



## **PRA-SKRIPSI**

# **Analisis Volatilitas Saham Initial Public Offering (IPO) Berdasarkan Data Prospektus Menggunakan Algoritma Extreme Gradient Boosting (XGBoost)**

**EKA MAURITA**

NPM 22081010119

## **DOSEN PEMBIMBING**

None

None

**KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI  
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR  
FAKULTAS ILMU KOMPUTER  
PROGRAM STUDI INFORMATIKA  
SURABAYA  
2025**

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

Pasar modal memegang peranan penting dalam perekonomian suatu negara termasuk Indonesia. Perkembangan kegiatan transaksi dalam bursa yang ditunjukkan oleh suatu indeks dapat menjadi ukuran bagi kondisi perekonomian suatu negara secara keseluruhan (Fitria Puteri Sholikah et al., 2022). Salah satu instrumen penting dalam pasar modal adalah Initial Public Offering (IPO). IPO Adalah penawaran umum saham yang ditawarkan oleh Perusahaan kepada Masyarakat untuk pertama kali, Perusahaan yang mencatatkan saham merupakan Perusahaan yang berkembang dalam mendapatkan dana untuk mengembangkan usaha (Sugiyanto et al., 2023).

Melalui kegiatan IPO, perusahaan dapat memperoleh tambahan modal, sedangkan investor memperoleh kesempatan untuk berinvestasi pada perusahaan yang baru melantai di bursa. Namun, saham yang baru melantai di pasar sering menunjukkan volatilitas harga yang tinggi, yang mencerminkan ketidakpastian dan risiko pasar yang terkait dengan saham IPO (Nugroho & Dewi, 2024). Fenomena ini disebabkan oleh berbagai faktor, termasuk lingkungan ekonomi makro, persepsi pasar terhadap prospektus perusahaan, dan psikologi investor di awal pembukaan pasar saham.

Prospektus adalah dokumen tertulis yang memuat informasi lengkap terkait proses penawaran umum efek, yang disusun untuk memberikan informasi kepada calon investor dengan tujuan mendorong mereka membeli efek yang ditawarkan (Hendrasyah, 2025). Investor menggunakan informasi dalam prospektus sebagai panduan untuk menilai risiko dan potensi imbal hasil sekuritas yang akan ditawarkan. Oleh karena itu, analisis data prospektif perusahaan dapat memberikan informasi tentang faktor-faktor yang memengaruhi volatilitas harga saham setelah IPO, serta memberikan kontribusi pada pengembangan model prediksi risiko investasi berbasis data terstruktur.

Dalam beberapa tahun terakhir, pendekatan berbasis *machine learning* banyak digunakan untuk menganalisis data keuangan termasuk memprediksi harga saham atau volatilitas. Salah satu algoritma *machine learning* yang memiliki performa tinggi dan sering digunakan dalam penelitian keuangan adalah *Extreme Gradient Boosting (XGBoost)*. Extreme Gradient Boosting (XGBoost), dipilih untuk memodel hubungan kompleks antara variabel-variabel prospektus dan volatilitas harga saham pada periode awal listing. XGBoost memiliki keunggulan dalam menangani hubungan nonlinier, interaksi antar variabel, serta missing

values, sehingga berpotensi menghasilkan pemahaman yang lebih akurat mengenai faktor-faktor yang mendorong fluktuasi harga IPO di tahap awal perdagangan. Dengan demikian, penelitian ini bertujuan menguji dan mengarahkan penggunaan data prospektus sebagai sumber informasi utama untuk memprediksi volatilitas saham IPO melalui algoritma XGBoost, serta mengidentifikasi variabel-variabel kunci yang berkontribusi pada volatilitas tersebut. Kontribusi penelitian ini meliputi peningkatan literatur keuangan berbasis data, dukungan praktik manajemen risiko IPO bagi investor dan emiten, serta metode analisis berbasis data yang dapat direplikasi pada konteks pasar modal lain.

## **1.2. Rumusan Masalah**

Berdasarkan latar belakang yang telah dijelaskan sebelumnya, maka perlu dirumuskan beberapa permasalahan yang akan menjadi fokus dalam penelitian ini. Berikut merupakan rumusan masalah pada penelitian ini :

1. Bagaimana pola volatilitas saham IPO selama 7 hari pertama setelah listing?
2. Variabel prospektus mana yang paling berpengaruh terhadap volatilitas 7 hari awal?
3. Seberapa akurat algoritma XGBoost dalam memprediksi volatilitas 7 hari awal dibandingkan baseline?

## **1.3. Tujuan Penelitian**

Berdasarkan rumusan masalah tersebut, maka tujuan dari penelitian ini adalah sebagai berikut :

1. Menganalisis pola volatilitas saham IPO selama 7 hari pertama setelah listing di Bursa Efek Indonesia untuk mengetahui karakteristik fluktuasi harga pada periode awal perdagangan. Membangun dan mengevaluasi model prediksi kinerja harga saham IPO menggunakan algoritma Extreme Gradient Boosting (XGBoost) berdasarkan data prospektus, untuk mengukur kemampuan model dalam memprediksi nilai return atau arah pergerakan harga saham IPO secara akurat.
2. Mengidentifikasi variabel-variabel dalam prospektus yang paling berpengaruh terhadap volatilitas harga saham selama 7 hari pertama pasca-IPO, sehingga dapat memberikan pemahaman mengenai faktor fundamental yang memicu ketidakstabilan harga.
3. Menerapkan dan mengevaluasi kinerja algoritma Extreme Gradient Boosting (XGBoost) dalam memprediksi volatilitas saham IPO selama 7 hari awal, serta membandingkan tingkat akurasinya dengan model baseline (misalnya regresi linier atau model statistik sederhana).

#### **1.4. Manfaat Penelitian**

Penelitian ini diharapkan dapat memberikan manfaat baik dari sisi teoritis maupun praktis, sebagai berikut:

##### **1. Manfaat Teoritis**

Penelitian ini diharapkan dapat memperkaya kajian literatur di bidang keuangan dan ilmu data, khususnya yang berkaitan dengan analisis volatilitas saham IPO menggunakan algoritma machine learning. Selain itu, penelitian ini juga memberikan kontribusi terhadap pengembangan model prediksi berbasis data prospektus, yang dapat digunakan sebagai referensi dalam penelitian selanjutnya mengenai perilaku harga saham di pasar modal.

##### **2. Manfaat Praktis**

- a. Bagi Investor, hasil penelitian ini dapat dijadikan acuan dalam memahami faktor-faktor prospektus yang berpengaruh terhadap volatilitas harga saham IPO, sehingga membantu dalam pengambilan keputusan investasi yang lebih rasional dan berbasis data.
- b. Bagi Emiten atau Perusahaan yang Akan IPO, penelitian ini dapat memberikan wawasan mengenai variabel prospektus yang perlu diperhatikan untuk mengurangi risiko volatilitas harga di awal perdagangan saham.
- c. Bagi Regulator dan Lembaga Keuangan, temuan penelitian ini dapat mendukung kebijakan pengawasan dan penilaian risiko terhadap perusahaan baru yang akan melantai di bursa.
- d. Bagi Akademisi dan Peneliti, penelitian ini dapat menjadi referensi metodologis dalam penerapan algoritma Extreme Gradient Boosting (XGBoost) untuk analisis data keuangan dan prediksi volatilitas saham di masa mendatang.

## **BAB II**

### **TINJAUN PUSTAKA**

#### **2.1. Penelitian Terdahulu**

Penelitian mengenai prediksi volatilitas dan harga saham dengan pendekatan *machine learning* telah banyak dilakukan, baik pada konteks saham IPO maupun saham perusahaan yang telah lama tercatat.

1. (Fauzi et al., 2025) melakukan penelitian dengan judul “*Analisis Volatilitas Saham IPO pada Hari Pertama dengan Metode Machine Learning*”. Penelitian ini menggunakan data IPO dari platform Kaggle dan menerapkan algoritma Random Forest untuk memprediksi volatilitas saham pada hari pertama pencatatan. Variabel yang digunakan mencakup harga final, return harian selama tujuh hari pertama, sektor industri, serta rasio waran. Hasil penelitian menunjukkan bahwa model Random Forest memiliki akurasi yang cukup tinggi dengan nilai  $R^2$  sebesar 0,85 dan MSE sekitar 0,0025, sehingga dapat memprediksi volatilitas secara efektif. Penelitian ini juga menegaskan pentingnya fitur harga penutupan, return harian, dan rasio waran dalam mempengaruhi volatilitas saham IPO
2. (Nabilah Selayanti et al., 2025) meneliti prediksi harga saham PT Bank Rakyat Indonesia (BBRI) menggunakan pendekatan *Hybrid LSTM–XGBoost*. Penelitian ini menunjukkan bahwa kombinasi LSTM sebagai *feature extractor* untuk pola temporal dengan XGBoost sebagai model prediksi mampu menghasilkan performa yang lebih baik dibanding model tunggal. Model hybrid mencapai RMSE 117,89, MAE 92,45, dan MAPE 2,21%, lebih unggul dibandingkan model LSTM atau XGBoost secara terpisah. Studi ini memperlihatkan bahwa XGBoost efektif dalam memodelkan hubungan non-linear pada data keuangan, sedangkan LSTM unggul dalam menangkap ketergantungan waktu

Penelitian-penelitian tersebut menunjukkan tren bahwa metode *machine learning*, khususnya algoritma berbasis ensemble seperti Random Forest dan XGBoost, dapat diterapkan untuk menganalisis volatilitas dan harga saham dengan hasil akurat. Namun, penerapan khusus XGBoost untuk volatilitas saham IPO berbasis data prospektus masih jarang dieksplorasi. Celah inilah yang menjadi dasar bagi penelitian ini.

#### **2.2. Landasan Teori**

##### **2.2.1 Initial Public Offering (IPO)**

IPO Adalah penawaran umum saham yang ditawarkan oleh Perusahaan kepada Masyarakat untuk pertama kali, Perusahaan yang mencatatkan saham merupakan Perusahaan yang berkembang dalam mendapatkan dana untuk mengembangkan usaha (Sugiyanto et al.,

2023).

Initial Public Offering (IPO) merupakan proses penawaran saham perdana suatu perusahaan kepada publik melalui pasar modal. IPO menjadi momen penting karena menentukan harga awal saham dan persepsi pasar terhadap perusahaan tersebut. Volatilitas harga saham pada hari-hari awal IPO cenderung tinggi karena adanya ketidakpastian informasi, tingginya permintaan spekulatif, dan minimnya data historis (Fauzi et al., 2025).

### **2.2.2 Volatilitas Saham**

Menurut (Fadilah & Wiharno, n.d.) Volatilitas harga saham merupakan ukuran ketidakpastian atau naik turunnya pergerakan saham di masa mendatang. Secara umum, volatilitas di pasar keuangan mencerminkan tingkat risiko yang dihadapi investor karena menggambarkan fluktuasi pergerakan harga saham dan ketidakpastian yang dimiliki investor. Semakin tinggi volatilitas harga saham, semakin besar fluktuasi harga saham dalam periode waktu tertentu. Banyak investor lebih memilih saham yang mudah diprediksi dan risikonya kecil karena semakin tinggi volatilitas harga saham maka kemungkinan semakin besar pula tingkat ketidakpastian untuk mendapatkan return.

### **2.2.3 Data Prospektus**

Prospektus adalah dokumen tertulis yang memuat informasi lengkap terkait proses penawaran umum efek, yang disusun untuk memberikan informasi kepada calon investor dengan tujuan mendorong mereka membeli efek yang ditawarkan (Hendrasyah, 2025). Informasi yang terdapat dalam prospektus mencakup antara lain: latar belakang dan profil perusahaan, rencana penggunaan dana hasil penawaran umum, struktur permodalan dan kepemilikan, informasi keuangan historis dan proyeksi, harga penawaran, jadwal penawaran umum, serta ketentuan terkait waran atau saham bonus. Data prospektus juga memuat informasi tentang risiko-risiko utama yang dihadapi perusahaan, termasuk kondisi industri dan faktor eksternal yang dapat mempengaruhi kinerja perusahaan.

### **2.2.4 Extreme Gradient Boosting (XGBoost)**

Extreme Gradient Boosting (XGBoost) merupakan algoritma *machine learning* berbasis pohon keputusan yang sangat populer untuk menyelesaikan tugas klasifikasi dan regresi, termasuk dalam prediksi harga dan volatilitas saham. XGBoost bekerja berdasarkan prinsip *boosting*, yaitu menggabungkan beberapa model pohon keputusan secara berurutan untuk memperbaiki kesalahan model sebelumnya dan meningkatkan akurasi prediksi secara keseluruhan. Algoritma ini memiliki kemampuan generalisasi yang baik, efisien dalam

komputasi, serta dapat dioptimasi melalui berbagai *hyperparameter* untuk menghasilkan performa yang optimal.

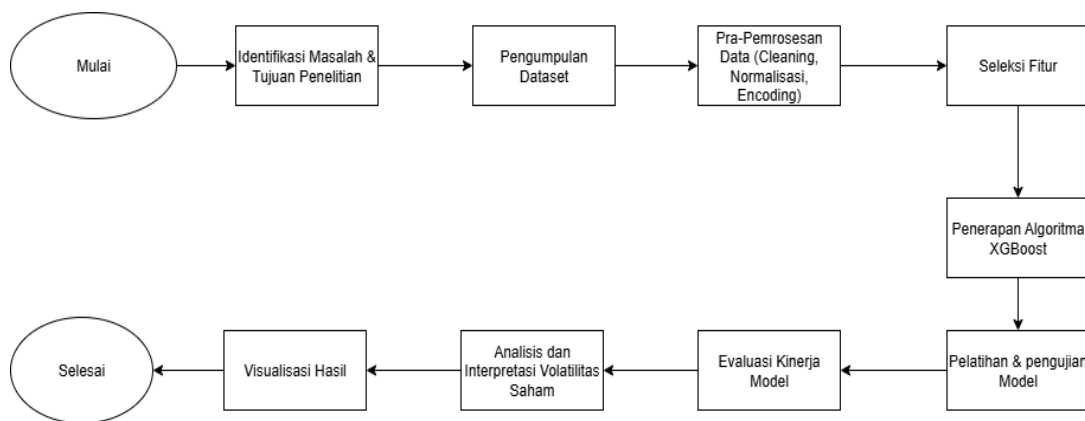
Dalam konteks prediksi saham, XGBoost sering digunakan karena kemampuannya dalam menangani data berukuran besar, mendeteksi hubungan non-linear antar variabel, serta menghasilkan kinerja prediksi yang kompetitif. XGBoost juga dilengkapi dengan mekanisme regularisasi yang berfungsi untuk mengontrol kompleksitas model dan mencegah *overfitting* (Herni Yulianti et al., 2022).

## BAB III

### DESAIN DAN IMPLEMENTASI SISTEM

#### 3.1. Metode Penelitian

Metode penelitian yang digunakan dalam studi ini adalah pendekatan kuantitatif dengan penerapan machine learning, khususnya algoritma *Extreme Gradient Boosting (XGBoost)* untuk memprediksi volatilitas harga saham IPO selama tujuh hari pertama setelah listing di Bursa Efek Indonesia.



Gambar 3. 1 Tahapan Penelitian

#### 3.2. Desain Sistem

Desain sistem menggambarkan alur proses penelitian dan implementasi sistem yang digunakan dalam menganalisis volatilitas saham IPO. Sistem dirancang untuk menerima input berupa data prospektus dan menghasilkan output berupa hasil analisis volatilitas serta faktor-faktor yang paling berpengaruh.

##### 3.2.1 Arsitektur Sistem

Arsitektur sistem penelitian ini terdiri dari beberapa komponen utama yang saling terhubung, yaitu:

##### 1. Input Data

Input data diambil dari data prospektus IPO (harga penawaran, nilai buku, total aset, laba bersih, umur perusahaan, sektor industri, dsb.) serta data return saham harian (hari ke-1 sampai ke-7 setelah IPO).



## 2. Preprocessing Module

Melakukan tahapan *data cleaning* (menghapus data kosong, duplikasi), *data transformation* (normalisasi atau standarisasi), serta *encoding* data kategorikal agar dapat diproses oleh model XGBoost.

## 3. Modeling Module (XGBoost)

Modul ini melakukan pemodelan menggunakan algoritma *Extreme Gradient Boosting (XGBoost)* untuk mengukur dan menganalisis volatilitas saham berdasarkan data prospektus.

## 4. Evaluation Module

Menghitung performa model dengan metrik evaluasi. Mengevaluasi performa model menggunakan metrik seperti: *Mean Squared Error (MSE)*, *Root Mean Squared Error (RMSE)* dan *R-squared ( $R^2$ )*

## 5. Visualization & Analysis Module

Menampilkan hasil analisis dalam bentuk grafik hubungan antar variabel, tingkat volatilitas saham, serta *feature importance* untuk memberikan insight yang mudah dipahami.

### 3.2.2 Teknologi dan Tools yang Digunakan

Untuk mendukung implementasi sistem ini, berbagai teknologi dan tools dipilih berdasarkan kebutuhan fungsional dan non-fungsional. Pemilihan teknologi ini bertujuan agar proses pengembangan, pengolahan data, serta visualisasi informasi dapat berjalan dengan efisien dan optimal.

Komponen	Tools / Library
Bahasa Pemrograman	Phyton
Analisis Data	Pandas, Numpy
Visualisasi Data	Matplotlib, Seaborn
Lingkungan Pengembangan	Google Colab
Penyimpanan Data	File CSV
Evaluasi Model	MSE, RMSE, $R^2$ Score
Machine Learning	XGBoost

### 3.2.3 Alur Kerja Sistem

1. Input Data → Sistem menerima data prospektus dan return saham.
2. Preprocessing → Data dibersihkan dan disiapkan untuk analisis.

3. Feature Selection → Dipilih fitur yang relevan terhadap volatilitas.
4. Modeling (XGBoost) → Model dilatih untuk menganalisis volatilitas.
5. Evaluation → Model dievaluasi menggunakan metrik performa.
6. Visualization → Hasil analisis divisualisasikan dalam bentuk grafik.
7. Output → Ditampilkan hasil akhir berupa tingkat volatilitas dan faktor dominan.

## DAFTAR PUSTAKA

- Fadilah, A., & Wiharno, H. (n.d.). Pengaruh Harga Saham, Return Saham, Volatilitas Harga Saham, Ukuran Perusahaan Dan Volume Perdagangan Saham Terhadap Bid-Ask Spread Saham.
- Fauzi, M. A., Suyahman, Prasetyo, D., Syafii, I., Tamtomo, A. P. S., & Ulfa, M. (2025). Analisis Volatilitas Saham IPO pada Hari Pertama dengan Metode Machine Learning. *Jurnal Teknologi*, 18(1), 16–22. <https://doi.org/10.34151/jurtek.v18i1.5236>
- Fitria Puteri Sholikah, Putri, W., & Rosalinda Maria Djangi. (2022). Peranan Pasar Modal Dalam Perekonomian Negara Indonesia. *ARBITRASE: Journal of Economics and Accounting*, 3(2), 341–345. <https://doi.org/10.47065/arbitrase.v3i2.496>
- Hendrasyah. (2025). Perlindungan Hukum Terhadap Investor dalam Initial Public Offering (IPO): Perspektif Aturan Hukum Pasar Modal di Indonesia. 2025.
- Herni Yulianti, S. E., Oni Soesanto, & Yuana Sukmawaty. (2022). Penerapan Metode Extreme Gradient Boosting (XGBOOST) pada Klasifikasi Nasabah Kartu Kredit. *Journal of Mathematics: Theory and Applications*, 21–26. <https://doi.org/10.31605/jomta.v4i1.1792>
- Nabilah Selayanti, Dwi Amalia Putri, Trimono Trimono, & Mohammad Idhom. (2025). PREDIKSI HARGA PENUTUPAN SAHAM BBRI DENGAN MODEL HYBRID LSTM-XGBOOST. *Informatika: Jurnal Teknik Informatika dan Multimedia*, 5(1), 52–64. <https://doi.org/10.51903/informatika.v5i1.1011>
- Nugroho, N. W., & Dewi, V. I. (2024). Analisis Forecasting Volatilitas Saham PT Goto GojekTokopedia Dengan Metode ARCH-GARCH. *INOBIIS: Jurnal Inovasi Bisnis dan Manajemen Indonesia*, 8(1), 48–57. <https://doi.org/10.31842/jurnalinobis.v8i1.352>