



**PENERAPAN *TIME SERIES CLUSTERING* DALAM PERAMALAN
HARGA SAHAM SEKTOR KEUANGAN DI INDONESIA
MENGUNAKAN METODE LSTM**

SKRIPSI

untuk memenuhi persyaratan
dalam menyelesaikan program sarjana Strata-1 Statistika

Oleh
HAFIFAH CICI AMALIA
NIM. 2211017320011

**PROGRAM STUDI S-1 STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU
MARET 2026**



**PENERAPAN *TIME SERIES CLUSTERING* DALAM PERAMALAN
HARGA SAHAM SEKTOR KEUANGAN DI INDONESIA
MENGUNAKAN METODE LSTM**

SKRIPSI

untuk memenuhi persyaratan
dalam menyelesaikan program sarjana Strata-1 Statistika

Oleh
HAFIFAH CICI AMALIA
NIM. 22110173320011

**PROGRAM STUDI S-1 STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU
MARET 2026**

SKRIPSI

PENERAPAN *TIME SERIES CLUSTERING* DALAM PERAMALAN HARGA SAHAM SEKTOR KEUANGAN DI INDONESIA MENGUNAKAN METODE LSTM

Oleh
Hafifah Cici Amalia
NIM. 2211017320011

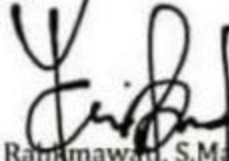
Telah dipertahankan pada hari Jumat, tanggal 13-03-2026 dan disetujui oleh dosen pembimbing dan dosen penguji sebagai berikut:

Pembimbing I



Dewi Sri Susanti, S.Si., M.Si.
NIP. 197305161999032002

Penguji I



Yeni Rahmawati, S.Mat., M.Si.
NIP. 199404032022032014

Pembimbing II



Hj. Maisarah, S.Pd.I., M.Pd.
NIP. 198907132025212053


Penguji II



Dewi Angraini, Ph.D.
NIP. 198303282005012001

Banjarbaru, 25 Maret 2026
Mengetahui,
Koordinator Program Studi Statistika
FMIPA ULM




Dewi Angraini, Ph.D.
NIP. 198303282005012001

PERNYATAAN

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Banjarbaru, 25 Maret 2026



Hafifah Cici Amalia
NIM. 2211017320011

PRODI STATISTIKA

ABSTRAK

Penerapan *Time Series Clustering* dalam Peramalan Harga Saham Sektor Keuangan Di Indonesia Menggunakan Metode LSTM (Oleh: Hafifah Cici Amalia; Pembimbing: Dewi Sri Susanti dan Maisarah, 2026; 76 halaman)

Clustering analysis merupakan metode analisis yang bertujuan mengelompokkan objek berdasarkan kesamaan karakteristiknya. *Time series clustering* merupakan pengembangan dari metode tersebut yang mengelompokkan data deret waktu berdasarkan kemiripan pola berubahannya sehingga setiap kluster merepresentasikan pola temporal yang serupa. Salah satu penerapannya diterapkan sebagai tahap pra-pemodelan dalam peramalan harga saham. Dengan banyaknya jumlah emiten, pendekatan ini memungkinkan peramalan dilakukan pada tingkat kluster sehingga lebih efisien dibandingkan pemodelan secara individual pada masing-masing saham. Oleh karena itu, penelitian ini menerapkan *time series clustering* dengan merepresentasikan setiap deret waktu menggunakan parameter *Autoregressive* orde satu sebagai input dalam algoritma *K-Error*, kemudian melakukan peramalan harga saham pada tingkat kluster untuk periode mendatang menggunakan LSTM. Hasil penelitian menunjukkan jumlah kluster optimal adalah sebanyak tiga dengan Klaster 1 memiliki harga yang relatif tinggi dan rentang pergerakan lebar, Klaster 2 memiliki harga yang menengah ke bawah dan pergerakan sempit, serta Klaster 3 didominasi oleh harga rendah dan pergerakan stabil dengan satu emiten berharga sangat tinggi. Kemudian, evaluasi peramalan menggunakan MAPE menunjukkan nilai pada data *training* dan *testing* seluruh kluster berada di bawah 10%, yang mengindikasikan kemampuan peramalan sangat baik. Hasil peramalan juga menunjukkan pola yang konsisten dengan karakteristik historis masing-masing kluster.

Kata Kunci: *Time Series Clustering*, *K-Error*, *Long Short-Term Memory*, *Harga Saham*.

ABSTRACT

Application of Time Series Clustering in Forecasting Financial Sector Stock Prices in Indonesia Using the LSTM Method (By: Hafifah Cici Amalia; Supervisors: Dewi Sri Susanti and Maisarah, 2026; 76 pages)

Clustering analysis is a statistical method used to group objects based on similarities in their characteristics. Time series clustering is an extension of this method that groups time series data according to similarities in their temporal patterns, so that each cluster represents a similar pattern of change over time. One of its applications is as a pre-modeling stage in stock price forecasting. With the large number of listed companies, this approach allows forecasting to be conducted at the cluster level, making the process more efficient compared to building individual models for each stock. Therefore, this study applies time series clustering by representing each time series using first-order Autoregressive (AR(1)) parameters as input to the K-Error algorithm. Stock price forecasting for future periods is then performed at the cluster level using the Long Short-Term Memory (LSTM) method. The results show that the optimal number of clusters is three. Cluster 1 consists of stocks with relatively high prices and wide movement ranges, Cluster 2 includes stocks with moderate to lower prices and narrower movements, and Cluster 3 is dominated by low-priced stocks with relatively stable movements, with one issuer having a very high price. Furthermore, the forecasting evaluation using Mean Absolute Percentage Error (MAPE) indicates that the MAPE values for both training and testing data in all clusters are below 10%, suggesting very good forecasting performance. The forecasting results also show patterns that are consistent with the historical characteristics of each cluster.

Keywords: Time Series Clustering, K-Error, Long Short-Term Memory, Stock Prices.

PRAKATA

Puji syukur atas kehadiran Allah Subhanahu Wa Ta'ala yang telah melimpahkan rahmat dan hidayah-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul "Penerapan *Time Series Clustering* dalam Peramalan Harga Saham Sektor Keuangan Di Indonesia Menggunakan Metode LSTM". Penyusunan skripsi ini untuk memenuhi salah satu syarat dalam rangka menyelesaikan program sarjana di Program Studi Statistika Fakultas MIPA Universitas Lambung Mangkurat.

Penulis menyadari bahwa dalam penyelesaian skripsi ini tidak lepas dari bantuan semua pihak. Oleh karena itu, penulis ingin mengucapkan terima kasih kepada:

1. Orang tua tercinta, Bapak Cipto Noermanto dan Ibu Astiyah, kakak saya Abdurrahman Halim, serta keluarga yang selalu memberikan dukungan, perhatian, do'a, dan semangat kepada penulis selama menempuh perkuliahan hingga penyelesaian skripsi ini.
2. Ibu Dewi Sri Susanti dan Ibu Hj. Maisarah selaku pembimbing tugas akhir yang telah bersedia memberikan bimbingan, arahan, serta motivasi selama proses penelitian dan penyusunan skripsi ini;
3. Ibu Yeni Rahkmawati dan Ibu Dewi Anggraini selaku penguji yang telah memberikan masukan, saran, dan bantuan dalam rangka perbaikan skripsi ini;
4. Bapak/Ibu dosen pengajar serta staf Program Studi Statistika FMIPA ULM yang telah memberikan ilmu dan bimbingan selama masa perkuliahan;
5. Sahabat "Asyik" dan Latifa Mutmainnah yang senantiasa memberikan semangat, dukungan, dan doa;
6. Seseorang yang telah menemani penulis selama setengah perjalanan perkuliahan, serta senantiasa memberikan dukungan, perhatian, semangat, dan doa hingga terselesaikannya skripsi ini;
7. Teman-teman angkatan 2022 Program Studi Statistika FMIPA ULM yang telah berjuang bersama dalam proses ini;
8. Semua pihak yang tidak bisa disebutkan satu per satu dalam tulisan ini.

Penulis menyadari bahwa penulisan skripsi ini masih terdapat banyak kekurangan. Oleh karena itu, penulis mengharapkan saran dan kritik yang membangun untuk menyempurnakan skripsi ini kedepannya. Akhir kata, penulis berharap skripsi ini dapat memberikan manfaat bagi pembaca dan bagi penulis sendiri.

Banjarbaru, 25 Maret 2026

Hafifah Cici Amalia

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PENGESAHAN	ii
PERNYATAAN	iii
ABSTRAK.....	iv
ABSTRACT	v
PRAKATA	vi
DAFTAR ISI.....	vii
DAFTAR GAMBAR	ix
DAFTAR TABEL	x
DAFTAR LAMPIRAN	xi
DAFTAR LAMBANG DAN SINGKATAN	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan Penelitian.....	3
1.4 Manfaat Penelitian.....	4
BAB II TINJAUAN PUSTAKA	5
2.1 Kajian Penelitian Terdahulu	5
2.2 Kajian Teori.....	7
2.2.1 Eksplorasi Data.....	7
2.2.2 <i>Clustering Analysis</i>	11
2.2.3 <i>Calinski-Harabasz Pseudo F-Statistic</i>	15
2.2.4 <i>Time Series Clustering Analysis</i>	16
2.2.5 <i>Feature Engineering</i>	17
2.2.6 <i>Time Series Split Cross Validation</i>	19
2.2.7 Normalisasi dan Denormalisasi Data.....	19
2.2.8 <i>Long Short-Term Memory (LSTM)</i>	20
2.2.9 <i>Adaptive Moment Estimation (ADAM)</i>	26
2.2.10 <i>Grid Search</i>	27
2.2.11 Evaluasi Performa Model Prediksi.....	27
2.2.12 Saham	28
2.2.13 Sektor Keuangan	29
BAB III METODE PENELITIAN	32
3.1 Sumber Data	32
3.2 Variabel Penelitian	32
3.3 Prosedur Penelitian	33
BAB IV HASIL DAN PEMBAHASAN	38
4.1 <i>Pre-processing</i> dan Eksplorasi Data.....	38
4.2 <i>Time Series Clustering Analysis</i>	42
4.2.1 Estimasi Parameter Model AR Orde 1	42
4.2.2 <i>Time Series Clustering</i> Menggunakan Algoritma <i>K-Error</i>	43
4.3 Identifikasi Hasil <i>Time Series Clustering Analysis</i>	48
4.4 Pelatihan Model LSTM Tingkat Kluster	52
4.4.1 Pelatihan Model LSTM pada Kluster 1	55

4.4.2 Pelatihan Model LSTM pada Klaster 2	63
4.4.3 Pelatihan Model LSTM pada Klaster 3	66
4.5 Evaluasi Model LSTM Tingkat Klaster	70
4.6 Peramalan Periode Mendatang	71
BAB V PENUTUP	75
5.1 Kesimpulan.....	75
5.2 Saran.....	75
DAFTAR PUSTAKA	77
LAMPIRAN.....	83

PRODI STATISTIKA

DAFTAR GAMBAR

Gambar 2.1 Visualisasi <i>Boxplot</i>	8
Gambar 2.2 Pola Horizontal	9
Gambar 2.3 Pola Musiman	10
Gambar 2.4 Pola Tren	10
Gambar 2.5 Pola Siklis	10
Gambar 2.6 Struktur Jaringan LSTM.....	21
Gambar 2.7 Struktur dalam Satu Sel LSTM.....	22
Gambar 2.8 Alur Informasi pada <i>Forget Gate</i>	22
Gambar 2.9 Alur Informasi pada <i>Input Gate</i>	23
Gambar 2.10 <i>Cell State</i>	24
Gambar 2.11 Alur Informasi pada <i>Output Gate</i>	25
Gambar 3.1 Diagram Alir Penelitian	37
Gambar 4.1 Visualisasi <i>Missing Value</i> Seluruh Data.....	38
Gambar 4.2 <i>Boxplot</i> Harga Penutupan Saham 41 Emiten Sektor Keuangan Periode Januari 2021 hingga Desember 2025	40
Gambar 4.3 <i>Time Series Plot</i> Harga Penutupan Saham 41 Emiten Sektor Keuangan Periode Januari 2021 hingga Desember 2025.....	41
Gambar 4.4 Nilai <i>Pseudo F Index</i> Algoritma <i>K-Error</i>	47
Gambar 4.5 <i>Time Series Plot</i> Nilai Residual Model AR(1) Masing-masing Emiten Setiap Klaster.....	49
Gambar 4.6 <i>Boxplot</i> Hasil Klaster Harga Penutupan Saham Masing-masing Emiten Setiap Klaster.....	51
Gambar 4.7 <i>Time Series Plot Prototype</i> dari Setiap Klaster	52
Gambar 4.8 Arsitektur Implementasi LSTM pada <i>Fold</i> 1 Klaster 1.....	61
Gambar 4.9 Grafik Perbandingan Prediksi dan Nilai Aktual Harga pada Data <i>Training</i> Klaster 1.....	62
Gambar 4.10 Grafik Perbandingan Prediksi dan Nilai Aktual Harga pada Data <i>Training</i> Klaster 2.....	66
Gambar 4.11 Grafik Perbandingan Prediksi dan Nilai Aktual Harga pada Data <i>Training</i> Klaster 3.....	70
Gambar 4.12 Pergerakan Data Historis dan Hasil Peramalan Harga Saham Klaster 1 Periode Ke-262 hingga Periode Ke-269	72
Gambar 4.13 Pergerakan Data Historis dan Hasil Peramalan Harga Saham Klaster 2 Periode Ke-262 hingga Periode Ke-269	73
Gambar 4.14 Pergerakan Data Historis dan Hasil Peramalan Harga Saham Klaster 3 Periode Ke-262 hingga Periode Ke-269	73

DAFTAR TABEL

Tabel 2.1 Kajian Penelitian Terdahulu	5
Tabel 2.2 Kriteria MAPE.....	28
Tabel 3.1 Data Kode Emiten Harga Saham Sektor Keuangan.....	32
Tabel 4.1 Statistika Deskriptif 41 Emiten Saham Sektor Keuangan	39
Tabel 4.2 Estimasi Parameter Koefisien Autoregresif dan Varian Residual....	42
Tabel 4.3 Keanggotaan Klaster Hasil Inisialisasi Awal	44
Tabel 4.4 Pusat Klaster pada Inisialisasi Awal Klaster	44
Tabel 4.5 Hasil Perhitungan Jarak <i>Mahalanobis</i> Iterasi Pertama.....	45
Tabel 4.6 Keanggotaan Klaster Iterasi Pertama	46
Tabel 4.7 Pusat Klaster Iterasi Pertama Hingga Iterasi Kelima.....	46
Tabel 4.8 Keanggotaan Klaster Iterasi Keempat	46
Tabel 4.9 Skema <i>Time Series Cross-Validation</i> untuk LSTM	54
Tabel 4.10 Konfigurasi Arsitektur dan <i>Hyperparameter</i> Model LSTM	54
Tabel 4.11 Dataset Mingguan dengan <i>Feature Engineering</i> pada Klaster 1 ...	56
Tabel 4.12 Hasil Normalisasi <i>Min-max</i> Data <i>Training Fold</i> 1 Klaster 1	57
Tabel 4.13 Hasil Normalisasi <i>Min-max</i> Data Validasi <i>Fold</i> 1 Klaster 1.....	57
Tabel 4.14 Hasil <i>Grid Search</i> Model LSTM Klaster 1	61
Tabel 4.15 Dataset Mingguan dengan <i>Feature Engineering</i> pada Klaster 2 ...	63
Tabel 4.16 Hasil Normalisasi <i>Min-max</i> Data <i>Training Fold</i> 1 Klaster 2	64
Tabel 4.17 Hasil Normalisasi <i>Min-max</i> Data Validasi <i>Fold</i> 1 Klaster 2.....	64
Tabel 4.18 Hasil <i>Grid Search</i> Model LSTM Klaster 2	65
Tabel 4.19 Dataset Mingguan dengan <i>Feature Engineering</i> pada Klaster 3 ...	67
Tabel 4.20 Hasil Normalisasi <i>Min-max</i> Data <i>Training Fold</i> 1 Klaster 3	68
Tabel 4.21 Hasil Normalisasi <i>Min-max</i> Data Validasi <i>Fold</i> 1 Klaster 3.....	68
Tabel 4.22 Hasil <i>Grid Search</i> Model LSTM Klaster 3	69
Tabel 4.23 Hasil Evaluasi Kinerja Model LSTM Berdasarkan Nilai MAPE pada Data <i>Training</i> dan <i>Testing</i> Tiap Klaster.....	71

DAFTAR LAMPIRAN

Lampiran 1. Data Harga Saham Mingguan Sektor Keuangan (Rupiah).....	83
Lampiran 2. Plot Data Deret Waktu Harga Saham Setiap Emiten Sektor Keuangan	84
Lampiran 3. Statistika Deskriptif Setiap Emiten Saham Sektor Keuangan.....	85
Lampiran 4. Ilustrasi Perhitungan <i>K-Error</i> Jarak <i>Mahalanobis</i>	86
Lampiran 5. <i>Feature Engineering</i> Masing-masing Klaster.....	87
Lampiran 6. Hasil Normalisasi Setiap <i>Fold</i> Masing-masing Klaster.....	88
Lampiran 7. Hasil Peramalan.....	93
Lampiran 8. <i>Syntax</i> RStudio (Analisis <i>Time Series Clustering</i>).....	94
Lampiran 9. <i>Syntax</i> Python (Analisis Lanjutan: Peramalan Tingkat Klaster)	100