

Implementasi Algoritma Regresi pada Machine Learning untuk Prediksi Indeks Harga Saham Gabungan

Dwi Eko Waluyo¹, Hayu Wikan Kinasih^{2*}, Cinantya Paramita³, Dewi Pergiwati⁴, Fauzi Adi Rafrastara⁵, Rajendra Nohan⁵

¹Program Studi Manajemen, Fakultas Ekonomi dan Bisnis, Universitas Dian Nuswantoro, Semarang

^{2*}Program Studi Akuntansi, Fakultas Ekonomi dan Bisnis, Universitas Dian Nuswantoro, Semarang

^{3,5,6}Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro, Semarang

⁴Program Studi Desain Komunikasi Visual, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro, Semarang
Jln. Imam Bonjol No. 207, Semarang, 50141, Indonesia

email: ¹dwi.eko.waluyo@dsn.dinus.ac.id, ^{2*}hayu.wikan@dsn.dinus.ac.id, ³cinantya.paramita@dsn.dinus.ac.id, ⁴dewiwati7@gmail.com, ⁵fauziadi@dsn.dinus.ac.id, ⁶111202113301@mhs.dinus.ac.id

Abstract – Indeks Harga Saham Gabungan (IHSG) or Indonesia Composite Index (ICI) is part of the macro indicators of a country that describes the economic condition of a country. ICI is an interesting study to research since its existence will be able to show market sentiment regarding an event that occurred in a country. This research tries to predict the ICI in the future based on historical data. The dataset used in this research is publicly available in Yahoo Finance. The experiment is conducted by implementing some regression machine learning algorithms, such as Decision Tree, Random Forest, k-Nearest Neighbor (kNN), and Linear Regression. As a result, Decision Tree has the lowest MSE value compared to other methods: 1268.242. In this research, a website-based application prototype was also developed that can be used to view IHSG graphs and make future predictions, using the 4 (four) tested algorithms.

Abstrak – Indeks Harga Saham Gabungan (IHSG) merupakan salah satu indikator makro sebuah negara yang dapat menggambarkan kondisi perekonomian suatu negara. IHSG merupakan topik yang menarik untuk diteliti mengingat keberadaannya akan mampu memperlihatkan sentiment pasar terhadap sebuah peristiwa yang terjadi pada sebuah negara. Penelitian ini mencoba memprediksi IHSG pada masa mendatang berdasarkan pada data historis. Dataset yang digunakan adalah dataset publik yang diperoleh dari Yahoo Finance. Eksperimen pada penelitian ini dilakukan dengan mengimplementasikan beberapa algoritma regresi pada machine learning, seperti: Decision Tree, Random Forest, k-Nearest Neighbor (kNN), dan Regresi Linier. Hasilnya, algoritma Decision Tree memiliki nilai MSE paling rendah, yaitu: 1268.242. Pada penelitian ini juga dikembangkan sebuah prototype aplikasi berbasis website yang dapat digunakan untuk melihat grafik IHSG dan melakukan prediksi di masa mendatang, menggunakan 4 (empat) algoritma yang diuji.

Kata Kunci – IHSG, prediksi, machine learning, decision tree, random forest, knn, regresi linear

I. PENDAHULUAN

Perkembangan perekonomian sebuah negara tidak lepas dari keberadaan pasar modal. Pasar modal sendiri merupakan lembaga yang memiliki dua fungsi yaitu fungsi ekonomi dan fungsi Investasi. Dalam menjalankan fungsi ekonomi, pasar

modal memiliki andil dalam mempertemukan antara pihak yang memiliki kelebihan dana dan pihak yang memerlukan dana. Pada fungsi keuangan, pasar modal menjalankan fungsi keuangan yaitu memberikan kesempatan bagi seseorang yang memiliki dana untuk dapat memperoleh return.

Menurut data statistik yang dikeluarkan oleh PT Kustodian Sentral Efek Indonesia (KSEI) pada awal tahun 2021, menunjukkan adanya peningkatan yang signifikan pada jumlah investor di pasar modal Indonesia. Pada tahun 2018 sampai dengan 2019, jumlah investor meningkat dari 1.619.372 menjadi 2.484.354, kemudian mengalami peningkatan lagi pada tahun 2020 menjadi sejumlah 3.880.753 [1]. Hal ini menunjukkan adanya peralihan pandangan bisnis dari para investor, yang semula lebih cenderung berinvestasi pada asset riil menjadi investasi pada efek. Peningkatan yang terjadi juga didorong oleh kondisi pandemi yang dialami oleh sebagian besar penduduk dunia, termasuk di Indonesia, dapat diduga kelesuan pada aktivitas riil mendorong masyarakat untuk merubah pandangannya dalam berbisnis.

Pada suatu perekonomian, pasar modal memiliki peran penting dalam hal pemerolehan dana bagi perusahaan serta sarana berinvestasi bagi masyarakat. Salah satu instrument yang diminati dalam pasar modal adalah saham. Saham dianggap sebagai salah satu instrument investasi yang menguntungkan dan menjanjikan tingkat pengembalian yang tinggi. Dalam melakukan investasi saham ini, calon investor dapat terlebih dahulu melakukan analisis atas kinerja sebuah saham, sehingga investor dapat mengambil keputusan untuk menjual atau membeli saham. Kesalahan dalam pengambilan keputusan pada akhirnya akan menyebabkan kerugian pada investor, sehingga perlu dilakukan analisis yang akurat [2]. Terdapat dua macam analisis yang biasa dilakukan dalam berinvestasi saham yaitu dengan menggunakan analisis teknikal (pergerakan saham) ataupun dengan menggunakan analisis fundamental (menggunakan data – data perekonomian) [3].

Dalam analisis teknikal, calon investor dapat mengamati pergerakan saham salah satunya dengan melihat grafik pergerakan harga saham. Sebagai penggambaran bentuk kondisi saham di Indonesia, investor dapat menggunakan Indeks Harga Saham Gabungan (IHSG) sebagai salah satu indikator untuk melihat perubahan harga dari suatu pasar

*) penulis korespondensi: Hayu Wikan Kinasih
Email: hayu.wikan@dsn.dinus.ac.id

dalam kegiatan perdagangan. Selain itu, IHSG juga mencerminkan saham secara keseluruhan di Indonesia [4]. IHSG sendiri dapat dipengaruhi berbagai faktor ekonomi diantaranya inflasi, BI rate dan kurs. Ketika kondisi perekonomian mengalami guncangan, IHSG pun juga mengalami penurunan, seperti halnya pada saat terjadi pandemi covid 19 yang mengakibatkan turunnya IHSG dan sempat dilakukan penutupan perdagangan selama 30 menit pada tanggal 12, 13, 17, 19, 23 dan 30 Maret tahun 2020. Setelah kondisi perekonomian membaik, maka IHSG kembali mengalami peningkatan [5].

Prediksi pergerakan IHSG dirasa memiliki andil dalam membantu investor untuk menentukan keputusan investasi yang akan mereka lakukan. Cara yang dapat dilakukan adalah dengan melakukan prediksi berbasis *machine learning* atas data historis jangka pendek berupa harga pembuka dan penutup IHSG harian yang diperoleh dari Yahoo Finance. Rentang waktu yang dipergunakan adalah periode pasca pandemi COVID 19 yang dianggap sebagai periode titik balik kinerja IHSG, utamanya di Indonesia. Beberapa penelitian telah dilakukan berkaitan dengan prediksi IHSG menggunakan algoritma *machine learning*, seperti Neural Network Backpropagation [6], ARIMA [7], juga Support Vector Machine dan PROPHET [8].

Pada penelitian ini, 4 algoritma dilibatkan guna mendapatkan skor evaluasi terbaik, yaitu Tree, Random Forest, k-Nearest Neighbor, dan Linear Regression. Evaluasi yang digunakan untuk mengukur performa dari ke-empat algoritma tersebut yaitu, Mean Squared Error (MSE) dan R-Squared. Di bagian akhir, akan dibahas produk yang telah dikembangkan, yaitu berupa aplikasi berbasis website yang dapat menampilkan hasil prediksi IHSG menggunakan 4 algoritma machine learning, yaitu Decision Tree, Random Forest, k-Nearest Neighbor (kNN), dan Regresi Linear.

II. PENELITIAN YANG TERKAIT

A. Analisis Teknikal

Analisis teknikal merupakan salah satu bentuk analisis penilaian saham yang dilakukan dengan menggunakan bantuan grafik atas data historis dari harga dan volume perdagangan saham. Analisis ini dimaksudkan untuk memberikan bantuan kepada calon investor ataupun investor dalam berinvestasi [4]. Terdapat beberapa hal yang perlu diperhatikan oleh investor dalam melakukan analisis teknikal, antara lain:

1. Keberadaan *support* dan *resistance level*. *Support* merupakan kondisi dimana saham akan berada pada level dimana permintaan cukup kuat untuk menahan harga mengalami penurunan, sedangkan *resistance* merupakan tingkatan dimana terjadi kecenderungan untuk menjual yang kuat dalam rangka menahan laju kenaikan harga.
2. Tren Pasar. Tren pasar ini memberikan kecenderungan bagi harga saham untuk bergerak dari waktu ke waktu ke arah tertentu. Terdapat tiga tren yaitu *uptrend*, *downtrend* dan *sideways*.

B. Data Mining

Data mining merupakan suatu proses mengekstrak pengetahuan dari tumpukan data guna menemukan pola atau hubungan pada data [9]. Dengan proses penambangan yang dilakukan terhadap data-data yang lampau (data historis),

banyak hal yang bisa dilakukan mulai dari kegiatan klasifikasi, klusterisasi, forecasting, dan juga estimasi. Salah satu penerapan data mining dalam dunia nyata adalah pada bidang ekonomi, yaitu untuk memprediksi pergerakan Indeks Harga Saham Gabungan (IHSG). Prediksi pergerakan IHSG dapat dilakukan dengan salah satu metode di dalam data mining, yaitu *Forecasting*. Ciri khas metode *forecasting* yaitu memiliki *class data* yang bernilai numerik dan data disusun secara rentet waktu (*time series*) [10]. Beberapa contoh algoritma *forecasting* yang populer adalah Neural Network, KNN, Decision Tree, Random Forest, dan Linear Regression.

Beberapa penelitian yang pernah dilakukan berkaitan dengan prediksi IHSG antara lain dilakukan oleh Barus dan Wijaya [6], yang menggunakan metode Neural Network Backpropagation dalam melakukan prediksi. Hasil pengujian dengan menggunakan algoritman Neural Network Backpropagation menunjukkan akurasi sebesar 63,06%. Penelitian yang dilakukan berkaitan dengan prediksi juga dilakukan oleh Siregar, Pangruruk, & Wijaya [8] yang menggunakan beberapa metode dalam melakukan prediksi terhadap IHSG antara lain dengan menggunakan metode ETS, LM, SVM, SVM_RBF, NNAR, dan PROPHET. Hasil percobaan yang dilakukan menunjukkan bahwa PROPHET memberikan hasil yang paling baik dalam memprediksi IHSG. Penelitian lain berkaitan dengan prediksi IHSG dilakukan oleh Susanti & Adji [7] yang menggunakan model ARIMA (Autoregressive Integrated Moving Average). Hasil menunjukkan bahwa model ARIMA terbaik yang digunakan dalam memprediksi adalah model ARIMA.

Dalam penelitian ini, empat buah algoritma *forecasting* akan diimplementasikan pada kasus prediksi pergerakan IHSG. Empat algoritma tersebut, yaitu: Decision Tree, Random Forest, kNN, dan Regresi Linear. Algoritma-algoritma tersebut akan diuji dan dibandingkan untuk mendapatkan algoritma terbaik dalam kasus prediksi pergerakan IHSG. Di bagian akhir, 4 algoritma tersebut akan dideploy ke dalam sebuah prototipe aplikasi berbasis website yang dapat digunakan untuk melihat hasil prediksi IHSG pada masa mendatang berdasarkan data historis yang telah disediakan.

III. METODE PENELITIAN

Dalam menyelesaikan penelitian ini, beberapa tahapan dilakukan, mulai dari Data Collection, Pre-Processing, Modelling, Evaluation, dan Deployment (Gambar 1).

A. Hardware dan Software

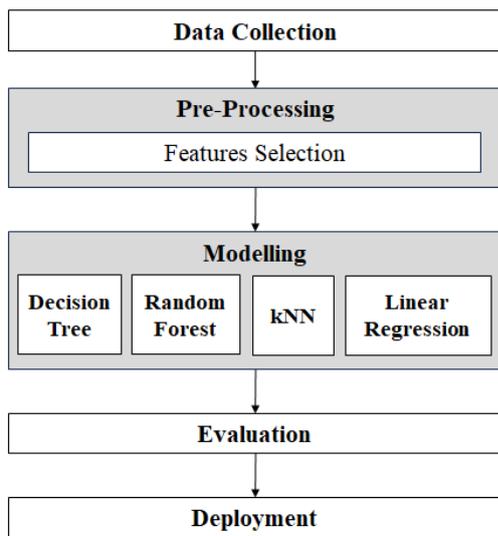
Tidak dapat dipungkiri bahwa salah satu faktor yang menentukan kelancaran proses penelitian adalah instrumen-instrumen pendukung, yang diantaranya adalah Hardware dan Software. Software yang baik tanpa dukungan dari perangkat keras yang memadai, tidak akan dapat berjalan dengan optimal. Sebaliknya, hardware dengan kualitas yang sangat baik tanpa diimbangi dengan software yang tepat juga tidak akan banyak membantu. Oleh karena itu, baik hardware maupun software memiliki peranan penting dalam mendukung kelancaran suatu penelitian.

Pada penelitian ini, perangkat komputer yang digunakan memiliki spesifikasi sebagai berikut:

- Processor : Intel Xeon E5620

- RAM : 16 GB
- Harddisk : 3 TB
- VGA : Radeon RX550

Sementara itu, software yang digunakan pada penelitian adalah Orange Data Mining (*orangedatamining.com*) dan Anaconda. Software Orange digunakan untuk melakukan *pre-processing* data sekaligus melakukan simulasi prediksi dan evaluasi secara cepat. Sedangkan Anaconda digunakan untuk melakukan simulasi berbasis *coding* dengan bahasa python, sekaligus *men-deploy*-nya hingga menjadi aplikasi berbasis website dengan dukungan dari GitHub (*github.com*) dan juga Streamlit (*streamlit.io*).



Gbr 1. Tahapan penelitian yang digunakan

B. Data Collection

Jenis data yang dipergunakan pada penelitian ini adalah data sekunder yang diperoleh dari Yahoo Finance. Data diambil dari rentang 21 Juni 2023 sampai dengan 31 Agustus 2023. Pemilihan batas atas rentang ini dikaitkan dengan pencabutan status pandemi oleh Presiden Joko Widodo pada tanggal 21 Juni 2023 menjadi endemi, yang mengakibatkan kinerja IHSG bergerak ke arah positif.

Dataset yang terdownload memiliki 847 baris data dan 7 fitur (Date, Open, High, Low, Close, Adj Close, dan Volume). Dari ke-tujuh atribut tersebut, kelak yang digunakan hanyalah fitur Date, Open dan Close, dimana fitur Date bertindak sebagai meta atribut.

C. Pre-processing

Pada tahap *pre-processing* ini, terdapat sebuah langkah penting agar dataset siap digunakan untuk modelling, yaitu *features selection* atau seleksi fitur. Seleksi fitur merupakan kegiatan memilih fitur-fitur yang dianggap memiliki kontribusi langsung dan signifikan pada kelas data atau variabel dependen. Pada penelitian ini, seleksi fitur juga dapat disebut dengan reduksi fitur, yaitu menghapus fitur-fitur yang dianggap tidak memberikan kontribusi atau tidak berkaitan langsung dengan tujuan pemodelan yang akan dilakukan. Dataset IHSG pada awalnya memiliki 5 fitur, yaitu *date*, *open*, *high*, *low*, dan *close*. Pada tahap reduksi fitur, fitur *high* dan

low dihapus sehingga menyisakan 3 fitur, yaitu *date*, *open*, dan *close*. Adapun fitur *date* selanjutnya diubah menjadi *meta-attribute* sehingga dapat diabaikan saat proses *modelling*.

D. Modelling

Modelling merupakan tahapan dimana algoritma data mining diimplementasikan untuk mencapai tujuan penelitian, yaitu prediksi IHSG. Terdapat 4 algoritma prediksi yang dilibatkan dalam eksperimen ini, yaitu Decision Tree, Random Forest, kNN, dan Regresi Linear. Ke-empat algoritma tersebut dapat menangani kasus regresi, yaitu memprediksi kelas dengan nilai numerik.

Decision Tree merupakan algoritma populer, baik untuk keperluan klasifikasi maupun regresi. Pada dasarnya, decision tree untuk klasifikasi memiliki kemiripan dengan regresi. Perbedaan utamanya ada pada penggunaan *Standard Deviation Reduction* (SDR) pada regresi, sementara *information gain* digunakan pada klasifikasi. Oleh karena itu, ada 3 formula dasar yang perlu diperhatikan saat menggunakan decision tree pada kasus regresi (lihat persamaan 1, 2, dan 3) [11], [12].

$$Variance = \frac{\sum_{i=1}^n (x_i - \mu)^2}{n-1} \tag{1}$$

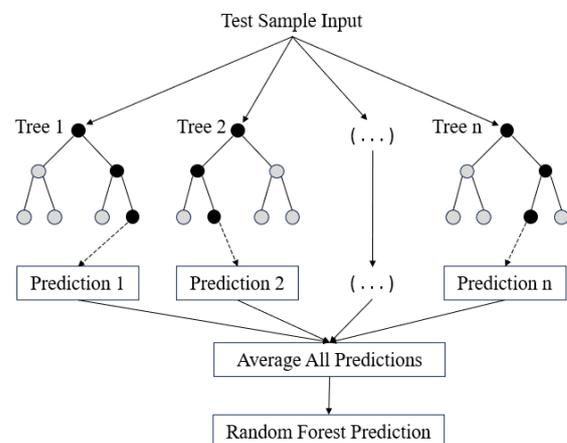
$$Standard\ Deviation = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n-1}} \tag{2}$$

$$\mu (mean) = \frac{\sum_{i=1}^n x_i}{\sum_{i=1}^n f_i} \tag{3}$$

Dimana:

- n* : dimensi data
- i* : variabel data
- x* : data
- μ : mean atau rata-rata

Standar deviasi dan varian digunakan untuk menunjukkan seberapa menyebar distribusi data yang dimiliki [13]. Nilai standar deviasi yang rendah bermakna bahwa pengamatan data cenderung sangat dekat dengan nilai *mean*. Sebaliknya, jika nilai standar deviasinya tinggi, maka dapat disimpulkan bahwa suatu data tersebut pada rentang nilai yang besar. Standar Deviasi diperoleh dengan menghitung nilai akar kuadrat dari varian. Pada persamaan 1, 2 dan 3, nilai *n* merujuk pada jumlah total keseluruhan data, sedangkan *x* merujuk pada data tertentu.



Gbr 2. Ilustrasi Random Forest pada kasus regresi

Random Forest merupakan algoritma *ensemble* yang terdiri dari gabungan beberapa decision tree [14], [15]. Pada kasus klasifikasi, voting digunakan untuk menentukan hasil kelas yang diprediksi. Sedangkan pada kasus regresi, hasil prediksi akan diperoleh dengan menghitung nilai rata-rata. Ilustrasi random forest pada kasus regresi dapat dilihat pada Gambar 2.

K-Nearest Neighbor (kNN) merupakan algoritma non-parametrik yang umum digunakan pada kasus klasifikasi maupun regresi. Untuk memprediksi data numerik dan *time series* menggunakan kNN, beberapa tahap perlu dilalui, seperti menentukan nilai *k*, pemilihan variabel input, strategi untuk memprediksi 1 atau lebih nilai di depan (*multi-step ahead*), metrik pengukuran, hingga bagaimana cara mengagregasi output yang diperoleh [16]. kNN menggunakan jarak sebagai acuan dalam menentukan tetangga terdekat. Metode pengukuran jarak yang sering digunakan di kNN adalah Euclidean Distance (Persamaan 4).

$$d(p, q) = \sqrt{\sum_{i=2}^n (q_i - p_i)^2} \tag{4}$$

Dimana:

- d* : distance atau jarak
- p* : data sample atau data training
- q* : data uji atau data testing
- i* : variabel data
- n* : dimensi data

Regresi linear merupakan algoritma regresi yang dapat digunakan untuk melakukan estimasi maupun *forecasting*. Secara umum, regresi linear dibagi menjadi 2, yaitu regresi linear sederhana dan regresi linear berganda. Regresi linear sederhana digunakan jika terdiri dari 1 variabel bebas (*independent*) dan 1 variabel terikat (*dependent*). Sedangkan regresi linear berganda digunakan jika memiliki lebih dari 1 variabel bebas. Mengingat penelitian ini hanya melibatkan 1 variabel bebas (variabel Open), maka jenis algoritma regresi linear yang digunakan adalah regresi linear sederhana. Formula dari regresi linear sederhana dapat dilihat pada persamaan 5 [17].

$$Y = a + bx \tag{5}$$

Keterangan:

- Y* : nilai yang diprediksi (variabel terikat/*dependent*)
- a* : konstanta
- x* : variabel bebas (*independent*)
- b* : koefisien regresi dari variabel *x*

E. Evaluation

Setelah 4 (empat) algoritma regresi diimplementasikan, berikutnya akan dievaluasi menggunakan matrix MSE. MSE atau Mean Square Error merupakan rataan selisih kuadrat antara nilai yang diprediksikan dengan diamati [17]. Semakin rendah nilai MSE menunjukkan semakin baik performanya. Formula untuk menghitung nilai MSE dapat dilihat pada persamaan 6.

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \tag{6}$$

Dimana:

- MSE* : Mean Squared Error

- n* : dimensi data
- Y_i* : nilai yang diobservasi
- Ŷ_i* : nilai yang diprediksi

Skor MSE dari 4 (empat) algoritma tersebut selanjutnya akan dibandingkan untuk mendapatkan algoritma regresi terbaik dalam kasus prediksi IHSG.

F. Deployment

Model Deployment merupakan langkah paling menantang dalam *machine learning life cycle* [18], karena pengembang harus mengintegrasikan model *machine learning* ke lingkungan sistem agar dapat diakses secara lebih sederhana oleh seluruh user yang berkepentingan. Pada tahap ini, bahasa pemrograman python, dan *tools* lain berbasis website seperti github.com dan streamlit.io digunakan untuk mendeploy model sehingga bisa diakses secara umum melalui aplikasi prototype berbasis website.

IV. HASIL DAN PEMBAHASAN

Data yang digunakan pada penelitian ini adalah data IHSG mulai tanggal 2 Maret 2020 hingga 29 Agustus 2023. Tabel 1 menunjukkan hasil pengujian performa 4 (empat) algoritma regresi, yaitu Decision Tree, Random Forest, kNN, dan Linear Regression menggunakan metrik MSE. Berdasarkan dari data yang diperoleh, Decision Tree mengungguli 3 algoritma yang lain dengan skor MSE yang paling kecil, yaitu 1268.343. Hal ini menunjukkan bahwa Decision Tree menghasilkan kesalahan peralaman yang lebih kecil daripada Random Forest, juga kNN dan Linear Regression, khususnya untuk memprediksi IHSG.

TABEL 1.
HASIL PERBANDINGAN PERFORMA ALGORITMA DECISION TREE, RANDOM FOREST, KNN, DAN LINEAR REGRESSION

Model	MSE
Decision Tree	1268.343
Random Forest	1344.187
kNN	2555.347
Linear Regression	3257.452

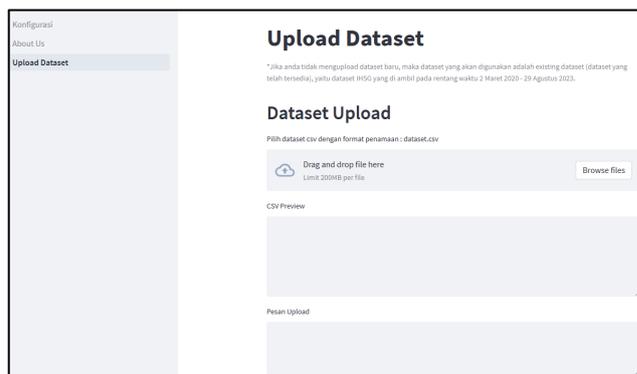
TABEL 2.
HASIL PREDIKSI PENUTUPAN IHSG PADA TANGGAL 30 AGUSTUS 2023 – 3 SEPTEMBER 2023 MENGGUNAKAN ALGORITMA DECISION TREE.

Date	Open	Prediksi Close (DT)
2023-08-30	6957.83	6951.12
2023-08-31	6913.66	6968.78
2023-09-01	6885.75	6988.16
2023-09-02	6859.31	6892.57
2023-09-03	6868.84	6818.75

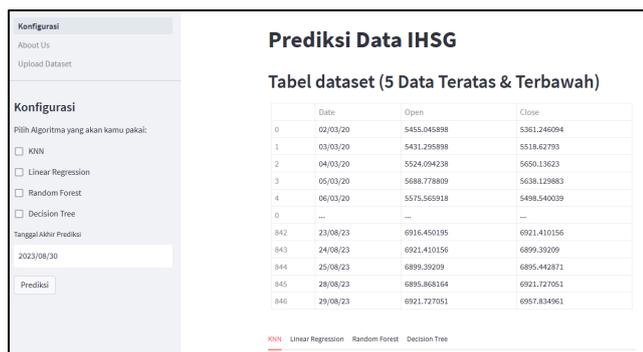
Mengingat Decision Tree menjadi algoritma terbaik di antara 4 algoritma yang dievaluasi, maka simulasi prediksi IHSG yang diterapkan adalah menggunakan algoritma Decision Tree. Hasilnya adalah sebagaimana terlihat pada Tabel 2.

Setelah model berhasil diimplementasikan, maka tahap akhir dari penelitian ini adalah melakukan *deployment* sehingga dapat diakses oleh *user* dengan lebih sederhana. Dengan menggunakan bahasa pemrograman Python serta *tools*

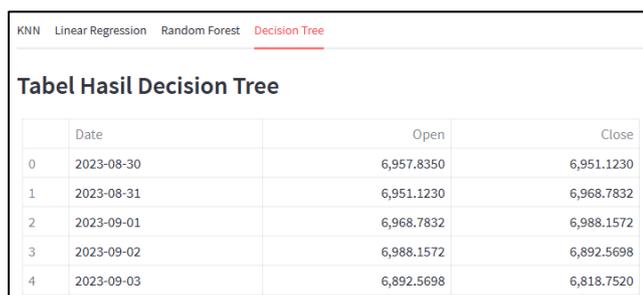
berbasis web, seperti github.com dan streamlit.io, sebuah prototype aplikasi prediksi IHSG berhasil dibuat dan dapat diakses di url: *prediction-ihsg.streamlit.io*. Gambar 3 menunjukkan halaman utama dari aplikasi tersebut. Pada aplikasi ini, user dapat membandingkan hasil prediksi IHSG dalam rentang waktu tertentu dengan menggunakan 4 algoritma sekaligus, yaitu Decision Tree, Random Forest, kNN, dan Regresi Linear (Gambar 4). Gambar 5 menunjukkan contoh hasil prediksi ke 4 algoritma untuk tanggal 30 agustus hingga 3 september 2023. Selain menggunakan dataset yang telah disediakan, user juga dapat mengganti datasetnya dengan dataset IHSG yang baru, untuk membuat skenario prediksi pada momentum yang lain (Gambar 6).



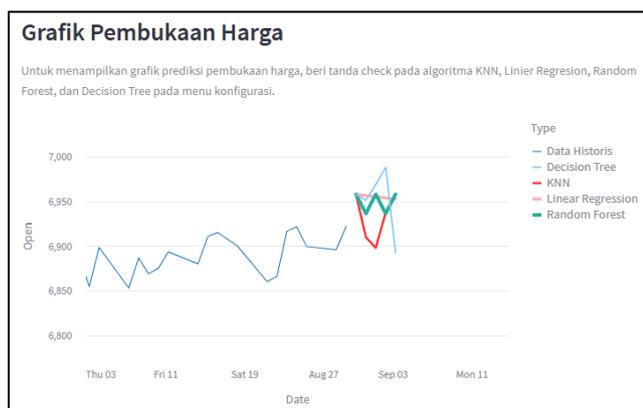
Gbr 6. Halaman untuk upload dataset.



Gbr 3. Halaman utama aplikasi prediksi data IHSG.



Gbr 4. Contoh hasil prediksi menggunakan algoritma Decision Tree.



Gbr 5. Hasil prediksi selama IHSG pada tanggal 3 Agustus 2023 – 3 September 2023 menggunakan 4 algoritma.

V. KESIMPULAN

Prediksi pergerakan IHSG dirasa memiliki andil dalam membantu investor untuk menentukan keputusan investasi yang akan mereka lakukan. Machine learning dapat digunakan untuk memprediksi pergerakan IHSG. Pada penelitian ini, 4 (empat) algoritma regresi diimplementasikan untuk memprediksi pergerakan IHSG pada tanggal 30 Agustus 2023 hingga 3 Oktober. Hasilnya, algoritma Decision Tree memiliki performa yang paling baik dibandingkan 3 (tiga) algoritma lain, seperti Random Forest, k-Nearest Neighbor, dan Linear Regression.

UCAPAN TERIMA KASIH

Penulis mengucapkan banyak terima kasih kepada seluruh pihak yang telah mendukung terwujudnya penelitian ini. Tak lupa penulis sampaikan ucapan terima kasih tak terhingga kepada Universitas Dian Nuswantoro dan juga LPPM Universitas Dian Nuswantoro.

DAFTAR PUSTAKA

- [1] S. R. Fadly, "Aktivitas Pasar Modal Indonesia di Era Pandemi," Kementerian Keuangan Republik Indonesia. Accessed: Apr. 11, 2023. [Online]. Available: <https://www.djkn.kemenkeu.go.id/kpkn-kupang/baca-artikel/13817/Aktivitas-Pasar-Modal-Indonesia-Di-Era-Pandemi.html>
- [2] A. Sadeq, "Analisis Prediksi Indeks Harga Saham Gabungan dengan Metode ARIMA (Studi pada IHSG di Bursa Efek Jakarta)," Universitas Diponegoro, Semarang, 2008.
- [3] E. F. Brigham and J. F. Houston, *Dasar-Dasar Manajemen Keuangan*, 4th ed. Jakarta: Salemba, 2019.
- [4] N. Hafizah, E. Noviani, and H. Perdana, "Analisis Teknikal Saham LQ-45 Menggunakan Indikator Bollinger Bands," *Bimaster*, vol. 8, no. 4, Oct. 2019, doi: 10.26418/bbimst.v8i4.36653.
- [5] E. L. Maretha and F. X. H. Prasetya, "Studi Peristiwa Masa Awal Pandemi Covid-19 dan Prediksi Analisis Teknikal Saham Perbankan Indonesia," *Jurnal Pasar Modal dan Bisnis*, vol. 3, no. 1, Feb. 2021.
- [6] O. Barus and C. Wijaya, "IMPLEMENTASI METODE NEURAL NETWORK BACKPROPAGATION DALAM PREDIKSI INDEKS HARGA SAHAM

- GABUNGAN (IHSG),” *seminastika*, vol. 3, no. 1, pp. 79–85, Nov. 2021, doi: 10.47002/seminastika.v3i1.252.
- [7] R. Susanti and A. R. Adji, “Analisis Peramalan IHSG dengan Time Seris Modelling ARIMA,” *J. Manaj. Kewirausahaan*, vol. 17, no. 1, p. 97, Jun. 2020, doi: 10.33370/jmk.v17i1.393.
- [8] B. Siregar, F. A. Pangruruk, and P. A. Widjaja, “Comparison of Various Forecasting Models for the Jakarta Stock Exchange Composite (JKSE) During the Covid-19 Pandemic,” *Jurnal Multidisiplin Madani (MUDIMA)*, vol. 2, no. 2, pp. 1035–1046, 2022.
- [9] J. Duque, F. Silva, and A. Godinho, “Data Mining applied to Knowledge Management,” *Procedia Computer Science*, vol. 219, pp. 455–461, 2023, doi: 10.1016/j.procs.2023.01.312.
- [10] E. Ernawati, S. S. K. Baharin, and F. Kasmin, “A review of data mining methods in RFM-based customer segmentation,” *J. Phys.: Conf. Ser.*, vol. 1869, no. 1, p. 012085, Apr. 2021, doi: 10.1088/1742-6596/1869/1/012085.
- [11] M. Martinez and M. Bartholomew, “What Does It ‘Mean’? A Review of Interpreting and Calculating Different Types of Means and Standard Deviations,” *Pharmaceutics*, vol. 9, no. 4, p. 14, Apr. 2017, doi: 10.3390/pharmaceutics9020014.
- [12] E. Windia Ambarsari, R. Avrizal, E. Doni Sirait, S. Dwiasnati, and R. Rahim, “Regression Tree Role for Interpret Monetizing of Game Live Streaming,” *J. Phys.: Conf. Ser.*, vol. 1424, no. 1, p. 012014, Dec. 2019, doi: 10.1088/1742-6596/1424/1/012014.
- [13] R. Fitriana, A. N. Habyba, and E. Febriani, *Data Mining dan Aplikasinya: Contoh Kasus di Industri Manufaktur dan Jasa*. Banyumas: Wawasan Ilmu, 2022.
- [14] Cinantya Paramita, Catur Supriyanto, Y. P. Astuti, Lukman Afi Syariffudin, and Fauzi Adi Rafrastara, “Performance Evaluation of Random Forest Algorithm in Cluster Environment,” *IJCSIS*, vol. 19, no. 12, pp. 9–12, Jan. 2022, doi: 10.5281/ZENODO.5852758.
- [15] F. A. Rafrastara, C. Supriyanto, C. Paramita, Y. P. Astuti, and F. Ahmed, “Performance Improvement of Random Forest Algorithm for Malware Detection on Imbalanced Dataset using Random Under-Sampling Method,” *JPIT*, vol. 8, no. 2, pp. 113–118, 2023.
- [16] F. Martínez, M. P. Frías, M. D. Pérez, and A. J. Rivera, “A methodology for applying k-nearest neighbor to time series forecasting,” *Artif Intell Rev*, vol. 52, no. 3, pp. 2019–2037, Oct. 2019, doi: 10.1007/s10462-017-9593-z.
- [17] G. N. Ayuni and D. Fitriannah, “Penerapan Metode Regresi Linear Untuk Prediksi Penjualan Properti pada PT XYZ,” *Jurnal Telematika*, vol. 14, no. 2, pp. 79–85.
- [18] H.-A. Goh, C.-K. Ho, and F. S. Abas, “Front-end deep learning web apps development and deployment: a review,” *Appl Intell*, vol. 53, no. 12, pp. 15923–15945, Jun. 2023, doi: 10.1007/s10489-022-04278-6.