

**ANALISIS FITUR EKSPONENTIAL MOVING AVERAGE  
DALAM TIME SERIES FORECASTING UNTUK  
MEMPREDIKSI KUNJUNGAN SITUS  
WEB DENGAN ALGORITMA  
XGBOOST**

**TESIS**

Disusun oleh: Jhiro Faran  
NIM: 227064518018



**PROGRAM STUDI S2 TEKNOLOGI INFORMASI  
FAKULTAS TEKNOLOGI KOMUNIKASI DAN INFORMATIKA  
UNIVERSITAS NASIONAL  
TAHUN 2**

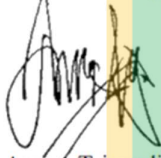
# PENGESAHAN

ANALISIS FITUR EKSPONENTIAL MOVING AVERAGE DALAM  
TIME SERIES FORECASTING UNTUK MEMPREDIKSI  
KUNJUNGAN SITUS WEB DENGAN  
ALGORITMA XGBOOST

Disusun Oleh:  
Jhiro Faran  
NIM: 227064518018

Tesis ini,  
Telah diperiksa dan disetujui oleh:

Dosen Pembimbing



Dr. Agung Triayudi, S.Kom., M.Kom  
NIDN. 0419068604

Ketua Program Studi



Ir. Asrul Sami, S.T., M.T., M.Kom., Ph.D  
NIDN. 0303067003

UNIVERSITAS NASIONAL

## PERNYATAAN ORISINALITAS

Saya menyatakan dengan sebenar-benarnya bahwa sepanjang pengetahuan saya, di dalam naskah Tesis ini tidak terdapat karya ilmiah yang pernah diajukan oleh orang lain untuk memperoleh gelar akademik di suatu perguruan tinggi, dan tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis disitasi dalam naskah ini dan disebutkan dalam daftar pustaka.

Apabila ternyata didalam naskah Tesis ini dapat dibuktikan terdapat unsur-unsur plagiasi, saya bersedia Tesis ini digugurkan dan gelar akademik yang telah saya peroleh (Magister) dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku (UU No. 20 Tahun 2003, Pasal 25 ayat 2 dan Pasal 70).



## KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa yang telah memberikan rahmat dan karunia sehingga penulis dapat menyelesaikan tesis dengan judul **“ANALISIS FITUR EKSPONENTIAL MOVING AVERAGE DALAM TIME SERIES FORECASTING UNTUK MEMPREDIKSI KUNJUNGAN SITUS WEB DENGAN ALGORITMA XGBOOST”** sebagai salah satu syarat kelulusan Program Studi Magister Ilmu Komputer, Fakultas Teknologi Informasi

Penelitian dan penulisan ini tidak terlepas dari bantuan berbagai pihak, oleh karena itu penulis menyampaikan banyak terima kasih terutama kepada dosen pembimbing Tugas Akhir Dr. (Nama Dosen) yang telah meluangkan banyak waktu, tenaga, pikiran, bimbingan, arahan, motivasi serta memaklumi segala kekurangan penulis selama penelitian tugas akhir dan penyusunan tesis. Penulis juga mengucapkan banyak terimakasih kepada:

1. Ayah dan ibu selaku orangtua dan kakak penulis yang telah banyak memberi dukungan dalam segala bentuk yang tak terhitung.
2. Rektor Universitas Nasional
3. Dekan Fakultas Teknologi Informasi:
4. Ketua Program Studi Magister Ilmu Komputer:
5. Seluruh dosen di Program Studi Magister Ilmu Komputer Fakultas Teknologi Informasi maupun dosen di Program Studi lain yang memberikan banyak ilmu.
6. Teman-teman seangkatan yang banyak membantu dan mendukung.
7. Teman-teman kantor yang memberikan banyak bantuan baik materil dan non-materil

Akhir kata, Semoga Tuhan Yang Maha Esa membalas kebaikan dan bantuan yang telah diberikan dengan hal yang lebih baik. Penulis mengharapkan kritik dan saran yang bersifat membangun dan semoga tesis ini memberikan manfaat di bidang Teknologi Informatika

Jakarta, 10 November 2022

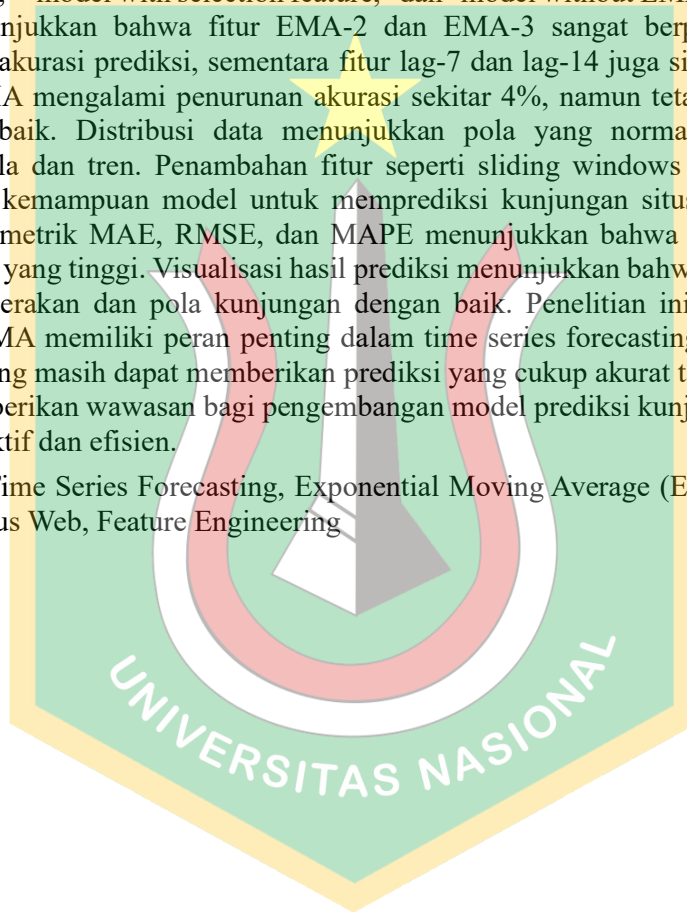


Jhiro Faran  
Jhiraf07@gmail.com

## ABSTRAK

Penelitian ini menganalisis peran Exponential Moving Average (EMA) dalam time series forecasting untuk memprediksi kunjungan situs web menggunakan algoritma XGBoost. Data yang digunakan terdiri dari 2167 baris dan 7 fitur, mencakup periode dari September 2014 hingga Agustus 2020. Tahapan penelitian meliputi eksplorasi data, persiapan data, dan pengembangan model machine learning, dengan pembagian data train (80%) dan data test (20%), serta scaling menggunakan min-max scaling. Tiga model dibangun: "model with all feature," "model with selection feature," dan "model without EMA feature." Hasil evaluasi menunjukkan bahwa fitur EMA-2 dan EMA-3 sangat berpengaruh dalam meningkatkan akurasi prediksi, sementara fitur lag-7 dan lag-14 juga signifikan. Model tanpa fitur EMA mengalami penurunan akurasi sekitar 4%, namun tetap menunjukkan kinerja yang baik. Distribusi data menunjukkan pola yang normal, memudahkan identifikasi pola dan tren. Penambahan fitur seperti sliding windows (lag) dan EMA meningkatkan kemampuan model untuk memprediksi kunjungan situs web. Evaluasi menggunakan metrik MAE, RMSE, dan MAPE menunjukkan bahwa model memiliki tingkat akurasi yang tinggi. Visualisasi hasil prediksi menunjukkan bahwa model mampu membaca pergerakan dan pola kunjungan dengan baik. Penelitian ini menyimpulkan bahwa fitur EMA memiliki peran penting dalam time series forecasting, namun model machine learning masih dapat memberikan prediksi yang cukup akurat tanpa fitur EMA. Hasil ini memberikan wawasan bagi pengembangan model prediksi kunjungan situs web yang lebih efektif dan efisien.

**Kata kunci:** Time Series Forecasting, Exponential Moving Average (EMA), XGBoost, Kunjungan Situs Web, Feature Engineering



# DAFTAR ISI

HALAMAN JUDUL .....	i
PENGESAHAN .....	ii
PERNYATAAN ORISINALITAS .....	iii
KATA PENGANTAR .....	iv
ABSTRAK .....	v
1 PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Identifikasi Masalah .....	4
1.3 Rumusan Masalah .....	4
1.4 Tujuan .....	4
1.5 Manfaat .....	4
1.6 Batasan Masalah .....	5
1.7 Sistematika Pembahasan .....	5
2 LANDASAN PERPUSTAKAAN .....	6
2.1 Data Mining .....	6
2.2 <i>Exploratory Data Analysis</i> (EDA) .....	7
2.3 Time-Series Forecasting .....	8
2.4 <i>Sliding Window</i> .....	8
2.5 <i>Exponential Moving Average</i> (EMA) .....	9
2.6 <i>Data Preprocessing</i> .....	9
2.6.1 <i>Outlier</i> .....	10
2.6.2 <i>Data Scaling</i> .....	10
2.6.3 Uji Korelasi .....	11
2.7 <i>Machine Learning</i> (ML) .....	12
2.7.1 XGBoost .....	13
2.7.2 Evaluasi ML .....	14
2.8 Penelitian Terkait .....	14
2.9 Tinjauan Obyek Penelitian .....	19
2.10 Kerangka Konseptual .....	20
2.11 Hipotesis .....	20
3 METODOLOGI PENELITIAN .....	21

3.1	Metode Penelitian .....	21
3.2	Tahapan Penelitian.....	21
3.3	Desain Sistem .....	23
3.3.1	Data Cleansing.....	23
3.3.2	Feature Engineering.....	24
3.3.3	Uji Korelasi .....	25
3.3.4	Handling Outlier .....	25
3.3.5	<i>Scaling Feature</i> .....	26
3.3.6	Modeling.....	26
3.3.7	Evaluasi ML dan Fitur Analisis .....	26
3.4	Perancangan Penelitian .....	27
3.5	Perangkat Sistem.....	28
4	HASIL PENELITIAN .....	29
4.1	Dataset .....	29
4.1.1	Tipe Data .....	29
4.1.2	Visualisasi Data Pergerakan Nilai .....	30
4.1.3	Distribusi Dataset Pada Setiap Fitur .....	31
4.2	Data Preparation .....	31
4.2.1	Total Kunjungan .....	32
4.2.2	Penyebaran Total Kunjungan.....	33
4.2.3	Rata-rata Kunjungan Harian.....	33
4.2.4	Penambahan Fitur <i>Sliding Window (Lag Feature)</i> .....	34
4.2.5	Penambahan Fitur <i>Exponential Moving Average (EMA)</i> .....	35
4.2.6	Manipulasi Data <i>Date</i> .....	35
4.3	<i>Exploratory Data Analysis (EDA)</i> .....	36
4.3.1	Korelasi Variabel <i>Pearson</i> .....	37
4.3.2	Korelasi Variabel Spearman .....	38
4.3.3	Distribusi Data Total Kunjungan, EMA, dan <i>Sliding Window</i> .....	39
4.3.4	Visualisasi Kuartal Setiap Tahun .....	44
4.3.5	Relasi Fitur <i>EMA</i> dan <i>Sliding Windows</i> Terhadap Target.....	45
4.4	Data Modeling .....	47
4.4.1	Scaling Data.....	47
4.4.2	Split Dataset.....	48
4.4.3	Model Satu ( <i>All Feature</i> ).....	49

4.4.4	Model Dua ( <i>Feature Selection</i> ) .....	52
4.4.5	Model Tiga ( <i>without EMA feature</i> ) .....	56
5	PEMBAHASAN PENELITIAN .....	60
5.1	Korelasi Hasil dan Tujuan Penelitian.....	60
5.2	Hasil Keseluruhan Dari Seluruh Model.....	60
5.3	Refleksi Penelitian .....	61
6	PENUTUP .....	62
6.1	Kesimpulan .....	62
6.2	Saran .....	62
	Daftar Pustaka .....	63





## DAFTAR GAMBAR

gambar 2.1 data mining .....	6
gambar 2.2 kerangka konseptual .....	20
gambar 3.1 langkah penelitian.....	21
gambar 3.2 desain sistem.....	23
gambar 3.3 feature engineering .....	24
gambar 3.4 uji korelasi fitur .....	25
gambar 4.1 visualisasi data dengan date.....	30
gambar 4.2 distribusi fitur menggunakan boxplot.....	31
gambar 4.3 total kunjungan website .....	32
gambar 4.4 penyebaran distribusi total data menggunakan boxplot .....	33
gambar 4.5 rata-rata kunjungan berdasarkan hari .....	33
gambar 4.6 visualisasi pergerakan data setiap satu tahunan.....	36
gambar 4.7 korelasi antar variabel pearson .....	37
gambar 4.8 korelasi antar variabel spearman .....	38
gambar 4.9 distribusi data total kunjungan website .....	39
gambar 4.10 distribusi data fitur lag 1 .....	39
gambar 4.11 distribusi data fitur lag 7 .....	40
gambar 4.12 distribusi data fitur lag 14.....	41
gambar 4.13 distribusi data fitur ema 2 .....	41
gambar 4.14 distribusi data fitur ema 3 .....	42
gambar 4.15 distrbusi data fitur ema 4 .....	43
gambar 4.16 distribusi data fitur ema 5 .....	43
gambar 4.17 visualisasi kunjungan pada setiap tahun berdasarkan kuartal .....	44
gambar 4.18 distribusi ema dan total kunjungan.....	45
gambar 4.19 distribusi sliding windows dan total kunjungan.....	46
gambar 4.20 data train test.....	48
gambar 4.21 visualisasi model satu .....	50
gambar 4.22 feature importance model satu.....	51
gambar 4.23 visualisasi model 2 .....	54
gambar 4.24 feature importance model 2 .....	55
gambar 4.25 visualisasi model 3 .....	57
gambar 4.26 feature importance model 3 .....	58

## DAFTAR TABEL

Tabel 2.1 Penelitian Terkait .....	16
Tabel 3.1 Rancangan penelitian .....	27
Tabel 3.2 Perangkat Sistem .....	28
Tabel 4.1 data kunjungan website.....	29
Tabel 4.2 tipe data pada data .....	29
Tabel 4.3 penambahan fitur lag .....	34
Tabel 4.4 penambahan fitur EMA .....	35
Tabel 4.5 manipulasi data date .....	35
Tabel 4.6 scaling data dengan min-max .....	47
Tabel 4.7 split dataset .....	48
Tabel 4.8 dataset model satu.....	49
Tabel 4.9 parameter XGBoost .....	49
Tabel 4.10 evaluasi model 1 .....	50
Tabel 4.11 perbandingan antara nilai aktual dan prediksi model satu.....	52
Tabel 4.12 dataset model 2 .....	52
Tabel 4.13 parameter XGBoost .....	53
Tabel 4.14 evaluasi model 2 .....	53
Tabel 4.15 perbandingan nilai aktual dan prediksi model dua .....	55
Tabel 4.16 dataset model 3 .....	56
Tabel 4.17 parameter XGboost.....	56
Tabel 4.18 evaluasi model 3 .....	57
Tabel 4.19 perbandingan nilai aktual dan prediksi model 3 .....	59
Tabel 5.1 hasil keseluruhan model .....	60



## DAFTAR RUMUS

2.1 sliding windows.....	8
2.2 exponential moving average.....	9
2.3 Interquartile relative .....	10
2.4 Batas toleransi .....	10
2.5 Standarization scaling.....	10
2.6 Normalizazion scaling.....	11
2.7 Pearson .....	11
2.8 Spearman .....	11
2.9 Spearman conditional .....	12
2.10 Leaf node .....	13
2.11 Objective theta .....	13
2.12 Root mean square error.....	14
2.13 Mean absolute error .....	14
2.14 Mean absoluter percentage error .....	14

