

PREDIKSI HARGA SAHAM DENGAN ALGORITMA REGRESI LINIER

Wiki Namansyah

System engineering faculty, Sumbawa University of Technology

Email: 1901013050.wiki.nasmansyah@uts.ac.id

Abstract: Investasi saham di pasar modal merupakan hal yang sangat penting bagi setiap perusahaan di dunia. Harga saham di pasar modal bergerak sangat acak, tinggi rendahnya harga saham dipengaruhi oleh banyak faktor. Oleh karena itu, perlu diprediksi harga saham sehingga dapat membantu para investor untuk melihat prospek investasi di masa yang akan datang. Dalam penelitian ini akan dilakukan prediksi harga saham, dengan menggunakan algoritma yaitu Regresi Linier pada rapid miner. *Algoritma Regresi Linier* ini adalah algoritma yang terbaik untuk digunakan karena yang paling kompleks dibandingkan dengan algoritma lain. Berdasarkan signaling theory yang merupakan sinyal informasi yang dibutuhkan oleh investor nilai hasil peramalan yang telah diperoleh dapat digunakan untuk mempertimbangkan keputusan investor bahwa saham tersebut memiliki risiko yang tinggi atau rendah di masa depan. Berdasarkan teori risiko, analisis peramalan ini membantu investor untuk meminimalisir kerugian. Prediksi saham merupakan salah satu analisis teknikal. Transaksi jual beli saham tanpa teknikal merupakan perilaku berjudi dan mengandung gharar atau ketidakjelasan. Dampak dari tidak menggunakan analisa teknikal ini jelas mengakibatkan transaksi yang mengandung maisir dan gharar yang jelas-jelas telah dilarang. Data historis saham yang digunakan dalam pengujian diperoleh dari halaman *web finance.yahoo.com*. Yang akan digunakan adalah data tahunan selama 5 tahun terakhir yang berbentuk time series yang disertai variabel *open*, *high*, *low* dan *volume* sebagai variabel *independent* dan *close* sebagai variabel *dependent*. Algoritma yang digunakan adalah dengan regresi linier berganda.

Keywords: Data mining, Stock Prediction, Linier Regression,

1 PENDAHULUAN

Saham merupakan instrumen pasar keuangan yang paling populer di kalangan masyarakat saat ini. Membangun saham merupakan salah satu pilihan sebuah perusahaan ketika mereka ingin memutuskan untuk pendanaan perusahaan. Selain itu, saham dapat dikatakan juga sebagai instrumen investasi yang banyak dipilih investor dikarenakan saham mampu memberikan keuntungan yang menarik. Perkembangan ekonomi di Indonesia melambat dalam beberapa tahun terakhir. Hal ini mengakibatkan pergerakan indeks untuk beberapa saham yang tercatat di BEI, khususnya LQ45 yang juga mengalami kenaikan dan penurunan. Oleh karena itu, perlu dilakukan analisis pergerakan harga saham agar hasil analisis dapat digunakan oleh investor untuk mengambil keputusan investasi. Salah satu metode yang biasa digunakan dalam menentukan prediksi harga saham adalah dengan metode *forecasting*. Hasil pengujian menggunakan metode *Neural Network* memiliki tingkat error yang lebih sedikit yang dapat dilihat dari nilai RMSE yaitu 0,034. Sehingga dapat dikatakan algoritma *Neural Network* memberikan prediksi lebih baik dari pada yang collected from previous development

algoritma Regresi Linier (Ramadhan, V. P., & Pamuji, 2022). Saham sebagai salah satu produk yang diperjual belikan dalam pasar modal merupakan suatu surat berharga yang berbentuk sertifikat, guna menunjukkan bukti kepemilikan suatu perusahaan. Semakin banyak saham yang dimiliki oleh seseorang di suatu perusahaan, berarti jumlah uang yang diberikan ke perusahaan itu semakin besar. Dalam berinvestasi pada dunia saham tentunya memiliki berbagai risiko. Risiko yang terjadi pada bursa saham seperti tidak adanya pembagian dividen, capital loss, risiko likuidasi, dan saham delisting dari bursa. Harga saham yang dinamis, cenderung membuat para praktisi memiliki pandangan yang berbeda satu sama lain akan sebuah saham sehingga tidak jarang hal ini menyebabkan seorang investor ragu untuk membeli, menjual atau mempertahankan saham. Saham memiliki sifat yang fluktuatif secara tidak menentu, hal ini yang membuat saham sulit diprediksi. Prediksi saham merupakan salah satu masalah penting di bidang *stock trading*. Dan bursa prediksi harga yang bervariasi tergantung pada waktu dan informasi masa lalu. Hal ini digunakan untuk menentukan nilai masa depan dari suatu saham perusahaan atau instrumen keuangan lainnya yang diperdagangkan di bursa keuangan. Akurasi hasil prediksi saham digunakan untuk banyak alasan, utamanya adalah untuk kebutuhan bagi investor untuk melindungi nilai terhadap risiko pasar yang potensial, dan kesempatan untuk spekulator pasar dan arbitrase untuk membuat keuntungan dari indeks perdagangan. Memprediksi harga merupakan salah satu tantangan utama yang dihadapi para *trader* dan *investor* saham di dunia. Hal ini disebabkan karena semakin akurat prediksinya, maka keuntungan yang diperoleh akan semakin besar. Dalam berinvestasi saham, keuntungan yang tinggi bisa didapat dengan analisa yang akurat serta pemahaman yang dalam atas kondisi pasar dan saham perusahaan itu sendiri. Dengan demikian, sebelum para investor membeli dan menjual saham, dilakukan analisa yang mendalam. Dalam dunia perdagangan saham dikenal dua jenis analisa, yaitu fundamental dan teknikal. Analisa *fundamental* bertujuan untuk mendapatkan nilai intrinsik dari saham suatu perusahaan. Sehingga penting bagi investor untuk mengevaluasi bisnis perusahaan tersebut seperti kinerja keuangan, operasional, bahkan kepemilikan dan manajemennya. Biasanya hal tersebut dapat dilihat melalui laporan keuangan dan laporan tahunan serta berita di media pasar modal terdapat sebuah pola yang menggambarkan pergerakan harga dari suatu saham pada periode tertentu, atau disebut juga dengan "*Trend Harga*".

Karakteristik yang paling utama dari sebuah trend harga adalah perubahan arah yang cepat dan tidak teratur. Karena sifatnya yang mudah berubah dan tidak teratur, maka munculah sebuah permasalahan untuk mengetahui kemana arah trend harga akan bergerak. Adanya kesalahan dalam memprediksi arah *trend* dapat menimbulkan kerugian. Penelitian ini mengimplementasikan algoritma random forest sebagai model solusi, dan *technical analysis* sebagai fitur

prediksi untuk meminimalkan kesalahan dalam memprediksi trend harga di masa depan (Bastian, 2021).

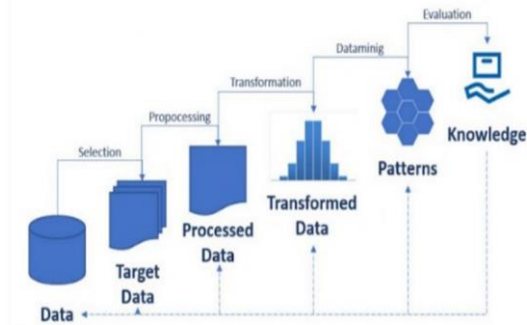
Model prediksi LR dapat menghasilkan hasil prediksi berkualitas tinggi dengan RMSE kurang dari 0,010 dan RMSE kurang dari 0,012, MAPE kurang dari 1,2% dan tertinggi kurang dari 1,9%, MAE kurang dari 0,006 dan tertinggi kurang dari 0,009, dan R2 kurang dari 99,8% dan kurang dari 99,6%. Hal ini menunjukkan bahwa, dalam penelitian ini, model *Linear Regression* mengungguli model RFR dan MLP dalam memprediksi data harga saham historis (Fitri, E., & Riana, 2022).

saham Bank BRI mengalami naik dan turun secara signifikan, hal tersebut yang menggambarkan sensitifnya saham terhadap suatu peristiwa. Oleh karena itu pentingnya melakukan prediksi harga saham agar bisa mengurangi resiko yang diterima para investor. Metode yang digunakan untuk data berdimensi time series yaitu salah satunya algoritma Linear Regression, karena metode ini bisa menangani data rentetan waktu. Perbedaan algoritma *linier regresi* yang peneliti gunakan dengan algoritma random forest yang dilakukan oleh penelitian terdahulu dilihat dari nilai RMSE.

2 METODE PENELITIAN

Penelitian ini peneliti menggunakan pendekatan deskriptif kuantitatif. Dalam pengertiannya penelitian deskriptif adalah penelitian yang menggambarkan suatu keadaan atau kondisi yang terjadi disuatu daerah, yang pemecahan masalahnya secara tersistem berdasarkan data-data yang bersikap fakta yang ada. dan data kualitatif adalah jenis data berupa angka yang dapat digunakan dalam proses operasi matematika, data ini berasal dari data fakta yang ada. Jenis data yang digunakan dalam penelitian ini bersifat rasio yang merupakan data yang berbentuk angka sebenarnya(Jange, 2021). akan diolah dengan menggunakan algoritma *Algoritma Random Forest*.

Analisis data adalah prosese penelitian yang dilakukan setelah semua data yang dibutuhkan telah terkumpul baik berupa data primer atau data sekunder. Analisis data di sini makasudnya untuk menemukan pola umum dari data yang terkumpul dengan cara mengolah atau meringkasnya. Metode analisa data yang digunakan dalam penelitian ini menggunakan *Knowledge Discovery in Database (KDD)* adalah proses yang bertujuan untuk menggali dan menganalisis data yang sangat besar menjadi informasi yang berguna untuk pengetahuan.



Gambar 1. Alur penelitian menggunakan KDD

3 HASIL DAN PEMBAHASAN

A. Data Selection

Pada tahapan ini data yang digunakan akan diseleksi dengan cara melihat, kesesuaian data dengan topik atau judul penelitian yang akan di teliti, dalam hal ini data yang

diperoleh dari data *public web archive.ics.edu* dengan jumlah data sebanyak 995 record sudah sesuai dengan format data dalam penelitian ini.

Gambar 3. Data Training

Data Training adalah data yang sudah dilakukan cleaning ,kemudian dilakukan pelabelan dan pemberian record sesuai bentuk data yang tertera didalam gambar diatas, sehingga data tersebut dapat dilakukan pengolahan susai dengan algoritma yang cocok dalam proses data ming untuk variberl label yang diambil adalah close.

B. Preprocessing Data

Tahapan *preprocessing* data dalam penelitian ini menggunakan *Operator replace missing value* pada *tool rapidminer*. operator replace missing value menangani nilai yang hilang dapat diganti dengan nilai minimum, maksimum, atau rata-rata dari Atribut tersebut. Nol juga dapat digunakan untuk menggantikan nilai yang hilang. Setiap nilai pengisian ulang juga dapat ditentukan sebagai pengganti nilai yang hilang.

C. Data Mining

Tahapan data mining pada penlitian ini menggunakan algoritma *linier regresi*. Data berasal dari data *public web finance.yahoo.com* kemudian data tersebut akan distandarisasi mengikuti proses tahapan data mining agar data tersebut layak dan dapat diolah menggunakan rapidminer dengan metode algoritma *linier regresi*

Gamber 4. Data Training

Data Training adalah data untuk mengetes performance dari data training sehingga data tersebut bias dilakukan testing dan menghasilkan output prediksi yang sesuai dengan yang diingikan.

D. Penerapan Algoritma linier regresi



Gambar. 5 Alur Proses Testing Data

Dari gambar diatas dapat penulis jelaskan bahwa data training yang sudah diberi label dan record di padukan dengan data testing, kemudian di hubungkan dengan Algoritma Regresi Linear , Apply Model dan Peformance serta dilakukan pengaturan mengitung nilai error dari setiap parameter dan performance dari data saham tersebut.

Row No.	Class	prediction	Date	Open	High	Low	Volume
1	1168.880	1167.877	Mar 19, 2010	1168.880	1168.200	1168.330	421240000
2	1168.870	1158.812	Mar 22, 2010	1157.250	1167.620	1152.880	421760000
3	1174.170	1165.239	Mar 23, 2010	1166.470	1174.720	1163.830	441760000
4	1167.720	1162.585	Mar 24, 2010	1172.700	1173.040	1166.870	470570000
5	1165.730	1168.817	Mar 25, 2010	1170.530	1169.080	1165.080	598880000
6	1169.580	1163.427	Mar 26, 2010	1167.580	1173.830	1161.480	470420000
7	1173.220	1165.488	Mar 29, 2010	1167.710	1174.890	1167.710	427090000

Gambar 6. Hasil Run Data

Dari gambar diatas menghasilkan nilai performance dari harga saham setelah dilakukan prediksi secara mendalam menggunakan Algoritma Linear Regression sehingga muncul sebuah pengetahuan yang dapat meprediksi Harga Saham kedepannya dari data yang telah dilakukan pengolahan.

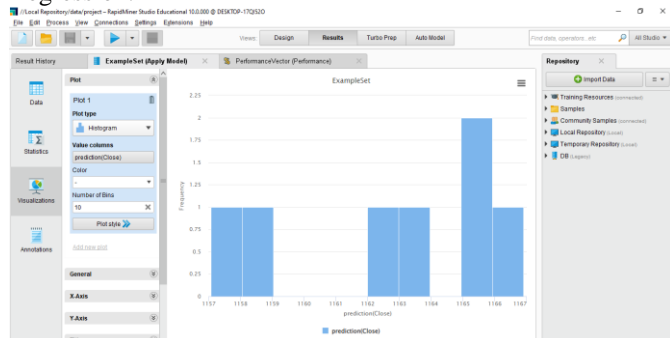
E. Hasil

Berdasarkan hasil pengujian algoritma prediksi Regresi Linier terhadap prediksi harga saham menggunakan aplikasi Rapid Miner. Maka dapat disimpulkan bahwa hasil pengujian menggunakan metode regresi linier memiliki root mean squard error

PerformanceVector
Root Mean Squared Error: 0.746
Absolute Error: 0.220
Relative Error: 0.498
Prediction Average: 0.228

Gambar 7. Performancevector

Dari Gambar diatas kita dapat mengetahui performance mulai dari root-mean, absolute_error, relative_error dari data saham yang telah dilkukan prediksi dengan Algoritma Linnear Regression.



Gambar . 8 Tabel Garfik Performance

Dari gambar diatas dapat kita ketahui Peformance dan error dari hasil pengolahan data.

F. Pembahasan

Berdasarkan hasil dari penelitian dalam melakukan prediksi terhadap harga saham dengan menerapkan komparasi metode regresi linier menghasilkan :

Analisa prediksi harga saham menggunakan metode *linier regresi* dengan rasio pembagian 60/40 serta beberapa jenis *feature selection* dan nilai *min-tolerance* 0.0 menghasilkan *performanceVector* prediksi harga saham didapat nilai *root mean squeared error* terbaik 27.780, *absolute error* 19.780, *relative error* 0,57% dan *prediction average* 3607.646.

KESIMPULAN

Berdasarkan hasil dari penelitian dalam melakukan prediksi terhadap harga saham dengan menerapkan komparasi metode regresi linier menghasilkan sehingga menghasilkan pengetahuan untuk mengetahui harga saham dari gunung data yang kita olah menggunakan metode Data Mining menggunakan Algoritma Linnear Regression.

DAFTAR PUSTAKA

- Bastian. (2021). *Prediksi Trend Harga Saham Jangka Pendek Berdasarkan Fitur Technical Analysis Dengan Menggunakan Algoritma Random Forest*. S1: Universitas Brawijaya,.
- Fitri, E., & Riana, D. (2022). *Analisa Perbandingan Model Prediction Dalam Prediksi Harga Saham Menggunakan Metode Linear Regression, Random Forest Regression Dan Multilayer Perceptron*. METHOMIKA JurnalManajemen Informatika Dan Komputerisasi Akuntansi.
- Jange, B. (2021). *Prediksi Harga Saham Bank BCA Menggunakan Prophet*. Journal of Trends Economics and Accounting.
- Ramadhan, V. P., & Pamuji, F. Y. (2022). *Jurnal Teknologi dan Manajemen Informatika Analisis Perbandingan Algoritma Forecasting dalam Prediksi Harga Saham LQ45 PT Bank Mandiri Sekuritas (BMRI)*.