

BAB II. TINJAUAN PUSTAKA

2.1 Landasan Teori

2.1.1 *Machine Learning*

Machine Learning adalah salah satu bagian dari ilmu komputer yang memungkinkan komputer untuk belajar dari data guna mengenali pola tertentu. Algoritma dan model yang digunakan untuk menemukan pola dalam data ini dapat dimanfaatkan untuk membuat prediksi atau keputusan berdasarkan informasi yang ada dalam data. Babu dkk menyatakan bahwa *Machine Learning* adalah proses pengembangan algoritma yang mampu belajar dan melakukan prediksi berdasarkan *Dataset*. (Babu dkk., 2016). Dengan semakin banyaknya data, penggunaan *Machine Learning* menjadi semakin meningkat, dan sudah banyak berbagai industri yang telah memanfaatkan *Machine Learning* untuk mengolah data yang ada.

2.1.2 *Stacking Ensemble Learning*

Ensemble Learning merupakan teknik pada *Machine Learning* yang dapat mengintegrasikan hasil ramalan dari berbagai model *Machine Learning* sehingga akan memperbaiki kinerja serta ketepatan dalam hasil prediksi. Di dalam *Ensemble Learning* ini, akan mengambil keuntungan dari kekuatan masing-masing model serta mengurangi kelemahannya dengan cara menggabungkan hasil prediksi dari tiap model.

Terdapat beberapa metode yang ada pada *Ensemble Learning*, diantaranya seperti metode *Bagging*, *Boosting*, dan *Stacking*, penelitian ini akan menggunakan metode *Ensemble Learning* yaitu *Stacking*. Konsep utama dari *Stacking* adalah mengurangi kesalahan generalisasi dengan menggabungkan hasil prediksi dari berbagai model dasar yang berbeda dan menggunakan model meta untuk menggabungkan hasil prediksi dari model-model dasar tersebut (Cui dkk., 2021) seperti pada Gambar 2.1 berikut.

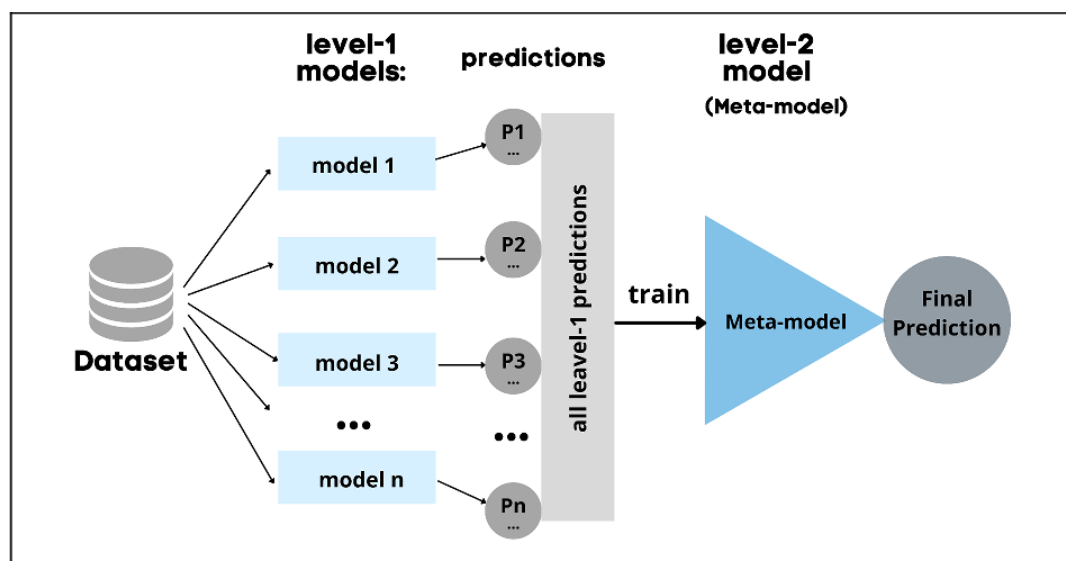


Figure 3. Stacking Structure.

Gambar 2. 1 Ilustrasi *Stacking Ensemble* (Hayat dkk., 2024)

2.1.3 *Linear Regression* dan Regularisasi

Linear Regression merupakan salah satu teknik analisis prediktif yang digunakan untuk memodelkan hubungan linear antara variabel independen dan variabel dependen. Namun, ketika data memiliki kompleksitas tinggi, regresi linier standar sering rentan terhadap *overfitting*. Oleh sebab itu, dikembangkan

pendekatan berbasis regularisasi seperti *Ridge*, *Lasso*, dan *Elastic Net*, yang memberikan penalti pada koefisien untuk meningkatkan stabilitas model (Ding, 2024; Roy dkk., 2015).

Pendekatan regularisasi secara khusus banyak digunakan dalam penelitian sebelumnya yang menekankan bahwa kombinasi regularisasi L1 dan L2 dapat menekan *overfitting* sekaligus mengatasi multikolinearitas dalam model regresi berganda (Mishra & Pandey, 2025). Dengan demikian, keempat teknik ini berperan penting dalam mengembangkan model prediksi yang lebih robust.

2.1.3.1. *Linear Regression*

Linear Regression bertujuan untuk memodelkan hubungan linear antara variabel input terhadap variabel output dengan persamaan yang diadaptasi dari penelitian Mishra & Padey (2025) berikut:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \varepsilon$$

Parameter estimasi diperoleh melalui metode *Ordinary Least Squares* (OLS) dengan meminimalkan total kuadrat error:

$$\min_{\beta} \sum_{i=1}^m (y_i - \hat{y}_i)^2$$

Meskipun sederhana, regresi linier memiliki keterbatasan pada dataset dengan multikolinearitas dan kecenderungan *overfitting*, sehingga diperlukan pendekatan regularisasi.

2.1.3.2. *Lasso Regression (L1 Regularization)*

Lasso Regression menggunakan penalti L1 yang memaksa sebagian koefisien menjadi nol sehingga bertindak sebagai seleksi fitur otomatis, *Lasso* ini sangat efektif dalam menyederhanakan model dengan memilih hanya prediktor yang paling relevan (Mishra & Pandey, 2025), dengan fungsi persamaan *loss Lasso* yang diadaptasi pada penelitian Mishra & Padey (2025) sebagai berikut :

$$\min_{\beta} \left(\sum_{i=1}^m (y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^n |\beta_j| \right)$$

Lasso membantu mengurangi *overfitting* dengan menghasilkan model yang lebih sederhana, yang sangat bermanfaat ketika terdapat banyak fitur tidak penting.

2.1.3.3. *Ridge Regression (L2 Regularization)*

Ridge Regression menambahkan penalti L2 ke dalam fungsi *loss* untuk mengatasi multikolinearitas. Ridge tidak menghapus variabel, tetapi mengecilkan nilai koefisien sehingga model menjadi lebih stabil. Mishra & Pandey (2025) menjelaskan bahwa *Ridge* sangat efektif ketika semua fitur dianggap relevan tetapi mengalami masalah kolinearitas, berikut persamaan fungsi *loss Ridge* yang diadaptasi pada penelitian Mishra & Padey (2025) :

$$\min_{\beta} \left(\sum_{i=1}^m (y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^n \beta_j^2 \right)$$

Ridge memberikan penalti merata terhadap seluruh koefisien, membuatnya cocok untuk model prediksi dengan banyak fitur yang saling berkorelasi

2.1.3.4. *Elastic Net Regression*

Elastic Net mengombinasikan penalti L1 dan L2 secara bersamaan. Metode ini dikembangkan untuk mengatasi kelemahan *Lasso* yang tidak stabil pada fitur berkorelasi tinggi dan kelemahan *Ridge* yang tidak dapat melakukan seleksi fitur. *Elastic Net* mengintegrasikan keunggulan *Lasso* dan *Ridge*, serta mampu mencapai keseimbangan antara kompleksitas model dan kinerja prediksi tanpa memerlukan sumber daya komputasi yang besar (Ding, 2024), berikut persamaan fungsi *Elastic Net* yang diadaptasi dari penelitian Ding (2024) :

$$\min_{\beta} \left(\sum_{i=1}^m (y_i - \hat{y}_i)^2 + \lambda_1 \sum_{j=1}^n |\beta_j| + \lambda_2 \sum_{j=1}^n \beta_j^2 \right)$$

Elastic Net menjadi pilihan ideal ketika dataset memiliki jumlah fitur banyak dan saling berkorelasi.

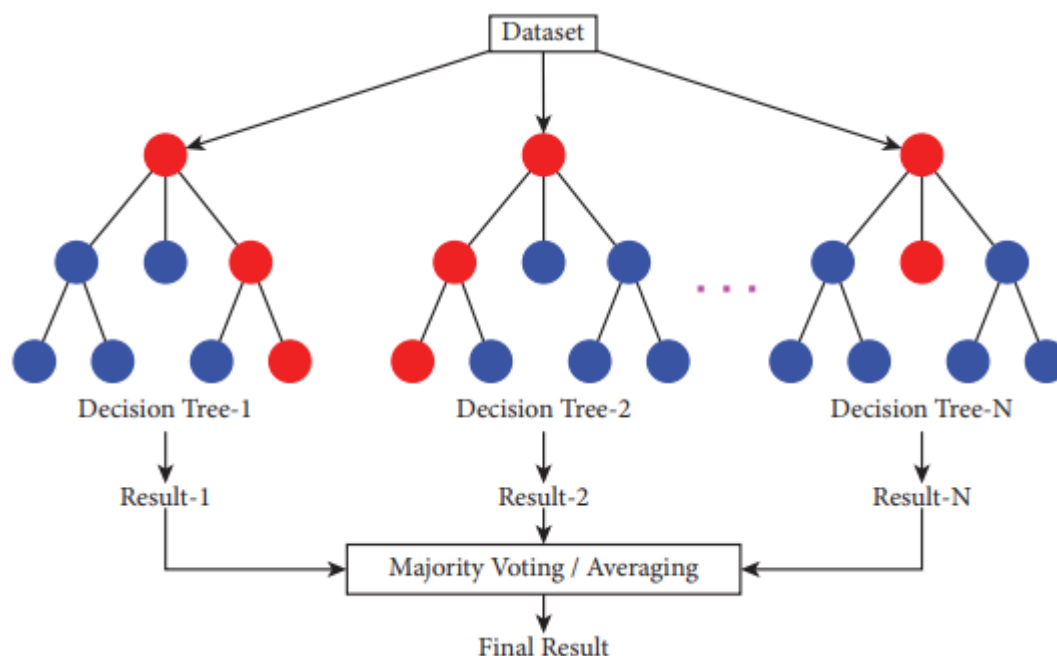
2.1.4 *Random Forest*

Dalam penelitian oleh Gupta (2024) menjelaskan bahwa Random Forest adalah algoritma pembelajaran ensemble yang membangun banyak pohon keputusan selama pelatihan dan menghasilkan prediksi rata-rata dari masing-masing pohon untuk tugas regresi (Gupta & Gupta, 2024). *Forest* atau ‘hutan’ terdiri dari banyak *tree* (pohon) yang terbentuk melalui suatu proses *Bagging* atau dikenal dengan *Bootstrap Aggregating*.

Model ini efektif menangani data yang tidak seimbang dan *missing value*, serta mengatasi *overfitting*, sehingga meningkatkan generalisasi terhadap data baru. *Random Forest* ini secara luas dianggap sebagai salah satu teknik yang paling

efisien dan kuat dalam *Machine Learning*, yang dapat diterapkan dalam kategorisasi otomatis, *data forecasting*, dan lainnya (Ahmed Salman dkk., 2024).

Berikut ilustrasi dari algoritma metode *Random Forest* pada Gambar 2.2.

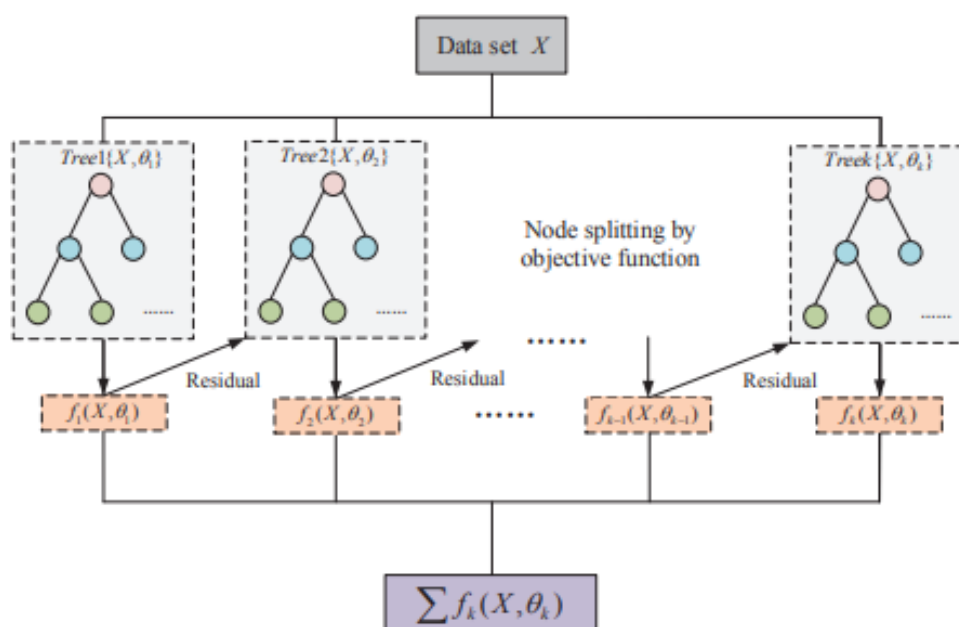


Gambar 2. 2 Ilustrasi *Random Forest* (Khan dkk., 2021)

2.1.5 *XGBoost*

XGBoost (*Extreme Gradient Boosting*) adalah algoritma machine learning berbasis ensemble yang mengintegrasikan sejumlah besar *Decision Tree* guna menghasilkan prediksi dengan akurasi tinggi. Algoritma ini dikembangkan dari konsep dasar *Gradient Boosting Machine* (*GBM*) pada tahun 2001 yang dipelopori oleh Friedman, dalam penelitiannya, ia menghubungkan *Boosting* dan optimasi dalam membangun *Gradient Boosting Machine* (*GBM*) (Ramadani, 2023). Algoritma ini sangat terkenal di dunia ilmu data dan sering digunakan dalam

kompetisi seperti Kaggle. *XGBoost* dapat digunakan untuk memecahkan berbagai masalah, termasuk regresi, klasifikasi, dan peringkat. Kelebihan *XGBoost* meliputi kemampuan dalam mendeteksi tren data yang rumit, menangani nilai yang hilang secara efisien (S. Sharma & Singh, 2024). Berikut merupakan ilustrasi dari algoritma *XGBoost* pada Gambar 2.3.



Gambar 2. 3 Ilustrasi *XGBoost* (Guo dkk., 2020)

2.1.6 Bayesian Optimization

Bayesian Optimization adalah teknik *Hyperparameter Tuning* yang memodelkan kinerja suatu model *Machine Learning* sebagai fungsi probabilistik. Teknik ini menyeimbangkan eksplorasi dan eksploitasi, sehingga menjadi metode yang efisien dalam mengoptimalkan *Hyperparameter*, terutama pada model yang kompleks dengan biaya evaluasi yang tinggi (Shaji dkk., 2024).

Selain itu, *Bayesian Optimization* menerapkan fungsi akuisisi untuk menentukan titik data berikutnya yang akan diuji, dengan mempertimbangkan keseimbangan antara eksplorasi (menguji kombinasi baru yang belum dievaluasi) dan eksploitasi (memilih kombinasi yang telah menunjukkan hasil terbaik sejauh ini). Hal ini menjadikannya sangat efektif dalam mengoptimalkan model dengan ruang *Hyperparameter* yang luas dan biaya komputasi yang tinggi.

2.2 *State of The Art* Bidang Penelitian

Saat ini, penelitian di berbagai bidang telah banyak mengadopsi metode *Machine Learning* untuk berbagai keperluan, seperti dalam sektor bisnis untuk memprediksi permintaan produk, di sektor kesehatan untuk menganalisis risiko penyakit, serta di sektor keuangan untuk memperkirakan pergerakan harga saham. Khusus dalam prediksi harga saham, berbagai algoritma *Machine Learning* telah digunakan, termasuk *Random Forest*, *XGBoost*, *K-Nearest Neighbors (KNN)*, *Support Vector Machine (SVM)*, dan *AdaBoost*.

Tabel 2.1 menyajikan perbandingan berbagai penelitian yang membahas prediksi harga saham menggunakan metode *Machine Learning*. Beberapa penelitian menunjukkan bahwa *Random Forest* dan *XGBoost* sering digunakan karena kemampuannya dalam menangani data keuangan yang kompleks. Studi sebelumnya telah membuktikan bahwa model *XGBoost* umumnya menghasilkan akurasi yang tinggi, seperti yang terlihat dalam penelitian Putratama dkk. (2024), Permana & Ernawati (2023), Kocaoğlu dkk (2023), dan Zhao Y (2024) yang menunjukkan bahwa *XGBoost* memiliki performa unggul dalam memprediksi harga saham dibandingkan dengan metode lainnya yang diuji, sementara itu pada penelitian lainnya yang dilakukan oleh Destriana R (2023), Wu Y (2023), didapati bahwa performa *Random Forest* unggul dalam memprediksi harga saham.

Selain itu, penelitian yang dilakukan oleh Şimsek A (2025) menunjukkan bahwa pendekatan *Stacking Ensemble Learning*, yang menggabungkan *Random Forest* dan *XGBoost*, dapat meningkatkan akurasi prediksi secara signifikan dibandingkan dengan penggunaan model tunggal. Hal ini juga didukung oleh

penelitian Haya & Ramme (2024) yang menerapkan metode *Stacking Ensemble Learning* berbasis pohon keputusan untuk meningkatkan akurasi prediksi dalam kasus klasifikasi. Penelitian lain yang dilakukan oleh Nti dkk (2020) serta Manjunath dkk (2024) juga menunjukkan bagaimana perbandingan metode Bagging, Boosting, dan Stacking dengan hasil bahwa metode Stacking mengungguli metode lainnya.

Dalam upaya lebih lanjut untuk meningkatkan akurasi prediksi, beberapa penelitian telah menerapkan *Bayesian Optimization* sebagai teknik *Hyperparameter Tuning*, seperti yang dilakukan oleh Liang dkk. (2024) dan Khurshid dkk. (2025). Hasil dari penelitian ini menunjukkan bahwa *Bayesian Optimization* dapat meningkatkan performa model *XGBoost* secara signifikan dalam berbagai kasus prediksi, termasuk di sektor finansial dan kesehatan.

Berdasarkan penelitian-penelitian tersebut, terlihat bahwa pendekatan *Stacking Ensemble Learning* dengan *Random Forest* dan *XGBoost* telah banyak digunakan dalam prediksi harga saham. Namun, masih terdapat peluang untuk meningkatkan performa model dengan mengoptimalkan *Hyperparameter* menggunakan *Bayesian Optimization*. Oleh karena itu, penelitian ini akan berfokus pada penerapan *Bayesian Optimization* dalam *Stacking Ensemble Learning* yang menggabungkan *Random Forest* dan *XGBoost*, dengan harapan dapat menghasilkan model yang lebih akurat dalam memprediksi pergerakan harga saham. Berikut ini adalah tabel perbandingan penelitian terkait yang telah dilakukan, sebagaimana disajikan dalam Tabel 2.1.

Tabel 2. 1 *State of The Art* Penelitian Terkait

<i>State of The Art</i>		
Penelitian		Hasil
1.	Judul : <i>Using Stacked Generalization Model in Stock Price Forecasting: A Comparative Analysis on BIST100 Indeks</i> (2025) (Şimşek, 2025) Peneliti : Şimşek A Metode : <i>Random Forest, XGBoost, Stacking</i>	Hasil penelitian menunjukkan bahwa model Stacked Generalization dengan menggabungkan model <i>Random Forest</i> dan <i>XGBoost</i> memberikan kinerja yang lebih baik dalam membuat prediksi dibandingkan dengan model lainnya. Model ini dievaluasi menggunakan indikator kinerja seperti <i>RMSE</i> dengan hasil 0.00002, <i>MSE</i> 0.00688, <i>MAE</i> 0.00385, dan <i>R2</i> 0.98024.
2.	Judul : Metode Random Forest Dan XGBoost: Studi Kasus Prediksi Arah Penutupan Harga Saham (2023) (Ramadani, 2023) Peneliti : Destriana Ramadani Metode : <i>Random Forest, XGBoost</i>	Memprediksi Arah Penutupan Harga Saham PT. Bank Rakyat Indonesia menggunakan metode <i>Random Forest</i> dan <i>XGBoost</i> serta <i>GridSearch</i> untuk membangun model prediksi arah harga saham BBRI. Hasil penelitian menunjukkan bahwa model prediksi menggunakan <i>Random Forest</i>

		menghasilkan nilai akurasi mencapai 90,01% dengan AUC sebesar 0,95, sedangkan model prediksi menggunakan <i>XGBoost</i> menghasilkan nilai akurasi 89,45% dengan AUC sebesar 0,94.
3.	Judul : Comparative Evaluation of Machine Learning Methods for Predicting Stock Price Changes (2024) (Putratama dkk., 2023) Peneliti : Putratama G, Fahreza S, Ramandhani Y Metode : <i>Random Forest</i>, <i>K-Nearest Neighbors (KNN)</i>, <i>XGBoost</i>	Hasil penelitian menunjukkan bahwa Random Forest mencapai skor ROC tertinggi, yaitu 0,98, sedangkan <i>XGBoost</i> menunjukkan kinerja yang unggul dalam hal akurasi, recall, dan presisi masing-masing sebesar 0,93, 0,69, 0,77. Metode <i>Windowing</i> juga diterapkan pada <i>Dataset</i> untuk mengatasi masalah <i>Overfitting</i>.
4.	Judul : Perbandingan Algoritma Extreme Gradient Boosting Dan Random Forest Untuk Memprediksi Harga Terendah Saham Dengan Index ISSI. (2023) (Permana & Ernawati, 2023) Peneliti : Permana F, Ernawati I Metode : <i>XGBoost</i>, <i>Random Forest</i>	Dalam penelitian ini, algoritma <i>Random Forest</i> dan <i>Extreme Gradient Boosting (XGBoost)</i> digunakan untuk memprediksi harga terendah saham dengan indeks ISSI. Model Random Forest mencapai nilai Mean Squared Error (MSE) sebesar 0,6458, Mean

		<i>Absolute Percentage Error</i> (MAPE) sebesar 0,0033, dan koefisien determinasi (R^2) sebesar 0,9985 untuk data uji. Sementara itu, model <i>XGBoost</i> menghasilkan nilai MSE sebesar 0,8019, MAPE sebesar 0,0037, dan R^2 sebesar 0,9982 untuk data uji.
5.	Judul : Machine Learning-Based Sales Prediction Using Bayesian Optimized XGBoost Algorithms (2024) (Liang dkk., 2024) Peneliti : Liang X, Lin Y, dkk Metode : <i>XGBoost</i>	Dalam penelitian ini, algoritma <i>Extreme Gradient Boosting</i> (<i>XGBoost</i>) digunakan untuk memprediksi penjualan produk menggunakan data pesanan "Zhenlong (Zhenpin)" pada tahun 2022. Model <i>XGBoost</i> yang dioptimalkan dengan Bayesian Optimization menghasilkan nilai <i>Mean Squared Error</i> (MSE) sebesar 2,3, <i>Root Mean Squared Error</i> (RMSE) sebesar 1,5, <i>Mean Absolute Error</i> (MAE) sebesar 0,78, dan koefisien determinasi (R^2) sebesar

		0,86. Hasil ini menunjukkan bahwa model <i>XGBoost</i> yang dioptimalkan dapat memprediksi penjualan produk dengan akurasi tinggi.
6.	Judul : Unveiling diabetes onset: Optimized XGBoost with Bayesian Optimization for enhanced prediction (2025) (Khurshid dkk., 2025) Peneliti : Khurshid M, Manzoor S, dkk Metode : <i>XGBoost</i>, <i>Bayesian Optimization</i>	Dalam penelitian ini, algoritma <i>XGBoost</i> yang dioptimalkan dengan <i>Bayesian Optimization</i> digunakan untuk memprediksi onset diabetes. Model <i>XGBoost</i> dengan <i>Bayesian Optimization</i> mencapai nilai akurasi sebesar 0,9726, F1-score sebesar 0,9572, dan <i>Matthews Correlation Coefficient</i> (MCC) sebesar 0,8118 untuk data uji. Algoritma ini menunjukkan performa yang unggul dalam memprediksi onset diabetes dibandingkan dengan algoritma Machine Learning tradisional lainnya.
7.	Judul : Sector-Based Stock Price Prediction with Machine Learning Models (2022) (Kocaoğlu dkk., 2022)	Dalam penelitian ini, algoritma <i>Random Forest</i> dan <i>Extreme Gradient Boosting</i> (<i>XGBoost</i>) digunakan untuk

	Peneliti : Kocaoğlu D, Turgut K, Konyar M Metode : <i>XGBoost</i>, <i>Support Vector Machine (SVM)</i>, <i>K-Nearest Neighbors (KNN)</i>, <i>Random Forest</i>	memprediksi harga saham dari perusahaan-perusahaan di sektor yang berbeda yang diperdagangkan di Indeks Bursa Istanbul (BIST) 30. Model <i>Random Forest</i> mencapai nilai <i>Mean Squared Error</i> (MSE) sebesar 0,233, <i>Mean Absolute Percentage Error</i> (MAPE) sebesar 0,031, dan koefisien determinasi (R^2) sebesar 0,987 untuk data uji. Sementara itu, model <i>XGBoost</i> menghasilkan nilai MSE sebesar 1,640, MAPE sebesar 0,029, dan R^2 sebesar 0,989 untuk data uji. Penelitian ini menemukan bahwa algoritma <i>XGBoost</i> memberikan hasil prediksi terbaik dibandingkan algoritma lainnya
8.	Judul : Penerapan Algoritma Stacking Ensemble Machine Learning Berbasis Pohon untuk Prediksi Penyakit Diabetes (2024) (Haya & Ramme, 2024) Peneliti : Haya A, Ramme M	Dalam penelitian ini, <i>Random Forest</i> dan <i>XGBoost</i> digunakan sebagai <i>Base Learner</i> dalam kerangka <i>Stacking Ensemble</i> untuk klasifikasi diabetes. Model <i>Random Forest</i> mencapai

	Metode : <i>CatBoost, XGBoost, Extra Tree, Decision Tree, Random Forest, Light Gradient Boosting Machine (LGBM)</i>	akurasi sebesar 84,82%, sementara <i>XGBoost</i> mencapai 75,71%. Model <i>Stacking Ensemble</i>, yang menggabungkan <i>Base Learner</i> ini dengan yang lain dan menggunakan <i>Light Gradient Boosting Machine (LGBM)</i> sebagai meta learner, menghasilkan akurasi yang lebih tinggi yaitu 87,47%, yang menunjukkan efektivitas teknik <i>Stacking</i> dibandingkan model tunggal.
9.	Judul : Stock Price Prediction Based on Simple Decision Tree Random Forest and XGBoost (2023) (Wu, 2023) Peneliti : Wu Y Metode : <i>XGBoost, Simple Decision Tree, Random Forest (RF)</i>	Hasil penelitian menunjukkan bahwa model Random Forest memberikan hasil terbaik dengan akurasi tertinggi dan nilai MSE terkecil, dengan hasil 99,80% untuk <i>Bank of America (BAC)</i>, 99,83% untuk <i>Chevron Corporation (CVX)</i>, dan 99,74% untuk <i>The Home Depot, Inc. (HD)</i>. Sebaliknya, model <i>XGBoost</i> menunjukkan performa terendah

		dengan akurasi 95,19% untuk BAC, 97,74% untuk CVX, dan 98,15% untuk HD, yang disebabkan oleh sensitivitas model terhadap skala variabel input. Nilai MSE untuk model <i>XGBoost</i> lebih tinggi dibandingkan dengan model <i>Random Forest</i>, dengan nilai 7,70 untuk BAC, 147,32 untuk CVX, dan 514,79 untuk HD.
10.	Judul : Analysis of LV Stocks Based on Linear Regression, Random Forest and XGBoost (2024) (Zhao, 2024) Peneliti : Zhao Y Metode : <i>Linear Regression, Random Forest, XGBoost</i>	Penelitian ini menggunakan beberapa model <i>Machine Learning</i> untuk memprediksi harga saham di masa depan, dengan fokus pada regresi linier, <i>Random Forest (RF)</i>, dan <i>XGBoost</i>. Dari hasil penelitian, model <i>XGBoost</i> terbukti sangat efisien dalam menghasilkan prediksi harga saham dengan nilai <i>Mean Squared Error (MSE)</i> yang rendah dan skor <i>R-squared (R²)</i> yang tinggi.

11.	Judul : <i>A comprehensive evaluation of ensemble learning for stock-market prediction</i> (Nti dkk., 2020) Peneliti : Nti I, Adekoya A, Weyori B Metode : <i>Decision Tree, SVM, Neural Network</i>	Dalam penelitian ini, <i>Decision Tree, SVM</i>, dan <i>Neural Network</i> diuji dengan teknik <i>Bagging, Boosting, Blending</i>, dan <i>Stacking</i> untuk prediksi pasar saham. Hasil menunjukkan bahwa <i>Stacking Ensemble</i> mencapai akurasi tertinggi hingga 100% pada data NYSE dan BSE, serta 90–98% pada data GSE dan JSE. Temuan ini menegaskan bahwa <i>Stacking</i> lebih unggul dibandingkan model tunggal maupun teknik <i>Bagging</i> dan <i>Boosting</i>.
12.	Judul : Stock market prediction employing ensemble methods: the Nifty50 index (Manjunath dkk., 2024) Peneliti : Manjunath C, Marimuthu B, Ghosh B Metode : <i>Random Forest, Gradient Boosting Regression, SVR, Decision Tree Regression, Linear Regression</i>	Penelitian ini menunjukkan bahwa teknik ensemble, khususnya <i>Bagging dengan Random Forest</i> dan <i>Stacking Ensemble</i>, mampu memprediksi indeks Nifty50 dengan akurasi hampir sempurna. Model <i>Bagging</i> mencapai R^2 0.9990 dengan <i>RMSE</i> 0.0084, sedangkan <i>Stacking Regresor</i> terbaik mencapai R^2 0.9999 dengan <i>RMSE</i> 0.0085. Hasil ini menegaskan bahwa

		metode <i>Stacking</i> mengungguli metode lainnya.
--	--	--

Berdasarkan penelitian sebelumnya, *Random Forest* dan *XGBoost* serta metode *Stacking*, secara terpisah telah terbukti sebagai metode yang efektif dalam memprediksi harga saham, dengan berbagai studi yang menunjukkan keunggulan masing-masing dalam akurasi dan kestabilan prediksi. Beberapa penelitian juga telah menerapkan pendekatan *Stacking Ensemble* dalam meningkatkan performa prediksi dengan menggabungkan keunggulan pada beberapa model. Selain itu, *Bayesian Optimization* telah digunakan dalam berbagai penelitian untuk mengoptimalkan parameter model guna meningkatkan akurasi lebih lanjut. Dengan itu penelitian ini dalam kebaruan penelitiannya akan **menggunakan kombinasi algoritma *Random Forest* dan *XGBoost* dengan eksplorasi *Meta Learner* dalam pendekatan *Stacking Ensemble* yang dioptimalkan dengan *Bayesian Optimization***, dengan harapan dapat menghasilkan model dengan prediksi harga saham yang lebih akurat dan andal dibandingkan dengan pendekatan sebelumnya.