DOI 10.2478/sbe-2024-0049
SBE no. 19(3) 2024

GOLD PRICE FORECASTING: A NOVEL APPROACH BASED ON TEXT MINING AND BIG-DATA-DRIVEN MODEL

Saeed KIAN POOR
Faculty of Economics, Payame Noor University, Tehran, Iran

Shahram FATTAHI
Faculty of Economics, Razi University, Kermanshah, Iran

Mohsen HAJIAN
Faculty of Economics, Payame Noor University, Tehran, Iran

Abstract:

This work uses Google Trends and online text mining to test a novel approach to gold price prediction. To explain the gold price prognosis, the convolutional neural network (CNN) also uses linguistic criteria for news items about gold. The findings of this study can be used to explain the theoretical underpinnings of information processing worldwide. According to experimental findings, big data forecasting and online data mining are more efficient than other methods. As a result, utilizing 19926 in sum news headlines acquired, the synergistic connection between news items and Google Trends is successful in accurately projecting the price of gold.

Key words: *gold price Prediction, Trends, Online News, Data Mining, CNN*

1. Introduction

It has received much research in this area, primarily because predicting changes in financial asset values is an indicator that enables us to look at the risk attached to the asset over time. One of the assets that are frequently used to counteract this volatility in a portfolio is gold. Forecasting the price of gold in the financial markets is essential due to its widespread use in the economic climate. Economic theory refers to gold as a Safe Haven since it attracts many investors during times of financial crisis, as has been widely reported in the financial press (Sanderson, 2015). For instance, the S & P 500 lost 10% during the 2015 global stock market decline.

Baur and Lucey (2010) created the concept of safe havens, which clarified the connection between gold and other assets in times of crisis. When stock indices drop below the 5 percent yield distribution values, they consider a market distressed (Baur et al., 2010).

For forecast variations, ARCH models are frequently explored (Trück et al., 2012). These models are effective at capturing short-term changes, but they struggle to identify long-term traits, one of their disadvantages. In the long and short runs, DNNs could pick up more intricate patterns., which allows them to improve the predictions produced by econometric models. Create a technique that can help in gold fluctuation prediction.

Selecting reliable predictors is one of the critical issues because some are challenging to extract. According to the development of big data technologies, there is enough internet data to reflect financial market shocks (Yu et al., 2019). Among search engines, Google's engine is the most useful for finding the most recent information on a particular topic.

In light of this, the current study considers Google Trends as one of the reliable predictors for gold. Google Trends has proven to be an effective predictor in the gold market. As an informational predictor of gold price expectations, A good indicator of drivers is provided by Google Trends. As a result, Google Trends is frequently cited as an example of a particular big data strategy with a wealth of data. Recent developments in quantitative evidence have included the qualitative examination of politics, emergencies, and disasters (Li et al.,2015).

Additionally, the amount of media content, including in-depth qualitative data that can be chosen to support projections of the price of gold, keeps growing. This study mixes qualitative data with statistical data to predict the price of gold to identify effective predictors. The following research question (R.Q.) is what this study aims to answer:
How much can gold price prediction be improved by combining Google Trends and internet media?

The aforementioned suggests that the best results for gold price prediction may come from mixing qualitative knowledge with statistical data.
According to the findings of our experiment, Google Trends and gold news features provide valid predictive data.

Following is a summary of the key sections:
(A)This is the first time we have used Google Trends and gold news to predict the price of gold weekly.
(B) To demonstrate the prediction power of gold news headlines, an in-depth learning system is utilized to derive textual attributes from online gold news.
(C) A fresh approach to gold price prediction was put out. With the addition of Google Trends and news text, the B combined test and Granger causality analysis were also used.

2. Related Works

This field includes many related papers on time series fluctuation prediction.

- Volatility forecasting methods

Many models in the literature can predict volatility, including ARCH models and combinations of GARCH-ISD figures (Kroner et al., 1995).

Aside from this, researchers used an AP-GARCH model to examine macroeconomic factors that affected gold between 1983 - 2003. Ultimately, the dollar emerged as the most significant explanatory factor (Tully et al., 2007). The TARCH model ultimately demonstrated

the most significant outcomes in further tests by the authors, where they assessed other models (GARCH, TARCH, TGARCH, ARMA) (Trück et al., 2012).
Strategies were used by Parisi, Parisi, and Dáz (2008) to forecast changes in the price of gold. Eventually examined Rolling and Recursive Neural Networks as changes to neural networks (Parisi et al., 2008). Time-series predictions have been made using machine learning approaches such as SVM (Choudhury et al., 2014), fuzzy logic (Dash et al., 2016), and ANFIS (Yazdani et al., 2012). (Kristjanpoller et al., 2018).

- Time series representation using visuals

Image categorization is a logical use for CNN. Convolutional network designs may be used to represent information and potentially deliver information based on pictures. However, financial time series suffer from this because no elements can recognize them(as in Gunduz, Yaslan, & Cataltepe, 2017). In Wang and Oates (2015), time series were encoded using a novel framework that permits the application of computer vision methods for categorization (Wang et al., 2015).

- Series data prediction with deep learning

The prediction of financial time series has attracted much interest thanks to deep learning models. In recent years, studies of the commodity market have extensively used models including DBN, GRU, and LSTM, DBN. Aside from improving the performance index, the hybrid model also (Kim et al., 2018). what is divided into two sections, and the combined use of both modules greatly enhanced the forecasts produced by the LSTM network. A DBN has been frequently utilized to implement various studies as a stock trading decision system (Zhu et al., 2014).

- Deep learning to classify images

This technology is state-of-the-art for numerous applications, including image identification, speech recognition, facial recognition, and others (Mitra et al., 2018). A stock market prediction model was created by researchers (Ding et al., 2015). A model that performed better than an SVM was put forth to forecast individual stocks. Accurately predicting the daily direction of 100 Stock Exchange stocks, it proposed a CNN architecture with ordered properties (Gunduz et al., 2017).

- Literature Review

Text mining has been the subject of several investigations. Del Vecchio, for example, researched smart tourism destinations (Del Vecchio et al., 2018). Kim and Ju examined key internet media headlines and blogs to forecast the rise (Hemmatian et al., 2019). Emerging artificial intelligence (A.I.) was summarized by Hemmatian and Sohrabi (Zhao et al., 2019). A deep learning technique called CNN has been widely used in research (Li et al., 2019). The CNN algorithm is used in this work to assess attitudes in addition to extracting related text characteristics from online gold news. (Nassirtoussi et al., 2015). Filtering raw texts, though, is very important. This study innovates using Google Trends and gold news to forecast gold prices. It is essential to investigate how Google Trends and gold news texts interact to affect the gold price.

3. Procedure

- obtaining and processing data

Gold News, Google Trends, and gold price data were gathered independently between 2011 and 2021. gold price data was received via the renowned international economic website "Investing. com."

From the "gold news" section of the well-known news website "Investing.com," we gathered all the available information on gold news. We used news headlines for our study rather than full articles since they capture the essence of the material and essential information (Ye et al., 2006).

The first crucial step in text mining is tokenization, which involves removing stop words and punctuation.

Search engines generally ignore stop words, But they are taken into account here. The length of each document should therefore be equal; we convert them using the padded sequence approach. In the following, some items' words have been deleted. Finally, To determine a word's importance within a corpus, we employ

TF-IDF. The following is a specific expression of TF-IDF:

$$tf - idf = \frac{n_{i,j}}{\sum_k n_{k,j}} \times \log \frac{|D|}{\left|\left\{J : \mathrm{t}_i \in D_j\right\}\right|} \tag{1}$$

$n_{i,j}$ = the number of times the term appears; $\sum_k n_{k,j}$ = the amount of words in existence; $\frac{n_{i,j}}{\sum_k n_{k,j}}$ represents TF, |D| = the number of papers overall; {j: ti∈dj} is the sum of documents

- Google Trends's preference

According to an inquiry on the gold market, the top three primary factors—price, inventory, and consumption—have a close relationship. Some studies have shown short-term gold inventories and gold prices to have a nonlinear relationship (Li et al., 2015). Some people chose five search terms using Google Trends (Cai et al., 2019). Finally, using the empirical methodology, our study chose four key search terms: "gold future," "gold data," and "gold price,." This research uses the cointegration test and Granger analysis to determine if and how Google Trends and gold prices are related. The cointegration test begins by examining the stationarity of time series due to non-stationary. For the static series data Xt and Yt, the Engel-Granger test is also used, and linear regression is performed as follows:

$$y_t = a_0 + a_1 \times x_t + \mu_t \tag{2}$$

The following is an expression for the Granger causality between yt and xt:

$$Pr\left(x_t \mid I_t - 1\right) = Pr\left(x_t \mid I_t - 1 - Y_{t-1}^n\right) \left(t = 1, 2, \cdots, T\right) \tag{3}$$

The following causal link is then modeled using the vector autoregression (VAR) model:

$$x_t = a_0 + a_1 x_{t-1} + \cdots + a_m x_{t-m} + b_1 y_{t-1} + \cdots + b_n y_{t-n} + \mu_t \tag{4}$$

$$y_t = c_0 + c_1 y_{t-1} + \cdots + c_n y_{t-n} + d_1 x_{t-1} + \cdots + d_m x_{t-m} + \gamma_t \tag{5}$$

The Granger causality moving from xt to yt can be demonstrated using Eq. (4) (or Eq. (5)).

- Neural Network Using Convolutions

Due to the features of CNNs constructed on deep learning networks and convolutional structures, this technique has concentrated on the affective interpretation of sentences. The CNN model's structure is shown in Figure 1. A tokenized sentence matrix serves as the basis of the CNN model. The sentence matrix is then convolved through some linear filters. The final classification is created by entering the result generated after applying a softmax function. Finally, To help us better comprehend the CNN algorithm, a clear and comprehensive flowchart is presented in Figure 2. It is important. TensorFlow is a Python library that we use.

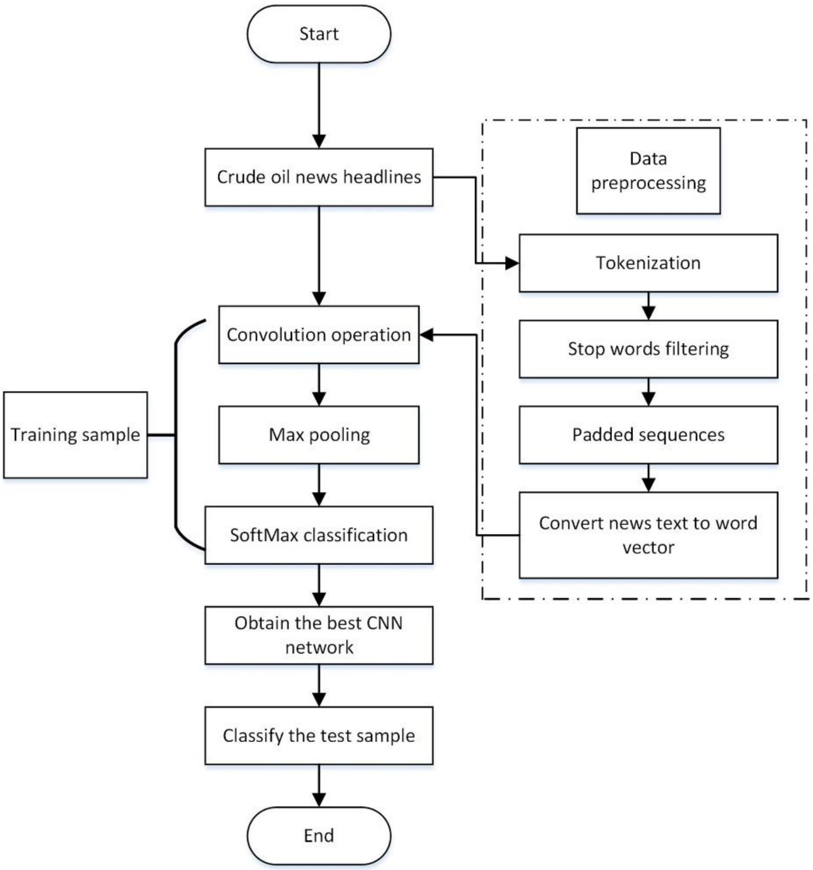

Fig. 1. Flowchart of CNN

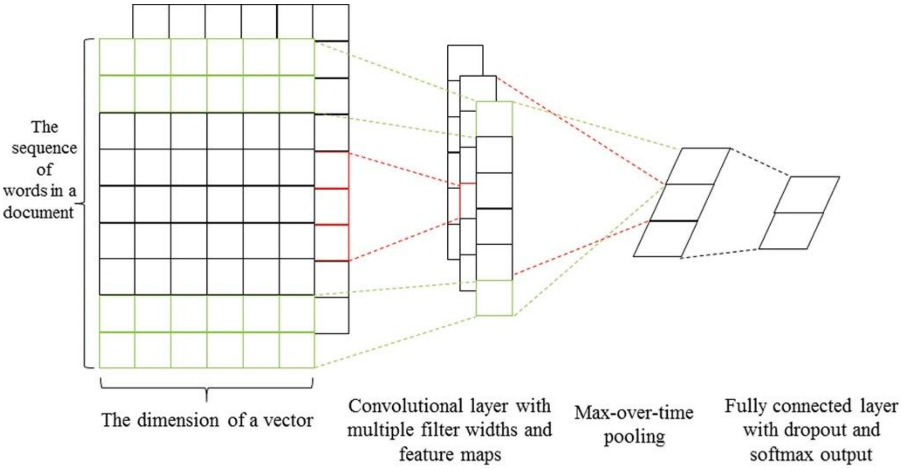

Fig. 2. complete flowchart of the CNN model

- Deep learning models

this study employs six in-depth learning models, including KNN, Decision Tree, CNN, SVM, RANDOM FOREST, and GRADIENTBOOSTING. The following provides a succinct overview of these six models: KNN is a nonparametric statistical technique used in pattern recognition for statistical classification and regression. The output of Ki depends on the type employed in the classification and regression, and it contains the closest instructional example in the data space. The SVM is one of the supervised learning algorithms utilized for both classifiers, in contrast to more conventional approaches for classification. Using a structure of numerous decision trees and depending on the training and output times of classes, random forest is a mixed learning method for classification and regression. The decision tree typically outperforms the random forest in terms of performance, although the data type also influences this performance improvement. Any ensemble approach that may combine some weak learners to produce a powerful learner is known as the gradient boosting method.

4. Experimental study

To demonstrate the superiority of the new method of gold price forecasting, a real-world example of the gold price is suggested below. Python, Jupiter Notebook, Google Colab, MATLAB, Ives, and Excel tools are all used to code the suggested algorithms. The computer used to perform this estimate has a 2.7GHz CPU, 8 G.B. of RAM, and a licensed copy of Windows 10.

- Data descriptions

The three sets of information are the historical gold price, news headlines, and sets of Google patterns. From 7 April 2011 to 30 June 2021, totaling 19926 news texts published in the "gold news" category of the well-known news portal "investing.com" were gathered. With 200 total observations, the weekly data span the time frame from 30 August 2017 to 30 June 2021. divides the data set used in this study into training and testing sets (see Fig. 3). Google trends for "gold price," "gold future," and "gold data" are available.

The training time for the CNN model is from 7 April 2011 to 30 August 2017, consisting of 334 weekly recordings and 11700 news headlines. There are 8226 articles and 200 weekly records over the test period, from 30 August 2017 to 30 June 2021.

Fig. 3. Training and testing

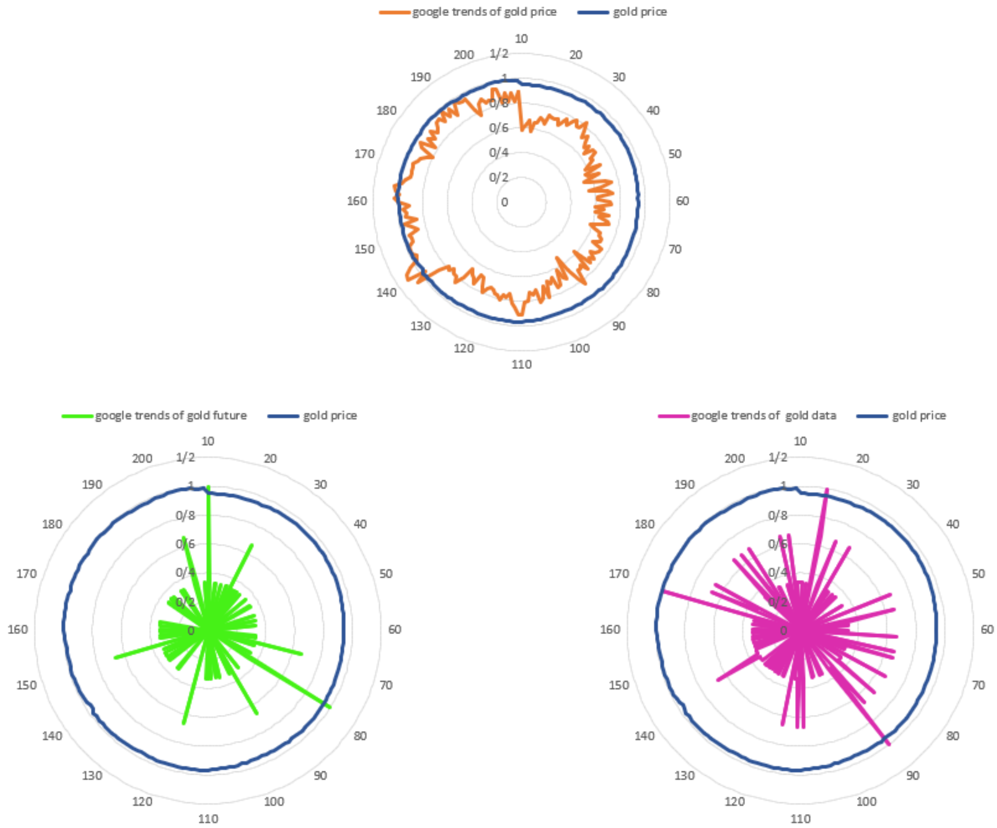

Fig. 4. Time series of the four Google Trends

- Google Trends's preference

The time-series data for the four Google Trends are displayed in Fig. 4. The three Google Trends for gold (gold price, gold future, and gold data) and the gold price are stationary, according to Table 1 (ADF test). As a result, the prerequisites for cointegration and Granger connections are correctly satisfied. The findings that demonstrate cointegration linkages are shown in Table 2. Table 5 (Granger causality) demonstrates that the Google trend of "gold price" Granger causes gold prices to vary over one and two orders as well as across one order and manifests as the Google trend of "gold future" The Granger causality of gold price does not exist.

Table 1. At the 1% level, check for stationarity.

Series	Level	First Difference
gold price	-3.4634 (0.8520)	-3.4634 (0.0000)
Google trends - “gold price.”	-3.4635 (0.1862)	-4.0055 (0.0000)
Google trends - “gold data.”	-3.4633 (0.0000)	-4.0063 (0.0000)
Google trends - “gold future.”	-4.0048 (0.0000)	3.4641(0.0000)-

Table 2. checks for cointegration at the 5% level.

Google Trends	First Difference
“gold price”	(0.0000) -15.3985
“gold data”	-14.6110 (0.0000)
“gold future”	-3.9044 (0.0368)

Table 3. Gold price descriptive statistics and "gold price" trend on Google

	Google trend of “Gold Price”	Gold Price
Mean	1.4171	0.9659
Median	1.3802	0.9618
Standard deviation	0.2264	0.0184
Skewness	0.2323	0.3918
Kurtosis	2.2804	1.7363
Jarque-bera	6.1129*	18.4241*

* significance=10%level.

Table 4. Choosing a lag order: Google trend of “gold price.”

	LR	FPE	AIC	HQ
0	Na	1.19e-08	-6.898401	-6.870915
1	813.7976*	1.81e-10*	-11.08359*	-10.94616*
2	21.45887	1.90e-10	-11.03419	-10.78682
3	11.74997	2.10e-10	-10.93316	-10.57585
4	24.34232	2.16e-10	-10.90560	-10.43834
5	15.76834	2.33e-10	-10.83114	-10.25394

Table 5. 5 percent level Granger causality analysis.

lags		
	1	2
Section a	Granger Does Not Cause Gold Prices, According to Google Trends	
Chi-sq	10.59943	10.71446
probability -value	0.0011	0.0047
Section b	Google trends of gold data do not Granger Cause Gold Price	
Chi-sq	0.302085	1.402580
probability -value	0.0582	0.4959
Section c	Google trends of gold future do not Granger Cause Gold Price	
Chi-sq	0.091561	1.837275
probability -value	0.7622	0.3991

Additionally, Fig. 5 shows the times series data for the price of gold and the "Gold Price" trend on Google.

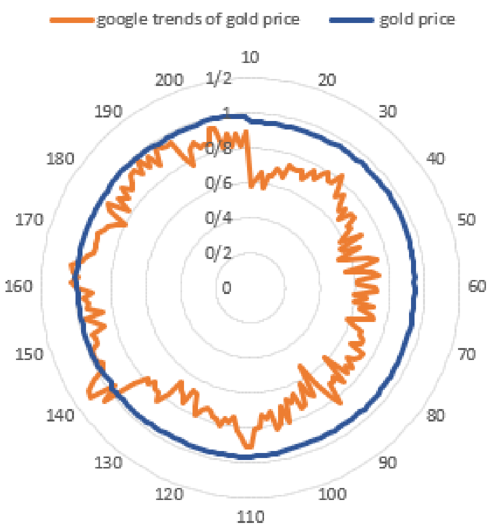

Fig. 5. Compare the "gold price" trend on Google with the historical gold price

- Text mining of online news

Word clouds with the TF-IDF are shown in Fig. 6 and include phrases like "dollar," "futures," "gold," "higher," "lower," and "prices," as well as "states," "stocks," "trade," and "united." The terms "gold," "global market," and "stocks" imply a close association with the price of gold, as seen in Fig. 6.

Fig. 6. Word cloud made out of the top 100 terms from the database

Following some hyperparameter experiments, the CNN hyperparameters are set via a grid search technique. The CNN model's findings are displayed in Table 6.

Table 6. outcomes of CNN's categorization.

Accuracy	Precision	Recall	F-measure
0.63	0.49	0.5	0.65

Because news items only provide a set of information that influences the price of gold. It notes that the CNN model falls short of expectations for incredible accuracy. However, other research shows that the research has a suitable accuracy(Wu, Hu et al., 2021, 2020).

- Performance assessment

MAE and RMSE assess forecasting accuracy (Hu, Lv et al., 2020, 2018). Small criterion values show high prediction accuracy.

To measure the additional explanatory power of Google Trends and text attributes, the improvement rate (IR) is introduced (Li et al., 2015):

$$IR_{RMSE} = \frac{RMSE_A - RMSE_B}{RMSE_B} \times 100 \quad (6)$$

$$IR_{MAE} = \frac{MAE_A - MAE_B}{MAE_B} \times 100 \quad (7)$$

- Hyperparameter tuning test

In this study, rolling-based prediction and one-step-ahead forecasting are both used. The details of each program using the grid search approach. Alternatively put, alternative parameter combinations are tested on the training set, and the effects of each combination of hyperparameters are assessed using 10-fold cross-validation (10- cv) (Ju et al., 2013).

Table 7. CNN's parameter combinations(10-cv MSE error)

Rate of Learning	Amount of hidden neurons				
	1	2	3	4	5
0.01	0.0056	0.0068	0.0058	0.0058	0.00062
0.05	0.0055	0.0055	0.0054	0.0065	0.0069
0.1	0.0061	0.0055	0.0061	0.0059	0.0064

Table 8. The Method Parameters

Method	hyperparameter
CNN	100 epochs, 0.05 learning rate, three hidden neurons
KNN	Completion is 0.01; the kernel is polynomial and radial; the epochs are 200.
SVM	C = 2.5; gamma = 0.205; kernel = "rbf"
Gradient Boosting	Max depth = 20, 30, and 40; stopping tolerance = 0.01; learning rate = 0.
Decision Tree	Epochs are 300, hidden neurons are 20, and batch size is 50.
Random Forest	,kernel = "rbf" ,**criterion{"gini", "entropy"}, default="gini**" $\delta^2 = [0,1]$

Table 9. Forecasting model performance comparison.

week	Actual	CNN	SVM	Random Forest	Decision Tree	KNN	GradientBoosting
.1-2	0/73272	0/7905	0/7905	0/7071	0/7071	0/7905	0/6123
.3-4	0/683833	0/8164	0/8164	0/6666	0/8819	0/7453	0/7453
.5-6	0/672821	0/5773	0/8164	0/4714	0/5773	0/3333	0/5773
.7-8	0/746905	0/5773	0/8164	0/4714	0/5773	0/3333	0/5773
.9-10	0/810017	0/7071	0/7071	0/3535	0/6123	0/3535	0/6123
.11-12	0/929404	0/7071	0/7071	0/3535	0/7071	0/5	0/5
.13-14	0/7613	0/7071	0/7071	0/6123	0/5	0/7071	0/6123
.15-16	1/017275	0/5	0/866	0/5	0/3535	0/5	0/3535
.17-18	0/682463	0/7908	0/525	0/5	0/3423	0/5666	0/4104
.19-20	0/790668	0/7905	0/625	0/7071	0/866	0/6123	0/866

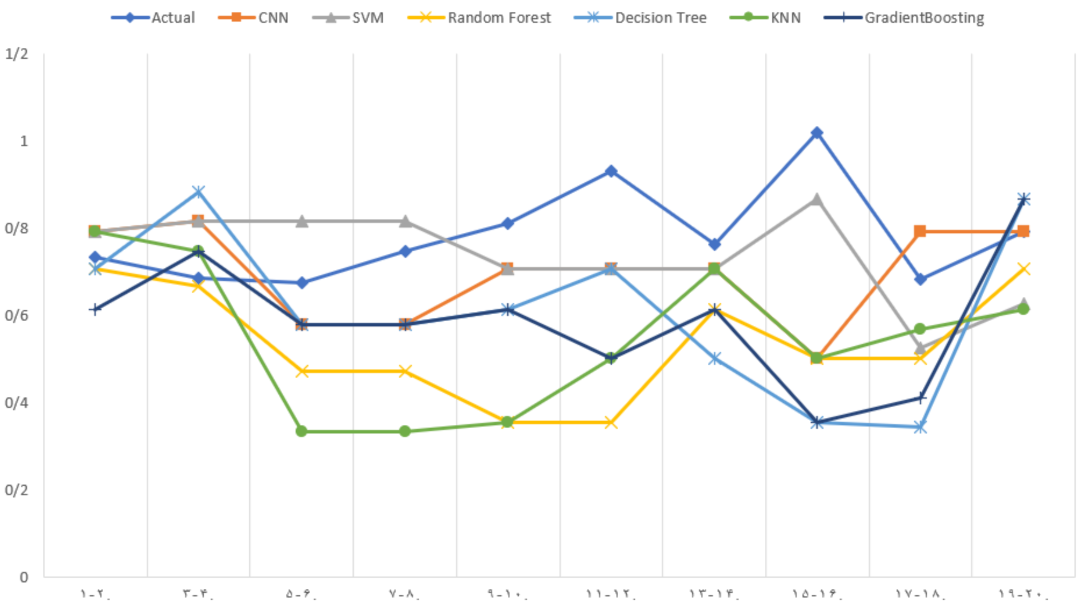

Fig. 7. Results of the different models

This study considers CNN, SVM, Random Forest, Decision Tree, KNN, and GradientBoosting to find the best approach for forecasting the price of gold. According to Table 9 and Fig. 7, the CNN model performs admirably and delivers the best RMSE criterion results. The Wilcoxon test, used to identify statistical differences, compares the results in Table 10. (Ju et al., 2013). The estimation's findings demonstrate the independence of the samples. Additionally, the mean rank demonstrates that compared to other neural networks, CNN exhibits a lower proportion of improvement.

Table 10.Wilcoxon test.

Models	Wilcoxon signed-rank test	Probability
CNN vs. Decision Tree	3.288291	0.0010
CNN vs. GRADIANTBOOST	2.759141	0.0058
CNN vs. KNN	2.078805	0.0376
CNN vs. SVM	1.020504	0.307
CNN vs. Random Forest	0.188982	0.0850

* Significant difference at α = 0.05

This study compared the effectiveness of text, Google Trends showcases, and all other chosen attributes, as stated in Table 12, to assess the relative estimates of the future. Results of the I.R.s are shown in Table 11. They suggest that Google Trend and news headlines work best together. They also show improvements in the IRMSE and IRMAE of 10.87 and 147.17 percent, respectively. Additionally, results that mix text and Google trend features are 158.04 percent more accurate than results that only use historical data and combine text with Google trend features. Note that Google trend features can significantly increase the precision of gold price predictions.

One of the better metrics for evaluating the allocated coefficient values is MIV. The mean impact value (MIV) technique in this investigation is shown in Table 12. Table 12 shows that the suggested strategy for predicting gold prices using relevant indicators like Google Trends and gold political news may be a helpful tool. We recognize that the Google Trend of "gold price" (0, 1) is a crucial indicator for predicting gold prices.

Table 11. CNN Model Efficiency Prediction.

	MAE	RMSE
Google trends features(1)	0.809	1.664
Text features(2)	0.3629	0.7352
Historical features(3)	0.897	1.669
IR from (3) to (1)	10.87%	0.3%
IR from (3) to (2)	147/17%	127.01%

Table 12. Input variables and MIV rankings for predicting the price of gold

Entry Parameter	MIV	Rank
Historical gold price(-1)	0.652837	1
Historical gold price(0)	0.652823	2
Google trends of gold (-1)	0.283243	3
Google trends of gold (0)	0.283439	4

5. Conclusion

To predict gold prices, this study combines qualitative and quantitative data. Relationship analysis, deep learning methods, and text mining are used to suggest a strategy for predicting the price of gold. The implications of forecasting the price of gold for marketers are numerous. Marketers can take these impacts into account because Google Trends and online media news have complementing effects on the financial market. A convolutional neural network model can show the patterns hidden in online gold news headlines. With the help of Google Trends characteristics and text, this research has produced a comprehensive prediction system. According to this study, it is possible to anticipate the price of gold with acceptable accuracy, precision, recall, and F-measure. The critical signal is the launch of a fresh text-based, big-data-driven idea that uses Google Trends and news text files.

Wu et al. (2021) results of their investigation were that the mean absolute percentage error of the proposed model is 0.0571 and 0.0459 for crude oil price forecasting of two cases, respectively. Accordingly, there is a complementary relationship between news headlines and Google Trends in accurately predicting crude oil prices. Li et al. (2019) study suggests that grouping news-text sentiment features and using Latent Dirichlet Allocation for topic modeling can improve forecasting accuracy. Empirical results indicate that the proposed models outperform older benchmark models, and combining text and financial features leads to more accurate price forecasts. Chen et al.'s (2016) empirical analysis showed that an artificial neural network forecasted gold futures prices better than ARIMA models. In addition, text mining reasonably explained the trend in gold futures prices. Livieris et al. (2020) also reached similar results and emphasized that the gold price prediction model based on deep learning can help to accurately predict the gold price in the future. Onsumran et al. (2015) succeeded in predicting the price of gold using the text mining method using news articles and proved how news can be a source for predicting the price of gold. Based on these studies, it can be said that the results of this research were in line with other researchers.

News articles might reflect social, cultural, or unforeseen economic or political events. These factors influence the volatility of the gold price, and basic information cannot be entirely gleaned from quantity data alone. Google Trends may both promote and disseminate news text content. Google Trends can proxy for investor attention, while gold price volatility and investor interest can interact. Thus, it is a fascinating and scientific way to account for and aggregate news text and Google Trends. Based on unstructured text input, CNN may be utilized to provide instructive dataset indicators. So, we can argue that this study adds to the methodology for predicting the price of gold. This study does have certain restrictions, it may be said. One involves selecting Google Trends, which is a difficult task. The number of online news data will also be increased, and deconstruction techniques will be used to increase accuracy, which is expected to boost the performance of deep learning models.

6. References

Baur, D. G., & Lucey, B. M. (2010). Is gold a hedge or a Safe Haven? An analysis of stocks, bonds, and gold. Financial Review, 45, 217–229. https://doi.org/10.32890/ijbf2012.9.1.8448

Cai, M., Pipattanasomporn, M., & Rahman, S. (2019). Day-ahead building-level load forecasts using deep learning vs. traditional time-series techniques. Applied energy, 236, 1078-1088. https://doi.org/10.1016/j.apenergy.2018.12.042

Choudhury, S., Ghosh, S., Bhattacharya, A., Fernandes, K. J., & Tiwari, M. K. (2014). A real-time clustering and SVM-based price-volatility prediction for optimal trading strategy. Neurocomputing, 131, 419–426. https://doi.org/10.1016/j.neucom.2013.10.002

Chen, H. H., Chen, M., & Chiu, C. C. (2016). The integration of artificial neural networks and text mining to forecast gold futures prices. Communications in Statistics-Simulation and Computation, 45(4), 1213-1225. https://doi.org/10.1080/03610918.2013.786780

Colladon, A. F. (2020). Forecasting election results by studying brand importance in online news. International Journal of Forecasting, 36(2), 414-427. https://doi.org/10.1016/j.ijforecast.2019.05.013

Dash, R., & Dash, P. (2016). An evolutionary hybrid fuzzy computationally efficient Egarch model for volatility prediction. Applied Soft Computing, 45, 40–60. https://doi.org/10.1016/j.asoc.2016.04.014

Del Vecchio, P., Mele, G., Ndou, V., & Secundo, G. (2018). Creating value from social big data: Implications for smart tourism destinations. Information Processing & Management, 54(5), 847-860. https://doi.org/10.1016/j.ipm.2017.10.006

Ding, X., Zhang, Y., Liu, T., & Duan, J. (2015). Deep learning for event-driven stock prediction. IJCAI, 2327–2333.

Gunduz, H., Yaslan, Y., & Cataltepe, Z. (2017). Intraday prediction of Borsa Istanbul using convolutional neural networks and feature correlations. Knowledge-Based Systems, 137, 138–148. https://doi.org/10.1016/j.knosys.2017.09.023

Hemmatian, F., & Sohrabi, M. K. (2019). A survey on classification techniques for opinion mining and sentiment analysis. Artificial intelligence review, 52(3), 1495-1545. https://doi.org/10.1007/s10462-017-9599-6

Hu, H., Wang, L., Peng, L., & Zeng, Y. R. (2020). Effective energy consumption forecasting using enhanced bagged echo state network. Energy, 193, 116778. https://doi.org/10.1016/j.energy.2019.116778

Hu, H., Wang, L., & Lv, S. X. (2020). Forecasting energy consumption and wind power generation using deep echo state network. Renewable Energy, 154, 598-613. https://doi.org/10.1016/j.renene.2020.03.042

Ju, F. Y., & Hong, W. C. (2013). Application of seasonal SVR with chaotic gravitational search algorithm in electricity forecasting. Applied Mathematical Modelling, 37(23), 9643-9651. https://doi.org/10.1016/j.apm.2013.05.016

Kim, H. Y., & Won, C. H. (2018). Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. Expert Systems with Applications, 103, 25–37. https://doi.org/10.1016/j.eswa.2018.03.002

Kim, L., & Ju, J. (2019). Can media forecast technological progress?: A text-mining approach to the on-line newspaper and blog's representation of prospective industrial technologies. Information Processing & Management, 56(4), 1506-1525. https://doi.org/10.1016/j.ipm.2018.10.017

Kristjanpoller, W., & Michell, K. (2018). A stock market risk forecasting model through integration of switching regime, ANFIS and GARCH techniques. Applied Soft Computing, 67, 106–116. https://doi.org/10.1016/j.asoc.2018.02.055

Kroner, K. F., Kneafsey, K. P., & Claessens, S. (1995). Forecasting volatility in commodity markets. Journal of Forecasting, 14, 77–95. https://doi.org/10.1002/for.3980140202

Livieris, I. E., Pintelas, E., & Pintelas, P. (2020). A CNN–LSTM model for gold price time-series forecasting. Neural computing and applications, 32, 17351-17360. https://doi.org/10.1007/s00521-020-04867-x

Li, X., Shang, W., & Wang, S. (2019). Text-based crude oil price forecasting: A deep learning approach. International Journal of Forecasting, 35(4), 1548-1560. https://doi.org/10.1016/j.ijforecast.2018.07.006

Li, X., Ma, J., Wang, S., & Zhang, X. (2015). How does Google search affect trader positions and crude oil prices?. Economic Modelling, 49, 162-171. https://doi.org/10.1016/j.econmod.2015.04.005

Mitra, V., Sivaraman, G., Nam, H., Espy-Wilson, C., Saltzman, E., & Tiede, M. (2018). Hybrid convolutional neural networks for articulatory and acoustic information-based speech recognition. Speech Communication, 89, 103–112. https://doi.org/10.1016/j.specom.2017.03.003

Nassirtoussi, A. K., Aghabozorgi, S., Wah, T. Y., & Ngo, D. C. L. (2015). Text mining of news-headlines for FOREX market prediction: A Multi-layer Dimension Reduction Algorithm with semantics and sentiment. Expert Systems with Applications, 42(1), 306-324. https://doi.org/10.1016/j.eswa.2014.08.004

Onsumran, C., Thammaboosadee, S., & Kiattisin, S. (2015). Gold price volatility prediction by text mining in economic indicators news (Doctoral dissertation, Mahidol University). https://doi.org/10.12720/jait.6.4.243-247

Parisi, A., Parisi, F., & Díaz, D. (2008). Forecasting gold price changes: Rolling and recursive neural network models. Journal of Multinational Financial Management, 18, 477–487. https://doi.org/10.1016/j.mulfin.2007.12.002

Sanderson, H. (2015). Gold rises amid expectations of ECB move on QE. The Financial Times. https://doi.org/10.1111/j.1540-6288.2010.00244.x

Trück, S., & Liang, K. (2012). Modeling and forecasting volatility in the gold market.International Journal of Banking and Finance, 9, 48–80. https://doi.org/10.32890/ijbf2012.9.1.8448

Tully, E., & Lucey, B. M. (2007). A power GARCH examination of the gold market. Research in International Business and Finance, 21, 316–325. https://doi.org/10.1016/j.ribaf.2006.07.001

Wang, Z., & Oates, T. (2015). Encoding time series as images for visual inspection and classification using tiled convolutional neural networks. In AAAI Publications, Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence.

Wu. B, Wang. L, Lv. Sh, Zeng Y.(2021). Effective crude oil price forecasting using new text-based and big-data-driven model, Elsevier, Volume 168, 108468. https://doi.org/10.1016/j.measurement.2020.108468

Yazdani-Chamzini, A., Yakhchali, S. H., Volungeviciene, D., & Zavadskas, E. K. (2012). Forecasting gold price changes by using an adaptive network fuzzy inference system. Journal of Business Economics and Management, 13, 994–1010. https://doi.org/10.3846/16111699.2012.683808

Ye, M., Zyren, J., & Shore, J. (2006). Forecasting short-run crude oil price using high-and low-inventory variables. Energy policy, 34(17), 2736-2743. https://doi.org/10.1016/j.enpol.2005.03.017

Yu, L., Zhao, Y., Tang, L., & Yang, Z. (2019). Online big data-driven oil consumption forecasting with Google trends. International Journal of Forecasting, 35(1), 213-223. https://doi.org/10.1016/j.ijforecast.2017.11.005

Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P., & Gao, R. X. (2019). Deep learning and its applications to machine health monitoring. Mechanical Systems and Signal Processing, 115, 213-237. https://doi.org/10.1016/j.ymssp.2018.05.050

Zhu, C., Yin, J., & Li, Q. (2014). A stock decision support system based on DBNs. Journal of Computational Information System, 10, 883–893.