SCIENTIFIC RESEARCH CENTER

International Journal of Social Science and Humanities Research
Vol. 3, No. 4, 2024, pp. 28-35.

ISSN 2348-2990

International Journal of Social Science and Humanities Research

www.scientificrc.com

Investigating the uncertainty of the predicted price of crude oil obtained from the machine learning model using Bayesian

Alireza Kamrani[1*], Shaghayegh Sahraei[2], Ramin Taghinejad[3]

** Master of Information Technology Management - Business Intelligence, Tarbiat Modares University, Tehran, Iran. E-Mail: a.kamrani@modares.ac.ir*
Assistant Professor, Department of Information Technology Management, Faculty of Management and Economics, Tarbiat Modares University, Tehran, Iran.
Industrial Engineering, Malek Ashtar University of Technology, Tehran, Iran.

Abstract

The expansion of artificial intelligence applications has led to an increase in the number of investors using machine learning models to predict and study the prices of commodities such as crude oil, gold, stocks, forex, and more. Recent research has shown that, unlike traditional statistical and econometric models, machine learning can not only predict prices but also describe the complex factors influencing this field. Among the predictive models based on machine learning, artificial neural network families have enjoyed widespread popularity. The analytical capabilities created by these models have increasingly facilitated human-machine interaction, such that organizations' and individuals' reliance on this type of analysis is clearly discernible. Furthermore, uncertainty analysis, as one of the most critical pillars of risk management, has always been of concern to investors, analysts, and financial and economic operators. Among the machine learning tools that effectively perform uncertainty analysis, Bayesian networks can be mentioned. Bayesian networks model the intensity of relationships between variables quantitatively, allowing conditional beliefs about them to be automatically updated with access to new information. In this research, a hybrid approach based on Bayesian networks was used to design 10 predictive pricing models. Unlike other typical price prediction research, this study focused on quantifying the uncertainty of the outcomes from machine learning models. Additionally, the performance of conventional machine learning models, before and after integration with the Bayesian method, was compared and analysed.

Keywords: Machine learning, Price prediction, Uncertainty, Bayesian networks.

Introduction

Crude oil, or as industry professionals call it, "black gold," is considered a vital element for the continuation of the global economic system. Throughout human civilization's history, humanity's dependence on this particular commodity is felt more today than ever before, and the extent of control and management of these energy-supplying resources has an undeniable connection with the level of countries' development. Therefore, given the importance of this issue for economic actors, accurate prediction of crude oil prices has become essential to ensure sustainable economic development. Having an accurate assessment of price fluctuation ranges allows suppliers to manage their risk, and the end consumer at the end of this supply chain will enjoy greater peace of mind. Financial and economic sector actors have always developed powerful models for predicting financial markets, and essentially, these models have played a very efficient role in optimizing investors' portfolios and managing their risks.

The expansion of artificial intelligence applications has led to an increase in the number of investors using machine learning models to predict and study stock and forex prices. It has been proven that stock fluctuations and forex prices can be predicted, with the difference that, unlike traditional statistical and econometric models, machine learning can describe complex influencing factors (Hu, Zhao & Khushi, 2021).

Since their initial introduction, machine learning systems have required large data sets for efficient training that are not available in other areas of economics. In addition, due to the relatively low processing power of computers of that time, training was very time-consuming. (Gogas & Papadimitriou, 2021)

In the framework of this scientific research, our main objective is to design and construct predictive models by employing a multifaceted set of machine learning techniques, including Long Short-Term Memory (LSTM) networks, Support Vector Machines (SVMs), linear regression, and Bayesian networks. Our research aims to harness the collective predictive power of these diverse methods to gain a comprehensive understanding of complex phenomena and deliver highly accurate predictions. It is anticipated that the integration of these models will provide a more accurate and robust perspective for prediction, as each model may excel in different aspects and under distinct conditions. Consequently, this integration holds the promise of creating a predictive framework that leads to greater adaptability and higher performance across a diverse range of scenarios.

In addition to building predictive models, the strength of our research is rooted in the critical task of assessing the inherent uncertainty of the predictive models themselves. This aspect of our study is important because it underscores the critical importance of acknowledging and quantifying the uncertainty that is an inherent feature of predictions made by machine learning models.

Research purposes:

The main goal of this research is to fill a vital gap in the current landscape of crude oil price prediction models based on machine learning, while simultaneously addressing the needs of business intelligence in organizations. Although significant efforts have been made to develop these models for making financial

and economic decisions, there is a noticeable oversight in evaluating and measuring the uncertainty associated with predictions. This deficiency is not specific to the financial sectors but is also an important issue in business intelligence, where making informed decisions based on data analysis and risk management is crucial.

The research objectives is to empower stakeholders in both the financial sector and business intelligence teams with a robust framework for assessing and managing the fluctuations and potential errors associated with predictions.

Methodology

This research effort entails the formulation of a specialized machine learning model specifically designed to forecast crude oil prices. Our model attempts to increase the accuracy of price prediction using established methods and available data sources. However, the hallmark of this research is its unwavering commitment to examining and reducing the inherent uncertainty underlying these predictions. In pursuit of this goal, we use machine learning methods, which is presented below a a diagram of the overall process of conducting this research.

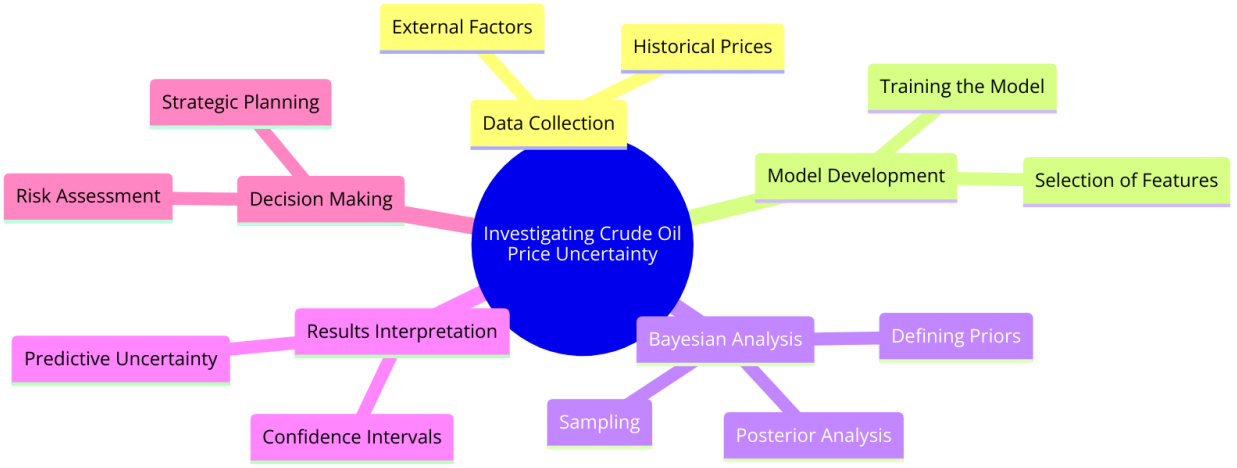

Figure 1: Mind map of the research process

Here's a breakdown of each part of the process:

1. Data Collection

• Historical Prices: Gathering data on past prices of crude oil which might include daily, monthly, or yearly averages.

• External Factors: Collecting information on variables that could influence oil prices, such as geopolitical events, changes in supply and demand, economic indicators, etc.

2. Model Development

• Selection of Features: Identifying and selecting the relevant features (variables) from the data that are most likely to influence predictions.

• Training the Model: Applying machine learning techniques to train the model on the historical data, ensuring it can make future predictions.

3. Bayesian Analysis

• Defining Priors: Establishing prior distributions for the parameters of the model, based on existing knowledge or assumptions.

• Sampling: Using techniques like Markov Chain Monte Carlo (MCMC) to sample from the posterior distribution of the model's parameters.

• Posterior Analysis: Analyzing the posterior distributions to understand the parameter estimates and the uncertainty around them.

4. Results Interpretation

• Predictive Uncertainty: Assessing how much uncertainty there is in the model's predictions. This might involve looking at confidence intervals or prediction intervals.

• Confidence Intervals: Generating intervals within which the true values of the predicted prices are expected to fall, with a certain level of confidence.

5. Decision Making

• Risk Assessment: Evaluating the risks associated with potential future prices and their uncertainties.

• Strategic Planning: Using the insights gained from the analysis to inform decisions and strategy, such as hedging strategies or investment decisions in the oil market.

This process leverages Bayesian methods to not just predict prices but to also understand the uncertainty in these predictions, which is crucial for making informed decisions in volatile markets like crude oil.

In this research endeavour, we employed a multi-faceted ensemble of machine learning techniques, including Long Short-Term Memory (LSTM) networks, Support Vector Machines (SVMs), and linear regression models, to generate predictions of the target variable, which in this case is the price of crude oil. The outputs obtained from these diverse predictive models were then utilized as input nodes for constructing a novel predictive framework based on Bayesian networks.

Specifically, we leveraged the predictive outputs from the LSTM, SVM, and linear regression models as evidence nodes within the formulated Bayesian network architecture. This approach allowed us to harness

the collective predictive capabilities of these individual models while accounting for their respective strengths and limitations through the principled probabilistic framework of Bayesian networks.

Furthermore, a significant advantage of adopting the Bayesian network paradigm is the inherent capability to quantify and characterize the uncertainty associated with the predicted outcomes. By incorporating the evidence from the machine learning models into the Bayesian network structure and conducting probabilistic inference, we were able to derive the posterior probability distribution over the target variable, which represents the predictive uncertainty.

Consequently, in addition to obtaining a consolidated prediction by integrating the outputs from multiple machine learning models, our proposed approach facilitated the determination of uncertainty ranges or credible intervals for the predicted crude oil prices. This uncertainty quantification is a valuable asset in decision-making processes, as it provides stakeholders with a comprehensive understanding of the potential variability and risk associated with the predictions, enabling more informed and robust decisions.

Results

In this research, we presented an integrated model that predicts the price. In picture below we can see the prediction results of LSTM model which is preparing to use for designing our Bayesian model Further.

Figure 2: LSTM Model

The following table shows the performance evaluation report of the model

Table 1: Performance report

Metric	Results
MAPE	25.29
RMSE	16.22
R-squared	56

Bayesain Networks

This model was created by combining all the other machine learning models as a new node in our Bayesian network, which consists of 8 nodes and 17 graphs. This model has a more complex structure compared to the introduced models and involves more conditional cause-and-effect relationships in its network. In this model, the ultimate effect is also the predicted price, and its results are as described in the following table:

Table 2: Performance Report of integrated Bayesian networks

Metric	**RMSE**	**MAPE**	**R-squared**
Calibration	4.17885	6.26727	0.982969
Validation	5.1382	9.06411	0.838692

In figure below you can see the you can see the components of this network.

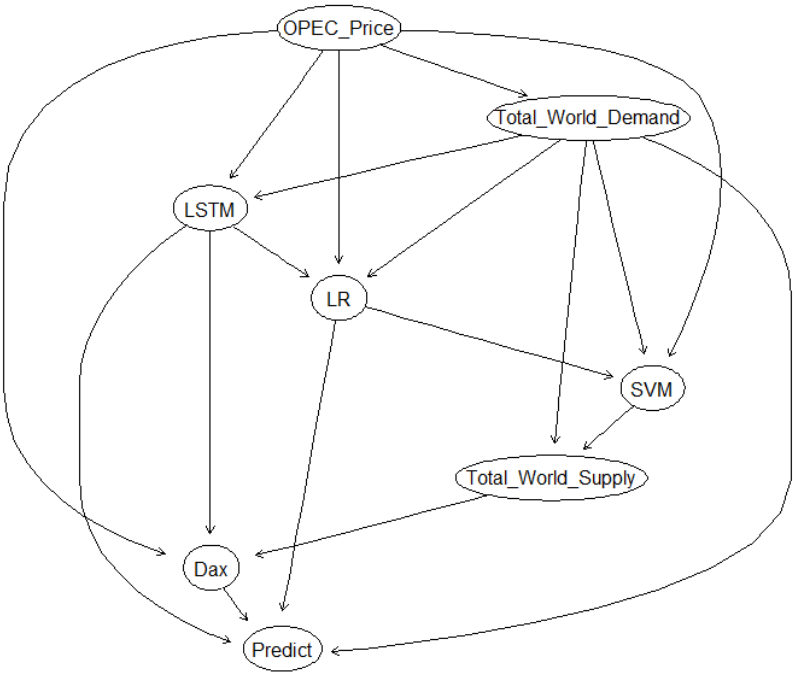

Figure 3: Integrated Bayesian networks

Also, in the figure below, the illustrated view of the uncertainty range of the predicted prices is displayed.

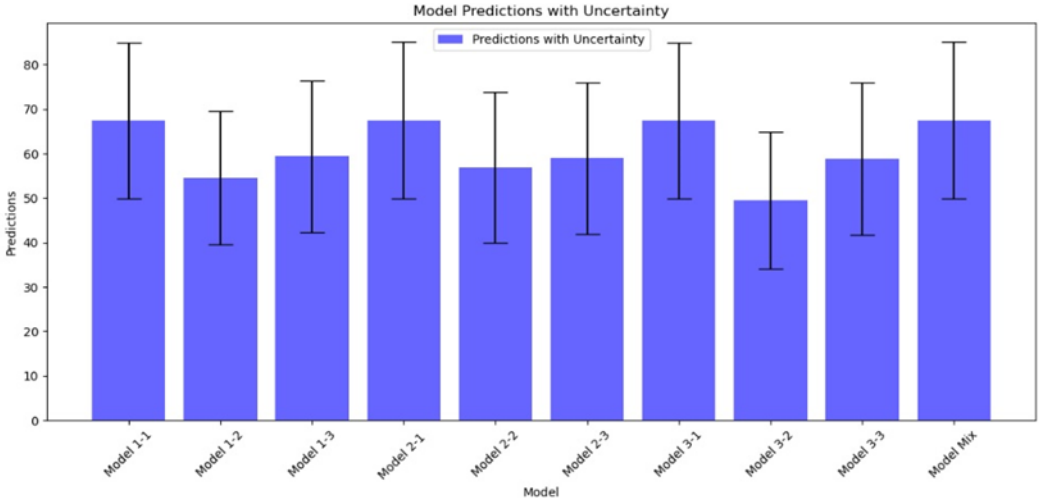

Figure 4: uncertainty range of the models

Conclusion

This research study has comprehensively examined the effectiveness of Bayesian networks and machine learning models in predicting crude oil prices. The analysis confirms that the hypotheses posited at the onset of the study are substantiated by the results obtained, thereby validating the pivotal role Bayesian networks play in enhancing the accuracy and reliability of predictive models.

Firstly, the use of Bayesian networks as tools for quantifying uncertainty in predictive models of crude oil prices has proven effective. The models trained with collected data on crude oil prices demonstrated a significant ability to quantify prediction uncertainties, thereby confirming the initial hypothesis.

Secondly, the integration of Bayesian networks into machine learning models for crude oil price prediction has resulted in noticeable improvements in the models' performance and reliability. This integration not only supported the original predictions but also provided a clearer understanding of the factors influencing price forecasts and the related uncertainties.

Furthermore, the application of Bayesian networks has contributed to increased transparency and interpretability of the models, allowing researchers to better understand the dynamics influencing predictions. This enhanced interpretability confirms the third hypothesis, underscoring the contribution of Bayesian networks to more comprehensible and accountable predictive analytics.

In summary, the findings of this research underscore the powerful role of Bayesian networks in refining machine learning models for crude oil price prediction. Future research can extend these methodologies to different types of data and incorporate geopolitical analysis using data from news and social media to

further improve predictive accuracy and reliability. This study's results advocate for the continued exploration and utilization of Bayesian networks in enhancing machine learning models across various domains, particularly in economic forecasting and market analysis.

References

Gogas, P. and T. Papadimitriou (2021). "Machine Learning in Economics and Finance." Computational Economics 57(1): 1-4.

Hu, Z., Zhao, Y., & Khushi, M. (2021). A survey of forex and stock price prediction using deep learning. Applied System Innovation, 4(1), 9.

Kumar, Y. J. N., Preetham, P., Varma, P. K., Rohith, P., & Kumar, P. D. (2020, July). Crude Oil Price Prediction Using Deep Learning. In 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA) (pp. 118-123). IEEE.

Lux, T., Segnon, M., & Gupta, R. (2016). Forecasting crude oil price volatility and value-at-risk: Evidence from historical and recent data. Energy Economics, 56, 117-133.

Moitra, N., Raj, P., Saxena, S., & Kumar, R. (2020). Crude oil prediction using LSTM. International Journal of Innovative Science and Research Technology, 5(2), 900-905.

Fazelabdolabadi, B. (2019). A hybrid Bayesian-network proposition for forecasting the crude oil price. Financial Innovation, 5(1), 1-21.