

SKRIPSI

PREDIKSI PERGERAKAN TREN HARGA SAHAM MELALUI  
ANALISIS TEKNIKAL MENGGUNAKAN *RANDOM FOREST*



FELIX ALPHA WINARTO

NPM: 6161901107

PROGRAM STUDI MATEMATIKA  
FAKULTAS TEKNOLOGI INFORMASI DAN SAINS  
UNIVERSITAS KATOLIK PARAHYANGAN  
2023

**FINAL PROJECT**

**PREDICTING THE TREND OF THE STOCK PRICE  
THROUGH TECHNICAL ANALYSIS USING RANDOM  
FOREST**



**FELIX ALPHA WINARTO**

**NPM: 6161901107**

**DEPARTMENT OF MATHEMATICS  
FACULTY OF INFORMATION TECHNOLOGY AND SCIENCES  
PARAHYANGAN CATHOLIC UNIVERSITY  
2023**

## LEMBAR PENGESAHAN

### PREDIKSI PERGERAKAN TREN HARGA SAHAM MELALUI ANALISIS TEKNIKAL MENGGUNAKAN *RANDOM FOREST*

Felix Alpha Winarto

NPM: 6161901107

Bandung, 3 Agustus 2023

Menyetujui,

Pembimbing 1



Liem Chin, M.Si.

Pembimbing 2



Dr. Andreas Parama Wijaya

Ketua Penguji



Agus Sukmana, M.Sc.

Anggota Penguji



Jonathan Hoseana, Ph.D.

Mengetahui,

Ketua Program Studi



Dr. Livia Owen

## PERNYATAAN

Dengan ini saya yang bertandatangan di bawah ini menyatakan bahwa skripsi dengan judul:

### **PREDIKSI PERGERAKAN TREN HARGA SAHAM MELALUI ANALISIS TEKNIKAL MENGGUNAKAN *RANDOM FOREST***

adalah benar-benar karya saya sendiri, dan saya tidak melakukan penjiplakan atau pengutipan dengan cara-cara yang tidak sesuai dengan etika keilmuan yang berlaku dalam masyarakat keilmuan.

Atas pernyataan ini, saya siap menanggung segala risiko dan sanksi yang dijatuhkan kepada saya, apabila di kemudian hari ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya saya, atau jika ada tuntutan formal atau non-formal dari pihak lain berkaitan dengan keaslian karya saya ini.

Dinyatakan di Bandung,  
3 Agustus 2023



Felix Alpha Winarto  
NPM: 6161901107

## ABSTRAK

Saham merupakan salah satu instrumen investasi yang dapat memberikan keuntungan yang cukup tinggi. Namun, tingkat keuntungan yang tinggi seringkali berhubungan dengan tingkat risiko yang tinggi pula. Perubahan arah tren yang cepat dan tidak teratur menjadi salah satu faktor yang menyebabkan risiko tersebut meningkat. Banyak penelitian yang sudah dilakukan untuk memprediksi harga saham. Dalam penelitian ini, prediksi akan dilakukan untuk mengidentifikasi kecenderungan atau tren harga saham di masa mendatang, bukan untuk memprediksi harga saham pada saat yang spesifik. Prediksi tren ini dilakukan menggunakan metode *random forest* dengan mengklasifikasikan tren menjadi tiga kelas, yaitu 1 untuk tren naik,  $-1$  untuk tren turun, dan 0 untuk tren tetap. Data saham yang digunakan dalam penelitian ini adalah data historis BBKA selama lima tahun terakhir. Data tersebut digunakan untuk menghitung berbagai nilai dari analisis teknikal yang berbeda. Nilai-nilai tersebut akan digunakan sebagai fitur prediksi dalam metode *random forest*. Hasil pengujian menunjukkan bahwa metode yang digunakan menghasilkan prediksi tren harga dengan tingkat akurasi sebesar 79,34%. Selain itu, penelitian ini juga menunjukkan adanya hubungan antara banyaknya pohon keputusan dalam model *random forest* dengan akurasi prediksi pergerakan tren harga saham. Namun, pengaruh banyaknya pohon keputusan terhadap akurasi dapat bervariasi, tergantung pada dataset dan skenario percobaan yang digunakan.

**Kata-kata kunci:** Pasar Saham, Tren Harga Saham, *Random Forest*, Analisis Teknikal

## ABSTRACT

Stocks are one of the investment instruments that could offer considerably high returns. However, high returns are often associated with high levels of risk. Rapid and irregular changes in trend direction are one of the factors that cause this risk to increase. Numerous studies have been conducted to predict stock prices. In this study, the prediction is performed to identify the tendencies or trends of stock prices in the future, rather than predicting stock prices at specific times. The prediction of these trends is done using the random forest method by classifying the trends into three classes, namely 1 for an upward trend, -1 for a downward trend, and 0 for a stagnant trend. The stock dataset used in this study is the historical dataset of BBCA over the past five years. This dataset is used to calculate various values from different technical analyses. These values will be used as predictive features in the random forest method. The test results demonstrate that the method used yields a prediction of stock price trends with an accuracy rate of 79,34%. Furthermore, this research also indicates a relationship between the number of decision trees in the random forest model and the accuracy of predicting stock price trend movements. However, the impact of the number of decision trees on accuracy can vary, depending on the dataset and experimental scenarios used.

**Keywords:** Stock Market, Stock Price Trend, Random Forest, Technical Analysis

## KATA PENGANTAR

Segala puji dan syukur ke hadirat Tuhan Yang Maha Esa atas berkat karunia-Nya sehingga penulis mampu menyelesaikan tugas akhir skripsi yang berjudul “Prediksi Pergerakan Tren Harga Saham Melalui Analisis Teknikal Menggunakan *Random Forest*” sebagai salah satu syarat kelulusan dari Program Studi Matematika dengan lancar dan tepat waktu. Selama tahap penyusunan skripsi ini, penulis banyak mendapatkan nasihat, dukungan, bantuan, dan bimbingan dari berbagai pihak. Pada kesempatan ini, penulis menyampaikan terima kasih setulus-tulusnya kepada:

- Orang tua penulis dan keluarga penulis yang selalu memberi dukungan dan mendoakan penulis dalam proses penyelesaian skripsi ini.
- Bapak Liem Chin, M.Si. dan Bapak Dr. Andreas Parama Wijaya selaku dosen pembimbing yang selalu memberikan bantuan dan saran sehingga penulis dapat menyelesaikan skripsi ini.
- Bapak Agus Sukmana, M.Sc. dan Bapak Jonathan Hoseana, Ph.D. selaku dosen penguji yang telah memberikan kritik dan saran sehingga skripsi ini dapat menjadi lebih baik
- Seluruh dosen Matematika dan dosen luar Matematika yang telah memberi ilmu kepada penulis selama perkuliahan.
- Evan Felix, Jovansen Hiustar, Raymond Susanto, Timothy Hoetama, dan Willy Zoe yang telah memberi semangat, dukungan, bantuan, dan hiburan kepada penulis selama perkuliahan dan penyusunan skripsi.
- Billy, Janice Kusuma, Joice Ivana, dan Vania Rosalie yang selalu memberikan informasi dan membantu penulis dalam proses penyelesaian skripsi ini.
- Teman-teman Matematika angkatan 2019 yang telah bersama-sama berjuang dalam perkuliahan.
- Semua pihak yang telah membantu penulis selama proses perkuliahan dan penyusunan skripsi.

Penulis menyadari bahwa masih terdapat beberapa kekurangan dalam penyusunan skripsi ini. Oleh karena itu, penulis menerima saran dan kritik yang membantu skripsi ini untuk menjadi lebih baik lagi.

Bandung, 3 Agustus 2023

Penulis

# DAFTAR ISI

|   |             |
|---|-------------|
| <b>KATA PENGANTAR</b>   | <b>viii</b> |
| <b>DAFTAR ISI</b>   | <b>ix</b>   |
| <b>DAFTAR GAMBAR</b>  | <b>x</b>    |
| <b>DAFTAR TABEL</b>   | <b>xi</b>   |
| <b>1 PENDAHULUAN</b>  | <b>1</b>    |
| 1.1 Latar Belakang . . . . .  | 1           |
| 1.2 Rumusan Masalah . . . . .   | 2           |
| 1.3 Tujuan . . . . .  | 3           |
| 1.4 <i>State of the Art</i> . . . . .                                     | 3           |
| <b>2 LANDASAN TEORI</b>   | <b>4</b>    |
| 2.1 Pembelajaran Mesin ( <i>Machine Learning</i> ) . . . . .              | 4           |
| 2.2 Pohon Keputusan ( <i>Decision Tree</i> ) . . . . .                    | 5           |
| 2.3 <i>Random Forest</i> . . . . .  | 7           |
| 2.4 Analisis Teknikal . . . . .   | 7           |
| 2.4.1 <i>Moving Average (MA)</i> . . . . .                                | 8           |
| 2.4.2 <i>Relative Strength Index (RSI)</i> . . . . .                      | 9           |
| 2.4.3 <i>Moving Average Convergence Divergence (MACD)</i> . . . . .       | 10          |
| 2.4.4 <i>Bollinger Bands</i> . . . . .                                    | 10          |
| 2.4.5 <i>Stochastic Oscillator</i> . . . . .                              | 11          |
| 2.5 <i>Confusion Matrix</i> . . . . .                                     | 12          |
| <b>3 <i>Random Forest</i> UNTUK PREDIKSI TREN HARGA SAHAM</b>             | <b>14</b>   |
| <b>4 HASIL PREDIKSI TREN HARGA SAHAM DENGAN <i>Random Forest</i></b>      | <b>22</b>   |
| 4.1 Pra-pemrosesan Data . . . . .   | 22          |
| 4.2 Metode <i>Random Forest</i> dengan Data Sebelum Penghalusan . . . . . | 24          |
| 4.3 Metode <i>Random Forest</i> dengan Data Sesudah Penghalusan . . . . . | 27          |
| <b>5 PENUTUP</b>  | <b>33</b>   |
| 5.1 Kesimpulan . . . . .  | 33          |
| 5.2 Saran . . . . .   | 33          |
| <b>DAFTAR REFERENSI</b>   | <b>35</b>   |



## DAFTAR GAMBAR

|     |   |    |
|-----|---|----|
| 2.1 | Ilustrasi cara kerja pohon keputusan . . . . .  | 5  |
| 2.2 | Ilustrasi cara kerja <i>random forest</i> . . . . .   | 8  |
| 2.3 | Contoh grafik RSI . . . . .   | 10 |
| 2.4 | Contoh <i>bollinger bands</i> dengan <i>upper band</i> , <i>middle band</i> , dan <i>lower band</i> . . . . . | 11 |
| 3.1 | Pohon keputusan melalui proses penghitungan manual . . . . .  | 20 |
| 3.2 | Pohon keputusan melalui proses penghitungan dengan bantuan perangkat lunak . . . . .                          | 21 |
| 4.1 | Data yang digunakan beserta dengan banyaknya data . . . . .   | 22 |
| 4.2 | Perbandingan grafik <i>adj close</i> antara data asli dan data yang dihaluskan . . . . .                      | 23 |
| 4.3 | Gambar <i>confusion matrix</i> skenario pertama dengan menggunakan data tanpa penghalusan . . . . .           | 25 |
| 4.4 | Gambar <i>confusion matrix</i> skenario kedua dengan menggunakan data tanpa penghalusan . . . . .             | 26 |
| 4.5 | Gambar <i>confusion matrix</i> skenario pertama dengan menggunakan data penghalusan . . . . .                 | 28 |
| 4.6 | Gambar <i>confusion matrix</i> skenario kedua dengan menggunakan data penghalusan . . . . .                   | 30 |
| 4.7 | Gambar <i>confusion matrix</i> skenario ketiga dengan menggunakan data penghalusan . . . . .                  | 31 |

## DAFTAR TABEL

|      |   |    |
|------|---|----|
| 2.1  | Contoh dari <i>confusion matrix</i> biner . . . . .   | 12 |
| 2.2  | Contoh dari <i>confusion matrix multiclass</i> . . . . .  | 13 |
| 3.1  | Data latih saham buatan untuk simulasi penghitungan . . . . .   | 14 |
| 3.2  | Penjelasan atribut data . . . . .   | 15 |
| 3.3  | Pelabelan pergerakan harga saham berdasarkan selisih harga adj . . . . .                                      | 15 |
| 3.4  | Data yang sudah dikelompokkan . . . . .   | 16 |
| 3.5  | Proses pemilihan atribut data . . . . .   | 18 |
| 3.6  | Data dengan kriteria $O \leq 99$ . . . . .  | 18 |
| 3.7  | Proses pencabangan simpul dengan kriteria data $O \leq 99$ . . . . .  | 19 |
| 3.8  | Data dengan kriteria $O \leq 99$ dan $vol \leq 100$ . . . . .   | 19 |
| 3.9  | Proses pencabangan simpul dengan kriteria $O \leq 99$ dan $vol \leq 100$ . . . . .                            | 19 |
| 4.1  | Penjelasan atribut data pada data historis saham . . . . .  | 23 |
| 4.2  | Data skenario pertama tanpa penghalusan data . . . . .  | 24 |
| 4.3  | Rasio klasifikasi target skenario pertama tanpa penghalusan data . . . . .                                    | 24 |
| 4.4  | Data skenario kedua tanpa penghalusan data . . . . .  | 26 |
| 4.5  | Rasio klasifikasi target skenario kedua tanpa penghalusan data . . . . .                                      | 26 |
| 4.6  | Data skenario pertama dengan data penghalusan . . . . .   | 27 |
| 4.7  | Rasio klasifikasi target skenario pertama dengan data penghalusan . . . . .                                   | 27 |
| 4.8  | Nilai perbandingan akurasi pada skenario pertama dengan beberapa penggunaan pohon keputusan berbeda . . . . . | 28 |
| 4.9  | Data skenario kedua dengan data penghalusan . . . . .   | 29 |
| 4.10 | Rasio klasifikasi target skenario kedua dengan data penghalusan . . . . .                                     | 29 |
| 4.11 | Nilai perbandingan akurasi pada skenario kedua dengan beberapa penggunaan pohon keputusan berbeda . . . . .   | 29 |
| 4.12 | Data skenario ketiga dengan data penghalusan . . . . .  | 30 |
| 4.13 | Rasio klasifikasi target skenario ketiga dengan data penghalusan . . . . .                                    | 30 |
| 4.14 | Nilai perbandingan akurasi pada skenario ketiga dengan beberapa penggunaan pohon keputusan berbeda . . . . .  | 31 |

# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang

Dalam suatu pasar saham terdapat sebuah pola yang disebut dengan tren harga saham. Tren merupakan arah kecenderungan pergerakan harga saham dalam periode waktu tertentu dan terbagi atas tren kenaikan (*bullish*), penurunan (*bearish*), dan mendatar (*sideways*). Periode waktu tren harga saham dibagi menjadi dua bagian, yaitu tren harga saham jangka pendek (harian, mingguan, dan bulanan) dan tren harga saham jangka panjang (tahunan). Tren harga saham memiliki suatu karakteristik, yaitu perubahan arah yang cepat dan tidak teratur. Perubahan yang terjadi dapat disebabkan oleh beberapa faktor, yang dibagi menjadi dua bagian, yaitu faktor eksternal dan internal. Faktor eksternal yang memengaruhi perubahan arah harga saham adalah kondisi ekonomi, kebijakan pemerintah, faktor panik, manipulasi pasar, dan lain-lain, sedangkan faktor internal terdiri dari fundamental perusahaan, proyeksi kinerja perusahaan pada masa mendatang, dan lain-lain. Karena sifat dari saham yang dapat merugikan dan tren harga saham yang tidak teratur, maka muncullah penelitian-penelitian untuk memprediksi pergerakan arah tren harga saham sehingga kerugian dapat diminimalkan dan keuntungan dapat dimaksimalkan.

Salah satu pendekatan yang telah menunjukkan keberhasilan dalam memprediksi tren harga saham adalah menggunakan metode pembelajaran mesin. Pembelajaran mesin menyediakan teknik-teknik terbaik dalam bidang kecerdasan buatan (*artificial intelligence*), seperti klasifikasi, regresi, dan pembelajaran dengan pengawasan (*supervised learning*), pembelajaran tanpa pengawasan (*unsupervised learning*), dan banyak teknik lainnya. Teknik-teknik ini digunakan untuk mengembangkan sistem kecerdasan buatan yang dapat melakukan tugas seperti klasifikasi data, prediksi, dan penemuan pola, tanpa perlu aturan atau petunjuk yang eksplisit. Terdapat beragam penelitian yang telah dilakukan untuk mengembangkan model prediksi harga saham dengan pembelajaran mesin. Sebagai contoh, penelitian oleh Patel [1] menerapkan empat model pembelajaran mesin seperti *Artificial Neural Networks* (ANN), *Support vector machine* (SVM), *Random forest*, dan *Naïve Bayes* untuk memprediksi arah pergerakan saham dan indeks harga saham. Percobaan yang dilakukan menunjukkan bahwa model *Naïve Bayes* (proses *Gaussian*) memiliki kinerja paling rendah dengan akurasi 73,3%, dan *random forest* memiliki kinerja tertinggi dengan akurasi sebesar 83,56%.

*Random Forest* (RF) merupakan salah satu metode pada pembelajaran mesin yang digunakan untuk klasifikasi, yang dilakukan melalui penggabungan pohon-pohon data dalam jumlah yang besar. Metode RF sudah digunakan oleh para peneliti dalam beberapa kasus, dan metode tersebut dianggap baik. Dalam penelitian yang dilakukan oleh Sadorsky [2], digunakan metode RF untuk

memprediksi pergerakan harga saham, dan didapatkan tingkat akurasi antara 85% sampai 90%. Nilai akurasi ini cukup tinggi jika dibandingkan dengan metode-metode yang digunakan dalam penelitian tersebut. Metode RF yang dibahas oleh Lohrmann dan Luukka [3] juga menunjukkan bahwa RF memiliki akurasi prediksi harga saham yang tinggi. Dari hasil yang didapatkan dari beberapa penelitian tersebut diketahui bahwa RF dapat menyelesaikan masalah dengan baik dan memiliki tingkat akurasi yang tinggi.

Dalam melakukan perdagangan saham, banyak orang yang memanfaatkan analisis teknikal untuk mengidentifikasi pola, tren, dan sinyal beli atau jual yang menjadi dasar pengambilan keputusan investasi. Penelitian yang dilakukan oleh Teixeira [4] menunjukkan bahwa pendekatan analisis teknikal dapat berhasil dalam kondisi pasar nyata. Hasil eksperimen menunjukkan bahwa metode yang diusulkan mencapai hasil yang baik dalam hal profitabilitas, bahkan mengungguli strategi *buy and hold* untuk sebagian besar saham yang dipertimbangkan. Oleh karena itu, dalam skripsi ini akan digunakan algoritma *random forest* yang dikombinasikan dengan analisis teknikal yang akan digunakan sebagai fitur prediksi. Analisis teknikal dipilih karena mampu memberikan informasi tentang tren dan pola yang muncul dari data historis harga saham, sedangkan RF sebagai metode klasifikasi dipilih untuk membantu dalam mengenali pola-pola tersebut dan membuat prediksi arah pergerakan harga saham berdasarkan fitur-fitur yang relevan. Penggabungan kedua pendekatan ini dapat memberikan keunggulan yang saling melengkapi dalam menganalisis pergerakan harga saham. Analisis teknikal lebih mengandalkan penilaian subjektif, sedangkan RF menghasilkan keputusan langsung dari data-data numerik berdasarkan data historis. Selain itu, penggunaan RF melalui analisis teknikal dipilih karena dalam penggunaannya, RF dapat memberikan beberapa keuntungan yang relevan untuk meningkatkan prediksi tren harga saham. Metode ini dapat mengurangi risiko *overfitting* pada data latihan, serta mampu mengatasi data yang tidak terstruktur dengan baik. Kelebihan ini berkontribusi pada peningkatan kemampuan analisis teknikal dalam memprediksi pergerakan harga saham di masa depan dengan akurasi yang lebih baik. Kombinasi ini diharapkan dapat meningkatkan akurasi prediksi dan membantu dalam pengambilan keputusan investasi yang lebih baik.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan terdapat beberapa permasalahan yang dapat menjadi pembahasan dalam masalah ini. Permasalahan-permasalahan tersebut dapat dirumuskan sebagai berikut.

1. Bagaimana hasil yang diperoleh dalam memprediksi pergerakan tren harga saham melalui analisis teknikal dengan menggunakan algoritma *random forest*?
2. Berapa tingkat persentase akurasi yang diperoleh dari proses prediksi tren harga saham selama satu hari ke depan dengan algoritma *random forest*?
3. Bagaimana pengaruh perubahan target prediksi terhadap hasil yang didapatkan dalam model *random forest*?
4. Bagaimana pengaruh banyaknya pohon keputusan dalam model *random forest* terhadap tingkat akurasi dalam memprediksi pergerakan tren harga saham?

### 1.3 Tujuan

Berdasarkan rumusan masalah yang telah diuraikan, adapun tujuan dari skripsi yang ingin dicapai adalah

1. memperoleh hasil prediksi pergerakan tren harga saham melalui analisis teknikal menggunakan algoritma *random forest*,
2. memperoleh persentase akurasi penggunaan algoritma *random forest* dalam memprediksi pergerakan tren harga saham,
3. menganalisis pengaruh perubahan target prediksi terhadap hasil prediksi yang didapatkan dalam model *random forest*,
4. mengeksplorasi pengaruh banyaknya pohon keputusan dalam model *random forest* terhadap tingkat akurasi dalam memprediksi pergerakan tren harga saham dan mengidentifikasi apakah peningkatan banyaknya pohon keputusan secara signifikan meningkatkan tingkat akurasi prediksi.

### 1.4 *State of the Art*

Dalam penelitian yang dilakukan pada skripsi ini, peneliti berfokus pada penelitian yang sudah dilakukan oleh Basak [5]. Dalam penelitian tersebut, digunakan metode *random forest* dengan masukan analisis teknikal sebagai fitur prediksi. Penggunaan metode ini memiliki keunggulan dalam mengatasi kompleksitas data saham dan variasi yang ada di berbagai pasar dan industri. Metode ini mampu secara akurat memahami pola dan tren dalam data besar yang dihasilkan oleh pasar saham, dan kemudian menggunakannya untuk mengambil keputusan prediksi. Model yang digunakan terbukti dapat secara akurat memperkirakan apakah harga saham akan mengalami kenaikan atau penurunan dalam periode yang ditentukan. Namun, dalam skripsi ini, akan dilakukan penelitian lebih mendalam dan pengembangan dalam beberapa aspek berikut.

1. Pemahaman tentang pengaruh berbagai target prediksi terhadap kinerja model menjadi hal yang penting. Dalam skripsi ini, dilakukan pengujian kinerja model dengan menggunakan beberapa target prediksi yang berbeda.
2. Setelah didapatkan hasil akurasi model, dilakukan perbandingan kinerjanya dengan variasi banyaknya pohon keputusan yang dibangun. Hal ini bertujuan memberikan wawasan tentang bagaimana performa model dapat ditingkatkan dan sejauh mana banyaknya pohon keputusan mempengaruhi akurasi.