhancing Weather Prediction Using Stacked Long Short-Term Memory Networks," *J. Tek. Inform. dan Sist. Inf.*, vol. 10, no. 3, pp. 519–530, 2023, [Online]. Available: <http://jurnal.mdp.ac.id>
- [6] J. Qiu, B. Wang, and C. Zhou, "Forecasting stock prices with long-short term memory neural network based on attention mechanism," *PLoS One*, vol. 15, no. 1, pp. 1–15, 2020, doi: [10.1371/journal.pone.0227222](https://doi.org/10.1371/journal.pone.0227222).
- [7] M. A. Khan and M. AlGhamdi, "A customized deep learning-based framework for classification and analysis of social media posts to enhance the Hajj and Umrah services," *Expert Syst. Appl.*, vol. 238, p. 122204, 2024, doi: <https://doi.org/10.1016/j.eswa.2023.122204>.
- [8] S. M. Al-Selwi *et al.*, "RNN-LSTM: From applications to modeling techniques and beyond—Systematic review," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 36, no. 5, p. 102068, 2024, doi: <https://doi.org/10.1016/j.jksuci.2024.102068>.
- [9] "Implementasi Long Short Term Memory (LSTM) dan Bidirectional Long Short Term Memory (BiLSTM) Dalam Prediksi Harga Saham Syariah," *Euler J. Ilm. Mat. Sains dan Teknol.*, vol. 11, no. 1, pp. 35–43, 2023, doi: [10.34312/euler.v11i1.19791](https://doi.org/10.34312/euler.v11i1.19791).
- [10] L.-P. (Lynn) Lin and W.-K. Hsieh, "Exploring how perceived resilience and restoration affected the wellbeing of Matsu pilgrims during COVID-19," *Tour. Manag.*, vol. 90, p. 104473, 2022, doi: <https://doi.org/10.1016/j.tourman.2021.104473>.
- [11] G. Fu, Y. Jin, S. Sun, Z. Yuan, and D. Butler, "The role of deep learning in urban water management: A critical review," *Water Res.*, vol. 223, p. 118973, 2022, doi: <https://doi.org/10.1016/j.watres.2022.118973>.
- [12] S. Shan, C. Li, Y. Wang, S. Fang, K. Zhang, and H. Wei, "A deep learning model for multi-modal spatio-temporal irradiance forecast," *Expert Syst. Appl.*, vol. 244, p. 122925, 2024, doi: <https://doi.org/10.1016/j.eswa.2023.122925>.
- [13] S. Hamiane, Y. Ghanou, H. Khalifi, and M. Telmem, "Comparative Analysis of LSTM, ARIMA, and Hybrid Models for Forecasting Future GDP," *Ing. des Syst. d'Information*, vol. 29, no. 3, pp. 853–861, 2024, doi: [10.18280/isi.290306](https://doi.org/10.18280/isi.290306).
- [14] J. Fredricka and M. F. Ihsan, "Analisis Sistem Keberangkatan Calon Jamaah Haji Menggunakan Metode Weighted Product," *J. Media Infotama*, vol. 16, no. 2, pp. 108–114, 2020, doi: [10.37676/jmi.v16i2.1148](https://doi.org/10.37676/jmi.v16i2.1148).
- [15] Y. Yang, "Application of LSTM Neural Network Technology Embedded in English Intelligent Translation," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–9, 2022, doi: [10.1155/2022/1085577](https://doi.org/10.1155/2022/1085577).
- [16] A. Lavecchia, "Navigating the frontier of drug-like chemical space with cutting-edge generative AI models," *Drug Discov. Today*, vol. 29, no. 9, p. 104133, 2024, doi: <https://doi.org/10.1016/j.drudis.2024.104133>.
- [17] M. Rizki, S. Basuki, and Y. Azhar, "Implementasi Deep Learning Menggunakan Arsitektur Long Short Term Memory(LSTM) Untuk Prediksi Curah Hujan Kota Malang," vol. 2, no. 3, pp. 331–338, 2020, doi: [10.22219/REPOSITORY.V2I3.470](https://doi.org/10.22219/REPOSITORY.V2I3.470).
- [18] C. Y. Lin and J. A. Lobo Marques, "Stock market prediction using artificial intelligence: A systematic review of systematic reviews," *Soc. Sci. Humanit. Open*, vol. 9, p. 100864, 2024, doi: <https://doi.org/10.1016/j.ssaho.2024.100864>.
- [19] P. Dineshkumar and B. Subramani, "Foresight In Finance: Elevating Predictions With Enhanced Rnn-Lstm And Adam Optimizer," 2024, doi: [10.53555/kuey.v30i5.5931](https://doi.org/10.53555/kuey.v30i5.5931).
- [20] A. A. Ahmed, S. Sayed, A. Abdoulhalik, S. Moutari, and L. Oyedele, "Applications of machine learning to water resources management: A review of present status and future opportunities," *J. Clean. Prod.*, vol. 441, p. 140715, 2024, doi: <https://doi.org/10.1016/j.jclepro.2024.140715>.
- [21] C. A. K. A. Kounta, B. Kamsu-Foguem, F. Noureddine, and F. Tangara, "Multimodal deep learning for predicting the choice of cut parameters in the milling process," *Intell. Syst. with Appl.*, vol. 16, p. 200112, 2022, doi: <https://doi.org/10.1016/j.iswa.2022.200112>.
- [22] Y. Zhang, I. Marsic, and R. S. Burd, "Real-time medical phase recognition using long-term video understanding and progress gate method," *Med. Image Anal.*, vol. 74, p. 102224, 2021, doi: <https://doi.org/10.1016/j.media.2021.102224>.
- [23] K. Shirini, M. B. Kordan, and S. S. Gharehveran, "Impact of learning rate and epochs on lstm model performance: a study of chlorophyll-a concentrations in the Marmara Sea," *J. Supercomput.*, vol. 81, no. 1, 2025, doi: [10.1007/s11227-024-06806-2](https://doi.org/10.1007/s11227-024-06806-2).
- [24] J. S. Hwang, S. S. Lee, J. W. Gil, and C. K. Lee, "Determination of Optimal Batch Size of Deep

- Learning Models with Time Series Data," *Sustain.*, vol. 16, no. 14, pp. 1–11, 2024, doi: 10.3390/su16145936.
- [25] K. Ijaz, Z. Hussain, J. Ahmad, S. F. Ali, M. Adnan, and I. Khosa, "A Novel Temporal Feature Selection Based LSTM Model for Electrical Short-Term Load Forecasting," *IEEE Access*, vol. 10, no. July, pp. 82596–82613, 2022, doi: 10.1109/ACCESS.2022.3196476.
- [26] S. Mehtab and J. Sen, "Stock Price Prediction Using CNN and LSTM-Based Deep Learning Models," *2020 Int. Conf. Decis. Aid Sci. Appl. DASA 2020*, pp. 447–453, 2020, doi: 10.1109/DASA51403.2020.9317207.
- [27] A. Al Hamoud, A. Hoenig, and K. Roy, "Sentence subjectivity analysis of a political and ideological debate dataset using LSTM and BiLSTM with attention and GRU models," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 10, Part A, pp. 7974–7987, 2022, doi: <https://doi.org/10.1016/j.jksuci.2022.07.014>.
- [28] H. Sokooti, S. Yousefi, M. S. Elmahdy, B. P. F. Lelieveldt, and M. Staring, "Hierarchical Prediction of Registration Misalignment Using a Convolutional LSTM: Application to Chest CT Scans," *IEEE Access*, vol. 9, pp. 62008–62020, 2021, doi: 10.1109/ACCESS.2021.3074124.
- [29] X. Wang, S. Dong, and R. Zhang, "An Integrated Time Series Prediction Model Based on Empirical Mode Decomposition and Two Attention Mechanisms," *Inf.*, vol. 14, no. 11, 2023, doi: 10.3390/info14110610.
- [30] P. Meel and D. K. Vishwakarma, "A temporal ensembling based semi-supervised ConvNet for the detection of fake news articles," *Expert Syst. Appl.*, vol. 177, p. 115002, 2021, doi: <https://doi.org/10.1016/j.eswa.2021.115002>.
- [31] F. Wanyoike, K. M. Rich, N. Mtimet, S. Bahta, and L. Godiah, "An assessment of small ruminant production, marketing, and investment options in Somaliland: A system dynamics approach," *Small Rumin. Res.*, vol. 218, p. 106882, 2023, doi: <https://doi.org/10.1016/j.smallrumres.2022.106882>.
- [32] D. Zhao *et al.*, "The research of ARIMA, GM(1,1), and LSTM models for prediction of TB cases in China," *PLoS One*, vol. 17, no. 2, pp. e0262734–e0262734, 2022, doi: 10.1371/journal.pone.0262734.
- [33] I. Fuentes, J. Padarian, and R. W. Vervoort, "Towards near real-time national-scale soil water content monitoring using data fusion as a downscaling alternative," *J. Hydrol.*, vol. 609, p. 127705, 2022, doi: <https://doi.org/10.1016/j.jhydrol.2022.127705>.
- [34] D. Iskandaryan, F. Ramos, and S. Trilles, "Comparison of Nitrogen Dioxide Predictions During a Pandemic and Non-pandemic Scenario in the City of Madrid using a Convolutional LSTM Network," *Int. J. Comput. Intell. Appl.*, vol. 21, no. 02, pp. 2250014:1-2250014:11, 2022, doi: 10.1142/s1469026822500146.
- [35] H. Zhu *et al.*, "Study on the influence of meteorological factors on influenza in different regions and predictions based on an LSTM algorithm," *BMC Public Health*, vol. 22, no. 1, 2022, doi: 10.1186/s12889-022-14299-y.