

## Prediksi Harga Saham Netflix Inc. (NFLX) di Nasdaq menggunakan Algoritma Machine Learning berbasis Data Mining

Waeisul Bismi<sup>1\*</sup>, Aris Parnius Siregar<sup>2</sup>, Muhammad Qomaruddin<sup>3</sup>

<sup>1,2</sup> Program Studi Informatika, Universitas Bina Sarana Informatika

<sup>3</sup> Program Studi Informatika, Universitas Nusa Mandiri

E-mail Korespondensi : [Waeisul.wbn@bsi.ac.id](mailto:Waeisul.wbn@bsi.ac.id)

History Artikel

Diterima : 5 Mei 2025 Disetujui : 10 September 2025 Dipublikasikan : 25 Oktober 2025

---

### ***Abstract***

The volatility of stock prices makes stock price prediction a crucial aspect in investment decision-making. Netflix Inc. (NFLX), as a leading technology company listed on NASDAQ, exhibits dynamic price movements influenced by both internal and external factors. This study aims to implement machine learning algorithms based on data mining to predict the stock price movements of Netflix Inc., using *Orange Data Mining* software, which offers a visual and user-friendly environment that supports a variety of classification methods. The research methodology consists of five main stages: data collection, data preprocessing, model implementation, model testing, and performance evaluation. The dataset was obtained from Investing.com, covering daily stock data from January 2022 to December 2023, totaling 501 entries with features such as opening price, closing price, high, low, and trading volume. The classification algorithms used in this study include Support Vector Machine (SVM), Neural Network, and Naive Bayes. The dataset was divided into 80% training data and 20% testing data. The experimental results show that the Neural Network algorithm achieved the best performance, with an AUC score of 0.937, an accuracy (CA) of 88%, and a Matthews Correlation Coefficient (MCC) of 0.759. The SVM algorithm achieved 84.3% accuracy and 0.690 MCC, while the Naive Bayes model reached 83% accuracy with a 0.660 MCC score. The confusion matrix evaluation demonstrated that the models were able to classify upward and downward price trends with a relatively low error rate. Thus, the application of machine learning algorithms in data mining proves effective in forecasting stock prices and can be utilized as a reliable tool to support investment decision-making based on historical data trends.

**Keywords:** *Stock Price Prediction, Machine Learning, Netflix, Orange Data Mining*

## ***Abstrak***

Pergerakan harga saham yang fluktuatif membuat prediksi harga saham menjadi aspek penting dalam pengambilan keputusan investasi. Netflix Inc. (NFLX), sebagai salah satu emiten teknologi terkemuka yang terdaftar di NASDAQ, menunjukkan pola pergerakan harga saham yang dinamis dan dipengaruhi oleh berbagai faktor internal dan eksternal. Penelitian ini bertujuan untuk menerapkan algoritma machine learning berbasis data mining dalam memprediksi harga saham Netflix Inc., dengan menggunakan perangkat lunak *Orange Data Mining* yang bersifat visual dan mendukung berbagai metode klasifikasi. Metode penelitian terdiri dari lima tahapan utama, yaitu: pengumpulan data, pra-pemrosesan data, implementasi model, pengujian model, dan evaluasi kinerja. Dataset yang digunakan diperoleh dari Investing.com dalam periode Januari 2022 hingga Desember 2023, dengan total 501 data harian yang mencakup fitur seperti harga pembukaan, harga penutupan, harga tertinggi, harga terendah, dan volume transaksi. Algoritma klasifikasi yang digunakan meliputi Support Vector Machine (SVM), Neural Network, dan Naive Bayes. Data dibagi menjadi 80% data pelatihan dan 20% data pengujian. Hasil pengujian menunjukkan bahwa algoritma Neural Network memberikan performa terbaik dengan nilai AUC sebesar 0,937, akurasi (CA) 88%, dan nilai MCC sebesar 0,759. Algoritma SVM menghasilkan akurasi 84,3% dan MCC 0,690, sedangkan Naive Bayes memiliki akurasi 83% dengan MCC 0,660. Evaluasi menggunakan confusion matrix memperlihatkan bahwa model mampu mengenali pola klasifikasi naik atau turun dengan tingkat kesalahan yang minim. Dengan demikian, penerapan algoritma machine learning pada data mining terbukti efektif dalam melakukan prediksi harga saham, dan dapat digunakan sebagai acuan dalam pengambilan keputusan investasi di pasar modal berbasis data historis.

**Kata Kunci:** *Prediksi Harga Saham, Machine Learning, Netflix, Orange Data Mining*

**How to Cite:** Waeisul Bismi (2025). Prediksi Harga Saham Netflix Inc. (NFLX) di Nasdaq menggunakan Algoritma Machine Learning berbasis Data Mining. *KOMPUTEK : Jurnal Teknik Universitas Muhammadiyah Ponorogo*, Vol 9 (2): Halaman 48-56

## PENDAHULUAN

Prediksi merupakan suatu proses sistematis untuk meramalkan kejadian yang paling mungkin terjadi di masa depan dengan menggunakan data historis dan kondisi saat ini, guna mengurangi selisih antara hasil prediksi dan kenyataan (Syahdan Nabil & Ramdhani, 2023). Prediksi bukan sekadar tebakan, melainkan didasarkan pada pemikiran teoritis yang kuat, yang bisa berupa teori ilmiah. Prediksi dalam konteks investasi merupakan suatu proses sistematis untuk memperkirakan kemungkinan pergerakan harga saham di masa depan dengan memanfaatkan data historis dan kondisi pasar saat ini (Irmayani & Jayanti, 2025). Tujuannya adalah untuk mengurangi selisih antara hasil prediksi dengan realitas yang terjadi.

Investasi merupakan suatu kinerja penempatan dana sebagai cara untuk mendapatkan keuntungan. *Investing* adalah kebijakan pasar modal dalam menyediakan data *real-time*, *quote* harga, perangkat keuangan, kabar terkini, dan analisis dari 250 pasar modal di seluruh dunia dengan 44 edisi internasional (B. Krishnan et al., 2020). Indikator pengetahuan investasi yaitu mengetahui tujuan investasi, mengetahui tentang tingkat imbal hasil, mengetahui tentang adanya risiko serta keuntungan yang didapatkan, mengetahui tentang alat investasi dalam pasar efek dan edukasi dasar mengenai investasi dalam pasar modal (Amelia Sekar Ayu Pramesti et al., 2024).

Salah satu investasi yang sangat populer dan menarik untuk dibahas adalah saham. Berbeda dengan investasi lainnya saham merupakan jenis investasi yang cukup fleksibel. Meskipun begitu para *investor* juga diharapkan dapat melakukan investasi dengan bijak dikarenakan tidak adanya jaminan dari kenaikan harga ataupun pengembalian dana saham, ada 2 jenis resiko dalam berinvestasi (Tambunan, 2020) yaitu:

1. Risiko Non Sistematis adalah cara diverifikasinya resiko ini dapat dieliminasi atau melakukan investasi

dalam berbagai jenis saham dari berbagai sektor karena berasal dari kondisi internal perusahaan. Dampak Risiko ini hanya terjadi pada suatu saham atau sektor tertentu;

2. Dikarenakan risiko sistematis berasal dari banyak kondisi ekonomi atau pasar maka resiko ini tidak dapat dieliminasi. Berbeda dengan resiko non sistematis resiko sistematis akan berdampak kepada semua jenis saham jika resiko ini muncul atau terjadi.

Perkembangan teknologi informasi telah memberikan kontribusi signifikan dalam dunia keuangan, khususnya dalam kegiatan analisis dan prediksi harga saham. yang di gunakan untuk menggambarkan pergerakan saham yaitu indeks harga saham (Eka Patriya, 2020). Salah satu perusahaan teknologi yang menjadi perhatian investor global adalah Netflix Inc. (NFLX), yang terdaftar di bursa saham NASDAQ. Harga saham perusahaan ini sangat fluktuatif dan dipengaruhi oleh berbagai faktor seperti laporan keuangan, tren pasar, sentimen publik, serta kompetisi industri hiburan digital. Oleh karena itu, diperlukan metode prediktif yang mampu menganalisis data historis saham secara akurat untuk mendukung pengambilan keputusan investasi (Rouf et al., 2021).

Software *Orange Data Mining* menjadi alat yang tepat untuk digunakan dalam penelitian ini karena bersifat open-source, visual, dan mendukung berbagai algoritma machine learning populer. Dengan tampilan antarmuka yang intuitif, Orange memfasilitasi proses eksperimen ML secara interaktif tanpa memerlukan coding yang kompleks. Selain itu, Orange telah digunakan dalam berbagai studi analisis prediktif di sektor keuangan dan kesehatan, menunjukkan kapabilitasnya sebagai platform pembelajaran mesin yang andal (Almaspoor et al., 2021).

## TINJAUAN PUSTAKA

Berdasarkan penelitian yang dilakukan oleh (Abdul Dwiyanto Suyudi et al., 2019) menggunakan analisis history harga saham dalam suatu perusahaan, dan Recurrent Neural Network (RNN) untuk melakukan prediksi terhadap nilai saham dari history harga saham. Fitur yang di indentifikasi yaitu harga terendah, harga tertinggi, harga buka, harga tutup, volume, rata-rata harga, dan pergerakan. Prediksi tujuh fitur variable dengan RNN menghasilkan akurasi sebesar 94% untuk data latih dan 55% untuk data uji. Akurasi diperoleh setelah pelatihan dengan menggunakan 1218 data.

Kemudian pada penelitian (Fadilah et al., 2020) mengenai “Analisis Prediksi Harga Saham PT. Telekomunikasi Indonesia Menggunakan Metode Support Vector Machine menghasilkan akurasi sebesar 0.9641 dan RMSE sebesar 0.0932. Pengujian juga dilakukan menggunakan algoritma k-Nearest Neighbors dengan tingkat akurasi sebesar 0.945 dan RMSE sebesar 0.1162.

Lalu dalam penelitian (Fatra, 2021) dengan judul “Analisis Prediksi Harga Saham Pt Blue Bird Tbk Ditengah Pandemi Covid-19 Dengan Metode Arima Dalam Perspektif Ekonomi Islam” mempunyai tingkat keakuratan peramalan sebesar 83.33% yang bertujuan membantu investor untuk meminimalisir kerugian dalam investasi.

Kemudian dalam penelitian (Hastomo et al., 2021) dengan Optimasi Deep Learning untuk Prediksi Saham di Masa Pandemi Covid-19 yaitu mengkombinasikan desain 4 hidden layer neural network menggunakan Long Short Term Memory (LSTM) dan Gated Recurrent Unit (GRU).

Dan penelitian yang dilakukan (Saputro & Swanjaya, 2023) mendapatkan prediksi harga saham yang akurat dengan pemilihan/ penggunaan variabel – variabel data transaksi saham (tanggal, open, high, low, close, volume dan net foreign flow) yang digabungkan dengan salah satu metode Neural Network.

## METODE PENELITIAN

Penelitian ini terdiri dari beberapa tahapan yang dapat dilihat pada gambar 1 berikut.



Gambar 1. Tahapan Penelitian

### 3.1 Pengumpulan Data

Dataset adalah objek yang merepresentasikan data dan relasinya di memory (Suhartini et al., 2020). Adapaun pengumpulan data merupakan tahapan awal dalam alur metodologi penelitian (Bismi et al., 2024).

### 3.2 Pra-Pemrosesan Data

Dalam tahapan ini dilakukan beberapa teknik, diantaranya *labelling*, yakni pemberian atribut kepada data yang dimiliki untuk menentukan nilai target pada data, kemudian melakukan transformasi data untuk mengubah data mentah menjadi data yang siap dipakai serta melakukan normalisasi data untuk mengetahui apakah semua fitur memiliki skala yang sama serta membuat fitur tambahan jika diperlukan. Dan teknik *splitting data*, yakni pembagian data menjadi dua set: data pelatihan 80% dan data pengujian 20%.

### 3.3 Implementasi Model

Model algoritma yang akan di implementasikan menggunakan algoritma klasifikasi *supervised learning* dengan menggunakan 3 algoritma yaitu:

1. Support Vector Machine (SVM), yang mempunyai aturan dasar *linier classifier*

yakni kasus klasifikasi yang secara linier dapat dipecah, akan tetapi sekarang *Support Vector Machine(SVM)* juga bisa menyelesaikan persoalan dari *non-linier* dengan menambahkan konsep kernel pada *workspace* berdimensi tinggi (Othchere et al., 2021).

- Neural Network, merupakan model sistem yang bisa beradaptasi dalam memecahkan strukturnya untuk menyelesaikan suatu masalah berdasarkan informasi internal ataupun eksternal yang dilalui jaringan tersebut(Leovinent & Yoannita, 2023).
- Naive Bayes, digunakan untuk mengklasifikasi sekumpulan dokumen metode yang dapat digunakan juga untuk menemukan peluang terbaik dalam metode ini memanfaatkan teorema probabilitas (Sumanto et al., 2021), dengan memprediksi probabilitas di masa yang akan datang berdasarkan informasi dari data di masa yang lampau (Veziroğlu et al., 2023)

### 3.4 Pengujian Model

Di dalam tahapan ini peneliti menggunakan *widjet test and score* untuk dapat menemukan hasil yang paling optimal dengan menggunakan metode *Random sampling*. Penggunaan dari *Test and Score* ini digunakan untuk mencari nilai seperti AUC, *Recall*, CA, F1, *Preccesion*, dan MCC.

### 3.5 Evaluasi Kinerja

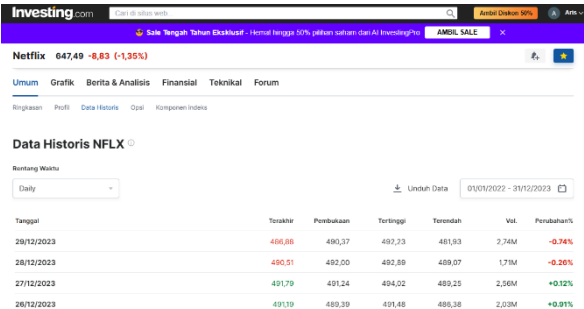
Dan untuk menentukan prediksi terbaik dari beberapa model algoritma yang digunakan pada tahapan evaluasi, peneliti menggunakan *Confusion Matrix* untuk melakukan analisisnya. Pada *Confusion Matrix* disini kita akan menentukan berapa banyak data yang bernilai *Actual* salah atau benar dari hasil prediksi yang dilakukan. *Confusion Matrix* akan menghasilkan nilai TN,FN, TP, dan FP seperti Tabel 1 berikut.

**Tabel 1.** Confusion Matrix

|                 |       | Observed            |                     |
|-----------------|-------|---------------------|---------------------|
|                 |       | True                | False               |
| Predicted Class | True  | True Positif (TP)   | False Positive (FP) |
|                 | False | False Negative (FN) | True Negative (TN)  |

## HASIL DAN PEMBAHASAN

### 4.1 Hasil Pengumpulan Data



**Gambar 2.** Data Historis NFLX

Dataset yang digunakan dalam penelitian ini berasal dari data historis harga saham Netflix Inc. (NFLX) yang diperoleh melalui platform keuangan Investing.com terlihat pada gambar 2. yang diambil dari jangka waktu 2 tahun dimulai dari tanggal 3 Januari 2022 sampai tanggal 29 Desember 2023 dengan jumlah banyaknya data berjumlah 501 data.

Dengan beberapa variabel dari history harga saham, termasuk tanggal, harga pembukaan, harga terakhir, harga tertinggi, harga terendah, volume perdagangan harian, dan perubahan harga yang terjadi. Pengumpulan data dilakukan dengan memperhatikan kelengkapan dan konsistensi data untuk memastikan kualitas analisis.

### 4.2 Hasil Pra-Pemrosesan Data

Tahapan pra-pemrosesan dilakukan untuk meningkatkan kualitas data dan kesiapan dataset sebelum dimasukkan ke dalam model prediksi. Langkah awal adalah *labelling*, yaitu menambahkan atribut target berupa kategori *Naik* atau *Turun* berdasarkan perbandingan harga penutupan hari ini dengan hari sebelumnya, seperti yang terlihat pada gamabr 3 berikut.

| Tanggal    | Terakhir | Pembukaan | Tertinggi | Terendah | Vol.  | Perubahan | harga baru |
|------------|----------|-----------|-----------|----------|-------|-----------|------------|
| 03/01/2022 | 597.37   | 605.61    | 609.99    | 590.56   | 3.07  | -0.84%    | 0          |
| 04/01/2022 | 591.15   | 599.91    | 600.41    | 581.6    | 4.39  | -1.04%    | 0          |
| 05/01/2022 | 567.52   | 592       | 592.84    | 566.88   | 4.15  | -4.00%    | 0          |
| 06/01/2022 | 553.29   | 554.34    | 563.36    | 542.01   | 5.71  | -2.51%    | 0          |
| 07/01/2022 | 541.06   | 549.46    | 553.43    | 538.22   | 3.38  | -2.21%    | 0          |
| 10/01/2022 | 539.85   | 538.49    | 543.69    | 526.32   | 4.49  | -0.22%    | 1          |
| 11/01/2022 | 540.84   | 536.99    | 543.91    | 530.07   | 3.08  | 0.18%     | 1          |
| 12/01/2022 | 537.22   | 544.27    | 544.27    | 532.02   | 3.79  | -0.67%    | 0          |
| 13/01/2022 | 519.2    | 537.06    | 540.79    | 518.26   | 4.48  | -3.35%    | 0          |
| 14/01/2022 | 525.69   | 517.6     | 538.37    | 511.88   | 7.86  | 1.25%     | 1          |
| 18/01/2022 | 510.8    | 520.08    | 521.75    | 508.68   | 4.84  | -2.83%    | 0          |
| 19/01/2022 | 515.86   | 515       | 523.21    | 510.51   | 4.35  | 0.99%     | 1          |
| 20/01/2022 | 508.25   | 517.75    | 526.64    | 506.93   | 12.66 | -1.48%    | 0          |
| 21/01/2022 | 397.5    | 400.42    | 409.15    | 379.99   | 58.9  | -21.79%   | 0          |
| 24/01/2022 | 387.15   | 383.91    | 387.26    | 351.46   | 32.39 | -2.60%    | 1          |
| 25/01/2022 | 366.42   | 379.14    | 387.71    | 365.13   | 15.15 | -5.35%    | 0          |
| 26/01/2022 | 359.7    | 378.27    | 382.66    | 356.62   | 12.68 | -1.83%    | 0          |
| 27/01/2022 | 386.7    | 382.06    | 394.8     | 378.1    | 24.32 | 7.51%     | 1          |
| 28/01/2022 | 384.36   | 386.76    | 387       | 372.08   | 11.97 | -0.61%    | 0          |
| 31/01/2022 | 427.14   | 401.97    | 427.7     | 398.2    | 20.05 | 11.13%    | 1          |
| 01/02/2022 | 457.13   | 432.96    | 458.48    | 425.54   | 22.57 | 7.02%     | 1          |

**Gambar 3.** Hasil *Labeling Dataset*

Selanjutnya dilakukan transformasi data dengan membersihkan kolom yang tidak relevan dan mengubah format tanggal menjadi indeks waktu. Normalisasi juga diterapkan agar setiap fitur memiliki skala yang sama, terutama pada fitur numerik seperti harga dan volume, seperti yang terlihat pada gambar 4 berikut.

| harga baru | Tanggal             | Terakhir | Terakhir (mean) | Pembukaan | mbukaan (harmoni) | Tertinggi | Tertinggi (max) | Terendah | Terendah (min) | Vol.  |
|------------|---------------------|----------|-----------------|-----------|-------------------|-----------|-----------------|----------|----------------|-------|
| 0          | 2022-01-04 00:00... | 591.15   | 594.26          | 599.91    | 602.7465          | 600.41    | 609.99          | 581.60   | 581.60         | 4.39  |
| 0          | 2022-01-05 00:00... | 567.52   | 579.3350        | 592.00    | 595.9288          | 592.84    | 600.41          | 566.88   | 566.88         | 4.15  |
| 0          | 2022-01-06 00:00... | 553.29   | 560.4050        | 554.34    | 572.5514          | 563.36    | 592.84          | 542.01   | 542.01         | 5.71  |
| 0          | 2022-01-07 00:00... | 541.06   | 547.1750        | 549.46    | 551.8892          | 553.43    | 563.36          | 538.22   | 538.22         | 3.38  |
| 1          | 2022-01-10 00:00... | 539.85   | 540.4550        | 538.49    | 543.9197          | 543.69    | 553.43          | 526.32   | 526.32         | 4.49  |
| 1          | 2022-01-11 00:00... | 540.84   | 540.3450        | 536.99    | 537.7390          | 543.91    | 543.91          | 530.07   | 526.32         | 3.08  |
| 0          | 2022-01-12 00:00... | 537.22   | 539.03          | 544.27    | 540.6055          | 544.27    | 544.27          | 532.02   | 530.07         | 3.79  |
| 0          | 2022-01-13 00:00... | 519.20   | 528.21          | 537.06    | 540.6410          | 540.79    | 544.27          | 518.26   | 518.26         | 4.48  |
| 1          | 2022-01-14 00:00... | 525.69   | 522.4450        | 517.60    | 527.1505          | 538.37    | 540.79          | 511.88   | 511.88         | 7.86  |
| 0          | 2022-01-18 00:00... | 510.80   | 518.2450        | 520.08    | 518.8370          | 521.75    | 538.37          | 508.68   | 508.68         | 4.84  |
| 1          | 2022-01-19 00:00... | 515.86   | 513.33          | 515.00    | 517.5275          | 523.21    | 523.21          | 510.51   | 508.68         | 4.35  |
| 0          | 2022-01-20 00:00... | 508.25   | 512.0550        | 517.75    | 516.3713          | 526.64    | 526.64          | 506.93   | 506.93         | 12.66 |
| 0          | 2022-01-21 00:00... | 397.50   | 452.8750        | 400.42    | 451.5884          | 409.15    | 526.64          | 379.99   | 379.99         | 58.90 |
| 1          | 2022-01-24 00:00... | 387.15   | 392.3250        | 383.91    | 391.9912          | 387.26    | 409.15          | 351.46   | 351.46         | 32.39 |
| 0          | 2022-01-25 00:00... | 366.42   | 376.7850        | 379.14    | 381.5101          | 387.71    | 387.71          | 365.13   | 351.46         | 15.15 |
| 0          | 2022-01-26 00:00... | 359.70   | 363.06          | 378.27    | 378.7045          | 382.66    | 387.71          | 356.62   | 356.62         | 12.68 |
| 1          | 2022-01-27 00:00... | 386.70   | 373.20          | 382.06    | 380.1556          | 394.80    | 394.80          | 378.10   | 356.62         | 24.32 |
| 0          | 2022-01-28 00:00... | 384.36   | 385.53          | 386.76    | 384.3956          | 387.00    | 394.80          | 372.08   | 372.08         | 11.97 |
| 1          | 2022-01-31 00:00... | 427.14   | 405.75          | 401.97    | 394.2183          | 427.70    | 427.70          | 398.20   | 372.08         | 20.05 |
| 1          | 2022-02-01 00:00... | 457.13   | 447.1350        | 432.96    | 416.8899          | 458.48    | 458.48          | 425.54   | 398.20         | 22.57 |

**Gambar 4.** Hasil Normalisasi *Dataset*

Terakhir, dilakukan splitting data menjadi dua bagian: 80% untuk data pelatihan dan 20% untuk data pengujian, yang dilakukan secara acak namun tetap mempertahankan proporsi kelas agar tidak terjadi ketidakseimbangan, adapun rinciannya dapat terlihat pada tabel 2 berikut.

**Tabel 2.** Hasil *Splitting Dataset*

| Jenis      | Persentase (%) | Jumlah |
|------------|----------------|--------|
| Data Train | 80%            | 401    |
| Data Test  | 20%            | 100    |

### 4.3 Hasil Implementasi Model

Penelitian ini mengimplementasikan tiga algoritma klasifikasi supervised learning menggunakan perangkat lunak Orange Data Mining, yaitu Support Vector Machine (SVM), Neural Network, dan Naive Bayes.

1. Model SVM menghasilkan pemisahan data dengan cukup baik pada data pelatihan karena mampu menangani data non-linear dengan kernel trick. Hasil visualisasi pada Orange menunjukkan margin pemisah yang optimal antara dua kelas (*Naik* dan *Turun*).
2. Model Neural Network menunjukkan hasil yang menjanjikan dengan kemampuan adaptif terhadap kompleksitas pola dalam data historis. Jumlah hidden layer dan neuron disesuaikan agar tidak terjadi overfitting.
3. Model Naive Bayes, meskipun sederhana, mampu memberikan hasil prediksi cepat dengan asumsi independensi antar fitur. Ini cocok diterapkan pada kasus prediksi cepat dan real-time, walaupun akurasi cenderung sedikit lebih rendah dibandingkan model lainnya.

### 4.4 Hasil Pengujian Model

Hasil Pengujian model dilakukan menggunakan widget Test and Score di Orange dengan teknik Random Sampling sebanyak 10 kali percobaan untuk mengukur konsistensi hasil serta 100 data acak. Parameter evaluasi yang digunakan mencakup:

1. Accuracy (CA)
2. Area Under Curve (AUC)
3. Precision
4. Recall
5. F1-Score
6. Matthews Correlation Coefficient (MCC)

Hasil rata-rata dari masing-masing algoritma terlihat pada gambar 5 berikut.

| Model          | AUC   | CA    | F1    | Prec  | Recall | MCC   |
|----------------|-------|-------|-------|-------|--------|-------|
| SVM            | 0.921 | 0.843 | 0.843 | 0.847 | 0.843  | 0.690 |
| Neural Network | 0.937 | 0.880 | 0.880 | 0.880 | 0.880  | 0.759 |
| Naive Bayes    | 0.903 | 0.830 | 0.830 | 0.830 | 0.830  | 0.660 |

**Gambar 5.** Train and Test Random Sampling

Model Neural Network menunjukkan performa paling konsisten, sedangkan Naive Bayes unggul dari segi kecepatan pemrosesan.

### 4.5 Evaluasi Kinerja

Untuk mengevaluasi kinerja akhir model, digunakan Confusion Matrix terlihat pada gambar 6.

|                              |          | Predicted |      |       |
|------------------------------|----------|-----------|------|-------|
|                              |          | 0         | 1    |       |
|                              |          | 0         | 1    | Σ     |
| Support Vector Machine (SVM) | Actual 0 | 4059      | 1041 | 5100  |
|                              | Actual 1 | 531       | 4369 | 4900  |
|                              | Σ        | 4590      | 5410 | 10000 |

|                |          | Predicted |      |       |
|----------------|----------|-----------|------|-------|
|                |          | 0         | 1    |       |
|                |          | 0         | 1    | Σ     |
| Neural Network | Actual 0 | 4526      | 574  | 5100  |
|                | Actual 1 | 630       | 4270 | 4900  |
|                | Σ        | 5156      | 4844 | 10000 |

|             |          | Predicted |      |       |
|-------------|----------|-----------|------|-------|
|             |          | 0         | 1    |       |
|             |          | 0         | 1    | Σ     |
| Naive Bayes | Actual 0 | 4270      | 830  | 5100  |
|             | Actual 1 | 871       | 4029 | 4900  |
|             | Σ        | 5141      | 4859 | 10000 |

**Gambar 6.** Hasil Confusion matrix 3 Model

Evaluasi dilakukan terhadap data uji menggunakan masing-masing model untuk melihat detail prediksi benar dan salah, diantaranya sebagai berikut.

1. *Support Vector Machine* (SVM) diperoleh nilai AUC sebesar 0,921 nilai CA sebesar 0,843 nilai F1 sebesar 0,843 nilai *precision* sebesar 0,847 nilai *Recall* sebesar 0,843 dan MCC 0,690.
2. *Neural Network* diperoleh nilai AUC sebesar 0,937 nilai CA sebesar 0,880 nilai F1 sebesar 0,880 nilai *precision* sebesar 0,880 nilai *Recall* sebesar 0,880 dan MCC 0,759.

3. *Naive Bayes* diperoleh nilai AUC sebesar 0,903 nilai CA sebesar 0,830 nilai F1 sebesar 0,830 nilai *precision* sebesar 0,830 nilai *Recall* sebesar 0,830 dan MCC 0,660.

## KESIMPULAN

Penelitian ini bertujuan untuk memprediksi harga saham Netflix Inc. (NFLX) di bursa NASDAQ menggunakan algoritma machine learning berbasis data mining. Dengan memanfaatkan data historis selama dua tahun (2022-2023) yang mencakup variabel seperti harga pembukaan, penutupan, tertinggi, terendah, dan volume perdagangan, penelitian ini mengimplementasikan tiga algoritma klasifikasi, yaitu Support Vector Machine (SVM), Neural Network, dan Naive Bayes.

Melalui perangkat lunak Orange Data Mining. Hasil evaluasi menunjukkan bahwa Neural Network memberikan performa terbaik dengan nilai akurasi (CA) sebesar 0,880, AUC 0,937, dan F1-Score 0,880, diikuti oleh SVM dengan akurasi 0,843 dan Naive Bayes dengan akurasi 0,830. Neural Network juga unggul dalam konsistensi dan kemampuan adaptasi terhadap pola data yang kompleks, sementara Naive Bayes menonjol dalam kecepatan pemrosesan meskipun akurasinya lebih rendah.

Penelitian ini membuktikan bahwa pendekatan machine learning, khususnya Neural Network, dapat menjadi alat yang efektif untuk memprediksi pergerakan harga saham dengan akurasi tinggi, sehingga dapat mendukung pengambilan keputusan investasi yang lebih informasional dan minim risiko.

## DAFTAR PUSTAKA

- Abdul Dwiyanto Suyudi, M., Djamal, E. C., Maspupah Jurusan Informatika, A., & Sains dan Informatika Universitas Jenderal Achmad Yani Cimahi, F. (2019). Prediksi Harga Saham menggunakan Metode Recurrent Neural Network. In *Seminar Nasional Aplikasi Teknologi Informasi (SNATi)*.
- Almaspoor, M. H., Safaei, A., Salajegheh, A., & Minaei-Bidgoli, B. (2021). *Support Vector Machines in Big Data Classification: A Systematic Literature Review*. <https://doi.org/10.21203/rs.3.rs-663359/v1>
- Amelia Sekar Ayu Pramesti, Dara Kesumadewi, & Zona Oktavia. (2024). Pemahaman Investasi, Motivasi Investasi Dalam Minat Masyarakat Berinvestasi Di Pasar Modal. *Mutiara : Jurnal Penelitian Dan Karya Ilmiah*, 2(1), 341–355. <https://doi.org/10.59059/mutiara.v2i1.1012>
- Bismi, W., Novianti, D., & Qomaruddin, M. (2024). Analisis Perbandingan Klasifikasi Citra Genus Panthera dengan Pendekatan Deep learning Model MobileNet. *Jurnal Informatika Dan Rekayasa Perangkat Lunak*, 6(1). <https://publikasiilmiah.unwahas.ac.id/JINRPL/article/view/9037/pdf>
- Eka Patriya, E. (2020). Implementasi Support Vector Machine Pada Prediksi Harga Saham Gabungan (Ihsg). *Jurnal Ilmiah Teknologi Dan Rekayasa*, 25(1), 24–38. <https://doi.org/10.35760/tr.2020.v25i1.2571>
- Fadilah, W. R. U., Agfiannisa, D., & Azhar, Y. (2020). Analisis Prediksi Harga Saham PT. Telekomunikasi Indonesia Menggunakan Metode Support Vector Machine. *Fountain of Informatics Journal*, 5(2), 45. <https://doi.org/10.21111/fij.v5i2.4449>
- Fatra, R. A. (2021). *ANALISIS PREDIKSI HARGA SAHAM PT BLUE BIRD TBK DITENGAH PANDEMI COVID-19 DENGAN METODE ARIMA DALAM PERSPEKTIF EKONOMI ISLAM [UNIVERSITAS ISLAM NEGERI]*. <https://repository.radenintan.ac.id/17424/2/SKRIPSI%20FULL.pdf>
- Hastomo, W., Satyo, A., Karno, B., Kalbuana, N., Nisfiani, E., & Etp, L. (2021). *Optimasi Deep Learning untuk Prediksi Saham di Masa Pandemi Covid-19*.
- Irmayani, W., & Jayanti, W. E. (2025). Visualisasi Prediksi Harga Saham Menggunakan Model Long Short-Term Memory (LSTM). In *Jurnal Sistem Informasi Akuntansi* (Vol. 6, Issue 1). <http://jurnal.bsi.ac.id/index.php/justian>
- Krishnan, B., O'Farrell, B., Delamaide, D., Smith, G., Burke, N., Cohen, P., & Gecgil,

- T. (2020). *Investing*. Fusion Media. <https://www.investing.com/>
- Leovinent, A., & Yoannita, Y. (2023). Klasifikasi Ras Anjing Berdasarkan Citra Menggunakan Convolutional Neural Network. *Jurnal Algoritme*, 3(2), 160–169. <https://doi.org/10.35957/algoritme.v3i2.3389>
- Otchere, D. A., Arbi Ganat, T. O., Gholami, R., & Ridha, S. (2021). Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: Comparative analysis of ANN and SVM models. *Journal of Petroleum Science and Engineering*, 200, 108182. <https://doi.org/10.1016/j.petrol.2020.108182>
- Rouf, N., Malik, M. B., Arif, T., Sharma, S., Singh, S., Aich, S., & Kim, H. C. (2021). Stock market prediction using machine learning techniques: A decade survey on methodologies, recent developments, and future directions. In *Electronics (Switzerland)* (Vol. 10, Issue 21). MDPI. <https://doi.org/10.3390/electronics10212717>
- Saputro, D. T., & Swanjaya, D. (2023). Analisa Prediksi Harga Saham Menggunakan Neural Network dan Net Foreign Flow. In *Generation Journal* (Vol. 7, Issue 2).
- Suhartini, Wijaya, L. kerta, & Pratiwi Nur Arini. (2020). Penerapan Algoritma K-Means Untuk Pendataan Obat Berdasarkan Laporan Bulanan Pada Dinas Kesehatan Kabupaten Lombok Timur. *Infotek : Jurnal Informatika Dan Teknologi*, 3, 147–157. [https://ejournal.hamzanwadi.ac.id/index.php/infotek/article/view/2315/pdf\\_36](https://ejournal.hamzanwadi.ac.id/index.php/infotek/article/view/2315/pdf_36)
- Sumanto, Sugiarti, Y., Supriyatna, A., Carolina, I., Amin, R., & Yani, A. (2021). Model Naïve Bayes Classifiers For Detection Apple Diseases. *9th International Conference on Cyber and IT Service Management (CITSM)*, 1–4. <https://doi.org/10.1109/CITSM52892.2021.9588801>
- Syahdan Nabil, P., & Ramdhani, Y. (2023). Application of bidirectional gated recurrent unit algorithm for rainfall prediction. In *Jurnal Teknik Informatika C.I.T Medicom* (Vol. 15, Issue 4).
- Tambunan, D. (2020). *Investasi Saham di Masa Pandemi COVID-19*. 4(2), 117–123.
- Veziroğlu, M., Veziroğlu, E., & Bucak, İ. Ö. (2023). Performance Comparison between Naive Bayes and Machine Learning Algorithms for News Classification. *Iğdır University & Kütahya Dumlupınar University, 11(tourism)*, 13.