International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Predictive Analytics Models for Commodity Market: A Literature Review

[1]Baljinder Kaur (Corresponding Author), [2]Brahmaleen K. Sidhu, [3]Gurjit Singh Bhathal

Submitted:06/07/2024 **Revised:**15/08/2024 **Accepted:**25/08/2024

Abstract: This research presents a literature review on the application of Artificial Intelligence (AI) and Machine Learning (ML) techniques in predictive analytics, with a main emphasis on the commodity market. Predictive analytics influences statistical models and Machine Learning approaches to analyze periodic data and predict future outcomes, providing substantial value across various domains such as finance, healthcare, retail, marketing, and commodities. The motive of this review is two-fold: (1) to survey AI-ML based approaches, techniques, and tools employed in predictive analytics, and (2) to inspect predictive analytics models precisely designed for the commodity market. The review integrates current methodologies, strategies, algorithms, and tools, while also surveying the types of commodities analyzed and their interdependencies. The major influence of this research lies in recognizing major gaps in existing studies, predominantly in model generalization, integration of heterogeneous data sources, and evaluation of real-world applicability. These gaps emphasize possibilities for future research and provide direction for developing more robust, accurate, and scalable predictive analytics frameworks.

Keywords: *Artificial Intelligence, Machine Learning, Predictive Analytics, Forecasting*

1. Introduction

Predictive analytics has appeared as a powerful tool in financial markets, particularly in the commodity market, which is characterized by volatile price fluctuations and complex interdependencies among various factors. Accurate prediction of commodity prices is vital for stakeholders, traders, producers, sponsors and decision-makers, to manage risk and make sound judgements. This has led to significant research efforts focused on developing and evaluating predictive models tailored for the commodity market. Based on our current research, this work is the first to do a literature analysis on predictive analytics in the commodity market, despite including gold, oil, silver, and other raw materials, these are influenced by the fact that there are many other literature review publications on the subject. On the other hand, commodity markets, including gold, oil, silver, and other raw materials, are dominated by various variables for example macroeconomic indicators, geopolitical events, supply-demand dynamics, and global trade patterns. The purpose of this Literature review is to provide a holistic understanding of various predictive analytics models applied in commodity markets. This study identifies key trends, methodologies, and performance metrics in commodity price prediction by reviewing existing research. The primary aim of this study is: (i) to examine the existing studies on predictive analytics and prominent methods for its implementation; (ii) to identify the research gaps and provided resultant methodology to bridge the gap.

[1]*baljinderkaur1057@gmail.com*
Department of Computer Science and Engineering, Punjabi University, Patiala, India
[2]*brahmaleen.ce@pbi.ac.in*
Department of Computer Science and Engineering, Punjabi University,Patiala, India
[3]*gurjit.ce@pbi.ac.in*
Department of Computer Science and Engineering, Punjabi University, Patiala, India

2. Review Process

This paper has followed the rule book given by Kitchenham and Charters. The key tasks in the planning phase involved identifying the study subjects to be tackled and formulating the protocol for the review procedures. Questions were framed by our research topic, and all the research papers were studied carefully based on below research questions.

RQ1. Which Artificial Intelligence and Machine Learning techniques, methods, algorithms, and models were used for predictive analytics or forecasting of commodities?

RQ2. Which AIML tools were used to implement/develop the ML model?

RQ3. Do prediction techniques work on single or multiple commodities?

i) Describe in brief which commodities were used.

RQ4. Did the research work analyze the interdependency among commodities?

RQ5.What parameters or metrics considered to examine the models?

In stage 1 a large chunk of data was received for review. The results of the query were run on IEEE, ACMDL, Wiley, Science Direct, Springer, and other sources, results are shown in Table 1. Titles obtained from the studies in stage 1 were again preprocessed and analyzed in stage 2 to determine their importance to the review's parameters. The abstracts of publications were studied and shortlisted in Stage 3. Papers that proposed AI-ML techniques, models, algorithms for predictive analytics of the commodity were considered in stage 4. The amount of data that was retrieved at each stage of the selection process is shown in Table 2. 1463 research data were found using the search keywords. A shortlist of 147 research works was created using their titles. Sixty-three research publications were chosen for full-text reading based on their abstracts. In the final stage 33 publications were short listed for extensive review after reading full texts, and were considered as primary studies.

Table 1: Search String for the Study

<table>
<tr><th>S.no</th><th>Digital library</th><th>Date of Search</th><th>Search String</th><th>Time period</th><th>Content Type</th><th>No. of matches</th></tr>
<tr><td>1</td><td>IEEE EXPLORE</td><td>24.8.2022</td><td rowspan="5">("Artificial Intelligence" OR "Machine learning") AND "Commodity Market"</td><td rowspan="5">Entire range</td><td rowspan="5">Conference Publications, Journals, magazines, books, eBooks</td><td>14</td></tr>
<tr><td>2</td><td>WILEY</td><td>05.09.2022</td><td>60</td></tr>
<tr><td>3</td><td>SPRINGER</td><td>07.05.2022</td><td>954</td></tr>
<tr><td>4</td><td>ACM DIGITAL LIBRARY</td><td>03.09.2022</td><td>33</td></tr>
<tr><td>5</td><td>SCIENCE DIRECT</td><td>05.09.2022</td><td>398</td></tr>
</table>

Table 2: Count of articles selected at every stage of the review process

Stage	**IEEE Explore**	**Wiley**	**Springer**	**ACM DL**	**Science Direct**	**Others**
1	14	60	954	33	398	4
2	8	8	76	3	50	2
3	7	6	26	2	20	2
4	7	3	10	0	13	0

During the data extraction process, the included elements were chosen to facilitate a thorough analysis of 33 original studies and to extract the data needed to address the research issues. Figure 1 illustrates the variety of research published in variety of databases over the years.

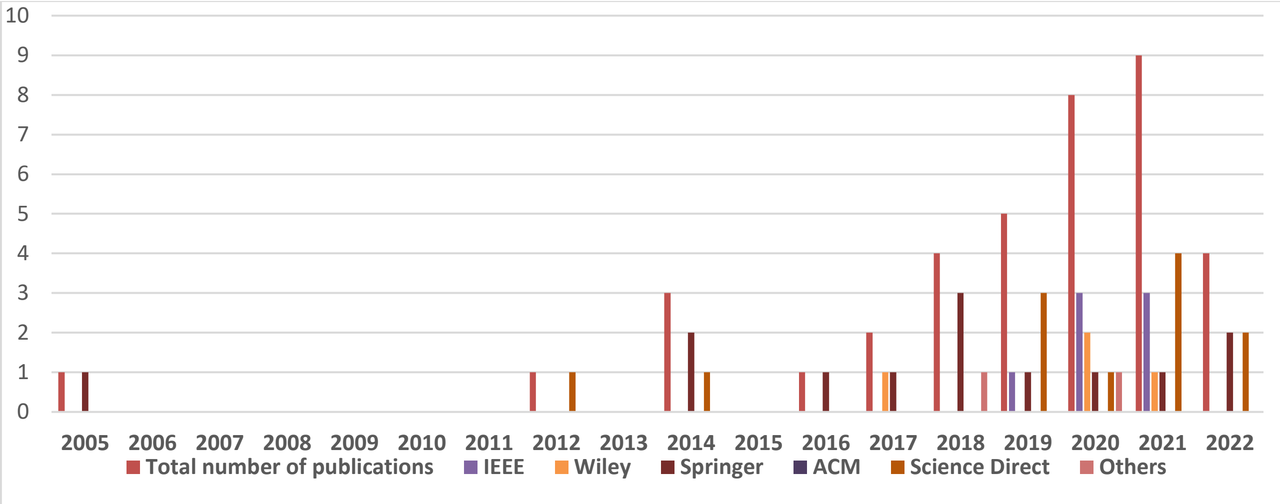

Fig. 1. Year and source-based count of publications

3. Results

Result section takes into consideration all five research question. All question were thoroughly studied and answered in this section with all the studies in detail.

3.1 AIML Techniques and Models

AI-ML refers to the broad set of techniques and models used to simulate intelligent behavior and enable machines to learn from data. A varied number of AIML techniques, models, methods, and algorithms have been used to predict the future values of commodities. This section showcases all the AIML techniques and methods covered during the process. The models, algorithms, modes, and AIML approaches implemented by the key papers under review are listed in Table 3.

Table 3: AIML techniques, models, algorithms, and methods

S. No.	Study	AIML techniques, models, algorithms, and modes
1	(Chevallier, Zhu, & Zhang, 2019)	Ensemble empirical mode decomposition (EEMD) Least-square support vector machine (LSSVM)
2	(Alameer, Elaziz, Ewees, & Ye, 2019)	Whale optimization algorithm (WOA), Neural Network (NN) ARIMA models
3	(Ali, Ahmed, Aliyuda, & Bello, 2022)	Deep Neural Network (DNN)
4	(Feng, Ji, Zhao, & Nian, 2005)	Neural Networks (NN)
5	(Wagner, Ramentol, Schirra, & Michaeli, 2022)	Dense Neural Network(DNN)
6	(Wang, Hong, Li, & Wang, 2020)	ANN, Linear Regression(LR), Suport Vector Regressor(SVR)
7	(Amin, 2020)	Support Vector Regressor (SVR), Random Forest (RF), Bagging (BG), GredientBoost (GB), Light Gradient Boosted Machine, Adaptive Boosting (AdaBoost)(AB), XGBoost (XGB)
8	(Das & Padhy, 2016)	USELM-SVR hybrid model
9	(Ghosh, Sanyal, & Jana, 2020)	EEMD, Singular Spectrum Analysis (SSA), Random Forest (RF), Bagging (BG)
10	(Gupta, Pierdzioch, & Salisu, 2022)	Random Forest (RF)
11	(He, Yu, & Lai, 2012)	Wavelet decomposed ensemble algorithm
12	(Ayyappa, Reddy, Vajha, & Venkat, 2021)	Time series analysis
13	(Cortez, Saydam, Coulton, & Sammut, 2018)	Time series

14	(Yousefi, Sianaki, & Sharafi, 2019)	Time Series, ARIMA
15	(Alameer, Elaziz, Ewees, & Ye, 2019)	ARIMA, Adaptive neuro-fuzzy inference system (ANFIS) Genetic Algorithm (GA), SVM, GARCH
16	(Bloznelis, 2017)	ARIMA, ARFIMA, ARARMA, VAR, VECM, KNN, ANN
17	(He, Tso, Zou, & Liu, 2018)	ARMA-GARCH, QRNN, VMD
18	(Huynh, Kreinovich, & Sriboonchitta, 2014)	AR-GARCH Models, Time series
19	(Jumoorty, Thoplan, & Narsoo, 2022)	Hybrid ANN-MC-GARCH model
20	(Xue and Sriboonchitta 2014)	GARCH Copula Approach
21	(Huang, Dai, Wang, & Zhou, 2021)	VMD-GARCH, LSTM
22	(Aziz, Abdullah, & Zaidi, 2020)	RNN-LSTM Neural Network
23	(Jiao, Song, Kong, & Tang, 2021)	PSO-LSTM model
24	(Lu, Sun, Duan, & Wang, 2021)	variable selection-LSTM integrated model
25	(Li, Zhu, & Wu, 2019)	VMD, SVM, NN, GA
26	(Li, Wang, Wei, & Zhu, 2021)	VMD, Cumulative sums of squares (ICSS)-bidirectional gated recurrent unit (BiGRU)
27	(Ferrari, Ravazzolo, & Vespignani, 2021)	Global VAR dataset sparse approach
28	(Ramyar & Kianfar, 2017)	ANN, VAR
29	(Gurzhiy, Paardenkooper, & Borremans, 2022)	Predictive Analytics
30	(Drachal K. , 2021)	Bayesian dynamic finite mixtures
31	(Drachal K. , 2019)	Dynamic Model Averaging (DMA)
32	(Lasheras, Nieto, Gonzalo, Valverde, & Krzemie´n, 2022)	Multivariate Adaptive Regression Splines (MARS)
33	(Yang, Guo, Sun, & Li, 2021)	K-means +, KPCA+, KELM

Prediction of the commodities prices involves a large number of ML algorithms, models, and approaches. Almost 11 out of 3l primary studies implemented NN in their research. The reviewed studies employed a diverse range of AIML techniques for commodity forecasting. Neural network models such as ANN, DNN, LSTM, and hybrid ANN–GARCH were the most widely used, effectively capturing non-linear and temporal patterns in commodities like oil, gas, electricity, and gold. Ensemble and tree-based models, including Random Forest, XGBoost, LightGBM, and EEMD-based ensembles, performed strongly for metal and agricultural commodities due to their robustness against noise. Several optimization-driven hybrid models—GA-ANFIS, PSO-LSSVM, WOA-NN, and VMD-GARCH/LSTM—showed superior performance by combining decomposition and learning techniques. Traditional time-series and econometric approaches such as ARIMA, GARCH, VAR, and copula models remained essential for modeling volatility and long-term co-movements. One study also integrated deep learning with sentiment analysis using PSO-LSTM. Overall, hybrid models consistently outperformed standalone algorithms across the literature.

3.2 Artificial Intelligence and Machine Learning Tools

This section discussed second research question about AIML tools used by the researchers.

In this study, it was found that nearly 16 out of 33 studies have disclosed the details of ML tools being used in their research, and out of which only 1 study has created there forecasting tool in the form of the e-commerce web application. Statistics as per studies inculcate that 3 out of 16 have used Python to cater their research, 5 out of 16 preferred R language for their research, 4 studies favored MATLAB software to accomplish their task and nearly 3 studies used the combination of MATLAB + R to achieve their desired research.

3.3 Commodities? Single commodity or Multiple commodities

This section explored third research question about commodities where researchers focused either on single or multiple commodities in their research and have applied predictive analytics accordingly and predicted their prices or growth or interdependencies. Findings are listed in the Table 4 below:

Table 4: Types of commodities

Single commodity	Multiple Commodity
(Alameer, Elaziz, Ewees, & Ye, 2019), (Alameer, Elaziz, Ewees, & Ye, 2019), (Ali, Ahmed, Aliyuda, & Bello, 2022), (Ayyappa, Reddy, Vajha, & Venkat, 2021) (Aziz, Aziz, Abdullah, & Zaidi, 2020), (Bloznelis, 2017), (Chevallier, Zhu, & Zhang, 2019), (Das & Padhy, 2016), (Feng, Ji, Zhao, & Nian, 2005), (Gupta, Pierdzioch, & Salisu, 2022), (He, Tso, Zou, & Liu, 2018), (He, Yu, & Lai, 2012), (Huang, Dai, Wang, & Zhou, 2021), (Jiao, Song, Kong, & Tang, 2021), (Lasheras, Nieto, Gonzalo, Valverde, & Krzemie´n, 2022), (Li, Zhu, & Wu, 2019), (Li, Wang, Wei, & Zhu, 2021), (Lu, Sun, Duan, & Wang, 2021), (Ramyar & Kianfar, 2017), (Wagner, Ramentol, Schirra, & Michaeli, 2022), (Wang, Hong, Li, & Wang, 2020), (Xue & Sriboonchitta, 2014), (Yang, Guo, Sun, & Li, 2021) and (Yousefi, Sianaki, & Sharafi, 2019)	(Amin, 2020), (Drachal K. , 2021), (Drachal K. , 2019), (Ferrari, Ravazzolo, & Vespignani, 2021), (Ghosh, Sanyal, & Jana, 2020), (Jumoorty, Thoplan, & Narsoo, 2022).

The reviewed studies focused on a wide range of commodities, grouped into four major sectors: metals, energy, agriculture, and seafood. Metal commodities—including gold, silver, copper, iron, zinc, lead, and nickel—were the most extensively examined across the literature. Energy commodities such as crude oil, natural gas, coal, electricity, and carbon prices also received significant attention, followed by agricultural commodities like wheat, cotton, avocado, and dairy products. Seafood commodities were the least represented, with only salmon studied. Notably, the majority of research (30 out of 33 studies) concentrated on forecasting a single commodity, while only a few studies explored multiple commodities simultaneously.

3.4 Interdependence among commodities? Name of commodity.

Commodity interdependency plays a great role in the price fluctuations thus affecting the stocks as well. One instance can be oil and gold are interdependent as the increase or decrease in the price of one affects the other. Table 5 reports all the details covered in this research question.

Table 5: Studies and Commodities Interdependency

S.No.	**Study**	**Commodity**	**Interdependence upon**
1.	(Alameer, Elaziz, Ewees, Ye, & Jianhua, 2019)	Gold	Copper, iron, silver, oil
2.	(Alameer, Elaziz, Ewees, & Ye, 2019)	Copper	Oil, iron, silver, gold
3.	(Ali, Ahmed, Aliyuda, & Bello, 2022)	Natural gas	Crude oil
4.	(Lasheras, Nieto, Gonzalo, Valverde, & Krzemie´n, 2022)	Gold	Aluminum, copper, iron, lead, nickel, platinum, potassium, nickel, silver, tin, zinc
5.	(Lu, Sun, Duan, & Wang, 2021)	Crude oil	Gold
6.	(Ramyar & Kianfar, 2017)	Crude oil	Natural gas
7.	(Xue & Sriboonchitta, 2014)	Coal	Agricultural Commodity
8.	(Yang, Guo, Sun, & Li, 2021)	Crude oil	Gold, Copper, Natural gas
9.	(Yousefi, Sianaki, & Sharafi, 2019)	Electricity	Natural gas

3.5 Metrics

This section focuses on Metrics that are used to check the performance of the model. Regression models have continuous output. Thus, we have MAE, MSE, RMSE, and R2 mostly. Some of the metrics used are discussed in detail below.

1. MAE(Mean Absolute Error) : It is the average between the real value and the predicted value as shown in equation 1. The lower the MAE more accurate the model.

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

2. MASE(Mean Absolute Scaled Error): It is the average absolute difference between the values fitted by the model and the observed historical data as shown in equation 2. The lower the MASE is, the better the model.

$$\text{MASE} = \frac{MAE}{\frac{1}{n-1} \sum_{i=2}^{n} | yi - yi-1 |} \quad (2)$$

3. MAPE(Mean Absolute Percentage Error): MAPE is the mean of all absolute percentage errors between the predicted and actual values as shown in equation 3. A low MAPE value is considered a good score.

$$\text{MAPE} = \frac{1}{Number\ of\ Predictions} \times \sum \left(\frac{| actual - prediction |}{actual} \right) \quad (3)$$

4. STD(Standard Deviation): Standard deviation is the measure of dispersion. It's represented by the sigma symbol and calculated by taking the square root of the variance is just the average of the squared differences from the mean as shown in equation 4. The smaller the standard deviation, the better the data.

$$\text{STD} = \frac{\sigma}{\sqrt{n}} \quad (4)$$

5. MSE (Mean Squared Error): MSE is a measure of the quality of the estimator. Mean or average of the difference between actual and estimated value as shown in equation 5. The lower the value, the better the MSE.

$$\text{MSE} = \frac{\sum(\text{yi} - \text{pi})^2}{\text{n}} \quad (5)$$

6. MPE (Mean Percentage error): MPE or Mean Percentage error (or deviation). It is a relative measure that essentially scales the ME (Mean error) in percentage units instead of percentage units instead of variable units as shown in equation 6. The main advantage of MPE is that it lets you compare variances between various scaled data. The lower the value of MPE, the better the model is.

$$\text{MPE} = \frac{100\%}{n} \sum_{t=1}^{n} \frac{a_t - f_t}{a_t} \quad (6)$$

7. RMSE (Root Mean Squared Error): RMSE or Root Mean Squared Error measures the average value difference between the value predicted by a model and the actual values as shown in equation 7. It provides an estimation of how well the model can achieve the target model value accurately. The lower the value of RMSE, the better it is.

$$\text{RMSE} = \frac{\sqrt{\sum_{i=1}^{n} ||\text{y(i)} - \widehat{y(i)}||^{2}}}{\sqrt{N}} \quad (7)$$

8. R^2 (R square): R^2 or the coefficient of determination is a statistical measure that determines the proportion of variance in the dependent variable that the independent variable can explain as shown in equation 8. The higher the R^2 value the better the model fits the data.

$$R^2 = 1 - \frac{\text{RSS}}{\text{TSS}} \quad (8)$$

All the metric used and their accuracy has been discussed in detail in Table 6.

Table 6: Studies, Metrics and Accuracy

S. No	Study	Metric	Proposed model accuracy
1.	(Alameer, Elaziz, Ewees, Ye, & Jianhua, 2019)	 RMSE MSE STD	WOA-NN 0.02131 0.00047 0.00340

		R^2	0.9989
2.	(Alameer, Elaziz, Ewees, & Ye, 2019)	 RMSE MSE MAE	GA-ANFIS 0.0813 0.0067 0.0644
3.	(Ali, Ahmed, Aliyuda, & Bello, 2022)	 RMSE MSE R^2	DNN 0.2440 0.0595 0.9937
4.	(Amin, 2020)	 MSE MAE R^2	Ensemble methods (RF,BG,AB, GB) 0.00 0.00 1.000
5.	(Aziz, Abdullah, & Zaidi, 2020)	MSE	WTI Brent 0.00167 0.00156
6.	(Bloznelis, 2017)	MAE MAPE MASE	2.6 2.8 2.6
7.	(Chevallier, Zhu, & Zhang, 2019)	RMSE	GA-LSSVM PSO-LSSVM 1.1075 1.1066
8.	(Das & Padhy, 2016)	 RMSE MAPE	USELM-SVR 0.0356 0.0356
9.	(Drachal K. , 2021)	RMSE MAE MASE	4.7359 3.4568 0.9750
10.	(Drachal K. , 2019)	 RMSE MAE MASE	lead nickel zinc 102.8000 1179.0000 112.0000 78.5800 865.3000 84.4000 0.9254 0.9278 0.9250
11.	(Feng, Ji, Zhao, & Nian, 2005)	MAPE	0.46
12.	(Ferrari, Ravazzolo, & Vespignani, 2021)	MSFE	coal gas oil 0.020 0.022 0.024
13.	(Ghosh, Sanyal, & Jana, 2020)	MSE	crude oil gold natural gas copper silver 0.00162 0.0016 0.0018 0.00228 0.0020
14.	(Gupta, Pierdzioch, & Salisu, 2022)	RMSFE MAFE	h=1 h=3 h=6 h=12 0.9903 0.9939 1.0028 1.0112 h=1 h=3 h=6 h=12 0.9920 0.9931 1.0020 1.0128
15.	(He, Tso, Zou, & Liu, 2018)	MSE	ARMA GARCH Crude oil 0.0019
16.	(He, Yu, & Lai, 2012)	MSE	8.9689
17.	(Huang, Dai, Wang, & Zhou, 2021)	 RMSE MAE MAPE	VMD-GARCH/LSTM Carbon 1 1 1
18.	(Jiao, Song, Kong, & Tang, 2021)	 MAE	1 step-size 2 step size 5 step-size 0.1556 0.1679 0.1973

		MSE RMSE MSLE	0.0667 0.0918 0.0904 0.2584 0.3031 0.3008 0.0113 0.0063 0.0692
19.	(Jumoorty, Thoplan, & Narsoo, 2022)	 RMSE(gold) RMSE(silver) MAE(gold) MAE(silver) R^2(gold) R^2(silver)	MC-GARCH ANN-MC-GARCH 1.682998e-04 1.131813e-04 2.920596e-04 2.508833e-04 8.928050e-05 6.440241e-05 1.791815e-04 1.543788e-04 0.2560 0.7399 0.3194 0.5111
20.	(Lasheras, Nieto, Gonzalo, Valverde, & Krzemie´n, 2022)	 MAD MSE RMSE MAPE	Pred12t Pred1t Pred12d Pred1d 293.4832 67.6022 207.3078 190.8743 284499.47 9403.18 50480.51 44362.32 533.3849 96.9700 224.6787 210.6237 15.7366 3.8803 12.1378 11.1630
21.	(Li, Zhu, & Wu, 2019)	 MAPE RMSE	WTI Brent 1.41(VMD-ARIMA) 1.28(VMD-GABP 3) 1.076(VMD-GASVM 2) 1.074(VMD-GARP 3)
22.	(Li, Wang, Wei, & Zhu, 2021)	 RMSE MAPE MAE DA	VMD-ICSS-BiGRU 0.0057 0.0062 0.0035 0.8351
23.	(Lu, Sun, Duan, & Wang, 2021)	 RMSE MAPE	LSTM 1.12 0.74
24.	(Ramyar & Kianfar, 2017)	R^2 MSE	0.99 0.0672
25.	(Wang, Hong, Li, & Wang, 2020)	 RMSE MAE MAPE	RR 5.42 4.13 5.40
26.	(Yang, Guo, Sun, & Li, 2021)	 MAPE RMSE DA	K-means+ KPCA+ KELM 5.44% 0.0311 90.91%
27.	(Yousefi, Sianaki, & Sharafi, 2019)	MSE MAE RMSE	0.015 0.096 0.125

The studies employed a variety of performance metrics to evaluate forecasting accuracy, with the most common being MAE, MSE, RMSE, MAPE, MASE, R^2, and MSLE. Overall, ensemble and hybrid models delivered the best results, often achieving extremely low error values, including RMSE scores below 0.01. Deep neural network models such as DNN and LSTM also demonstrated strong predictive performance, particularly for energy commodities. In contrast, traditional econometric models showed moderate accuracy but remained highly effective for capturing volatility dynamics in commodity markets. Despite significant advancements in predictive analytics for commodity markets, several important research gaps remain.

- Limited Generalization Across Commodities and Contexts: Many models are trained on limited datasets focusing on a single commodity (e.g., crude oil or gold) or specific regional markets, restricting their applicability to broader, global contexts. The paper notes that 24 out of 33 studies focused on single commodities, leading to poor.
- Lack of Scalable Frameworks: Existing models often fail to scale to multi-commodity scenarios or handle dynamic, non-linear market patterns effectively. The review emphasizes the need for more robust, accurate, and scalable predictive analytics frameworks.
- Insufficient Integration of Heterogeneous Data Sources: Studies rarely incorporate diverse data types, such as macroeconomic indicators, geopolitical events, investor sentiment, social media trends, high-frequency trading data, or alternative unstructured sources (e.g., online news or search trends). This limits the holistic understanding of market behavior.
- Absence of Case Studies: Few studies include relevant case studies on predictive analytics usage in commodity organizations, limiting insights into practical implementation.

The paper focus specifically on predictive analytics in commodity markets, filling a gap in comprehensive review. However, it stresses that these identified gaps particularly in generalization, data integration, and real-world testing emphasize opportunities for future research. To address these gaps, we suggest directions like developing adaptive trading agents, incorporating alternative data sources, and focusing on multi-commodity, explainable models.

5. Conclusion

This systematic literature review examined 33 studies on AI and ML-based predictive analytics for commodity markets, covering metals, energy, agriculture, and seafood. The findings show a clear shift from traditional econometric models like ARIMA and GARCH toward advanced machine learning, deep learning, and hybrid ensemble methods, which consistently deliver higher forecasting accuracy. Neural networks, LSTM variants, Random Forest, SVM, and decomposition-based hybrids were the most effective techniques, while Python, R, and MATLAB were the commonly used tools. Most studies focused on single-commodity forecasting, highlighting a gap in multi-commodity and interdependency modeling. Key limitations included limited datasets, lack of tool transparency, and minimal use of heterogeneous data sources such as sentiment or macroeconomic indicators. Future research should integrate reinforcement learning, multi-commodity analysis, alternative data sources, and explainable AI to build more adaptive, interpretable, and real-time forecasting systems. Overall, the review provides a comprehensive foundation for advancing intelligent and data-driven predictive analytics in commodity markets.

References

[1] Chevallier, J., Zhu, B., & Zhang, L. (2019). Forecasting Infection Points: Hybrid Methods with Multiscale Machine Learning Algorithms. *Computational Economics*, 39.

[2] Akulwar, P. (2020). A Recommended System for Crop Disease detection and yield prediction using machine learning approach. *Recommender System with Machine Learning and Artificial Intelligence: Practical Tools and Applications in Medical, Agricultural and Other Industries*, 23.

[3] Alameer, Z., Elaziz, M. A., Ewees, A. A., Ye, H., & Jianhua, Z. (2019). *Natural Resources Research*.

[4] Alameer, Z., Elaziz, M. A., Ewees, A. A., Ye, H., & Jianhua, Z. (2019). Forecasting gold price fluctualtions using improved multilayer perceptron neural network and whale optimization algorithm. *Resources policy*, 11.

[5] Ali, A., Ahmed, M. K., Aliyuda, K., & Bello, A. M. (2022). Deep Neural Network Model for Improving price prediction of natural gas. *2021 Internation conference on Data Analytics for Business and Industry* (p. 5). IEEE Explore.

[6] Amin, M. N. (2020). Predicting Price of Daily Commodities using Machine Learning. *2020 International Conference on Innivation and Intelligence for Informatics, Computing and Technlogies* (p. 6). Saint Etienne: IEEE explore.
[7] Arai, K. (2021). Combined Non-Parametric and Parametric Classification Method Depending on Normality of PDF of Training Samples. *Int. J. Adv. Comput. Sci. Appl*, 310-316.
[8] Ayyappa, P., Reddy, P., Vajha, A., & Venkat, S. (2021). Cotten price Prediction: An Artifical Intelligence based solution. *Proceeding of the 5th International conference on I-SMAC* (p. 5). IEEE Explore.
[9] Aziz, Aziz, N., Abdullah, M. A., & Zaidi, A. N. (2020). Predictive Analytics for Crude oil Price using RNN-LSTM neural network. *2020 International conference on computational Intelligence* (p. 5). IEEE Explore.
[10] Bloznelis, D. (2017). Short-term salmon price forecasting. *Journal Of Forecasting*, 19.
[11] Cortez, C. T., Saydam, S., Coulton, J., & Sammut, C. (2018). Alternative techniques for forecating mineral commodity prices. *International Journal of Mining Science and technology*, 14.
[12] Das , S. P., & Padhy, S. (2016). Unsupervised extreme learning machine and support vector regression hybrid model for predicting energy commodity futures index. *Memetic Comp.*, 14.
[13] Drachal, K. (2019). Forecating prices of selected metals with bayesian data rich models. *Resouces policy*, 15.
[14] Drachal, K. (2021). Forecasting selected energy commodities prces with bayesian dynamic finite mixtures. *Energy ecomonics*, 14.
[15] Feng, C., Ji, G., Zhao, W., & Nian, R. (2005). The Prediction of the Financial Time Series Based on Correlation Dimension. *International Conference on Natural Computation* (p. 10). Springer.
[16] Ferrari, D., Ravazzolo, F., & Vespignani, J. (2021). Foreccasting Energy commodity prices: A large global dataset sparse approach. *Energy economics*, 12.
[17] Ghosh, I., Sanyal, M. K., & Jana, R. K. (2020). An Ensemble of emsemble Framework for predictive analytics of commodity market. *2020 4th International Conference on Computaional Intelligence and Network* (p. 6). IEEE Explore.
[18] Ghule, R., Gadhave, A., Dubey, M., & Kharade, J. (2022). Gold Price Prediction using Machine Learning. *International Journal of Environmental Engineering*.
[19] Gupta, r., Pierdzioch, C., & Salisu, A. A. (2022). Oil price uncertanity and U.K. unemployment rate: A forecating experiment with random forest using 150 years of data. *Resources Policy*, 7.
[20] Gurzhiy, A., Paardenkooper, K., & Borremans, A. (2022). Predictive Analytics at an Oil and Gas Company: The Rosneft Case. *Algorithms and Solutions Based on Computer Technology*, 13.
[21] Haarnoja, T., Zhou, A., Abbeel, P., & Levine, S. (2018). Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor. *Proceedings of the 35th International Conference on Machine Learning*, (pp. PMLR 80, 1861-1870.).
[22] He, K., Tso, G. K., Zou, Y., & Liu, J. (2018). Crude Oil Risk Forecasting: New Evidence from Multiscale Analysis Approach. *Energy Economics*, 36.
[23] He, K., Yu, L., & Lai, K. K. (2012). Crude oil price analysis and forecasting using wavelet decomposed ensemble model. *Energy*, 11.
[24] Hu, J., Wang, J., & Ma, K. (2015). A Hybrid Technique for Short-Term Wind Speed Prediction. *Energy*, 563-574.
[25] Hu, Y., Liu, K., Zhang, X., Su, L., Ngai, E., & Liu, M. (2015). Application of evolutionary computation for rule discovery in stock algorithmic trading: A literature review. *Applied Soft Computing*, 534-551.
[26] Huang, Y., Dai, X., Wang, Q., & Zhou, D. (2021). A hybrid model for carbon price forecasting using GARCH and long short term memory network. *Applied Energy*, 18.
[27] Huynh, V. N., Kreinovich, V., & Sriboonchitta, S. (2014). *Modeling Dependence in Econometrics.* Springer.
[28] Jiao, X., Song, Y., Kong, Y., & Tang, X.

(2021). Volatility forecasting for crude oil based on text. *Journal of Forecasting*, 12.

[29] Jumoorty, A. F., Thoplan, R., & Narsoo, J. (2022). High frequency volatility forecasting: A new approach using a hybrid ANN-MC-GARCH Model. *IJFE*, 20.

[30] Lasheras, F. S., Nieto, P. J., Gonzalo, E. G., Valverde, G. F., & Krzemie´n, A. (2022). Time Series Forecasting of Gold Prices with the Help of Its Decomposition and Multivariate Adaptive Regression Splines. *16th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2021)* (p. 10). Springer.

[31] Li, J., Zhu, S., & Wu, Q. (2019). Monthly crude oil spot price forecating using variational mode decomposition. *Energy economics*, 14.

[32] Li, Y., Wang, S., Wei, Y., & Zhu, Q. (2021). A New Hybrid VMD-ICSS-BiGRU Approach for Gold Futures Pricec Forecasting and Algoritm Trading. *IEEE Transactions on Computational Social Systems* (p. 12). IEEE Explore.

[33] Lu, Q., Sun, S., Duan, H., & Wang, S. (2021). Analysis and forecasting of crude oil price based on the variable selection-LSTM integrated model. *Energy Informatics*, 20.

[34] Ramyar, S., & Kianfar, F. (2017). Forecasting Crude Oil Prices: A Comparison Between Artificial Neural Networks and Vector Autoregressive models. *Comput Econ*, 19.

[35] Wagner, A., Ramentol, E., Schirra, F., & Michaeli, H. (2022). Short and long term forecating of electricity prices using embedding of calender information in neural networks. *Journal of commodity markets*, 18.

[36] Wang, J., Hong, T., Li, X., & Wang, S. (2020). A multi-granularity heterogeneous combination approach to crude oil price forecasting. *energy Economics*, 28.

[37] Xue, G., & Sriboonchitta, S. (2014). Co-movement of Prices of Energy and Agricultural Commodities in Biofuel Era: A Period-GARCH Copula Approach. *Advances in Intelligent Systems and Computing*, 15.

[38] Yang, Y., Guo, J., Sun, S., & Li, Y. (2021). Forecasting crude oil prices with a new hybrid approach and multi souce data. *Engneering Applications of Artificial Intelligence*, 10.

[39] Yousefi, A., Sianaki, O. A., & Sharafi, D. (2019). Long term electricity price forecast using machine learning techniques. *2019 IEEE PES innivative Smart Grid Technologies Asia* (p. 5). IEEE.