

TESIS
PREDIKSI HARGA SAHAM BERDASARKAN ANALISIS
TEKNIKAL DAN SKOR SENTIMEN

*Predicting Stock Prices Utilizing Technical Analysis And Sentiment
Scores*

Disusun dan diajukan oleh

ANGGA KURNIAWAN

D082211018



PROGRAM STUDI MAGISTER TEKNIK INFORMATIKA
UNIVERSITAS HASANUDDIN
GOWA
2024

PENGAJUAN TESIS

**PREDIKSI HARGA SAHAM BERDASARKAN ANALISIS
TEKNIKAL DAN SKOR SENTIMEN**

Tesis

Sebagai Salah Satu Syarat Untuk Mencapai Gelar Magister
Program Studi Teknik Informatika

Disusun dan diajukan Oleh

ANGGA KURNIAWAN

D082211018

Kepada

**FAKULTAS TEKNIK
UNIVERSITAS HASANUDDIN
GOWA
2024**

TESIS

PREDIKSI HARGA SAHAM BERDASARKAN ANALISIS TEKNIKAL DAN SKOR SENTIMEN

ANGGA KURNIAWAN
D082211018

Telah dipertahankan di hadapan Panitia Ujian Tesis yang dibentuk dalam rangka penyelesaian studi pada Program Magister Teknik Informatika Fakultas Teknik Universitas Hasanuddin
Pada tanggal 18 September 2024
dan dinyatakan telah memenuhi syarat kelulusan

Menyetujui,

Pembimbing Utama



Prof. Dr. Indrabayu, S.T., M.T., M.Bus.Sys., IPM., ASEAN.Eng.
NIP. 19750716 200212 1 004

Pembimbing Pendamping



Mukarramah Yusuf, B.Sc., M.Sc. Ph.D
NIP. 19831008 201212 2 003

Dekan Fakultas Teknik
Universitas Hasanuddin



Prof. Dr.Eng. Ir. Muhammad Isran Ramli, M.T. IPM., ASEAN.Eng.
NIP. 19730926 200012 1 002

Ketua Program Studi
S2 Teknik Informatika



Dr. Ir. Zahir Zainuddin, M.Sc.
NIP. 19640427 198910 1 002

**PERNYATAAN KEASLIAN TESIS
DAN PELIMPAHAN HAK CIPTA**

Yang bertanda tangan dibawah ini

Nama : Angga Kurniawan
Nomor Mahasiswa : D082211018
Program Studi : S2 Teknik Informatika

Dengan ini menyatakan bahwa tesis berjudul “Prediksi Harga Saham Berdasarkan Analisis Teknikal Dan Skor Sentimen” adalah benar karya Saya dengan arahan dari komisi pembimbing Prof. Dr. Ir. Indrabayu, ST., MT., M.Bus.Sys., IPM, ASEAN. Eng dan Mukarramah Yusuf, B.Sc., M.Sc. Karya ilmiah ini belum diajukan dan tidak sedang diajukan dalam bentuk apapun kepada perguruan tinggi manapun. Sumber Informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka tesis ini. Sebagian dari isi tesis ini telah dipublikasikan di Prosiding 2024 IEEE Internasional Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT), DOI : 10.1109/IAICT62357.2024.10617768 sebagai artikel dengan judul “Stock Price Prediction Using Technical Data and Sentiment Score”.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya berupa tesis ini kepada Universitas Hasanuddin

Gowa, 16 Oktober 2024

Yang Menyatakan



Angga Kurniawan

KATA PENGANTAR

Puji Syukur penulis panjatkan ke hadirat Allah SWT, berkat rahmat, petunjuk, bimbingan dan karunia-Nya sehingga Tesis dengan judul “Prediksi Harga Saham Berdasarkan Analisis Teknikal Dan Skor Sentimen” ini dapat selesai.

Penulisan tesis ini merupakan salah satu syarat untuk lanjut ke tahap penyusunan tesis dan sekaligus sebagai syarat untuk mendapatkan gelar Magister pada Program Studi Teknik Informatika, Departemen Teknik Informatika, Fakultas Teknik Universitas Hasanuddin.

Penulis menyadari bahwa selama penulisan dan penyusunan Tesis ini banyak pihak yang telah membantu dan memberikan dukungannya baik secara material maupun moril. Demikian pula segala bantuan yang penulis peroleh selama perkuliahan sehingga dapat membantu penulis dalam menyusun Tesis ini.

Oleh karena itu Penulis menyampaikan ucapan Terimakasih yang tulus kepada keluarga penulis, terkhusus kepada kedua orang tua, Ayahanda ku (Alm) Abdul Salam Adam dan Ibunda ku Erny Rosmawaty serta kedua Kakek Nenekku tercinta yang menjadi alasan penulis untuk tetap bertahan hingga saat ini. Tidak lupa pula penulis ucapkan terima kasih yang setulusnya dan penghormatan kepada bapak Prof. Dr. Ir. Indrabayu, ST., MT., M.Bus.Sys., IPM, ASEAN. Eng. selaku Pembimbing I dan Ibu Mukarramah Yusuf, B.Sc., M.Sc selaku Pembimbing II, yang telah meluangkan waktu, tenaga, dan pikiran serta memberikan arahan-arahan yang sangat berharga bagi penulis, sehingga penulis dapat menyelesaikan penelitian ini. Selanjutnya, penghargaan dan rasa Terimakasih yang setinggi-tingginya juga Penulis haturkan kepada:

1. Bapak Dr. Ir. Zahir Zainuddin, M.Sc. selaku Ketua Prodi Magister Teknik Informatika yang selalu memberikan arahan kepada Penulis agar cepat menyelesaikan segala rangkaian proses akademik di Departemen Teknik Informatika.
2. Bapak Dr. Eng. Ady Wahyudi Paundu, S.T., M.T selaku Penguji I, Bapak Dr. Adnan, S.T., M.T selaku Penguji II, dan Bapak Prof. Dr. Ir. Ansar Suyuti, M.T. selaku Penguji III yang senantiasa meluangkan waktunya untuk memberikan kritik serta saran dalam pembuatan tesis ini.

3. Bapak dan Ibu Dosen Departemen Teknik Informatika yang telah mengajarkan Ilmu dan Pengetahuan bagi Peneliti selama proses perkuliahan.
4. Seluruh staff tata usaha yang telah membantu proses administrasi bagi Penulis saat berkuliah di Fakultas Teknik, Universitas Hasanuddin.
5. Teman-teman Mahasiswa Angkatan 4 serta seluruh Pengurus dan Anggota Himpunan Mahasiswa Magister Teknik Informatika yang senantiasa memberikan dukungan dan semangat kepada Penulis dalam menyelesaikan penelitian.
6. Kepada semua pihak yang tidak bisa disebutkan satu persatu, terima kasih atas semua dukungan yang telah diberikan dalam penyelesaian skripsi ini. Kalian akan selalu menjadi kenangan yang penulis rindu kan untuk terulang kembali. Terima kasih atas segalanya.

Penulis menyadari masih banyak kekurangan yang terdapat pada Skripsi ini. Saran dan kritik penulis harapkan dari pembaca untuk perbaikan-perbaikan di masa yang akan datang. Semoga karya kecil ini dapat menjadi suatu kontribusi bagi perkembangan ilmu pengetahuan terkhusus pada bidang Informatika. *Amin Allahumma Amin*. Sekali lagi, tanpa segala bantuan dari kalian penulis bukanlah apa-apa.

Akhirul Qalam, Wassalamu'Alaikum Warahmatullahi Wabarakatuh

Gowa, September 2024

ANGGA KURNIAWAN

ABSTRAK

ANGGA KURNIAWAN. Prediksi Harga Saham Berdasarkan Analisis Teknikal Dan Skor Sentimen (dibimbing oleh **Indrabayu** dan **Mukarramah Yusuf**).

Dalam memprediksi harga saham, metode analisis yang sering digunakan oleh seorang investor sebelum menentukan keputusan investasi adalah analisis teknikal dan analisis sentimen. Analisa teknikal umumnya mampu memprediksi harga saham short-term, tetapi tidak mampu memprediksi harga saham long-term salah satunya karena tidak mempertimbangkan sentimen pasar. Penggabungan analisa teknikal dan sentimen adalah cara yang tepat untuk mengatasi masalah ini, dan telah dicoba di dalam beberapa penelitian, tetapi metode-metode yang diusulkan masih memiliki kekurangan dan masih dapat dikembangkan. Paper ini mengusulkan prediksi harga saham long-term berdasarkan analisis teknikal dan analisis sentimen menggunakan algoritma BERT yang dimodifikasi pada lapisan feed forward layer menggunakan sigmoid activation, sedangkan analisis teknikal menggunakan algoritma LSTM dan XGBoost. Untuk mengetahui keberhasilan penggunaan skor sentimen dalam prediksi saham, model yang diajukan dibandingkan dengan model penelitian sebelumnya. Model BERT yang telah dimodifikasi memperoleh akurasi sebesar 92.73%. Di sisi lain, Model LSTM, XGBoost, BERT-LSTM dan BERT-Xgboost dievaluasi menggunakan metrik MAE, RMSE, MSE, MAPE memperoleh hasil berturut-turut sebesar {0.008, 0.010, 0.001, 1.13%}, {0.010, 0.014, 0.002, 1.40%}, {0.018, 0.025, 0.006, 2.24%} dan {0.011, 0.014, 0.002, 1.40%}. Berdasarkan hasil tersebut dapat disimpulkan bahwa model LSTM memperoleh kesalahan prediksi paling sedikit dibandingkan model yang lain. Akan tetapi jika data teknikal digabungkan dengan skor sentimen dari BERT maka model BERT-XGBoost lebih baik. Pada prediksi harga saham jangka panjang disarankan untuk menggunakan model BERT-XGBoost karena terbukti efektif dalam menangani data non-linear dan memiliki kemampuan untuk menangkap pola yang lebih kompleks dalam data finansial seperti harga saham jika dibandingkan LSTM.

Kata Kunci : LSTM, BERT, XGBoost, *Stock prediction model*, *sentimen scoring*, *Natural Language Processing*.

ABSTRACT

ANGGA KURNIAWAN. Stock Price Prediction using Technical Data and Sentiment Score (supervised by **Indrabayu dan Mukarramah Yusuf**).

In predicting stock prices, the analysis methods often used by an investor before making an investment decision are technical analysis and sentiment analysis. Technical analysis is generally able to predict short-term stock prices, but is unable to predict long-term stock prices partly because it does not consider market sentiment. Combining technical analysis and sentiment analysis is a good way to solve this problem, and has been tried in several studies, but the proposed methods still have shortcomings and can still be developed. This paper proposes a long-term stock price prediction based on technical analysis and sentiment analysis using BERT algorithm modified in the feed forward layer using sigmoid activation, while technical analysis uses LSTM and XGBoost algorithms. To determine the success of using sentiment scores in stock prediction, the proposed model is compared with previous research models. The modified BERT model obtained an accuracy of 92.73%. On the other hand, LSTM, XGBoost, BERT-LSTM and BERT-Xgboost models evaluated using MAE, RMSE, MSE, MAPE metrics obtained results of {0.008, 0.010, 0.001, 1.13%}, {0.010, 0.014, 0.002, 1.40%}, {0.018, 0.025, 0.006, 2.24%} dan {0.011, 0.014, 0.002, 1.40%}. Based on these results, it can be concluded that the LSTM model obtained the least prediction error compared to other models. However, if technical data is combined with sentiment scores from BERT, the BERT-XGBoost model is better. In long-term stock price prediction, it is recommended to use the BERT-XGBoost model because it has proven effective in handling non-linear data and has the ability to capture more complex patterns in financial data such as stock prices when compared to LSTM.

Keywords : LSTM, BERT, XGBoost, *Stock prediction model, sentiment scoring, Natural Language Processing.*

DAFTAR ISI

HALAMAN SAMPUL	i
KATA PENGANTAR.....	ii
ABSTRAK	iv
ABSTRACT	v
DAFTAR ISI	vi
DAFTAR TABEL	viii
DAFTAR GAMBAR.....	ix
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	3
1.3 Hipotesa Penelitian	3
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian	4
1.6 Batasan Masalah	5
BAB II TINJAUAN PUSTAKA	6
2.1 Saham.....	6
2.2 Analisis Teknikal	6
2.3 Analisis Sentiment	7
2.4 BERT (Bidirectional Encoder Representations from Transformers).....	8
2.5 LSTM (Long Short Term Memories)	11
2.6 XGBoost	12
2.7 Ensemble Learning	13
2.8 Penelitian yang Terkait	16
2.9 Kerangka Pikir Penelitian	19
BAB III METODE PENELITIAN	20
3.1 Jenis Penelitian.....	20
3.2 Tahapan Penelitian.....	20
3.3 Sumber Data	21
3.4 Rancangan Sistem.....	22
3.5 Pengujian Sistem.....	33
3.6 Waktu dan Lokasi Penelitian	33

BAB IV HASIL DAN PEMBAHASAN.....	34
4.1 Pengembangan Model.....	34
4.2 Evaluasi Model	37
BAB V KESIMPULAN DAN SARAN	45
5.1 Kesimpulan	45
5.2 Saran	46
DAFTAR PUSTAKA.....	47

DAFTAR TABEL

Tabel 3. 1 Data Skor sentiment setelah penggabungan.....	25
Tabel 3. 2 Data hasil scraping	26
Tabel 3. 3 Data teknikal setelah smooting.....	26
Tabel 3. 4 Dataset setelah penggabungan.....	26
Tabel 3. 5 Dataset setelah normalisasi	27
Tabel 3. 6 Hyperparameter Tuning BERT	29
Tabel 3. 7 Hyperparameter Tuning LSTM dan BERT-LSTM.....	30
Tabel 3. 8 Konfigurasi Hyperparameter tuning XGBoost.....	32
Tabel 3. 9 Hyperparameter Tuning XGBoost	32
Tabel 4. 1 Konfigurasi Hyperparameter Algoritma BERT	34
Tabel 4. 2 Konfigurasi Hyperparameter algoritma LSTM.....	36
Tabel 4. 3 Konfigurasi hyperparameter XGBoost.....	37
Tabel 4. 4 Hasil Evaluasi BERT Modifikasi	39
Tabel 4. 5 Interpretasi skor sentiment	39
Tabel 4. 6 Hasil evaluasi masing-masing model	42
Tabel 4. 7 Skor sentiment bulanan	42

DAFTAR GAMBAR

Gambar 4. 1 Confusion Matrix BERT Modifikasi	38
Gambar 4. 2 Training dan Validation loss model LSTM.....	40
Gambar 4. 3 Training dan Validation loss model BERT-LSTM	40
Gambar 4. 4 Grafik perbandingan prediksi masing-masing model dengan nilai aktual	41

BAB I PENDAHULUAN

1.1 Latar Belakang

Menurut data statistik publik yang dikeluarkan oleh PT Kustodian Sentral Efek Indonesia (KSEI) pada bulan Desember 2022, terjadi peningkatan signifikan jumlah investor pasar modal. Dalam rentang waktu dari akhir tahun 2018 hingga akhir tahun 2022, jumlah investor meningkat dari 1,6 juta investor menjadi 10,3 juta investor, yang artinya terjadi kenaikan sebesar 543.75% atau meningkat hampir 5,5 kali lipat. Dapat disimpulkan bahwa bisnis di pasar modal terutama saham diminati oleh masyarakat beberapa tahun belakangan ini karena pengurusan nya yang lebih mudah, memiliki peluang meraup untung yang lebih besar dalam waktu yang lebih singkat dibandingkan dengan bisnis *real estate* yang lain. Dalam menentukan saham yang potensial diperlukan analisis serta prediksi terhadap harga saham di masa yang akan datang. Prediksi harga saham adalah bidang penelitian yang sudah lama dan masih dikaji hingga hari ini. Memprediksi harga saham adalah hal yang menantang karena bergantung pada banyak faktor seperti politik, ekonomi global, laporan keuangan dan kinerja perusahaan [1]. Dalam memprediksi harga saham, metode analisis yang sering digunakan oleh seorang investor sebelum menentukan keputusan investasi adalah analisis teknikal dan analisis sentiment [2].

Analisis teknikal adalah analisis yang melibatkan penggunaan grafik dan data historis harga saham untuk mengidentifikasi tren dan pola harga saham. Terdapat beberapa indikator yang perlu diperhatikan seperti *Moving Average*, *Relative Strength Index (RSI)*, *Stochastic Oscillator* Dan *Volume*. Sedangkan analisis sentiment adalah sebuah teknik yang digunakan untuk mengidentifikasi sebuah teks atau mengekstrak sentiment pada teks seperti ulasan produk, laporan media sosial ataupun tweet pada twitter. Dalam konteks prediksi harga saham, analisis sentiment digunakan untuk mengukur sentiment pasar terhadap suatu saham. Sentiment pasar dapat mempengaruhi penawaran (supply) dan permintaan (demand). Jika jumlah saham yang dijual lebih banyak daripada jumlah saham yang ingin dibeli oleh investor maka penawaran akan lebih besar sehingga harga saham akan turun. Sebaliknya, jika jumlah investor yang ingin membeli suatu saham lebih

banyak dibandingkan dengan saham yang tersedia maka permintaan akan lebih besar sehingga harga saham akan naik.

Analisa teknikal digunakan untuk prediksi harga saham short-term dengan timeframe seperti 1 menit, 5 menit atau 30 menit yang biasanya digunakan oleh trader. Analisis ini memiliki kelemahan yaitu tidak mampu memprediksi harga saham jangka panjang karena tidak mempertimbangkan peristiwa tidak terduga, faktor fundamental perusahaan serta sentiment pasar. Penggabungan analisa teknikal dan sentiment adalah cara yang tepat untuk mengatasi masalah tersebut [3]. Dengan penggabungan ini, akan dibuat model yang dapat memprediksi harga saham secara long-term dan adaptive bagi investor.

Analisis sentiment biasanya menggunakan algoritma Support Vector Machine, Random Forest, dan Naïve Bayes. Akan tetapi algoritma tersebut memiliki beberapa kekurangan seperti setiap kata dianggap independent dan konteks kalimat tidak diperhitungkan, hal ini membuat algoritma tersebut tidak dapat mengenali sarkasme. Selain itu algoritma tersebut hanya menghitung frekuensi kemunculan data sehingga kata dengan frekuensi yang banyak memiliki bobot lebih besar. kekurangan ini akan sangat mempengaruhi hasil analisis terlebih jika dataset yang digunakan terdiri dari bahasa formal maupun non-formal seperti tweet. Algoritma analisis sentiment yang tepat digunakan untuk menangani kelemahan tersebut adalah algoritma BERT [4].

Untuk prediksi saham long-term, Analisa sentiment menggunakan BERT telah coba digabungkan dengan Analisa teknikal menggunakan LSTM pada penelitian yang dilakukan oleh X. Weng. etc [3], tetapi hasilnya belum optimal dan masih perlu ditingkatkan terlebih pada akurasi analisis sentiment yaitu 66%. Pada penelitian ini diusulkan prediksi harga saham long-term menggunakan algoritma BERT yang dimodifikasi pada lapisan *feed forward layer* menggunakan sigmoid activation, sedangkan analisis teknikal menggunakan algoritma LSTM dan XGBoost. Penggunaan sigmoid activation akan menghasilkan analisa sentiment dalam bentuk skor sehingga data lebih bervariasi dan membuat model lebih adaptif dibandingkan jika hanya dalam bentuk positif, negatif dan netral. Untuk mengetahui keberhasilan penggunaan sentiment dalam bentuk skor dalam prediksi saham, model yang diajukan akan dibandingkan dengan penelitian yang telah dilakukan oleh X. Weng etc [3]. Data yang digunakan untuk

mengevaluasi metode yang diusulkan adalah data saham BRI (Bank Rakyat Indonesia). Berdasarkan studi literatur, BRI dipilih pada penelitian ini karena BRI memiliki pengaruh yang besar terhadap perkembangan ekonomi masyarakat berdasarkan prestasi yang diraih pada tahun 2022 sehingga populer di masyarakat baik secara langsung ataupun tidak langsung.

Penelitian ini terbagi atas dua tahap. Tahap pertama analisis sentiment dan tahap kedua analisis teknikal. Awalnya data dikumpulkan dari platform twitter dengan hashtag #BBRI dengan cara *scrapping*. Selanjutnya dilakukan pre-processing data kemudian model dibuat menggunakan algoritma BERT yang telah dimodifikasi sehingga outputnya berupa skor sentiment dengan nilai range 0 -1, semakin tinggi skor maka semakin positif tweet tersebut. Data teknikal diperoleh menggunakan API yahoo finance dengan beberapa fitur yaitu *open price, close price, highest price, lowest price, RSI, MA, Stochastic oscillator, and volume*. Data teknikal dan data analisis sentiment yaitu skor sentiment selanjutnya digabung untuk digunakan pada algoritma LSTM dan XGBoost. Evaluasi hasil prediksi akan dilakukan menggunakan RMSE, MSE dan MAE.

1.2 Rumusan Masalah

Rumusan masalah dalam penelitian ini adalah bagaimana pengaruh skor sentimen terhadap prediksi harga saham jangka panjang menggunakan algoritma XGBoost (*Extreme Gradient Boosting*), LSTM (*Long Short-Term Memory*), dan BERT (*Bidirectional Encoder Representations form Transformers*).

1.3 Hipotesa Penelitian

Pada penelitian ini diusulkan prediksi harga saham menggunakan skor sentiment dan analisis teknikal. Adapun beberapa hipotesa pada penelitian ini sebagai berikut :

1. Penggunaan skor sentimen yang dihasilkan dari model BERT (*Bidirectional Encoder Representation Transformers*) dapat meningkatkan akurasi prediksi harga saham jangka panjang.

2. Model LSTM yang menggunakan skor sentimen menunjukkan performa prediksi yang lebih baik dibandingkan dengan model LSTM yang tidak menggunakan data sentimen
3. Algoritma XGBoost yang menggabungkan skor sentimen dengan fitur teknikal akan menghasilkan prediksi harga saham yang lebih akurat.
4. Algoritma XGBoost, LSTM, dan BERT dipilih karena masing-masing memiliki keunggulan spesifik yang dapat meningkatkan akurasi prediksi harga saham jangka panjang. XGBoost dipilih karena kemampuannya dalam menangani data yang tidak seimbang dan kecepatan dalam menghasilkan model yang efisien. LSTM dipilih karena kemampuannya dalam menganalisis urutan data waktu, memungkinkan model untuk memahami pola temporal yang kompleks dalam data. Sementara itu, BERT dipilih karena efektivitasnya dalam memahami konteks dan nuansa sentimen dari teks, memberikan insight tambahan yang dapat memengaruhi harga saham.

1.4 Tujuan Penelitian

Adapun tujuan pada penelitian ini adalah sebagai berikut :

1. Membangun *prototype*
2. Menggunakan analisis teknikal dan skor sentiment untuk prediksi harga saham jangka panjang
3. Menggunakan output analisis sentiment dalam bentuk nilai kontinu atau scoring.
4. Mengevaluasi model yang dikembangkan
5. Mengetahui seberapa besar pengaruh analisis sentiment terhadap prediksi harga saham

1.5 Manfaat Penelitian

Adapun manfaat penelitian ini adalah sebagai berikut :

1. Bagi investor atau *trader* dapat menjadi referensi dalam memilih saham yang potensial sehingga meminimalisir kerugian.
2. Bagi Peneliti dapat menambah pengetahuan dalam bidang investasi modern khususnya dalam hal investasi saham
3. Bagi Universitas dapat menjadi referensi bagi peneliti lainnya

1.6 Batasan Masalah

Adapun Batasan masalah pada penelitian ini adalah sebagai berikut :

1. Saham yang diteliti adalah saham Bank Rakyat Indonesia (BBRI)
2. Pergerakan harga saham ditinjau menggunakan *timeframe* 1 Days.
3. Dataset yang digunakan pada penelitian ini adalah dataset teknikal yaitu *datetime*, *open*, *close*, *high*, *low*, *market cap*, *volume*, *RSI*, *moving average* dan *stochastic oscillator* dari Yahoo Finance sedangkan dataset sentiment yaitu *datetime* dan *tweet* dari Twitter.
4. Dataset yang digunakan dalam analisis sentiment menggunakan Bahasa Indonesia formal maupun non formal.
5. Hasil analisis sentiment dalam bentuk scoring dengan range 0-1.

BAB II TINJAUAN PUSTAKA

2.1 Saham

Saham adalah salah satu instrumen pasar uang yang paling populer saat ini. Saham, juga dikenal sebagai efek atau stock, merupakan surat berharga yang menunjukkan bukti kepemilikan dalam sebuah perusahaan atau badan usaha. Saham juga dapat diartikan sebagai tanda penyertaan modal oleh seseorang atau badan usaha dalam sebuah perusahaan atau perseroan terbatas.

Saham merupakan salah satu instrumen investasi yang populer saat ini. Saham merupakan surat berharga yang menandakan kepemilikan seseorang dalam suatu perusahaan sehingga pihak tersebut memiliki klaim atas pendapatan dan aset perusahaan, serta berhak hadir dalam rapat umum pemegang saham (RUPS) yang dihitung berdasarkan jumlah lembar kepemilikan. Biasanya, saham dapat dibeli per satu lot atau 100 lembar.

Harga saham adalah nilai suatu saham yang ditentukan selama berlangsungnya pasar saham berdasarkan permintaan dan penawaran terhadap saham tersebut. Harga saham di pasar modal biasanya ditentukan oleh pelaku pasar yang sedang memperdagangkan saham mereka. Dalam transaksi saham di Bursa Efek Indonesia (BEI), para investor mengenal beberapa jenis harga pasar, seperti harga pembukaan (preopening) dan harga penutupan (closing). Perubahan harga saham dapat terjadi karena adanya tawar-menawar antara penjual dan pembeli saham. Proses tawar-menawar ini berlangsung terus menerus hingga jam perdagangan saham berakhir. Harga yang terbentuk pada akhir jam perdagangan disebut sebagai harga penutupan [5].

2.2 Analisis Teknikal

Analisis teknikal adalah studi tentang perilaku pasar yang ditampilkan melalui grafik untuk memprediksi tren harga di masa depan. Namun, analisis teknikal tidak hanya didasarkan pada pembacaan pola grafik harga; ada banyak metode lain untuk melakukan analisis teknis, termasuk menghitung berbagai indeks saham menggunakan rumus perhitungan modern. Dalam literatur, indikator analisis teknikal seperti Moving Average Convergent and Divergent (MACD), Relative Strength Index (RSI), dan stochastic

oscillator digunakan untuk menentukan tren pasar. Pada penelitian ini, data teknikal seperti data *open price*, *close price*, *high price*, *low price*, *date*, *RSI*, *Moving Average Convergent and Divergen* dan *stochastic oscillator* akan digunakan. Berikut penjelasan terkait data-data tersebut.

- a. *Open Price* harga awal pada hari itu saat pasar saham dibuka.
- b. *Close Price* atau harga penutupan adalah harga terakhir yang muncul pada saham sebuah perusahaan sebelum ditutupnya bursa saham
- c. *High price* merupakan istilah yang menunjukkan harga tertinggi suatu saham pada satu hari.
- d. *Low price* Menunjukkan Harga Terendah Yang Pernah Terjadi Dalam Satu Hari Perdagangan Di Bursa Saham.
- e. *Date* adalah data fitur yang menunjukkan tanggal suatu data. Pada penelitian ini, *date* menunjukkan tanggal data harga saham harian.
- f. RSI adalah alat analisis yang mengukur kecepatan perubahan kenaikan dan penurunan harga saham. Indikator RSI membantu menentukan apakah harga pasar sudah mencapai kondisi *overbought* atau *oversold*. Nilai RSI berkisar antara 1 hingga 100 [6].
- g. *Moving Average Convergent and Divergen* (MACD) adalah indikator yang menunjukkan tren yang sedang terjadi dalam perdagangan saham. Indikator ini mudah dibaca karena hanya menggunakan dua garis yang saling berpotongan. [6].
- h. *Stochastic oscillator* adalah indikator analisis pasar keuangan, khususnya dalam analisis teknikal harga saham. Penggunaan indikator ini dapat mengidentifikasi kondisi *overbought* (jenuh beli) dan *oversold* (jenuh jual).

2.3 Analisis Sentiment

Analisa sentiment atau penambangan opini adalah studi komputasi tentang opini, sentiment, emosi, penilaian dan sikap orang terhadap entitas seperti produk, layanan, organisasi, individu, masalah, peristiwa, topik ataupun entitas yang lain [7]. Analisis sentiment adalah interpretasi dan klasifikasi emosi menggunakan teknik pemrosesan bahasa alami (NLP). Teknik ini banyak digunakan untuk mengubah opini menjadi skor

sentiment. Ini memungkinkan bisnis atau organisasi untuk mengidentifikasi sentiment pelanggan terhadap produk, merek, atau layanan dalam percakapan dan umpan balik online [8].

Ada banyak aplikasi dan peningkatan pada algoritma analisis sentiment dalam beberapa tahun terakhir. Berdasarkan penelitian yang dilakukan Medhat dkk [9] menjelaskan bahwa analisis sentiment dapat dilakukan menggunakan 2 pendekatan yaitu machine learning dan lexicon-based. Pendekatan machine learning menerapkan algoritma machine learning yang terkenal dan menggunakan fitur linguistik seperti Support Vector Machine, Neural Network, Naïve Bayes dan Bayesian Network. Pendekatan Machine learning sangat bergantung pada kecocokan algoritma dengan jenis data yang digunakan serta pre-processing. Pendekatan Berbasis Leksikon bergantung pada leksikon sentiment, kumpulan istilah sentiment yang diketahui dan disusun sebelumnya sehingga hasil dari pendekatan ini sangat bergantung pada seberapa banyak kosa kata yang dimiliki atau disimpan.

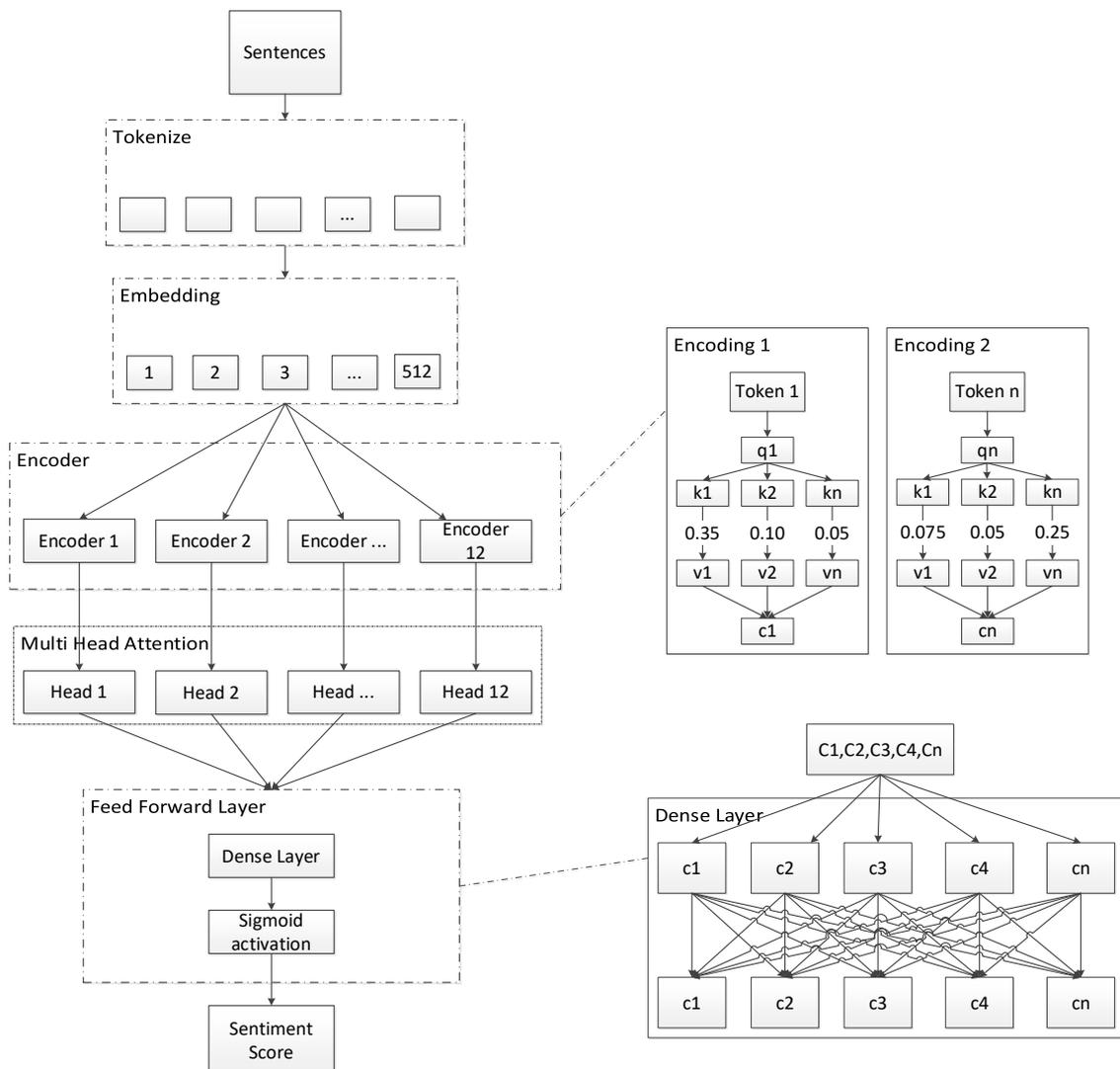
Analisis sentiment menggunakan algoritma machine learning tradisional seperti Naïve Bayes dan SVM tidak dapat memahami konteks komentar secara mendalam terkait semantik kata karena algoritma tersebut hanya mempelajari pola-pola seperti frekuensi kemunculan kata. Penelitian ini menggunakan algoritma BERT yang merupakan pengembangan dari pendekatan Machine Learning dan Lexicon.

2.4 BERT (*Bidirectional Encoder Representations from Transformers*)

BERT singkatan dari Bidirectional Encoder Representations from Transformers adalah algoritma yang dapat diakses dan digunakan secara gratis karena bersifat *opensource* berbasis jaringan neural yang digunakan untuk pre-training dalam pemrosesan bahasa alami (NLP) yang dikembangkan oleh google untuk search engine. Teknologi ini akan membuat sistem lebih mudah memahami konteks pencarian yang dimaksud oleh pengguna dengan menganalisis korelasi antara setiap kata kunci yang diketik. Secara sederhana, BERT memungkinkan sistem komputer memahami bahasa lebih mirip dengan cara manusia memahaminya. Sebagai contoh, dalam kalimat "nine to five" dan "a quarter to five," kata "to" memiliki arti yang berbeda. "Nine to five" berarti

jam 9 sampai jam 5, sedangkan "a quarter to five" berarti jam 5 kurang 15 menit. Bagi manusia, dua kalimat tersebut mudah dibedakan, tetapi bagi mesin pencari, kedua kalimat tersebut bisa membingungkan. Oleh karena itu, teknologi seperti BERT digunakan.

BERT adalah model deep learning yang terdiri dari 6 lapisan Transformer yang ditumpangkan pada encoder dan decoder, yang menghasilkan proses pelatihan yang sangat kompleks, konfigurasi tinggi, waktu pelatihan yang lama, dan biaya yang sangat mahal. Namun, Google telah membuka akses ke model pretrained BERT, yang dapat digunakan tanpa perlu membuat model dari awal. Proses BERT dimulai dari sebuah kata dengan representasi embedding dari lapisan embedding. Setiap lapisan melakukan beberapa penghitungan multi-headed attention pada representasi kata dari lapisan sebelumnya untuk menghasilkan representasi perantara baru. Semua representasi perantara ini memiliki ukuran yang sama [10]. Algoritma BERT pada awalnya dikembangkan menggunakan Bahasa Inggris, namun saat ini algoritma BERT telah dapat digunakan menggunakan Bahasa Indonesia dan disebut algoritma Indo BERT. Algoritma BERT sendiri terdiri dari beberapa tahapan proses seperti yang digambarkan pada gambar 2.1 dibawah ini.

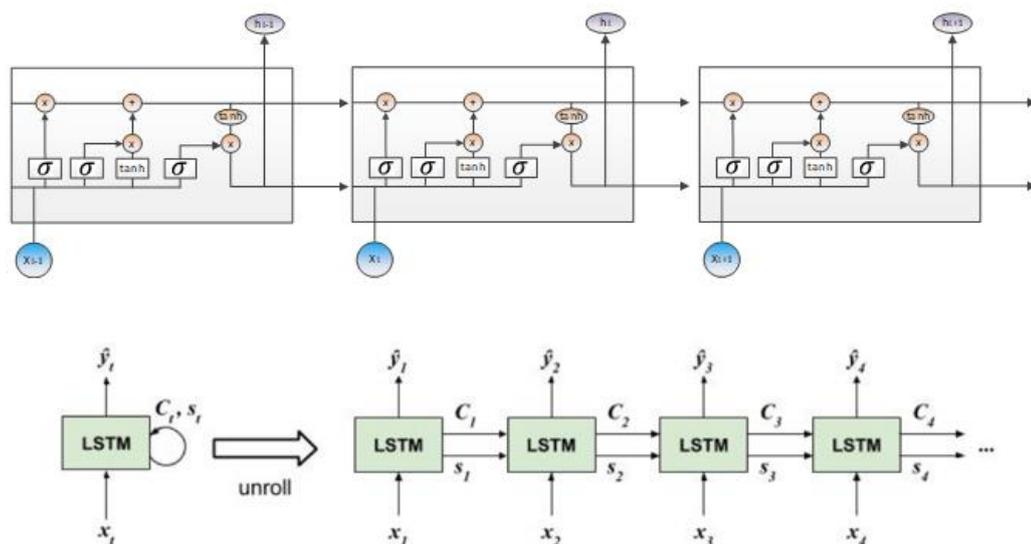


Gambar 2. 1 Struktur Algoritma BERT

Setiap kalimat akan dipisahkan terlebih dahulu menjadi susunan kata kemudian diberi nomor sesuai urutan kata tersebut dalam korpus, proses I ini dinamakan tokenize dan outputnya berupa token. Setelah itu, masing-masing token di proses di lapisan encoding dengan cara mengubah token tersebut menjadi query dan digunakan untuk menghitung nilai k dan v sehingga menghasilkan nilai *attention weight*. Gabungan nilai *attention weight* masing-masing token menghasilkan nilai *multi head attention*. Nilai ini kemudian dijadikan input pada *feed forward layer* dan menghasilkan output berupa skor sentiment.

2.5 LSTM (Long Short Term Memories)

Long Short Term Memory (LSTM) adalah salah satu variasi dari Recurrent Neural Network (RNN) yang dirancang untuk mengatasi masalah pada hidden layer. Inti dari algoritma LSTM adalah mengintegrasikan kontrol non-linear dan dependensi ke dalam sel RNN, sekaligus memastikan bahwa gradient fungsi tujuan tidak menghilang dengan memperhatikan sinyal. LSTM digunakan untuk mengatasi masalah vanishing gradient atau situasi di mana gradient bernilai 0 atau mendekati 0 melalui mekanisme gate. Algoritma LSTM menggabungkan state sebelumnya dengan memori saat ini serta nilai input. Untuk efisiensi, LSTM mampu merekam long-term dependencies. [11]. Struktur algoritma LSTM ditunjukkan seperti pada gambar 2.2 dibawah ini.



Gambar 2. 2 Struktur Algoritma LSTM (Long Short Term Memory)

Berdasarkan gambar 2.2 terdapat beberapa simbol yang digunakan sebagai variabel dalam persamaan untuk menghitung output dan memperbarui nilai memori jangka panjang pada algoritma LSTM. Simbol $C(t-1)$ adalah nilai memori sebelumnya nilai memori jangka panjang pada *timestep* sebelumnya, simbol $s(t-1)$ adalah nilai hidden state pada step sebelumnya, $x(t)$ adalah input pada *timestep* saat ini dan $y(t+1)$ adalah output pada *timestep* berikutnya. Dalam algoritma LSTM, informasi disimpan dalam dua jenis memori, yaitu memori jangka pendek (short-term memory) dan memori jangka

panjang (long-term memory). Pada setiap timestep, LSTM menghitung output dan memperbarui nilai memori jangka panjang dengan mempertimbangkan input, output, serta nilai memori sel dan hidden state pada timestep sebelumnya.

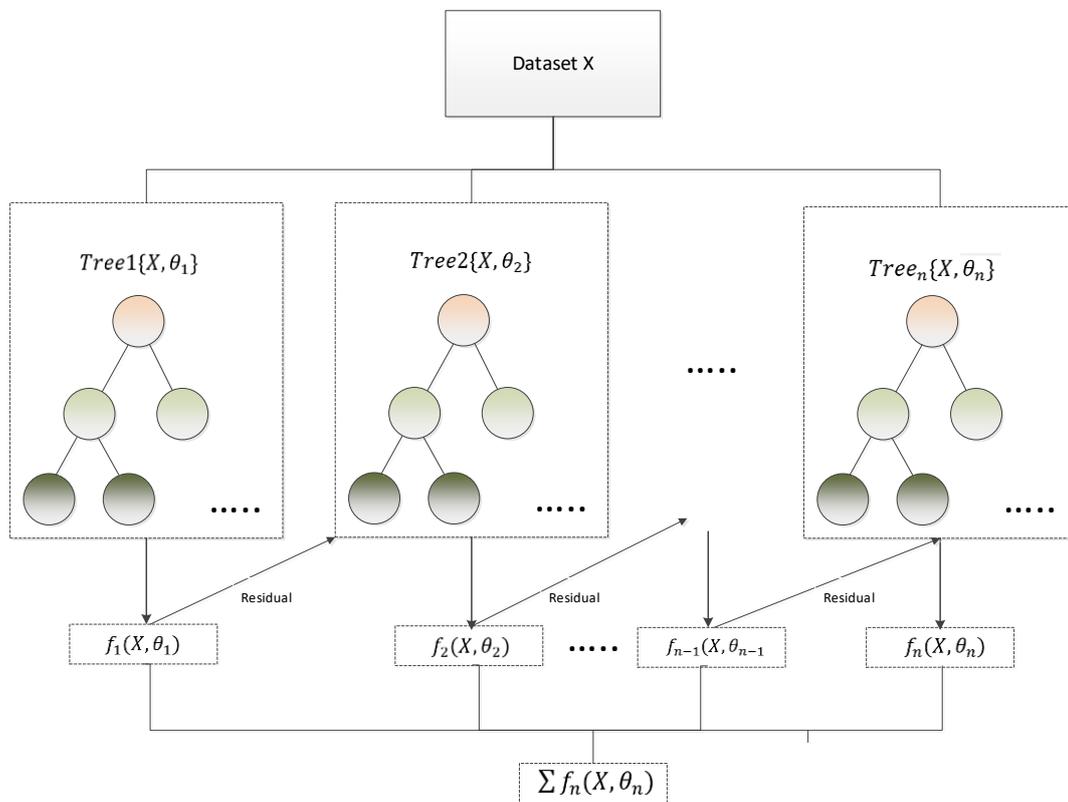
LSTM memproses data pada beberapa bagian. Bagian-bagian tersebut memiliki fungsi dan tugas masing-masing. Awalnya LSTM memutuskan informasi apa yang harus tetap utuh dan apa yang harus dibuang dari cell state. Lapisan sigmoid yang bertanggung jawab untuk membuat keputusan ini. Selanjutnya, LSTM menentukan informasi baru apa yang harus disimpan dan menggantikan yang tidak relevan yang berhasil diidentifikasi pada lapisan sigmoid tadi. Fungsi tanh dan sigmoid sangat penting dalam menentukan informasi yang relevan. Output kemudian ditentukan dengan menggunakan cell state yang telah difilter oleh fungsi sigmoid dan tanh.

2.6 XGBoost

XGBoost adalah salah satu algoritma machine learning yang diklasifikasikan sebagai algoritma boosting pada model ensemble learning. Algoritma ini ditemukan oleh Friedman. Metode ini merupakan pengembangan dari algoritma GBDT (Gradient Boosting Decision Tree). XGBoost merupakan algoritma yang digunakan untuk memprediksi atau mengklasifikasi kan berbasis pohon keputusan [12]. Performa algoritma ini lebih cepat 10kali dalam melakukan Optimasi dibandingkan dengan algoritma GBM lainnya. XGBoost mendominasi kumpulan data terstruktur atau tabular dalam masalah pemodelan prediktif untuk klasifikasi dan regresi, yang terbukti sering digunakan oleh pemenang kompetisi di platform ilmu data kompetitif seperti Kaggle. Meskipun XGBoost dirancang untuk klasifikasi dan regresi pada data tabular, algoritma ini juga dapat digunakan untuk peramalan deret waktu. [13].

XGBoost termasuk dalam algoritma ansambel learning teknik Boosting. Berdasarkan gambar 2.6 algoritma ini memproses dataset secara sequential artinya output dari *tree1* atau iterasi 1 akan dievaluasi, jika terdapat output yang error maka output yang error tersebut akan diteruskan pada iterasi 2 atau *tree 2* sebagai input. Output *tree1* atau iterasi 1 yang bernilai true atau benar maka tidak lanjutkan pada iterasi berikutnya. Proses inilah

yang membuat algoritma XGBoost sangat baik dalam proses peningkatan akurasi dari hasil yang sebelumnya.



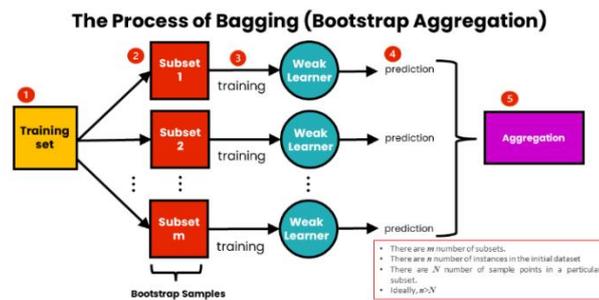
Gambar 2. 3 Proses XGBoost

2.7 Ensemble Learning

Ensemble learning adalah teknik dalam machine learning yang menggabungkan beberapa model atau algoritma pembelajaran mesin untuk membentuk semua model baru yang lebih kuat dan akurat. Dalam *ensemble learning*, model-model individu digabungkan untuk meningkatkan performa dan mengurangi varians dalam prediksi. Terdapat beberapa teknik ensemble learning yang umum digunakan seperti *bagging*, *boosting* dan *stacking*.

Bagging (*Bootstrap Aggregating*) adalah salah satu metode dalam ensemble learning yang menggabungkan hasil beberapa model dengan variasi yang berbeda untuk mengurangi overfitting dan meningkatkan akurasi prediksi. Berdasarkan gambar 2.3 di bawah ini dijelaskan bahwa bagging membangun beberapa model dengan

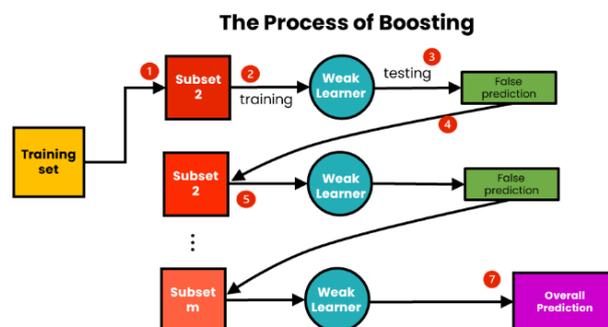
menggunakan subset data latih yang diambil secara acak dan kemudian menggabungkan hasil prediksi dari masing-masing model untuk menghasilkan prediksi akhir dengan teknik seperti majority voting. Bagging dapat diterapkan pada berbagai algoritma machine learning seperti decision tree, k-nearest neighbours, dan neural network.



Sumber (Brijesh Soni, 2023; et al.,)

Gambar 2. 4 Proses Bagging

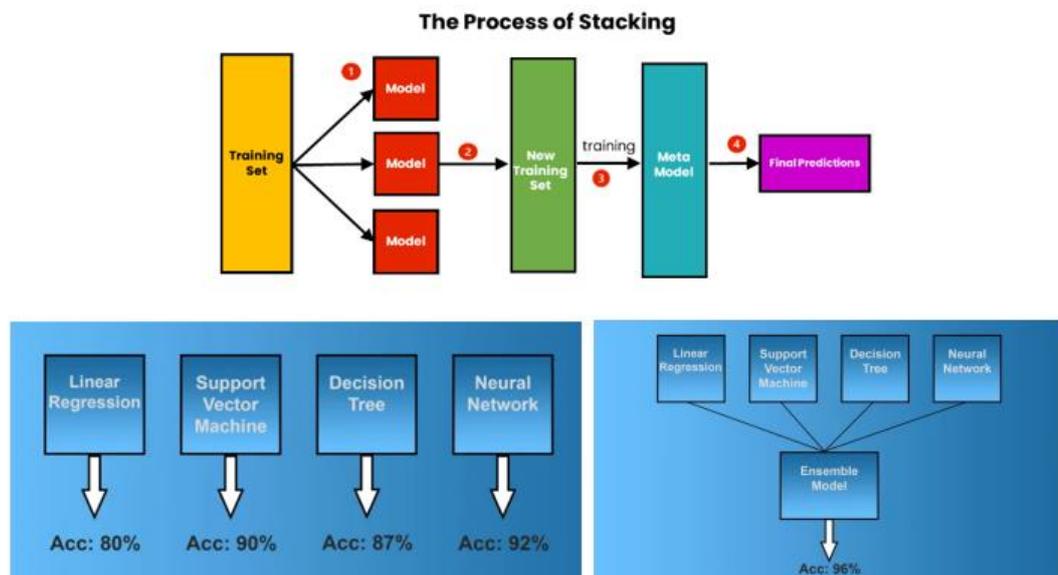
Boosting adalah sebuah teknik ensemble learning di mana model-model lemah (weak learners) digabungkan menjadi sebuah model yang lebih kuat. Teknik ini melatih model secara berulang dengan memberikan bobot yang berbeda pada setiap iterasi seperti yang ditunjukkan pada gambar 2.4. Data yang salah diklasifikasikan akan diberikan bobot yang lebih tinggi untuk dilatih pada iterasi berikutnya. Dengan cara ini, model dapat fokus pada data yang lebih sulit untuk diklasifikasikan. Model-model yang dihasilkan pada setiap iterasi digabungkan untuk membentuk model yang lebih kuat. Boosting dapat digunakan untuk masalah klasifikasi dan regresi. Contoh algoritma yang menggunakan teknik boosting adalah Adaboost dan XGBoost.



Sumber (Brijesh Soni, 2023; et al.,)

Gambar 2. 5 Proses Boosting

Stacking merupakan salah satu metode ensemble learning yang menggabungkan beberapa model machine learning, seperti decision tree, SVM, dan regresi logistik, menjadi satu model yang lebih kuat dengan menggunakan model-model tersebut sebagai input dan menggunakan model pembelajaran yang lebih tinggi sebagai output. Pada metode stacking, output dari beberapa model pembelajaran digunakan sebagai fitur untuk model pembelajaran yang lebih tinggi, yang bertujuan untuk meningkatkan performa dan keakuratan prediksi. Proses stacking melibatkan pembagian data menjadi beberapa bagian untuk dilatih pada setiap model yang berbeda, kemudian model-model tersebut diuji pada data uji yang berbeda dan diintegrasikan untuk membuat prediksi akhir. Pada gambar 2.5 dijelaskan misalkan terdapat output dari masing-masing algoritma berturut-turut yaitu linear regression, SVM, decision tree dan neural network memiliki akurasi sebesar 80%, 90%, 87%, 92%, kemudian output dari masing-masing algoritma tersebut dijadikan input menggunakan teknik stacking maka outputnya dapat menjadi 96%.



Sumber (Brijesh Soni, 2023; et al.,)

Gambar 2. 6 Proses Stacking

Penelitian ini menggunakan teknik *stacking*, dimana output algoritma BERT yaitu sentiment skor digabung dengan data teknikal sebagai input pada algoritma *predictor* yaitu LSTM dan XGBoost. Teknik stacking dipilih karena pada penelitian ini menggabungkan analisis teknikal dan analisis sentiment untuk prediksi harga saham.

2.8 Penelitian yang Terkait

Pada penelitian yang dilakukan oleh [14] dan [15], yang pertama [14] menggunakan pendekatan teknikal, fundamental dan sentiment untuk memprediksi harga saham sedangkan penelitian yang kedua [15] hanya fokus pada pendekatan fundamental dan sentiment saja. Penelitian pertama menggunakan algoritma J48 With Bagging, SVM dan Decision Tree untuk mengklasifikasikan keputusan antara “Buy”, “Sell” atau “hold”. Hasil penelitian menunjukkan model memiliki akurasi sekitar 91% dalam uji coba *Realtime*. Penelitian ini memperoleh hasil 283 saham diklasifikasikan “Buy”, 180 mencapai target harga, sementara lebih dari 70 tidak mencapai *stoploss*. Namun, pendekatan sentiment penelitian ini juga mendeteksi 36 berita buruk, 27 saham telah mencapai *stoploss*. Di sisi lain, penelitian kedua [15] hanya menggunakan algoritma Naïve Bayes, SVM, *Grey Relational Analysis (GRA)*, *Queen Genetic algorithm* dan *ensemble classifier* untuk memprediksi apakah harga saham akan “Up”, “Flat”, atau “Down”. Penelitian ini tidak menggunakan pendekatan teknikal sehingga hasil prediksi sangat dipengaruhi oleh opini publik. Berdasarkan kedua penelitian tersebut terlihat bahwa pendekatan teknikal dapat memberikan hasil yang objektif tetapi juga perlu memperhatikan faktor sentiment yang dapat mempengaruhi pasar secara signifikan.

Pada penelitian yang dilakukan oleh [16][17][18] mengeksplorasi pendekatan yang berbeda dalam memprediksi harga saham. Penelitian pertama [16] mengintegrasikan analisis fundamental dan teknikal menggunakan algoritma Multi-layer perceptron dengan *pre-processing* menggunakan data *discretization* dan *feature selection*. Hasil menunjukkan bahwa pendekatan ini berhasil memprediksi pergerakan harga saham dengan akurasi arah rata-rata sebesar 65,87% yang lebih baik daripada hasil yang diperoleh dari analisis fundamental dan teknis secara terpisah. Di sisi lain, penelitian [17] membahas perbandingan antara prediksi harga saham menggunakan pendekatan teknikal saja dengan pendekatan gabungan teknikal dan sentiment. Penelitian ini menggunakan algoritma RNN dan LSTM dan menemukan bahwa prediksi menggunakan dua pendekatan menghasilkan prediksi yang lebih akurat berdasarkan plot diagram. Namun, evaluasi model hanya berdasarkan plot diagram tanpa menggunakan metrik evaluasi yang lengkap seperti RMSE, MSE, MAE dan lainnya. Sedangkan penelitian [18] fokus pada

analisis teknikal saja, penelitian menggunakan dataset harga saham dengan fitur *open* dan *close* dan menguji dengan beberapa model algoritma. Hasilnya menunjukkan bahwa jaringan PSO-Elman Network memiliki performa yang lebih baik dibandingkan dengan model lainnya berdasarkan nilai matriks MSE dan MAE. Dari perbandingan ketiga penelitian tersebut, terlihat bahwa pendekatan gabungan antara analisis fundamental dan teknikal seperti pada penelitian [16] mampu memberikan hasil prediksi yang kuat. Di sisi lain, penelitian [17] menyoroti pentingnya menggabungkan pendekatan teknikal dan sentiment untuk meningkatkan akurasi, meskipun model tersebut masih bisa ditingkatkan. Sementara penelitian [18] menunjukkan bahwa fokus pada analisis teknikal saja dapat memberikan hasil yang baik dengan pemilihan model algoritma yang tepat. Namun, perlu diingat bahwa kekurangan penelitian [18] adalah fokusnya terbatas yang hanya menggunakan analisis teknikal saja sehingga tidak dapat digunakan untuk prediksi harga saham jangka panjang.

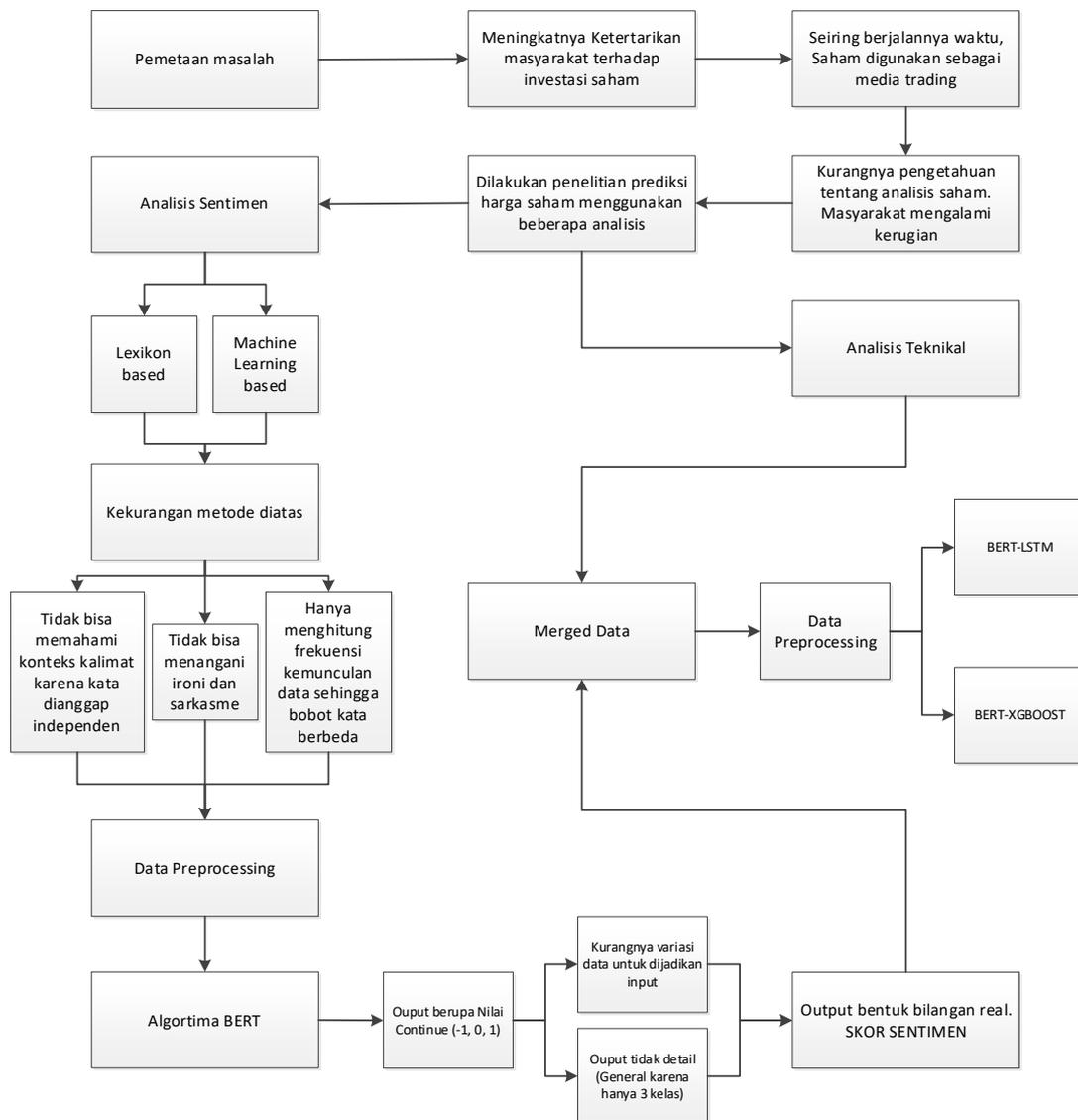
Pada penelitian yang dilakukan oleh [19][20][1], tiga penelitian tersebut menggambarkan beragam pendekatan dalam memprediksi pergerakan harga saham. Penelitian pertama [19] menggunakan kombinasi analisis teknikal dan algoritma gradient boosting untuk memprediksi indeks harga saham gabungan (IHSG). Hasilnya menunjukkan perbedaan signifikan dalam akurasi antara model yang menggunakan kombinasi analisis teknikal dan algoritma gradient boosting (86.96%) dengan yang hanya menggunakan algoritma gradient boosting saja (74.2%). Hal ini menegaskan bahwa analisis teknikal memiliki pengaruh yang besar terhadap model. Di sisi lain, penelitian [20] fokus pada prediksi efek sentiment berita terhadap pasar saham. Penelitian ini mengembangkan kamus analisis sentiment atau korpus mengenai keuangan untuk memprediksi tren pergerakan harga saham jangka pendek. Meskipun hanya menggunakan sentiment berita, penelitian ini mencapai akurasi sebesar 70.59%, menunjukkan bahwa analisis sentiment juga dapat memberikan kontribusi yang signifikan dalam prediksi saham. Lalu pada penelitian [1] lebih berfokus pada pendekatan teknikal menggunakan algoritma *random forest regression* dan *artificial neural network* untuk memprediksi harga penutupan saham. Hasil analisis komparatif menunjukkan bahwa *artificial neural network* memberikan prediksi yang lebih baik dibandingkan dengan *random forest* dengan nilai RMSE : 0.42, MAPE : 0.77 dan MSE : 0.013.

Penelitian ini juga menjelaskan bahwa model *deep learning* dapat dikembangkan dengan mempertimbangkan artikel berita keuangan dan parameter keuangan lainnya untuk hasil yang lebih baik di masa mendatang. Dari ketiga penelitian tersebut, terlihat bahwa pendekatan yang menggabungkan analisis teknikal dengan algoritma seperti pada penelitian [19] dapat memberikan akurasi yang tinggi. Namun, analisis sentiment seperti pada penelitian [20] juga memberikan kontribusi yang signifikan, terutama dalam prediksi tern pergerakan harga saham jangka pendek. Sementara itu, penggunaan model *deep learning* seperti *artificial neural network* dalam penelitian [1] menunjukkan potensi untuk hasil prediksi yang lebih baik di masa mendatang dengan mempertimbangkan lebih banyak parameter dan informasi yang relevan.

Adapun beberapa penelitian yang menggunakan algoritma BERT untuk klasifikasi sentiment. Penelitian pertama [4] bertujuan untuk mendeteksi berita palsu atau *hoax* menggunakan algoritma BERT. Dataset berjumlah 3.465 berita palsu dan 766 berita asli atau nyata. Dataset tersebut digunakan untuk menguji 3 model yaitu TF-IDM + SVM, TF-IDM + Naïve Bayes dan indobert. Hasilnya menunjukkan bahwa algoritma BERT memiliki akurasi yang sangat tinggi dengan *precision*, *recall*, *F1-Score*, dan akurasi sebesar 94.66%. Hal ini mengkonfirmasi bahwa algoritma BERT sangat baik untuk digunakan dalam text mining dan analisis sentiment. Pada penelitian [21] menggunakan algoritma BERT untuk menganalisis sentiment review film dalam bahasa inggris. Mereka membagi data menjadi data training dan data testing dan menggunakan *convolution matrix* untuk validasi. Hasil evaluasi mencapai akurasi sebesar 73%. Meskipun akurasi ini tidak sebaik pada penelitian pertama, tetapi tetap menunjukkan kegunaan algoritma BERT dalam analisis sentiment. Selanjutnya penelitian [8] menggunakan algoritma BERT untuk analisis sentiment pasar saham berdasarkan data dari Twitter. Penelitian ini mengklasifikasikan sentiment menjadi positif atau negatif, dan hasilnya menunjukkan bahwa model BERT dapat mengenali sentiment investor dengan akurasi lebih dari 87.3%. Terakhir penelitian [22] membandingkan algoritma hybrid BERT dan CNN dalam analisis sentiment dan prediksi harga saham. Hasilnya menunjukkan bahwa algoritma BERT-CNN memiliki akurasi yang tinggi dengan nilai *accuracy* 0.905, *recall rate* 0.901, dan *F1-Score* 0.903 yang lebih baik dibandingkan dengan algoritma lain yang diuji. Berdasarkan perbandingan keempat penelitian tersebut,

terlihat bahwa algoritma BERT memang sangat efektif dalam berbagai jenis analisis teks dan sentiment. Namun, tingkat akurasi dan kinerja dapat bervariasi tergantung pada konteks dan dataset yang digunakan serta penggunaan algoritma tambahan seperti CNN dalam beberapa kasus dapat meningkatkan performa analisis sentiment dan prediksi harga saham.

2.9 Kerangka Pikir Penelitian



Gambar 2. 7 Kerangka Berfikir