



**PERBANDINGAN ALGORITMA *EXTREME GRADIENT BOOSTING*
DAN *RANDOM FOREST* UNTUK MEMPREDIKSI HARGA TERENDAH
SAHAM DENGAN INDEX ISSI**

SKRIPSI

**FRANSISCO READY PERMANA
NIM. 1910511045**

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN” JAKARTA
2023**



**PERBANDINGAN ALGORITMA *EXTREME GRADIENT BOOSTING* DAN
RANDOM FOREST UNTUK MEMPREDIKSI HARGA TERENDAH SAHAM
DENGAN INDEX ISSI**

SKRIPSI

**Diajukan sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana
Komputer**

**FRANSISCO READY PERMANA
NIM. 1910511045**

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN” JAKARTA
2023**

LEMBAR PERSETUJUAN

Yang bertanda tangan di bawah ini:

Nama : Fransisco Ready Permana

NIM. : 1910511045

Program Studi : Informatika/~~Sistem Informasi Program Sarjana/Diploma 3~~
(*Coret yang tidak perlu)

Judul Skripsi/TA. : Perbandingan Algoritma Extreme Gradient Boosting dan
Random Forest Untuk Memprediksi Harga Terendah Saham
Dengan Index ISSI

Dinyatakan telah memenuhi syarat dan menyetujui untuk mengikuti ujian sidang skripsi.

Jakarta, 1 Juli 2023

Mengetahui


Ketua Program Studi,



Dr. Widya Cholil, M.I.T

Menyetujui,

Dosen Pembimbing



Iin Ernawati S.Kom., M.Si

PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya sendiri dan semua sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar.

Nama : Fransisco Ready Permana

NIM : 1910511045

Tanggal : 27 Juni 2023

Judul Skripsi : **Perbandingan Algoritma Extreme Gradient Boosting dan Random Forest Untuk Memprediksi Harga Terendah Saham Dengan Index ISSI**

Bilamana pada kemudian hari ditemukan ketidaksesuaian dengan pernyataan saya ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 27 Juni 2023

Yang Menyatakan,



Fransisco Ready Permana

PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademika Universitas Pembangunan Nasional “Veteran” Jakarta, saya yang bertanda tangan di bawah ini:

Nama : Fransisco Ready Permana
NIM : 1910511045
Fakultas : Ilmu Komputer
Program Studi : Informatika

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan karya ilmiah saya kepada Universitas Pembangunan Nasional “Veteran” Jakarta Hak Bebas Royalti Non-Eksklusif (*Non-Exchange Royalty Free Right*) untuk dipublikasikan dengan judul:

Perbandingan Algoritma Extreme Gradient Boosting dan Random Forest Untuk Memprediksi Harga Terendah Saham Dengan Index ISSI

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional “Veteran” Jakarta berhak menyimpan, mengalih media atau memformatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan skripsi saya selama tetap mencantumkan nama saya sebagai penulis atau pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta
Pada tanggal : 27 Juni 2023

Yang Menyatakan,



Fransisco Ready Permana

LEMBAR PENGESAHAN

Tugas Akhir ini diajukan oleh:

Nama : Fransisco Ready Permana
NIM : 1910511045
Program Studi : S1 Informatika
Judul Tugas Akhir : Perbandingan Algoritma Extreme Gradient Boosting dan Random Forest Untuk Memprediksi Harga Terendah Saham Dengan Index ISSI

Telah berhasil dipertahankan dihadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.



Dr. Widya Cholil, M.I.T..

Penguji I



Yuni Widiastiwi, S.Kom, M.Si.

Penguji II



Iin Ernawati S.Kom., M.Si

Pembimbing



Dr. Ermatita, M.Kom.

Dekan



Dr. Widya Cholil, M.I.T

Kepala Program Studi

Ditetapkan di : Jakarta

Tanggal Ujian : Selasa, 27 Juni 2023



**PERBANDINGAN ALGORITMA *EXTREME GRADIENT BOOSTING*
DAN *RANDOM FOREST* UNTUK MEMPREDIKSI HARGA TERENDAH
SAHAM DENGAN INDEX ISSI**

FRANSISCO READY PERMANA

ABSTRAK

Saham dengan indeks ISSI (Indeks Saham Syariah Indonesia) adalah saham yang bisa dijadikan pilihan untuk berinvestasi karena saham ini memiliki tingkat stabilitas yang cukup baik dibandingkan dengan indeks saham lainnya. Maka dari itu, penelitian ini ingin membuat sebuah model *machine learning* yang bisa memprediksi harga terendah saham ISSI sebagai nilai ambang bawah serta membandingkan dua algoritma andal yaitu algoritma *random forest* dan *extreme gradient boosting* (XGBoost) menggunakan data saham yang diambil dari situs *google finance*. Tahapannya meliputi identifikasi masalah, studi literatur, menyiapkan data, *load* dataset, *exploratory data analysis*, *preprocessing*, pembagian data, pelatihan data dan evaluasi model. Untuk mengetahui algoritma mana yang lebih baik, kedua algoritma tersebut dibandingkan menggunakan tiga metrik penilaian seperti *Mean Squared Error* (MSE), *Mean Absolute Percentage Error* (MAPE), dan R^2 . Hasilnya, didapatkan nilai rata-rata MSE uji *random forest* sebesar 0.6458 dan nilai rata-rata MSE uji XGBoost sebesar 0.8019; Nilai rata-rata MAPE uji *random forest* sebesar 0.0033 dan nilai rata-rata MAPE uji XGBoost sebesar 0.0037; Nilai rata-rata R^2 uji *random forest* ada diangka 0.9985 dan nilai rata-rata R^2 uji ada diangka 0.9982. Dari nilai tersebut bisa dilihat bahwa *random forest* memberikan nilai MSE dan MAPE yang lebih kecil serta nilai R^2 yang lebih besar dibandingkan XGBoost. Sehingga berdasarkan ketiga metrik penilaian tersebut, dapat disimpulkan bahwa, *random forest* dapat memprediksi harga terendah saham dengan indeks ISSI lebih akurat dibandingkan dengan XGBoost.

Kata Kunci: Saham, ISSI, *Random Forest*, *Extreme Gradient Boosting*.

**COMPARISON OF EXTREME GRADIENT BOOSTING AND RANDOM
FOREST ALGORITHMS FOR PREDICTING THE LOWEST PRICE OF
STOCK WITH THE ISSI INDEX**

FRANSISCO READY PERMANA

ABSTRACT

Stocks with the ISSI index (Indonesian Sharia Stock Index) are stocks that can be used as an investment choice because these stocks have a fairly good level of stability compared to other stock indices. Therefore, this research wants to create a machine learning model that can predict the lowest price of ISSI shares as a lower threshold value and compare two reliable algorithms, namely the random forest algorithm and extreme gradient boosting (XGBoost) using stock data taken from the Google Finance website. The stages include problem identification, literature study, data preparation, dataset loading, exploratory data analysis, preprocessing, data sharing, data training, and model evaluation. To find out which algorithm is better, the two algorithms are compared using three assessment metrics such as Mean Squared Error (MSE), Mean Absolute Percentage Error (MAPE), and R2. As a result, the average value of the MSE random forest test was 0.6458 and the average MSE value for the XGBoost test was 0.8019; The average value of the MAPE random forest test was 0.0033 and the average value of the MAPE XGBoost test was 0.0037; The average R2 value of the random forest test is 0.9985 and the average R2 test value is 0.9982. From these values, the random forest gives smaller MSE and MAPE values and a larger R2 value compared to XGBoost. So based on these three assessment metrics, it can be concluded that random forest can predict the lowest price of stocks with the ISSI index more accurately than XGBoost.

Keywords: *Stocks, ISSI, Random Forest, Extreme Gradient Boosting.*

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas segala karunia-Nya sehingga skripsi ini berhasil diselesaikan dengan baik dan tepat waktu. Judul yang dipilih dalam penelitian yang dilaksanakan sejak Oktober 2022 ini adalah “Perbandingan Algoritma *Extreme Gradient Boosting* dan *Random Forest* Untuk Memprediksi Harga Terendah Saham Dengan Index ISSI”. Penulisan skripsi ini bertujuan untuk memenuhi salah satu persyaratan untuk memperoleh gelar Sarjana Komputer Jurusan Informatika.

Peneliti menyadari dalam penyusunan laporan skripsi ini tidak akan selesai tanpa bantuan dari berbagai pihak. Oleh karena itu, pada kesempatan ini peneliti ingin mengucapkan terimakasih kepada:

1. Orang tua dan keluarga peneliti, yang telah memberikan dukungan dan doa untuk kesuksesan peneliti dalam menyelesaikan skripsi.
2. Ibu Iin Ernawati, S.Kom., M.Si selaku dosen pembimbing skripsi yang membantu peneliti dalam penyusunan skripsi dan memberikan saran dengan baik.
3. Ibu Dr. Widya Cholil, M.I.T. selaku Kepala Program Studi Informatika.
4. Ibu Dr. Ermatita, M.Kom selaku Dekan Fakultas Ilmu Komputer.
5. Seluruh jajaran Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jakarta yang telah membantu dalam perizinan dan administrasi.
6. Sahabat-sahabat perjuangan peneliti yaitu Kemmy, Berli, Jovan yang selalu membantu dan memberikan masukan serta doa kepada peneliti agar terselesaikannya skripsi ini dengan baik.
7. Teman-teman Program Studi Informatika Angkatan 2019 Fakultas Ilmu Komputer yang telah memberikan dukungan dan doa satu sama lain.
8. Seluruh pihak yang terlibat dalam kelancaran pembuatan skripsi ini dan yang belum disebutkan di atas, peneliti ucapkan terimakasih.

Peneliti menyadari bahwa penyusunan skripsi ini masih jauh dari sempurna. Oleh karena itu, peneliti mengharapkan kritik dan saran yang bersifat membangun untuk kesempurnaan skripsi ini yang dapat dikirimkan pada email peneliti yaitu fransiscorp@upnvj.ac.id. Akhir kata peneliti berharap semoga skripsi ini dapat bermanfaat bagi semua pihak.

Jakarta, 3 Juni 2022

A handwritten signature in black ink, consisting of several overlapping, stylized strokes that form a unique, cursive-like mark.

Fransisco Ready Permana

DAFTAR ISI

	Halaman
HALAMAN JUDUL	ii
LEMBAR PERSETUJUAN	iii
PERNYATAAN ORISINALITAS	iv
PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS	v
LEMBAR PENGESAHAN	vi
ABSTRAK.....	iii
ABSTRACT.....	iv
KATA PENGANTAR	v
DAFTAR ISI.....	vii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xiii
DAFTAR LAMPIRAN.....	xv
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Ruang Lingkup.....	3
1.5 Manfaat Penelitian	4
1.6 Luaran Yang Diharapkan	4
1.7 Sistematika Penulisan	4
BAB 2 TINJAUAN PUSTAKA	6
2.1 Saham.....	6

2.1.1	Harga Saham	6
2.1.2	Indeks Saham	7
2.2	Indeks Saham Syariah Indonesia (ISSI).....	7
2.3	<i>Data mining</i>	8
2.3.1	Fungsi <i>Data mining</i>	8
2.3.2	Ruang Lingkup <i>Data mining</i>	8
2.3.3	Proses <i>Data Mining</i>	10
2.4	<i>Exploratory Data Analysis (EDA)</i>	12
2.5	<i>Forecasting</i>	15
2.6	<i>Forecasting Time Series</i>	16
2.7	<i>Machine learning</i>	17
2.8	<i>Supervised Learning</i>	18
2.9	Regresi.....	18
2.10	<i>Ensemble learning</i>	18
2.11	<i>Bootstrap Aggregating (Bagging)</i>	19
2.12	<i>Boosting</i>	20
2.13	Pohon Keputusan (<i>Decision Tree</i>)	20
2.14	<i>Random Forest</i>	21
2.15	<i>Extreme Gradient Boosting (XGBoost)</i>	22
2.15.1	Fungsi Objektif.....	23
2.15.2	<i>Decision Tree Ensemble</i>	24
2.15.3	<i>Feature Importance</i>	27
2.16	<i>K-Fold Cross Validation (CV)</i>	28
2.17	R^2 , MSE dan MAPE.....	29
2.18	<i>Hyperparameter</i>	30

2.19	<i>Overfitting</i>	32
2.20	<i>Review Penelitian Terdahulu</i>	33
BAB 3 METODOLOGI PENELITIAN		35
3.1	Tahapan Penelitian	35
3.2	Identifikasi Masalah	35
3.3	Studi Literatur	36
3.4	Menyiapkan Data	36
3.5	<i>Load Dataset</i>	36
3.6	<i>Exploratory Data Analysis (EDA)</i>	37
3.6.1	Melihat Statistik Deskriptif	37
3.6.2	Melihat Bentuk Persebaran Data	37
3.7	<i>Preprocessing</i>	37
3.7.1	<i>Data Cleaning</i>	38
3.7.2	<i>Data Integration</i>	38
3.7.3	<i>Data Selection</i>	38
3.7.4	<i>Data Transformation</i>	38
3.8	<i>Data mining</i>	38
3.8.1	<i>Data Splitting</i>	39
3.8.2	Pelatihan Data.....	39
3.8.3	Memprediksi Data	39
3.8.4	Melihat Hasil Prediksi Model.....	40
3.9	Evaluasi	40
3.10	Kesimpulan	40
3.11	Waktu dan Tempat Penelitian	40
3.12	Alat Bantu Penelitian	41

3.12.1	<i>Hardware</i>	41
3.12.2	<i>Software</i>	41
3.13	Jadwal Penelitian.....	41
BAB 4 HASIL DAN PEMBAHASAN		43
4.1	Data	43
4.2	<i>Preprocessing</i>	50
4.2.1	<i>Data Cleaning</i>	51
4.2.2	<i>Data Integration</i>	51
4.2.3	<i>Data Selection</i>	52
4.2.4	<i>Data Transformation</i>	53
4.3	<i>Data mining</i>	53
4.3.1	Pembagian Data / <i>Dataset splitting</i>	54
4.3.2	Pelatihan Data.....	57
4.3.3	Memprediksi Data	74
4.3.4	Melihat Hasil Prediksi Model.....	78
4.4	Evaluasi	86
BAB 5 PENUTUP		93
5.1	Kesimpulan	93
5.2	Saran.....	94
DAFTAR PUSTAKA		95
RIWAYAT HIDUP		100
LAMPIRAN.....		101
LAMPIRAN 1. TURNITIN.....		101

DAFTAR TABEL

	Halaman
Tabel 2.1 Daftar <i>Hyperparameter</i>	30
Tabel 3.1 Jadwal Penelitian	41
Tabel 4.1 Data Saham ISSI (data asli).....	44
Tabel 4.2 Data Saham ISSI Setelah Kolom <i>Date</i> Menjadi <i>Index</i>	44
Tabel 4.3 Statistik Deskriptif Data.....	45
Tabel 4.4 Data Saham ISSI yang Diurutkan Berdasarkan Kolom <i>Low</i>	47
Tabel 4.5 Informasi <i>Missing Value</i> Pada Data.....	51
Tabel 4.6 Data Setelah Kolom <i>Volume</i> Dihapus	52
Tabel 4.7 Data latih Satu.....	54
Tabel 4.8 Data Uji Satu.....	55
Tabel 4.9 Data Latih Dua.....	55
Tabel 4.10 Data Uji Dua	56
Tabel 4.11 Data Latih Tiga	56
Tabel 4.12 Data Uji Tiga	57
Tabel 4.13 Lima Baris Data Pertama Data Untuk Contoh Pemrosesan	59
Tabel 4.14 Contoh Bostraped Data dari Model <i>Random Forest</i>	60
Tabel 4.15 Nilai <i>Threshold Root Node</i>	61
Tabel 4.16 Nilai <i>Threshold</i> yang Memenuhi Kondisi	64
Tabel 4.17 Nilai <i>Threshold</i> yang Tidak Memenuhi Kondisi	64
Tabel 4.18 Perhitungan Residual	69
Tabel 4.19 Penentuan <i>Threshold</i> XGBoost	69
Tabel 4.20 Contoh Lima Data Hasil Prediksi Model.....	74
Tabel 4.21 Nilai <i>Output</i> dari Setiap Pohon Keputusan <i>Random Forest</i>	75

Tabel 4.22 Nilai <i>Output</i> dari Setiap Pohon Keputusan <i>XGBoost</i>	77
Tabel 4.23 Nilai Metrik Penilaian Model	88
Tabel 4.24 Penyederhanaan Nilai Metrik Penilaian Berdasarkan Tabel 4.21	91

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Contoh Anomali	9
Gambar 2.2 Diagram KDD	10
Gambar 2.3 Contoh Histogram	13
Gambar 2.5 Contoh <i>Box Plot</i>	14
Gambar 2.6 Contoh <i>Violin Plot</i>	14
Gambar 2.7 Contoh <i>Heatmap</i>	15
Gambar 2.8 Diagram Dekomposisi <i>Data Time Series</i>	16
Gambar 2.9 <i>Family Tree</i>	24
Gambar 2.10 Ilustrasi <i>K-Fold Cross Validation</i>	28
Gambar 3.1 Alur Penelitian	35
Gambar 4.1 Fluktuasi Harga Saham Indeks ISSI	46
Gambar 4.2 Fluktuasi Harga Saham Indeks ISSI Berdasarkan Tanggal	48
Gambar 4.3 Fluktuasi Harga Saham Indeks ISSI Berdasarkan Bulan	48
Gambar 4.4 Histogram Bentuk Persebaran Data	49
Gambar 4.5 Heatmap Korelasi Variabel pada Data	50
Gambar 4.6 Contoh pohon keputusan berdasarkan tabel 4.14	65
Gambar 4.7 Pohon Keputusan XGBoost Pertama Berdasarkan Tabel 4.13	72
Gambar 4.7 Visualisasi Hasil Perdiksi Data Latih <i>Random Forest</i> Pertama	79
Gambar 4.8 Visualisasi Hasil Perdiksi Data Uji <i>Random Forest</i> Pertama	79
Gambar 4.9 Visualisasi Hasil Perdiksi Data Latih <i>Random Forest</i> Kedua	80
Gambar 4.10 Visualisasi Hasil Perdiksi Data Uji <i>Random Forest</i> Kedua	80
Gambar 4.11 Visualisasi Hasil Perdiksi Data Latih <i>Random Forest</i> Ketiga	81
Gambar 4.12 Visualisasi Hasil Perdiksi Data Uji <i>Random Forest</i> Ketiga	81

Gambar 4.13 Visualisasi Hasil Perdiksi Data Latih XGBoost Pertama	83
Gambar 4.14 Visualisasi Hasil Perdiksi Data Uji XGBoost Pertama.....	83
Gambar 4.15 Visualisasi Hasil Perdiksi Data Latih XGBoost Kedua	84
Gambar 4.16 Visualisasi Hasil Perdiksi Data Uji XGBoost Kedua	84
Gambar 4.17 Visualisasi Hasil Perdiksi Data Latih XGBoost Ketiga.....	85
Gambar 4.18 Visualisasi Hasil Perdiksi Data Uji XGBoost Ketiga	85

DAFTAR LAMPIRAN

	Halaman
LAMPIRAN 1. TURNITIN.....	101