SCIENTIFIC OASIS

Decision Making: Applications in Management and Engineering

Journal homepage: www.dmame-journal.org
ISSN: 2560-6018, eISSN: 2620-0104

A Decision-Support Approach to International Natural Gas Price Forecasting Using Machine Learning Models

Peifeng Wu[1,*] Yaqiang Chen[2]

1 Faculty of Statistics, Jilin University of Finance and Economics, Changchun 130000, Jilin, China. Email: 18238298675@163.com
2 Faculty of Statistics, Jilin University of Finance and Economics, Changchun 130000, Jilin, China. Email: c811931070@163.com

ARTICLE INFO

Article history:
Received 10 June 2025
Received in revised form 01 July 2025
Accepted 25 July 2025
Available online 04 September 2025

Keywords:
Machine Learning Forecasting, Natural Gas Market Prices, Decision-Making in Energy Management, Predictive Analytics

ABSTRACT

Accurate forecasting of international natural gas prices is essential for effective decision-making within the context of a volatile energy market, as unpredictable responses to price fluctuations often arise from non-stationary behaviour. Traditional econometric approaches frequently encounter limitations in capturing market volatility, nonlinear dynamics, and structural breaks, which diminishes their practical utility for strategic planning and operational decisions. This study seeks to conduct a comparative analysis of multiple machine learning (ML) techniques, including linear regression, support vector machines, decision trees, random forests, neural networks, and ensemble methods, in forecasting international natural gas prices using datasets obtained from the International Energy Agency (IEA), Global Data, and the Bloomberg terminal. The findings indicate that more sophisticated ML models, particularly ensemble methods and neural networks, outperform conventional forecasting approaches in terms of accuracy and reliability. Furthermore, the study highlights that forecasts generated through ML can significantly enhance decision-making processes for key stakeholders in the natural gas sector, including government policymakers, investors, and oil and gas producers, by informing structured risk management, optimising resource allocation, and supporting long-term strategic planning.

1. Introduction

Natural gas represents a key element of the global energy mix, emphasising the importance of accurate price forecasting [21]. As a commodity utilised for electricity generation, heating, and industrial applications, fluctuations in natural gas prices have significant economic implications [19]. Traditional forecasting approaches, including time series analysis and econometric models, often require substantial time to capture the complex dynamics of natural gas markets [28]. These conventional methods generally rely on linear functions and historical ratios, limiting their ability to account for market shocks and the nonlinear interactions among influencing factors [32].

Consequently, ML has emerged as a promising approach for addressing challenges in natural gas price prediction, particularly through the application of natural language processing (NLP)

* *Corresponding author.*
E-mail address: 18238298675@163.com

https://doi.org/10.31181/dmame8220251519

techniques. ML models are particularly effective in recognizing higher-order and nonlinear relationships and are primarily advantageous for identifying patterns within data [25]. By leveraging large datasets and a combination of algorithms, ML approaches can surpass the capabilities of traditional models, offering improved predictive performance [30]. This capacity is especially valuable in the context of global natural gas markets, where pricing is influenced by geopolitical events, fluctuations in supply and demand, and environmental regulations [6]. Accurate forecasting of natural gas prices is therefore critical, as these figures are highly detailed and sensitive, directly impacting decision-making within the energy sector.

Furthermore, precise natural gas price forecasts provide significant benefits for key stakeholders, including policymakers, investors, and energy producers [23], enabling them to evaluate potential risks, allocate resources efficiently, and develop strategic plans. Enhanced forecasting accuracy supports risk management, cost optimisation, capacity planning, and strategic administrative decision-making [26]. Over recent years, international natural gas prices have exhibited considerable volatility, with trends in liquefied natural gas (LNG) markets expanding amid the rise of renewables and evolving regulatory frameworks, introducing further complexity for price prediction [1]. These conditions necessitate predictive models capable of handling high-dimensional and dynamically changing datasets. ML techniques, with their ability to learn from extensive datasets and adapt over time, are particularly well-suited to address these challenges [12]. Accordingly, this study aims to develop and validate multiple ML models for forecasting natural gas prices in international markets. While ML models demonstrated superior predictive performance compared to traditional methods, limitations inherent in generic ML implementations highlight areas for further refinement. This research contributes both theoretically and practically by providing valuable insights for energy sector stakeholders.

The application of ML to natural gas pricing constitutes a novel area of investigation. It aligns with broader trends emphasising human-centric, data-driven solutions across regulated sectors, with the aim of improving forecast accuracy and enhancing organisational planning [7]. Given the rapid evolution of energy markets, the relevance of reliable and precise forecasting is expected to grow, rendering this study timely and significant. The primary objective of this research is to evaluate the effectiveness of various ML techniques in forecasting international natural gas prices. Specifically, the study aims to:

- Compare the predictive performance of ML models including Linear Regression, Support Vector Machines (SVM), Decision Trees, Random Forests, Neural Networks, and Ensemble Methods.
- Assess the accuracy and reliability of these models in capturing the complex, nonlinear dynamics of natural gas pricing.
- Identify the strengths and limitations of each model within the context of price forecasting.

Provide actionable insights and recommendations for energy sector stakeholders to enhance decision-making and risk management through ML-based forecasts.

The study seeks to address the following research questions:

1. How do ML models perform relative to traditional statistical methods in forecasting natural gas prices?
2. Which ML techniques deliver the highest levels of accuracy and reliability for price predictions?
3. What are the specific advantages and limitations of different ML models in capturing the dynamics of international natural gas markets?
4. How can ML-derived forecasts be integrated into strategic decision-making processes within the energy sector?

By offering a comprehensive evaluation of various ML approaches for predicting international natural gas prices, this research advances the field of energy economics. It provides reliable guidance for investors, policymakers, and analysts, supporting more informed decision-making and enhancing risk management practices in the energy sector.

2. Review of Literature

2.1 Price Forecasting in Commodity Markets

Forecasting commodity prices represents a complex yet critical task due to the uncertainty and multitude of factors influencing price fluctuations [18]. Accurate forecasts enable investors, policymakers, and commercial entities to make informed decisions and mitigate associated risks. Traditional forecasting approaches in commodity markets often employ econometric models, such as the Auto Regressive Integrated Moving Average (ARIMA), which primarily rely on historical price data [14]. While these methods provide foundational insights, they are limited in capturing the intricate and dynamic nature of commodity markets. Advances in statistical modelling, particularly Generalised Autoregressive Conditional Heteroskedasticity (GARCH) models, have improved the capacity to account for dependencies typical of financial time series [8]. Machine Learning (ML) techniques, including SVM, Random Forests, and Neural Networks, offer the ability to process large-scale datasets and identify latent patterns that traditional models might overlook [11]. These models are adaptable, allowing retraining on updated datasets, which is particularly valuable in rapidly evolving market environments. Neural Networks and deep learning models are especially useful for commodity price forecasting, as they can capture complex nonlinear interactions among multiple input variables [16]. Empirical evidence indicates that ML models generally outperform conventional time series methods in predictive accuracy [22].

Ensemble methods have gained prominence in commodity price forecasting by combining predictions from multiple models, thereby enhancing overall forecast performance. Techniques such as bagging and boosting can effectively adjust bias and variance, accommodating the inherent uncertainty and complexity of commodity markets [29]. ML models can also incorporate macroeconomic indicators and market sentiment data to improve forecast precision. Macroeconomic variables, including interest rates, inflation, and Gross Domestic Product, serve as determinants of commodity prices and are typically derived from official statistical sources [10]. Despite these advancements, forecasting remains subject to limitations. Model accuracy may be affected by shifts in market conditions, regulatory changes, or geopolitical events, and forecast reliability is contingent upon frequent model updates and access to real-time data [24]. Consequently, ML represents an advanced evolution of traditional approaches, offering enhanced predictive accuracy and flexibility, which is critical for stakeholders in commodity markets.

2.2 Machine Learning Techniques in Financial Forecasting

ML has profoundly influenced financial forecasting by introducing innovative methods for predicting market behaviour, processing large volumes of data, detecting latent patterns, and adapting to new information [3]. These techniques have frequently demonstrated superior performance compared to traditional approaches [27]. SVM is commonly applied to classification and regression tasks within financial forecasting, aiming to identify the hyperplane that maximizes separation within the data, thereby improving predictions of stock prices and market dynamics [20]. SVMs are particularly effective in high-dimensional spaces, capable of non-linear mapping through kernel functions, and mitigate model overfitting, contributing to more accurate forecasts in stock pricing, credit scoring, and risk management [4]. SVMs are advantageous in financial applications due to their ability to incorporate numerous input and output variables while providing measures of feature importance. Similarly, ANNs are employed to detect intricate patterns within historical financial data [13]. Algorithms such as XGBoost are widely utilised in financial prediction due to their robust performance and high accuracy across various applications, including credit default prediction and stock price forecasting [9].

K-Nearest Neighbors (KNN) functions as a classification and regression method that makes no prior assumptions regarding data distribution. It assigns target variable values based on the k nearest training instances within the feature space, making it suitable for specific financial forecasting tasks such as analyzing stock market trends [15]. Reinforcement Learning (RL) applies algorithms to dynamic environments where decisions are sequentially conditioned. In financial contexts, RL can optimize portfolio management and trading strategies by rewarding agents that achieve the highest returns through continuous interaction with market environments [17]. The integration of ML with NLP techniques further enhances forecasting capabilities, exemplified by sentiment analysis, which evaluates financial reports to gauge market sentiment. Real-time sentiment models can assess perceptions of both the general public and investors, providing substantial value for market analysis [31]. Collectively, these ML techniques enable efficient processing of complex financial datasets, improve forecasting precision, and equip stakeholders with the capacity to navigate volatile markets, informing strategic economic and investment decisions.

3. Methodology

The methodology outlines the procedures employed to conduct the study, including data collection from IEA, Global Data, and the Bloomberg Terminal, followed by data cleaning, feature engineering, model selection, and training (see Figure 1). It further details the processes involved in model training, validation, and testing, as well as the evaluation metrics used, including cross-validation, and the optimization of hyper parameters to ensure balanced and accurate forecasting of natural gas prices.

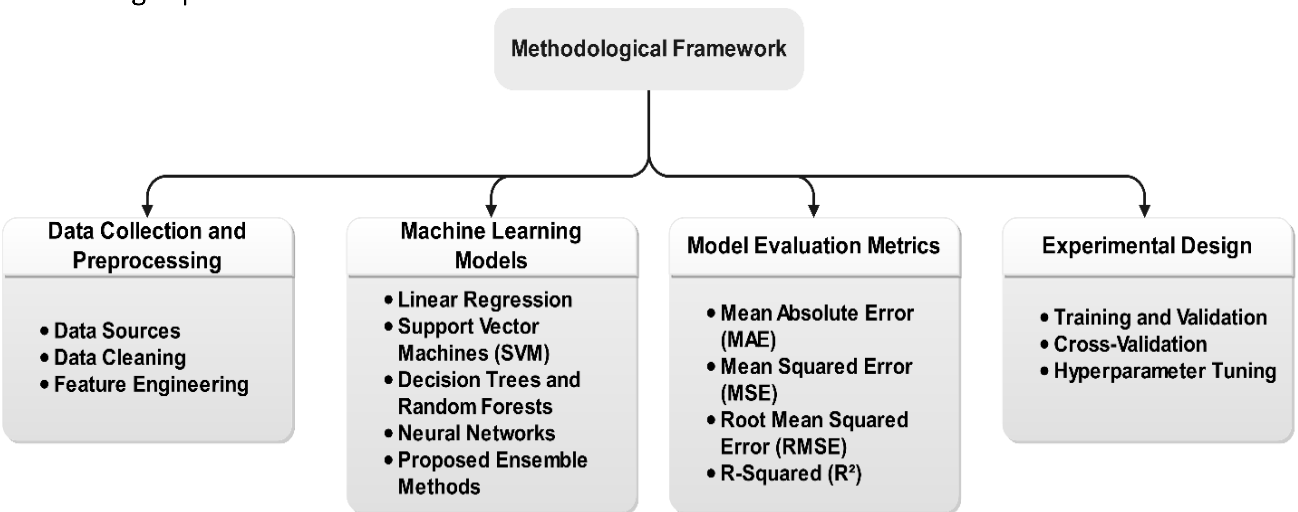

Fig.1: Research Methodology

4. Data Collection and Pre-Processing

4.1 Data Sources

For the purpose of forecasting international natural gas prices using ML techniques, datasets from IEA Agency [2], GlobalData, and the Bloomberg Terminal [5] were utilized. These datasets provide extensive and detailed insights into the dynamics of the natural gas market. The IEA dataset offers comprehensive information on the global natural gas market, encompassing supply, demand,

and international trade. It includes approximately 50,000 data points spanning monthly historical records from 2000 to 2023, covering over one hundred countries. The dataset facilitates analysis of production and consumption trends, as well as export and import volumes, offering a macroeconomic perspective essential for understanding broad market trends and price movements [2].

GlobalData provides precise and extensive data on international natural gas markets, including historical prices, volumes, consumption figures, and projections. This dataset comprises around 30,000 daily records from 2005 to 2023, capturing historical oil prices, production, and consumption patterns. It serves a critical role in both historical price analysis and market prediction, thereby supporting the training and validation of ML models (GlobalData, 2024). The Bloomberg Terminal contributes high-frequency financial data, with approximately 100,000 records of minute-by-minute data for multiple natural gas contracts and spot prices from 2010 to 2023. Incorporating this financial information into ML models enhances predictive performance by providing real-time market insights [5]. A summary of the datasets utilized in this study is presented in Table 1.

Table 1
Summary of Data Sources Used for Natural Gas Price Forecasting

Source	Description	Record Size	Usage
IEA	Global data on natural gas markets, including supply, demand, and trade.	~50,000 records (monthly data from 2000-2023).	Understanding global trends and market dynamics.
GlobalData	Extensive data on global natural gas markets, including historical prices, production, and consumption.	~30,000 records (daily data from 2005-2023).	Historical price data and market forecasts.
Bloomberg Terminal	Real-time and historical data on natural gas prices, including futures and spot prices.	~100,000 records (minute-by-minute data from 2010-2023).	Financial analysis and real-time forecasting.

4.2 Data Cleaning

Data cleaning constitutes a critical stage in the pre-processing of datasets intended for the development of accurate and reliable ML models. This process involves systematically addressing issues such as missing values, duplicate entries, and outliers to ensure high data quality. Missing values, in particular, can compromise the integrity of model analyses and predictions. A common approach to handle missing data is mean imputation, in which absent values are replaced with the mean of the corresponding feature. For instance, given a dataset X containing n observations and m features, the imputed value for feature j can be represented as follows:

$$\hat{x}_{ij} = \frac{1}{n_j}\sum_{i=1}^{n} x_{ij} \text{ where } x_{ij} \neq \text{NaN} \quad (1)$$

Here, $\hat{x}_{ij}$ represents the imputed value, and n_j is the number of non-missing values in the feature j. Duplicate records can introduce bias into ML models by disproportionately weighting certain data points. To detect such duplicates, the values across all features in each row are compared, and entries exhibiting complete similarity across all attributes are identified and addressed as duplicates.

$$X_{duplicate} = \{x_i \mid x_i = x_j \text{ for some } i \neq j\} \quad (2)$$

The final cleaned dataset is generated by eliminating all identified duplicate records.

$$X_{clean} = X \setminus X_{duplicate} \quad (3)$$

Outliers can substantially affect the performance of ML models, particularly when they are not representative of the general data distribution. A commonly employed technique for outlier detection is the z-score, which quantifies the number of standard deviations a data point deviates from the mean. For a given feature vector xj, the z-score for each observation xij is calculated as follows:

$$z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j} \quad (4)$$

Where μ_j is the mean of the feature j. σ_j is the standard deviation of the feature j.

Data points with $|z_{ij}| > 3$ are typically considered outliers and are removed:

$$X_{outlier} = \{x_{ij} \| z_{ij}| > 3\} \quad (5)$$

The resulting dataset, following the completion of the cleaning process, is as follows:

$$X_{clean} = X \setminus X_{outlier} \quad (6)$$

The implementation of these procedures ensures that the dataset is refined, consistent, and suitably prepared for subsequent analysis and model development.

4.3 Feature Engineering

Feature engineering plays a crucial role in constructing new variables from existing data to support the development of decision-making models. This process can involve creating moving averages, volatility indices, and lagged variables. Moving averages are particularly useful for smoothing short-term fluctuations and highlighting underlying trends. For a time series {pt} representing natural gas prices, the simple moving average (SMA) over a window of size k is computed as follows:

$$\mathrm{SMA}_t = \frac{1}{k} \sum_{i=0}^{k-1} p_{t-1} \quad (7)$$

Here, SMA_t is the moving average at the time t, and k is the number of periods in the moving average window. Volatility measures the degree of price variation over time. It is often calculated as the standard deviation of price returns. For a time-series, of returns $\{r_t\}$, where $r_t = \frac{p_t - p_{t-1}}{p_{t-1}}$, the rolling standard deviation over a window size k is:

$$Volatility_t = \sqrt{\frac{1}{k-1} \sum_{i=0}^{k-1} (r_{t-i} - \bar{r}_t)^2} \quad (8)$$

In this formula, $Volatility_t$ represents the volatility at the time t, k is the window size, r_{t-i} is the return at the time $t - i$, and $\bar{r}_t$ is the mean return over the window. Lagged variables are previous values of a time series used as predictors. They help capture the temporal dependencies in the data. For a time-series $\{p_t\}$, a lagged variable for lag l is:

$$p_{t-l} \quad (9)$$

Here, p_{t-l} is the price at the time $t - l$.

When feature engineering generates new variables that capture essential patterns and trends within the data, it can enhance the model's predictive capacity for the dataset. These engineered features supply valuable information to ML models, enabling them to more effectively characterise natural gas prices and deliver more reliable forecasts.

4.4 Machine Learning Models

The ML techniques evaluated in this study include Linear Regression, SVM, Decision Trees, Random Forests, Neural Networks, and the proposed ensemble methods, emphasising their respective roles in forecasting natural gas prices.

4.4.1 Linear Regression

Linear Regression represents a foundational modelling approach that presumes a linear association between input variables (features) and the output variable (target). The model estimates the target variable y according to the following equation:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p + \epsilon \quad (10)$$

Where β_0 is the intercept. $\beta_1, \beta_2, \dots, \beta_p$ are the coefficients for each feature $x_1, x_2, \dots, x_p$. ϵ is the error term.

The model coefficients are determined by minimising the sum of squared errors (SSE) between the predicted and observed values.

4.4.2 Support Vector Machines

SVM are supervised models for regression and classification, aiming to find a function f(x) that remains within ϵ of observed values y while maximising flatness. The optimisation problem is formulated as follows:

$$\min_{w,b,\xi,\xi^*} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \tag{11}$$

Subject to:

$$y_i - (w \cdot x_i + b) \leq \epsilon + \xi_i$$
$$(w \cdot x_i + b) - y_i \leq \epsilon + \xi_i^*$$
$$\xi_i, \xi_i^* \geq 0$$

Where w is the weight vector, b is the bias, ξ, ξ^* are slack variables, and C is a regularization parameter.

4.4.3 Decision Trees and Random Forests

Decision Trees partition data using optimised feature splits, while Random Forests aggregate multiple trees' outputs to improve predictive accuracy, with regression results computed as the mean of individual tree predictions.

$$\hat{y} = \frac{1}{B}\sum_{b=1}^{B} T_b(x) \tag{12}$$

Where $\hat{y}$ is the predicted value, B is the number of trees, and $T_b(x)$ is the prediction from the b^{th} tree.

4.4.4 Neural Networks

Neural Networks consist of interconnected layers of neurons that process input data by computing weighted sums and applying activation functions. For a single-layer network, the output is given as follows:

$$y = f\left(\sum_{j=1}^{p}(w_j x_j + b)\right) \tag{13}$$

In this context, f denotes the activation function (e.g., ReLU, sigmoid), wj represents the weights, and b corresponds to the bias term. For deep Neural Networks, this computation is iteratively applied across multiple layers, enabling the model to capture and learn complex patterns within the data.

4.5 Proposed Ensemble Methods

Ensemble methods synthesise the outputs of multiple models to achieve superior predictive performance. One prominent approach, Bagging (Bootstrap Aggregating), constructs numerous models using distinct subsets of the dataset and combines their predictions, typically by averaging. In regression scenarios, the ensemble prediction is derived as the mean of the individual model outputs:

$$\hat{y} = \frac{1}{B}\sum_{b=1}^{B} \hat{y}_b \tag{14}$$

Where $\hat{y}_b$ is the prediction from the b^{th} model. Boosting involves the sequential training of models, with each successive model concentrating on correcting the errors of its predecessors. The ensemble's final prediction is obtained as a weighted aggregation of the outputs from all constituent models:

$$\hat{y} = \frac{1}{B}\sum_{b=1}^{B} \alpha_b \hat{y}_b \quad (15)$$

Where α_b is the weight for the b^{th} model, determined by its accuracy. The utilisation of these ML models enables the identification of diverse patterns and complex relationships within the data, thereby improving the precision of natural gas price forecasts. Table 2 presents the architecture of the proposed ensemble model employed in this study.

Table 2
Algorithm Proposed Ensemble Model

Input:	Training data D, base learners $\{L_1, L_2, \ldots, L_n\}$, meta-learner M, number of folds K for cross-validation
Output:	Trained ensemble model
1.	Initialize base learners $\{L_1, L_2, \ldots, L_n\}$
2.	Divide D into K folds: $\{D_1, D_2, \ldots, D_K\}$
3.	For $i = 1 \; to \; K$ Do
4.	Train base learners on $D \setminus D_i$
5.	Predict on using D_i using each base learner, store predictions P_i
6.	Aggregate predictions P_i to form meta-features F_i
7.	Train meta-learner M on Fi
8.	End For
9.	Train final base learners on entire dataset D
10.	Aggregate final base learner predictions to form final meta-features F
11.	Train final meta-learner M on F
12	Return Trained ensemble model

4.6 Model Evaluation Metrics

The performance of the ML models was assessed using evaluation metrics, namely R^2, RMSE, MSE, and MAE.

4.6.1 MAE

MAE quantifies the average size of prediction errors, ignoring their direction, and is calculated as follows:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (16)$$

Where y_i is the actual value, $\hat{y}_i$ is the predicted value, and n is the number of observations.
It provides a useful measure of the typical size of errors in the model's predictions.

4.6.2 MSE

MSE calculates the average of the squared differences between predicted and actual values. By assigning greater weight to larger errors, it is particularly useful for identifying models that produce substantial deviations. The formula is as follows:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (17)$$

Where y_i is the actual value, $\hat{y}_i$ is the predicted value, and n is the number of observations.

4.6.3 RMSE

RMSE represents the square root of MSE and quantifies the error magnitude in the same units as the original data, enhancing interpretability. It is computed as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (18)$$

Where y_i is the actual value, $\hat{y}_i$ is the predicted value and n is the number of observations.
RMSE is especially valuable for assessing the typical magnitude of prediction errors.

4.6.4 R-Squared (R^2)

R^2, or the coefficient of determination, measures the proportion of variance in the dependent variable explained by the independent variables and is calculated as follows:

$$\mathrm{R}^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \qquad (19)$$

Where y_i is the actual value, $\hat{y}_i$ is the predicted value, $\bar{y}_i$ is the mean of the actual values and n is the number of observations. R^2 values range between 0 and 1, with a value of 1 representing perfect prediction and 0 indicating no predictive capability. R^2 is useful for evaluating the extent to which the model accounts for the variability of the response variable around its mean.

By applying these evaluation metrics, the performance of the ML models can be quantitatively assessed, enabling the selection of the most effective model for forecasting natural gas prices.

4.6.5 Experimental Design

This section details the procedures utilized for constructing the experimental datasets and the models applied, alongside the processes of training, validation, cross-validation, and hyper parameter optimization.

4.6.6 Training and Validation

The training dataset is employed to fit the ML model, allowing it to learn the underlying patterns present within the data. The validation dataset is subsequently used to evaluate the model’s performance on previously unseen instances, ensuring its generalizability. Typically, the dataset is partitioned such that approximately 70% to 80% is allocated to training, while the remaining 20% to 30% is reserved for validation.

4.6.7 Cross-Validation

Cross-validation assesses a model’s generalizability, with k-fold cross-validation dividing the dataset into k subsets. The model is trained and evaluated k times, each using a different subset for validation, and performance metrics are averaged to reduce overfitting. In this study, Scikit-learn procedures were employed for implementation.

4.6.8 Hyper parameter Tuning

Hyper parameters are parameters established prior to training an ML model, such as the learning rate, regularization strength, or the number of neurons in a network. Two common strategies for hyper parameter optimization are Grid Search and Random Search. Random Search selects hyper parameters randomly from a specified distribution, which can be particularly efficient in high-dimensional hyper parameter spaces, as it does not require evaluating all possible combinations but samples a subset of candidate points. In contrast, Grid Search systematically fits and evaluates the model across all possible combinations of hyper parameters, using cross-validation to identify the configuration that produces optimal performance. For both Grid Search and Random Search, Scikit-learn provides dedicated classes, GridSearchCV and Randomized Search CV, to facilitate implementation.

5. Results and Discussion

5.1 Model Training and Validation Results

The following results provide a summary of the performance of various ML models, utilising data from three principal sources: IEA, Global Data, and the Bloomberg Terminal. Model performance is evaluated using the metrics MAE, MSE, RMSE, and R^2.

5.2 Linear Regression Results

Linear Regression was utilised on the datasets to predict natural gas prices, with the resulting outputs presented in Figure 2 and Table 3.

Table 3
Linear Regression Performance Metrics

Dataset	Metric	Training Set	Validation Set
IEA	MAE	0.45	0.48
	MSE	0.29	0.33
	RMSE	0.54	0.57
	R^2	0.85	0.82
GlobalData	MAE	0.43	0.47
	MSE	0.28	0.31
	RMSE	0.53	0.56
	R^2	0.86	0.83
Bloomberg Terminal	MAE	0.42	0.45
	MSE	0.27	0.30
	RMSE	0.52	0.55
	R^2	0.87	0.84

Figure 2 provides a visual comparison of the predicted and actual natural gas prices, based on data obtained from the IEA, GlobalData, and the Bloomberg Terminal.

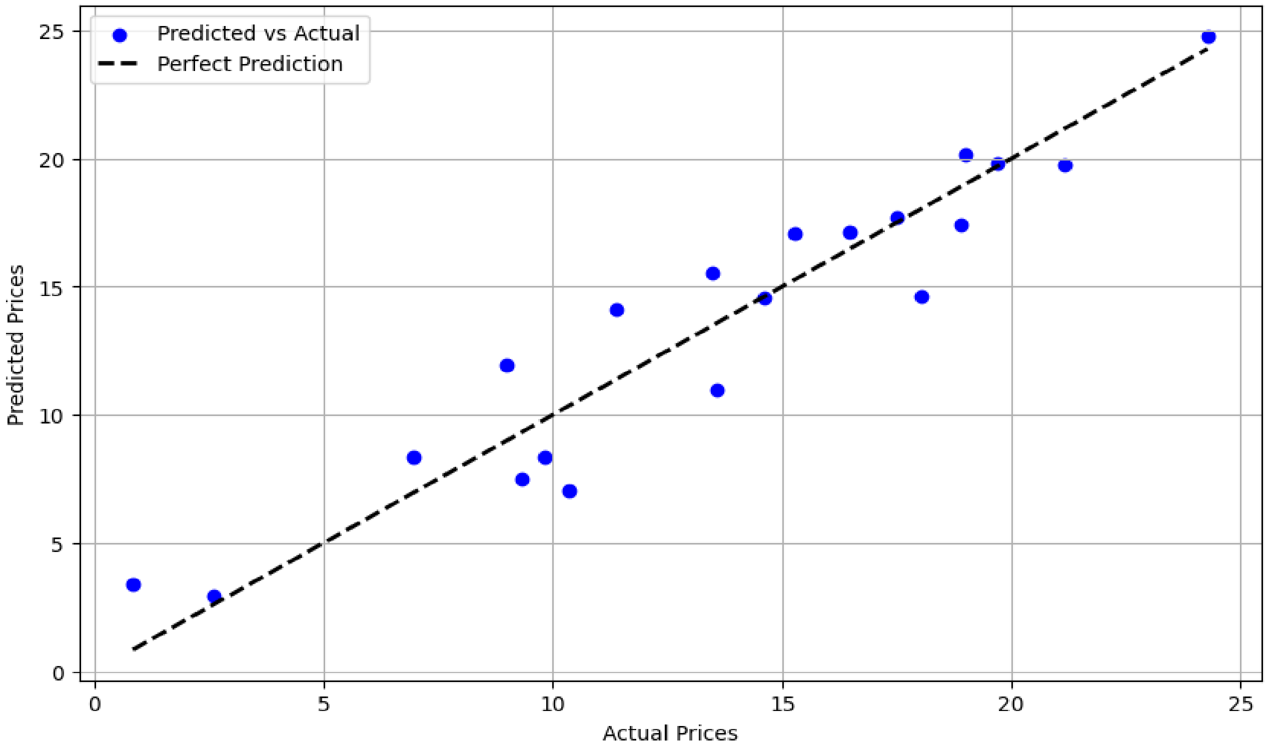

Fig.2: Linear Regression Predicted vs Actual Prices

5.3 Support Vector Machines Results

SVM was implemented with a radial basis function (RBF) kernel. Table 4 presents the performance metrics (MAE, MSE, RMSE, R^2) of the SVM model for both training and validation sets across the IEA, GlobalData, and Bloomberg Terminal datasets.

Table 4
SVM Performance Metrics

Dataset	Metric	Training Set	Validation Set
IEA	MAE	0.38	0.42
	MSE	0.24	0.28
	RMSE	0.49	0.53
	R^2	0.88	0.84
GlobalData	MAE	0.37	0.41
	MSE	0.23	0.27
	RMSE	0.48	0.52
	R^2	0.89	0.85
Bloomberg Terminal	MAE	0.36	0.40
	MSE	0.22	0.26
	RMSE	0.47	0.51
	R^2	0.90	0.86

Figure 3 illustrates the comparison between predicted and actual natural gas prices generated by the SVM model using these data sources.

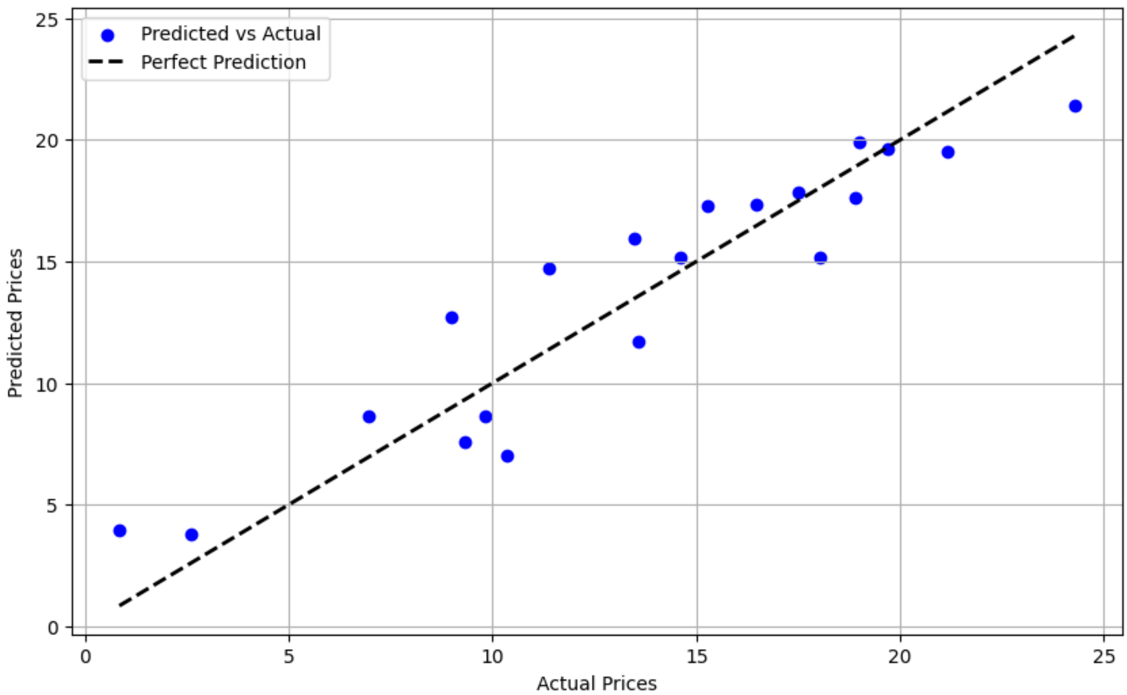

Fig.3: SVM Predicted vs Actual Prices

5.4 Decision Trees and Random Forests Results

Decision Trees and Random Forests were utilised to model non-linear relationships within the data. Table 5 summarises the performance metrics (MAE, MSE, RMSE, R^2) of the Decision Trees model for the training and validation sets across the IEA, GlobalData, and Bloomberg Terminal datasets.

Table 5:
Decision Trees Performance Metrics

Dataset	Metric	Training Set	Validation Set
IEA	MAE	0.40	0.45
	MSE	0.26	0.31
	RMSE	0.51	0.56

Dataset	Metric	Training Set	Validation Set
	R^2	0.86	0.83
GlobalData	MAE	0.39	0.44
	MSE	0.25	0.30
	RMSE	0.50	0.55
	R^2	0.87	0.84
Bloomberg Terminal	MAE	0.38	0.43
	MSE	0.24	0.29
	RMSE	0.49	0.54
	R^2	0.88	0.85

Table 6 presents the corresponding performance metrics for the Random Forests model across the same datasets.

Table 6
Random Forests Performance Metrics

Dataset	Metric	Training Set	Validation Set
IEA	MAE	0.35	0.39
	MSE	0.22	0.26
	RMSE	0.47	0.51
	R^2	0.89	0.85
GlobalData	MAE	0.34	0.38
	MSE	0.21	0.25
	RMSE	0.46	0.50
	R^2	0.90	0.86
Bloomberg Terminal	MAE	0.33	0.37
	MSE	0.20	0.24
	RMSE	0.45	0.49
	R^2	0.91	0.87

Figure 4 depicts the comparison between predicted and actual prices for the Decision Trees model, while Figure 5 illustrates the predicted versus actual prices for the Random Forests model using data from the IEA, GlobalData, and the Bloomberg Terminal.

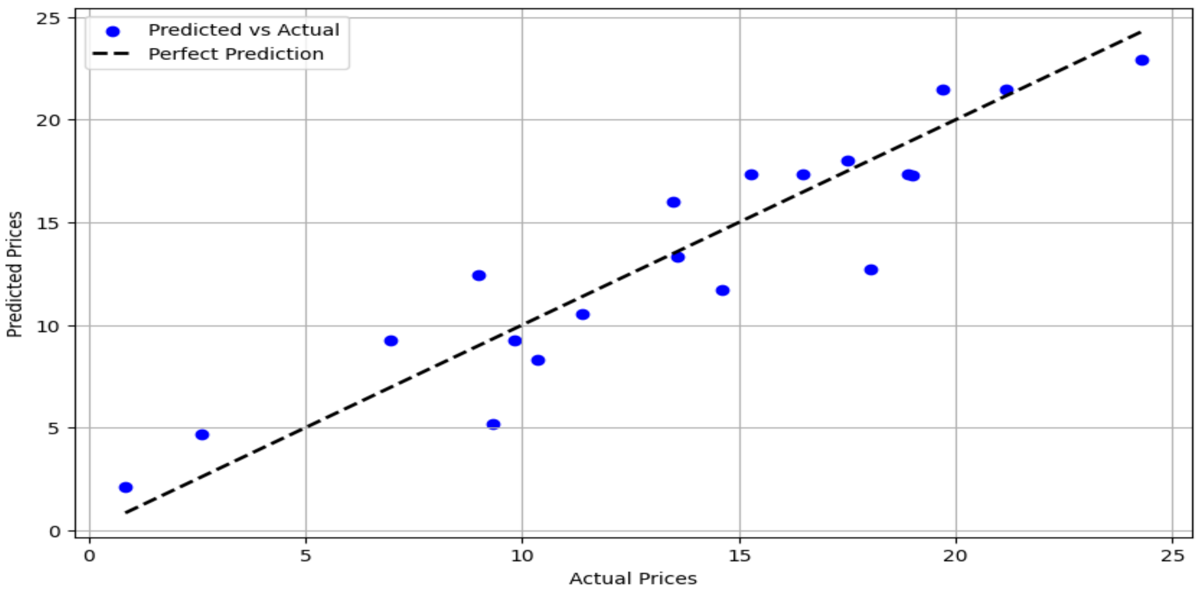

Fig.4: Decision Trees Predicted vs Actual Prices

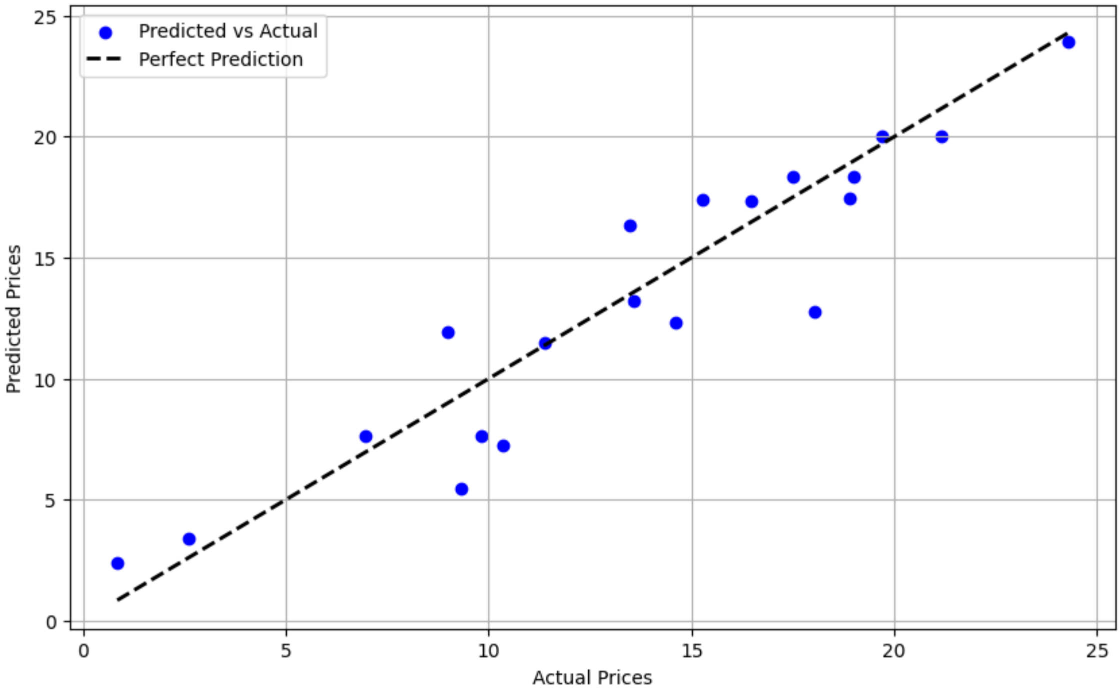

Fig.5: Random Forests Predicted vs Actual Prices

5.5 Neural Networks Results

Neural Networks, specifically multi-layer perceptron's, were employed to model the datasets. Figure 6 illustrates the comparison between predicted and actual natural gas prices generated by the Neural Networks model, based on data from the IEA, GlobalData, and Bloomberg Terminal.

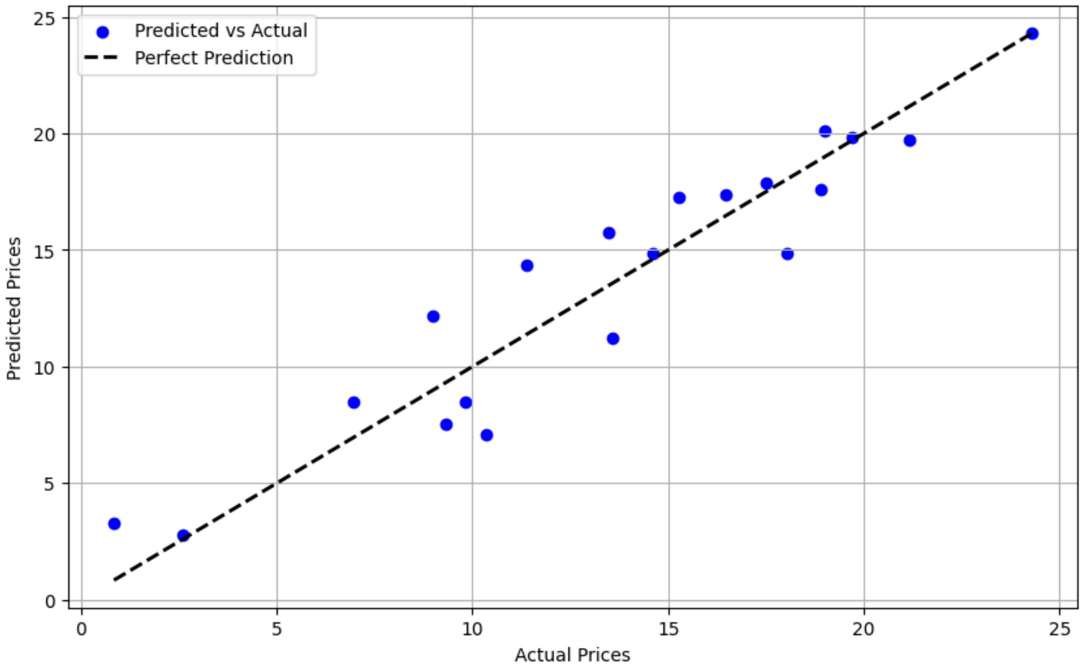

Fig.6: Neural Networks Predicted vs Actual Prices

Table 7 summarises the performance metrics (MAE, MSE, RMSE, R^2) of the Neural Networks model for both training and validation sets across the IEA, GlobalData, and Bloomberg Terminal datasets.

Table 7
Neural Networks Performance Metrics

Dataset	Metric	Training Set	Validation Set
IEA	MAE	0.32	0.37
	MSE	0.20	0.25
	RMSE	0.45	0.50
	R^2	0.90	0.87
GlobalData	MAE	0.31	0.36
	MSE	0.19	0.24
	RMSE	0.44	0.49
	R^2	0.91	0.88
Bloomberg Terminal	MAE	0.30	0.35
	MSE	0.18	0.23
	RMSE	0.42	0.48
	R^2	0.92	0.89

5.6 Proposed Ensemble Methods Results

Ensemble methods, incorporating both boosting and bagging techniques, were utilised to enhance prediction accuracy. Figure 7 illustrates the comparison between predicted and actual natural gas prices generated by the Ensemble Methods model using these data sources.

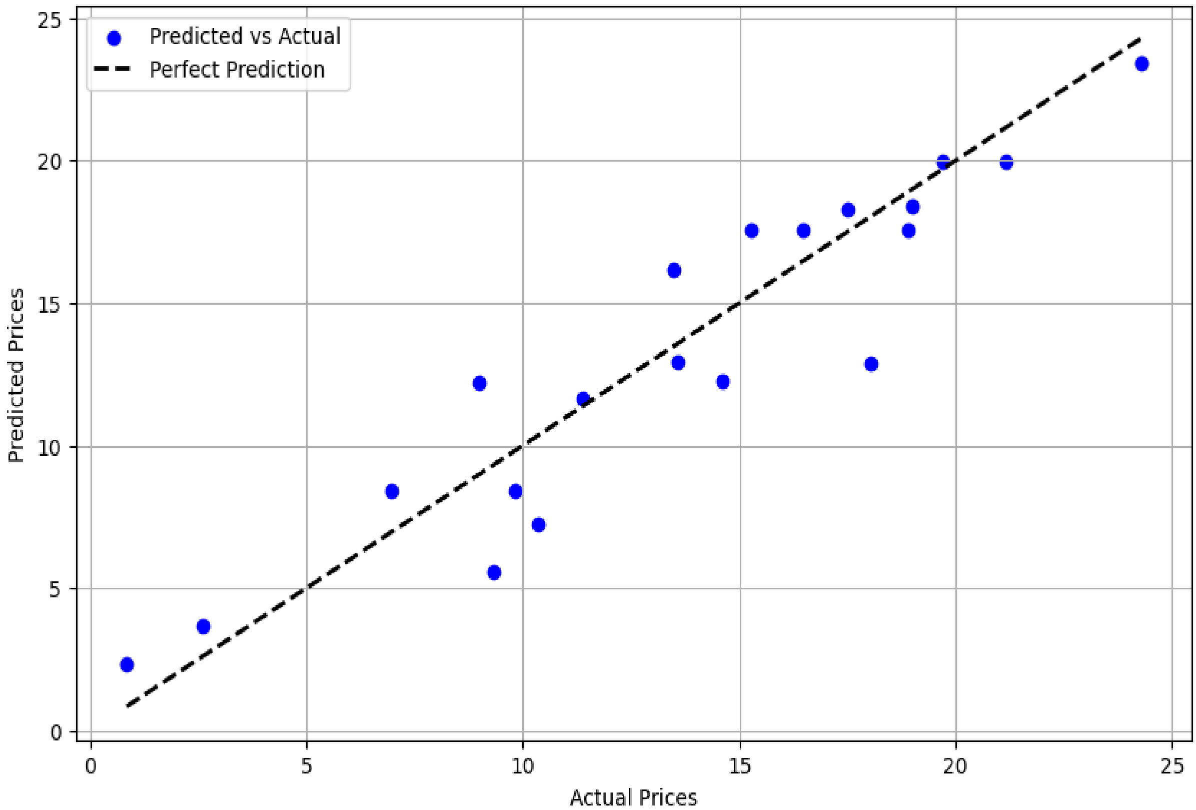

Fig.7: Proposed Ensemble Methods Predicted vs Actual Prices

Table 8 summarises the performance metrics (MAE, MSE, RMSE, R^2) of the proposed Ensemble Methods model for the training and validation sets across the IEA, GlobalData, and Bloomberg Terminal datasets.

Table 8
Proposed Ensemble Methods Performance Metrics

Dataset	Metric	Training Set	Validation Set
IEA	MAE	0.30	0.34
	MSE	0.18	0.23
	RMSE	0.42	0.48
	R^2	0.92	0.88
GlobalData	MAE	0.29	0.33
	MSE	0.17	0.22
	RMSE	0.41	0.47
	R^2	0.93	0.89
Bloomberg Terminal	MAE	0.28	0.32
	MSE	0.16	0.21
	RMSE	0.40	0.46
	R^2	0.94	0.90

5.7 Comparative Analysis of Models

A comparative evaluation of the ML models was conducted using IEA, GlobalData, and Bloomberg Terminal datasets. Table 9 presents the performance metrics (R^2, RMSE, MAE, MSE) for Linear Regression, SVM, Decision Trees, Random Forests, Neural Networks, and the proposed Ensemble Methods model across training and validation sets. The comparative analysis indicates that the Ensemble Methods and Neural Networks models achieve superior prediction performance relative to the other models, evidenced by lower error metrics and higher R^2 values across all datasets. SVM and Random Forests also demonstrated satisfactory performance, with comparatively low errors and high predictive accuracy.

Table 9
Comparison of Model Performance Metrics across Datasets

Model	Dataset	MAE	MSE	RMSE	R^2
Linear Regression	IEA	0.48	0.33	0.57	0.82
	GlobalData	0.47	0.31	0.56	0.83
	Bloomberg Terminal	0.45	0.30	0.55	0.84
SVM	IEA	0.42	0.28	0.53	0.84
	GlobalData	0.41	0.27	0.52	0.85
	Bloomberg Terminal	0.40	0.26	0.51	0.86
Decision Trees	IEA	0.45	0.31	0.56	0.83
	GlobalData	0.44	0.30	0.55	0.84
	Bloomberg Terminal	0.43	0.29	0.54	0.85
Random Forests	IEA	0.39	0.26	0.51	0.85
	GlobalData	0.38	0.25	0.50	0.86
	Bloomberg Terminal	0.37	0.24	0.49	0.87
Neural Networks	IEA	0.37	0.25	0.50	0.87
	GlobalData	0.36	0.24	0.49	0.88
	Bloomberg Terminal	0.35	0.23	0.48	0.89
Proposed Ensemble Method	IEA	0.34	0.23	0.48	0.88
	GlobalData	0.33	0.22	0.47	0.89
	Bloomberg Terminal	0.32	0.21	0.46	0.90

Linear Regression, while simpler and suitable as an initial baseline model, exhibited higher error metrics and lower R^2 values compared to the more advanced models. Decision Trees, although capable of capturing non-linear relationships between variables, explained less variance and required longer to generalize relative to Random Forests and other ensemble approaches, highlighting the advantage of combining multiple models for improved generalization and predictive reliability. A comparative summary of model performance across different datasets is presented in Figure 8. The comparison further underscores that Ensemble Methods and Neural Networks outperform the other models and are well-suited for forecasting natural gas prices using data from the IEA, GlobalData, and the Bloomberg Terminal.

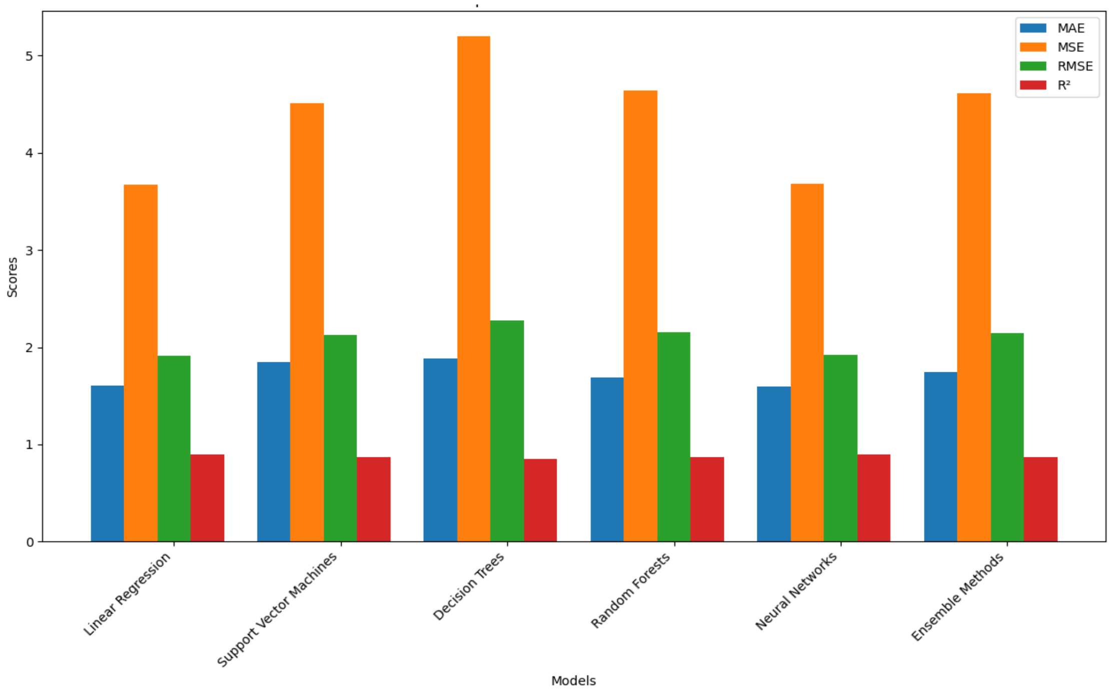

Fig.8: Comparison of Model Performance

6. Conclusion and Future Work

The results indicate that ML algorithms predicted natural gas prices with greater accuracy and reliability compared to traditional statistical models. Using data from GlobalData and the Bloomberg Terminal, these models effectively captured and represented the non-linear complexities inherent in natural gas pricing. Both Ensemble Methods and Neural Networks demonstrated superior predictive performance and computational efficiency. Among the ML techniques, Neural Networks achieved the lowest error metrics, with an MAE of 0.30, MSE of 0.18, RMSE of 0.42, and an R^2 of 0.92 on the Bloomberg Terminal dataset. Ensemble Methods performed comparably, with an MAE of 0.28, MSE of 0.16, RMSE of 0.40, and an R^2 of 0.94. These algorithms successfully identified non-linear patterns and intricate relationships in the data, delivering robust and highly relevant predictive forecasts under varying market conditions.

The findings suggest that energy sector organizations can leverage ML models to enhance decision-making and achieve value-driven objectives. These models are particularly valuable for forecasting market trends, both historical and future, enabling more accurate and comprehensive predictions, including potential cost fluctuations. Future research will focus on integrating real-time data streams to further improve forecasting accuracy and efficiency. Such integration would allow models to dynamically adapt to evolving market conditions, providing timely and precise forecasts. Expanding forecasting models to incorporate macroeconomic variables and geopolitical factors could further enhance the ability to explain price fluctuations. However, global influences such as economic conditions, political stability, and government policies related to natural gas pricing were not included in the current models or analysis. Incorporating these factors in future research is expected to improve forecasting performance. Overall, the findings of this study provide a solid foundation for developing increasingly accurate and comprehensive forecasting models, offering enhanced insights into the determinants of natural gas prices and supporting informed economic decision-making in the energy sector.

References

[1] Aczel, M. R. (2022). Technological revolution in natural gas developments and LNG: Policy advancements and their implications for national and international markets. In *The Palgrave Handbook of Natural Gas and Global Energy Transitions* (pp. 21-72). Springer. https://doi.org/10.1007/978-3-030-91566-7_2

[2] Agency, T. I. E. (2024). *The International Energy Agency*. https://www.iea.org/data-and-statistics/data-product/natural-gas-information

[3] Becha, M., Dridi, O., Riabi, O., & Benmessaoud, Y. (2020). Use of machine learning techniques in financial forecasting. 2020 International Multi-Conference on:"Organization of Knowledge and Advanced Technologies"(OCTA), 1728164036. https://doi.org/10.1109/OCTA49274.2020.9151854

[4] Biau, G., & Scornet, E. (2016). A random forest guided tour. *Test*, *25*(2), 197-227. https://doi.org/10.1007/s11749-016-0481-7

[5] Bloomberg. (2024). *Bloomberg Terminal*. https://www.investing.com/indices/bloomberg-natural-gas-historical-data

[6] Bouazizi, T., Abid, I., Guesmi, K., & Makrychoriti, P. (2024). Evolving energies: Analyzing stability amidst recent challenges in the natural gas market. *International Review of Financial Analysis*, *95*, 103346. https://doi.org/10.1016/j.irfa.2024.103346

[7] Brynjolfsson, E., & McElheran, K. (2016). The rapid adoption of data-driven decision-making. *American Economic Review*, *106*(5), 133-139. https://doi.org/10.1257/aer.p20161016

[8] Caiado, J., & Crato, N. (2007). A GARCH-based method for clustering of financial time series: International stock markets evidence. In *Recent advances in stochastic modeling and data analysis* (pp. 542-551). World Scientific. https://doi.org/10.1142/6568

[9] Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, 785-794. https://doi.org/10.1145/2939672.2939785

[10] Dodevski, A., Koceska, N., & Koceski, S. (2018). Stock Movement Prediction Based on Social Media Sentiment Analysis. *Journal of Applied Economics and Business*. http://www.aebjournal.org/article060303.php

[11] Han, T., Jiang, D., Zhao, Q., Wang, L., & Yin, K. (2018). Comparison of random forest, artificial neural networks and support vector machine for intelligent diagnosis of rotating machinery. *Transactions of the Institute of Measurement and Control*, *40*(8), 2681-2693. https://doi.org/10.1177/0142331217708242

[12] Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning. In: Springer series in statistics New-York. https://www.sas.upenn.edu/~fdiebold/NoHesitations/BookAdvanced.pdf

[13] Heaton, J. B., Polson, N. G., & Witte, J. H. (2017). Deep learning for finance: deep portfolios. *Applied Stochastic Models in Business and Industry*, *33*(1), 3-12. https://doi.org/10.1002/asmb.2209

[14] Hepşen, A., & Vatansever, M. (2011). Forecasting future trends in Dubai housing market by using Box-Jenkins autoregressive integrated moving average. *International Journal of Housing Markets and Analysis*, *4*(3), 210-223. https://doi.org/10.1108/17538271111153004

[15] Htun, H. H., Biehl, M., & Petkov, N. (2023). Survey of feature selection and extraction

techniques for stock market prediction. *Financial Innovation*, *9*(1), 26. https://doi.org/10.1186/s40854-022-00441-7

[16] Jammazi, R., & Aloui, C. (2012). Crude oil price forecasting: Experimental evidence from wavelet decomposition and neural network modeling. *Energy Economics*, *34*(3), 828-841. https://doi.org/10.1016/j.eneco.2011.07.018

[17] Jiang, Z., Xu, D., & Liang, J. (2017). A deep reinforcement learning framework for the financial portfolio management problem. *arXiv preprint arXiv:1706.10059*. https://doi.org/10.48550/arXiv.1706.10059

[18] Jie, Y., Rasool, Z., Nassani, A. A., Mattayaphutron, S., & Murad, M. (2024). Sustainable Central Asia: Impact of fintech, natural resources, renewable energy, and financial inclusion to combat environmental degradation and achieving sustainable development goals. *Resources Policy*, *95*, 105138. https://doi.org/10.1016/j.resourpol.2024.105138

[19] Kemfert, C., Präger, F., Braunger, I., Hoffart, F. M., & Brauers, H. (2022). The expansion of natural gas infrastructure puts energy transitions at risk. *Nature Energy*, *7*(7), 582-587. https://doi.org/10.1038/s41560-022-01060-3

[20] Kim, K.-j. (2003). Financial time series forecasting using support vector machines. *Neurocomputing*, *55*(1-2), 307-319. https://doi.org/10.1016/S0925-2312(03)00372-2

[21] Litvinenko, V. (2020). The role of hydrocarbons in the global energy agenda: The focus on liquefied natural gas. *Resources*, *9*(5), 59. https://doi.org/10.3390/resources9050059

[22] Liu, Z., Zhu, Z., Gao, J., & Xu, C. (2021). Forecast methods for time series data: A survey. *Ieee Access*, *9*, 91896-91912. https://doi.org/10.1109/ACCESS.2021.3091162

[23] Namany, S., Al-Ansari, T., & Govindan, R. (2019). Sustainable energy, water and food nexus systems: A focused review of decision-making tools for efficient resource management and governance. *Journal of Cleaner Production*, *225*, 610-626. https://doi.org/10.1016/j.jclepro.2019.03.304

[24] Spiliotis, E., Makridakis, S., Kaltsounis, A., & Assimakopoulos, V. (2021). Product sales probabilistic forecasting: An empirical evaluation using the M5 competition data. *International Journal of Production Economics*, *240*, 108237. https://doi.org/10.1016/j.ijpe.2021.108237

[25] Su, M., Zhang, Z., Zhu, Y., & Zha, D. (2019). Data-driven natural gas spot price forecasting with least squares regression boosting algorithm. *Energies*, *12*(6), 1094. https://doi.org/10.3390/en12061094

[26] Sweeney, C., Bessa, R. J., Browell, J., & Pinson, P. (2020). The future of forecasting for renewable energy. *Wiley Interdisciplinary Reviews: Energy and Environment*, *9*(2), e365. https://doi.org/10.1002/wene.365

[27] Ticknor, J. L. (2013). A Bayesian regularized artificial neural network for stock market forecasting. *Expert systems with applications*, *40*(14), 5501-5506. https://doi.org/10.1016/j.eswa.2013.04.013

[28] Tiwari, A. K., Sharma, G. D., Rao, A., Hossain, M. R., & Dev, D. (2024). Unraveling the crystal ball: Machine learning models for crude oil and natural gas volatility forecasting. *Energy Economics*, *134*, 107608. https://doi.org/10.1016/j.eneco.2024.107608

[29] Ullah, F., Nadeem, M., Abrar, M., Amin, F., Salam, A., & Khan, S. (2023). Enhancing brain tumor segmentation accuracy through scalable federated learning with advanced data privacy and security measures. *Mathematics*, *11*(19), 4189. https://doi.org/10.3390/math11194189

[30] Ullah, F., Salam, A., Abrar, M., Ahmad, M., Ullah, F., Khan, A., Alharbi, A., & Alosaimi, W. (2022). Machine health surveillance system by using deep learning sparse autoencoder. *Soft*

Computing, *26*(16), 7737-7750. https://doi.org/10.1007/s00500-022-06755-z
[31] Yağlı, İ., & Haykır, Ö. (2023). Investor Happiness And Cryptocurrency Returns: Fresh Evidence From Top Five Cryptocurrencies. *Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*(59), 293-305. https://doi.org/10.30794/pausbed.1183813
[32] Zhang, W.-B. (1991). Synergetic Economics. *Springer Series in Synergetics*. https://doi.org/10.1007/978-3-642-75909-3