Expert Systems With Applications 193 (2022) 116487

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Fuzzy Rule-Based Prediction of Gold Prices using News Affect

Petr Hajek [a,*], Josef Novotny [b]

[a] Science and Research Centre, Faculty of Economics and Administration, University of Pardubice, Studentsk á 84, 532 10 Pardubice, Czech Republic
[b] Institute of Business Economics and Management, Faculty of Economics and Administration, University of Pardubice, Studentská 84, 532 10 Pardubice, Czech Republic

ARTICLE INFO

Keywords:
Gold price
Prediction
Fuzzy rule-based system
News
Sentiment

ABSTRACT

Because of gold's value, systems for predicting its price have attracted extensive interest in the scientific and industrial communities. Diverse artificial intelligence methods outperform traditional statistical methods in predicting short- and long-term gold price. However, previous research has neglected the transparency of these systems, nor have these systems incorporated the potentially important effect of media sentiment on investment decisions. Therefore, we here propose a fuzzy rule-based prediction system with a component that processes various aspects of news stories. This system is trained on historical data to provide investors with one- and five-days-ahead gold price predictions while achieving a highly interpretable trading strategy in terms of rule complexity. We demonstrate that the proposed system is effective in terms of both prediction accuracy and interpretability compared with state-of-the-art models, such as extreme learning machines and neural networks with deep learning. Our findings suggest that the component of news affect is particularly important for one-day-ahead predictions. We also show that the proposed system performs well in terms of average annual return while providing an interpretable set of linguistic trading rules. This has important implications for investors.

1. Introduction

Gold has had gaining share of investment portfolios of both retail and institutional investors over the last decade (Smales, 2014). In the global economy, gold has been widely utilized as a 'store of value' and 'safe haven', as well as a derivative instrument and risk-diversification security, amongst other uses (Ntim, English, Nwachukwu, & Wang, 2015). Therefore, factors affecting gold prices have been widely investigated, with a focus on financial and macroeconomic variables (Qian, Ralescu, & Zhang, 2019). Linkages among gold and equity markets have also been assessed to demonstrate the gold's capacity to hedge equity losses (El Hedi Arouri, Lahiani, & Nguyen, 2015), its characteristic as a 'safe haven' (Creti, Joëts, & Mignon, 2013). Causalities have also been found between gold and other precious metal markets (Bhatia, Das, Tiwari, Shahbaz, & Hasim, 2018), and significant interconnections have been found between the gold market and various macroeconomic variables, such as inflation (Gangopadhyay, Jangir, & Sensarma, 2016).

In the past decade, interest has increased in statistical and artificial intelligence methods to predict the price of gold and other precious metals. Distinct characteristics of gold, including its function as a financial asset, store of monetary value, and supply accumulation, make predicting gold prices a challenging task as real-world gold price data violate assumptions of traditional statistical methods (e.g., ARIMA, GARCH) (Qian et al., 2019). Therefore, artificial intelligence prediction models, such as neural networks (NNs) (Y. Liu, Yang, Huang, & Gui, 2019) and decision trees (García & Kristjanpoller, 2019), dominate traditional statistical models in terms of accuracy. Despite their accuracy, however, existing prediction models that use artificial intelligence methods have several weaknesses. First, these models' lack of transparency and interpretability prevent investors from adopting them as a tool for decision support. Gold price time series also have large inherent uncertainty associated with gold's diverse functions and with the strong interactions between gold market and other financial and commodity markets. Another limitation of existing prediction models is that they do not adequately consider the effect of news events on gold price. As indicated below, while extant literature concerns the effect of news sentiment on financial markets, little attention has been paid to commodity markets (Smales, 2014).

The current study, seeking to address these limitations, proposes an interpretable, automatic gold-price prediction system that utilizes polarity and sentiment intensity in financial news. The proposed prediction system incorporates several components: (1) technical indicators to consider the effectiveness of technical trading strategies, (2) prices of other commodities, stock market index and foreign exchange, and (3)

* Corresponding author.
E-mail addresses: petr.hajek@upce.cz (P. Hajek), josef.novotny@upce.cz (J. Novotny).

https://doi.org/10.1016/j.eswa.2021.116487
Received 16 August 2021; Received in revised form 20 October 2021; Accepted 30 December 2021
Available online 7 January 2022

linguistic indicators extracted from affective news. These components are used as inputs to FURIA (Fuzzy Unordered Rule Induction Algorithm) (Hühn & Hüllermeier, 2009), a fuzzy modification of an unordered version of the RIPPER algorithm. This study uses FURIA because of its advantages over alternative fuzzy rule-based methods, including its high computational effectiveness and the linguistic quality of the induced rules (Chen, Shang, Su, & Shen, 2018). We here show that the proposed fuzzy rule-based prediction system is not only highly competitive with state-of-the-art artificial intelligence methods but also provides investors with an interpretable decision-support tool comprising linguistic trading rules. Trading strategy based on the proposed prediction system demonstrably outperforms alternative baseline strategies in terms of profit. This prediction system can therefore effectively mitigate uncertainty in gold price fluctuations and support future investment decisions.

The remainder of this paper is organized as follows. Section 2 overviews earlier research both on predicting gold prices and investigating the relationship between commodity prices and news affect. Section 3 introduces the proposed fuzzy rule-based prediction model, while Section 4 presents the data used for empirical analysis. Section 5 examines the performance of the proposed model compared with state-of-the-art gold price prediction models. Finally, Section 6 discusses the results and concludes.

2. Related literature on gold price prediction and the effect of news affect

This section reviews previous work on gold price prediction and provides the theoretical justification for incorporating news affect into the gold price prediction model.

2.1. Gold price prediction

In today's globalized world, it is becoming more difficult to estimate future prices for many commodities, including gold. After all, gold is valuable and forms the strategic reserves of most central banks (Wen, Yang, Gong, & Lai, 2017). The difficulty of forecasting prices also affects mining companies, which must consider future prices in their business activities. On the one hand, negative forecasts may limit, cancel, temporarily suspend or slow down mining activities. On the other hand, positive events and forecasts accelerate or expand mining activities compared to a company's initial assumptions. Therefore, an accurate prediction model is essential to the investment decisions of both the investor community and the mining companies. Previous authors have underlined this (Kristjanpoller & Minutolo, 2015), adding that predicting gold price fluctuations with high accuracy is important for both commodity markets and the global economy.

Earlier research on predicting commodity prices has focused on two main categories of method: (1) mathematical or statistical methods and (2) artificial intelligence methods. Table 1 overviews previous models used to predict gold prices.

In the mathematical or statistical category, traditional GARCH and ARIMA time series prediction models have been used to model fluctuations in gold price. A GARCH model with mixed data sampling was applied to demonstrate how global policy uncertainty can help to predict volatility in the gold futures (Fang, Chen, Yu, & Qian, 2018). A VAR-GARCH model was used to investigate volatility spillovers between the stock market and the price of gold in China (El Hedi Arouri et al., 2015). Significant volatility and return cross-effects were found, and gold was identified as a hedge for stocks. The safe-haven position of gold was also demonstrated using a dynamic conditional correlation GARCH model (Creti et al., 2013), supporting the notion that there are significant links between equity and commodity markets. To model the conditional volatility in gold price, short- and long-term volatility components were captured using GARCH-MIDAS models (Li et al., 2021).

A modified random walk model was used to show the relationship

Table 1
Summary of previous studies on gold price prediction.

Study	Methods	Input variables	Performance
(Shafiee & Topal, 2010)	random walk, ARIMA	oil price, inflation rate	RMSE = 116.52, MAE = 88.26
(Yazdani-Chamzini et al., 2012)	ANFIS, neural network	oil price, silver price	RMSE = 29.48, MAE = 0.029, R^2 = 0.971
(Kristjanpoller & Minutolo, 2015)	neural network + GARCH	USD exchange rate, DJIA index, FTSE index, oil price	MAPE = 0.649
(Gangopadhyay et al., 2016)	vector error correction model	stock market index, exchange rate, CPI, US bond rate, oil price	R^2 = 0.386
(Sivalingam et al., 2016)	extreme learning machine	silver price, oil prices, S&P500 index, foreign exchange rate	Acc = 0.938
(Guha & Bandyopadhyay, 2016)	ARIMA	previous gold price	RMSE = 716.35, MAPE = 3.135, MAE = 463.15, R^2 = 0.993
(Sharma, 2016)	ARIMA	previous gold price	R^2 = 0.462
(Baur, Beckmann, & Czudaj, 2016)	dynamic model averaging	MSCI stock price index, S&P500 index, CRB commodity price index, silver price, US consumer price index, global composite price index, US dollar index, euro index, US treasury bill and bond yield, US aggregate bond index, global foreign currency reserves index	RMSE = 5.88, MAE = 4.10, Acc = 0.59
(D. Liu & Li, 2017)	random forest	USD index, oil price, DJIA index, US consumer price index, Hang Seng index, US bond futures, S&P500 index	Acc = 0.769
(He et al., 2017)	multivariate empirical mode decomposition + ARMA	silver price, palladium price, platinum price	MSE = 1.222, Acc = 0.520
(Kristjanpoller & Hernández, 2017)	neural network + GARCH	SZSE component index, DJIA index, FTSE index, SBSE Sensex index, USD exchange rates, oil price	HMSE = 1.367
(Fang et al., 2018)	GARCH	COMEX gold price returns, GEPU index	RMSE = 3.392, MAE = 1.1789
(Kia, Haratizadeh, & Shouraki, 2018)	support vector machine + semi-supervised approach	previous gold price	Acc = 0.568
(Alameer et al., 2019)	neural network + whale optimization algorithm	silver price, copper price, iron price, oil price, exchange rates, inflation rates	RMSE = 0.021, R^2 = 0.999
(F. Weng et al., 2020)	online extreme learning machine	silver price, oil price, S&P500 index, previous gold price	RMSE = 7.520, MAPE = 0.004, MAE = 5.681
(Zhang & Ci, 2020)	deep belief network	DJIA index, exchange rate, oil	RMSE = 0.056, MAPE = 0.006,

(continued on next page)

Table 1 (*continued*)

Study	Methods	Input variables	Performance
		price, federal funds rate, US consumer price index	MAE = 0.046, Acc = 0.574
(Livieris et al., 2020)	CNN + LSTM	previous gold price	RMSE = 0.008, MAE = 0.008, Acc = 0.516, AUC = 0.519
(Sadorsky, 2021)	decision trees	gold price technical indicators	Acc = 0.870
(Mohtasham Khani et al., 2021)	LSTM	COVID-19 cases, eleven market sectors data	RMSE = 0.025, MAE = 0.018, R^2 = 0.858
(Jabeur, Mefteh-Wali, & Viviani, 2021)	XGBoost	silver price, oil price, foreign exchange rate, inflation, S&P500 index, iron price	RMSE = 34.921, MAE = 21.948, R^2 = 0.994
this study	fuzzy rule-based system FURIA with evolutionary tuning	gold price technical indicators, oil price, silver price, copper price, foreign exchange rate, DJIA index, emotion indicators in news	

Legend: Acc – accuracy, AUC – area under the receiver operating characteristic curve, HMSE – Heteroskedasticity-adjusted MSE, MAE – mean absolute error, MAPE – mean absolute percentage error, MSE – mean square error, R^2 – the coefficient of determination, and RMSE – root mean square error.

between the price of gold and its determinants, such as oil price (Shafiee & Topal, 2010). This model outperformed the ARIMA model in terms of forecasting errors. Although several other studies have applied ARIMA to predict gold prices (Guha & Bandyopadhyay, 2016; Sharma, 2016), significant improvement was achieved through modifications such as multivariate empirical mode decomposition for identifying noise factors (He, Chen, & Tso, 2017).

The Granger causality approach was used to find evidence of bidirectional causality among the prices of gold, palladium, platinum and silver (Bhatia et al., 2018). Earlier empirical evidence had also suggested that precious metal prices respond nonlinearly to changes in crude oil price and that bidirectional causality exists between oil and gold prices (Bildirici & Turkmen, 2015; Chai, Zhao, Hu, & Zhang, 2021). The existence of this relationship was later confirmed (Bouri, Jain, Biswal, & Roubaud, 2017), and exchange rate was found to have significant effects on gold price (Jain & Biswal, 2016). Similarly, a causal effect of the stock market and inflation hedge on gold price has been demonstrated (Gangopadhyay et al., 2016). Overall, these mathematical and statistical models revealed important insights into the determinants of fluctuations in gold price. However, assumptions such as the homoscedasticity and stationarity of time series limit these models' performance, because these restrictions are difficult to impose on real-world data.

Artificial intelligence models have proven more effective in predicting gold prices than the mathematical and statistical models. For example, adaptive neuro-fuzzy inference system (ANFIS) outperformed ARIMA and NN (Yazdani-Chamzini, Yakhchali, Volungevičiene, & Zavadskas, 2012). A hybrid model combined GARCH and NN to predict gold price return volatility, with the incorporation of the NN significantly reducing the error compared with GARCH alone (Kristjanpoller & Hernández, 2017; Kristjanpoller & Minutolo, 2015). To avoid the poor prediction performance of traditional gradient-based training methods due to their convergence to local minima of the error function, the whale optimization algorithm has been used to train NN (Alameer, Elaziz, Ewees, Ye, & Jianhua, 2019). The resulting model outperformed NNs trained with alternative evolutionary algorithms, as well as traditional NN and ARIMA models. Long short-term memory (LSTM) NNs have shown superior performance in recent studies (Livieris, Pintelas, & Pintelas, 2020; Mohtasham Khani, Vahidnia, & Abbasi, 2021) due to their ability to capture complex high-level temporal features from time-series data.

Earlier studies have also proposed models to predict other metal prices, such as aluminium, nickel or copper. A graphical prediction model called grey wave forecasting was developed to model irregular fluctuations in the prices of aluminium and nickel (Chen, He, & Zhang, 2016). Empirical results indicated this model was more effective compared with a random walk and ARMA. Copper prices have been predicted using several artificial intelligence methods, including a time series function optimized using the bat algorithm (Dehghani & Bogdanovic, 2018), a combination of NN and a price volatility network (Wang, Zhang, Wang, Lim, & Ghadimi, 2019), decision trees (Liu, Hu, Li, & Liu, 2017), a set of ARIMA, GARCH, NN and fuzzy inference system models (García & Kristjanpoller, 2019) and a LSTM NN (Liu et al., 2019).

Taken together, earlier related research on predicting gold and other metal prices has shown that hybrid systems combining time series analysis with artificial intelligence have improved prediction accuracy compared to mathematical or statistical models alone. The main problem with these hybrid prediction models is their lack of transparency. None of the above-mentioned artificial intelligence methods offer investors an interpretable decision-support system. Despite extensive research on gold price prediction, to the best of our knowledge, no study has adequately utilized the advantages of fuzzy rule-based systems.

2.2. *Effect of news affect*

The effectiveness of the market largely depends on the availability of information (Fama, 1965). In today's globalized world, information, whether positive or negative, spreads quickly across the Internet and social networks. Dzielinski reported that business news may resolve asymmetric information, especially on news (announcement) days (Dzielinski, 2012). Interestingly, on average, only neutral and negative news resolved asymmetric information; since positive news did not, this indicates a "positive bias" in news. Institutional and retail investors' processing and response to the news are crucial, as their response greatly impacts the overall development of financial markets (Balduzzi, Elton, & Green, 2001; Henry, 2008; Long, Song, & Tian, 2019; Piñeiro-Chousa, López-Cabarcos, Pérez-Pico, & Ribeiro-Navarrete, 2018; Tetlock, 2007; Weng, Ahmed, & Megahed, 2017). Liebmann, Hagenau, and Neumann (2012) concluded that, on one the hand, investors utilize news to implement their investment strategies and business transactions. On the other hand, financial analysts are waiting for these investors to react, evaluating their behaviour after some delay (Liebmann et al., 2012). The effect of media sentiment on returns increased for investors paying close attention (Siering, 2013), confirming this model for how investors react to news.

Understanding news events and how they affect commodity markets is critical for investors to make rational decisions. Several earlier studies have analysed the impact of macroeconomic news announcement on the gold and silver markets (Christie-David, Chaudhry, & Koch, 2000; Cornell & French, 1986; Frankel & Hardouvelis, 1985). Results indicated strong and positive effects of information thereafter.

The effect of news sentiment on commodity prices was investigated recently (Brandt & Gao, 2019; Clements & Todorova, 2016; Maslyuk-Escobedo, Rotaru, & Dokumentov, 2017; Sadik, Date, & Mitra, 2019), with causality empirically confirmed (Borovkova & Mahakena, 2015); in other words, investors trade commodity futures based on an aggregated news function. Negative sentiment reportedly has the greatest impact on gold futures returns (Smales, 2014). Similarly, changes in news sentiment promote changes in oil prices (Feuerriegel, Heitzmann, & Neumann, 2015). Both sentiment and appraisal effects are also significant determinants of commodity returns in crude oil and gold spot markets (Shen, Najand, Dong, & He, 2017).

Unlike previous studies that have investigated the relationship between news sentiment and commodity prices, in this study, we integrate news affect into a model for predicting gold prices.

3. Fuzzy Rule-Based prediction system FURIA with evolutionary tuning

In general, any fuzzy rule-based system comprises three components; fuzzification, an inference system and defuzzification (Herrera, 2008). In the fuzzification process, membership functions are used to convert numeric input variables into fuzzy sets. The inference system evaluates rules, and, in the final step, fuzzy sets are converted into output variable values. The inference system consists of a rule base and a data base. The data base comprises linguistic terms and their semantics defined by membership functions. For an n-dimensional classification problem, an if-then rule R_j in the rule base may be given as:

$$R_j:\ if\ x_1\ is\ A_{1,j}[a_1,\ b_1,\ c_1,\ d_1]\ and\ x_2\ is\ A_{2,j}[a_2,\ b_2,\ c_2,\ d_2]\ and\ \ldots\ and\ x_i\ is\ A_{i,j}[a_i,\ b_i,\ c_i,\ d_i]\ and\ \ldots\ and\ x_n\ [a_n,\ b_n,\ c_n,\ d_n]\ is\ A_{n,j}\ then\ class\ c_j\ with\ CF_j, \tag{1}$$

where R_j is the j-th rule ($j = 1, 2, \ldots, N$); $x_1, \ldots, x_n$ are input variables; $A_{1,j}$ [a_i, b_i, c_i, d_i], …, $A_{1,n}$ [a_n, b_n, c_n, d_n] denote antecedent fuzzy sets represented by trapezoidal membership functions defined by parameters a, b, c, and d (see eq. (2)); c_j represents the consequence class and CF_j denotes the grade of certainty for the j-th rule.

FURIA (Hühn & Hüllermeier, 2009) is a state-of-the-art fuzzy rule-based system that extends the crisp rule-based RIPPER algorithm. FURIA has outperformed competitive fuzzy rule-based systems due to its computational efficiency and the interpretability of its rule base (Chen et al., 2018). FURIA keeps the classification performance of the Boosting algorithm while producing a reasonable number of rules (Palacios, Sánchez, Couso, & Destercke, 2016), which is achieved as follows. First, a class is selected. Crisp rules are trained to discriminate this class from the others, a decomposition that prevents systematic bias in favour of one class. A crisp rule is defined using an antecedent (numerical attribute) A_i on the crisp interval I as follows: if $A_i \leq v$ then $I = (-\infty, v)$, if $A_i \geq u$ then $I = (u, \infty)$, and $I = [u, v]$ otherwise. Antecedents are added gradually, using a modification of RIPPER that considers information gain to be the objective function. The objective of the fuzzification process is to obtain the maximum support bound; thus, the antecedent with the greatest rule purity is selected. Crisp intervals obtained by RIPPER are replaced by fuzzy intervals $I^F(v) = (a, b, c, d)$ represented by trapezoidal membership functions:

$$I^F(v) = \begin{cases} 1, b \leq x \leq c \\ \dfrac{v-a}{b-a}, a \leq x \leq b \\ \dfrac{d-v}{d-c}, c \leq x \leq d \\ 0, otherwise \end{cases} \tag{2}$$

where b and c are the lower and upper bounds of the core of the fuzzy set, respectively, and a and d denote the lower and upper bounds of the support of the fuzzy set, respectively. Note that the fuzzy sets obtained by FURIA correspond to fuzzy half-intervals open to one side, i.e., $a = b = -\infty$, or $c = d = \infty$. In other words, the number of fuzzy sets is fixed to two in FURIA.

Rules are then pruned is performed to minimize the descriptive length of the rule base. Redundant antecedents are removed, and the rules are fuzzified by maximizing the purity (pur) of each rule:

$$pur = \frac{\sum_{x \in D_+^i} \mu_{Ai}(x)}{\sum_{x \in D_+^i} \mu_{Ai}(x) + \sum_{x \in D_-^i} \mu_{Ai}(x)} \tag{3}$$

where D_+^i and D_-^i are training data covered by the rule antecedent A_i, and $\mu^{Aj}(x)$ is the membership degree of instance × for antecedent A_i. A fuzzy rule R_j with n antecedents $A_i \in I_i^F$, $i = 1, 2, \ldots, n$, covers an instance × to the degree calculated as follows:

$$\mu_{Rj}(x) = \prod_{i=1}^{n} \mu_{Ai}(x_i) \tag{4}$$

Then, the grade of certainty can be calculated for the j-th rule R_j and class c_j as follows:

$$CF(cj) = \frac{2\frac{|D(cj)|}{|D|} + \sum_{x \in D(cj)} \mu^{Rj}(x)}{2 + \sum_{x \in D} \mu^{Rj}(x)} \tag{5}$$

where $D(cj)$ is are training instances in class c_j. Rule stretching is applied only if there are uncovered instances.

Inspired by Asadi (Asadi, 2019), in the last step, we propose to use evolutionary tuning in order to optimize the values of the parameters a_i, b_i, c_i, and d_i used to construct the shape and position of the trapezoidal membership functions. Among evolutionary algorithms, genetic algorithms (GAs) have been particularly successful in tuning fuzzy rule-based systems, giving rise to genetic fuzzy systems (Fernandez, Herrera, Cordon, Jose Del Jesus, & Marcelloni, 2019). GA is a robust global search technique capable of exploring complex search spaces while requiring only a measure of performance (Asadi, 2019). Genetic coding also allows them to incorporate prior knowledge. Therefore, we used a GA to tune the parameters of trapezoidal membership functions as follows:

1. A real-coded GA was used with the chromosome structure illustrated in Fig. 1. Two membership functions are generated for each input variable using the FURIA algorithm. To allow the GA to reach the optimum more easily, the values of parameters c_i and d_i were set so that the upper bounds of the core and support of the fuzzy set are equal to the corresponding values of the lower bounds of the support and core, respectively, i.e. $a_i = c_i$ and $b_i = d_i$. In step 1, possible bounds of a crisp interval are examined, and parameters $b_i = d_i$ (and $a_i = c_i$) are chosen such that the interval's purity pur is maximized. In step 2, parameters $a_i = c_i$ and $b_i = d_i$ are coded in a chromosome and optimized using GA. This is, a tuning of the two definition points of the trapezoidal membership functions is performed.
2. As indicated above, the initial population (chromosome) is represented by the parameters of the membership functions $a_1, b_1, \ldots, a_i, b_i, \ldots, a_n, b_n$, where n is the number of input variables, as generated by the FURIA algorithm.
3. For chromosome evaluation, we adopted the fitness function recommended for fuzzy rule-based classifier systems dealing with class-imbalance problems (Sanz, Bernardo, Herrera, Bustince, & Hagras, 2015): Fitness = (TPR + TNR)/2.
4. Following (Asadi, 2019), a standard GA was used with crossover, mutation and elitism genetic operators. To select the chromosomes, binary tournament selection was employed. In agreement with (Asadi, 2019), the settings of GA was as follows: population size = 100, crossover rate = 0.9, mutation rate = 0.1, and the 10% best individuals (elitism set) were copied to the next generation

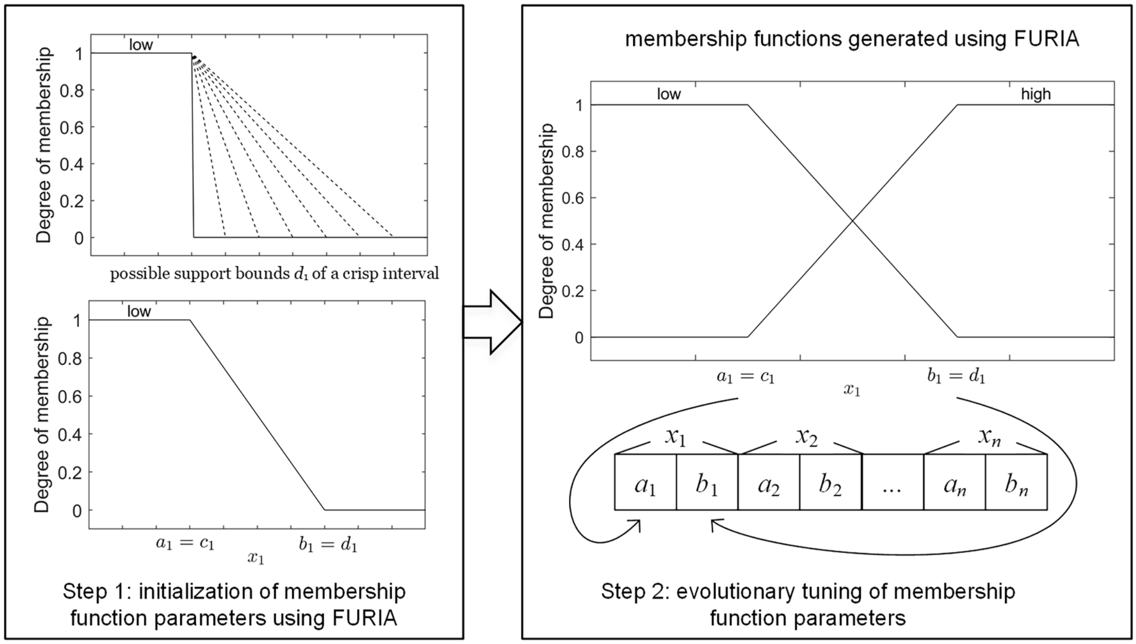

Fig. 1. Evolutionary tuning of membership functions generated by FURIA.

Fig. 2. COMEX Gold futures, 2007–2017.

Table 2
Input and output variables used in this study.

Category	Variable
Gold price technical indicators	20-day SMA, EMA, RSI, STDEV and ROC
Other commodities' prices	Oil, silver and copper price
Exchange rate and stock market	Exchange rate US Dollar/Chinese Yuan, DJIA index
News sentiment	positive/negative indicators from Bing Liu's and OpinionFinder dictionaries
Intensity of news sentiment	positive/negative indicators from FINN, S140, SentiWordNet and NRC Hashtag dictionaries
Other news emotions	Indicators of anger, anticipation, disgust, fear, joy, sadness, surprise and trust
Target class	Positive/negative movement of gold price

Table 3
Descriptive statistics on input and output financial variables.

Input variable	Calculation/description	Domain of variable	Mean ± St. Dev.
20-day SMA	$SMA_t = \sum_{k=1}^{20} P_{t-k}$	[609.3, 1883.2]	1214.1 ± 287.1
20-day EMA	$EMA_t = \frac{2}{21}(SMA_t - EMA_{t-1}) + EMA_{t-1}$	[612.7, 1862.6]	1213.8 ± 287.2
20-day RSI	$RSI_t = 100 - \frac{100}{1+RS}$	[0, 100]	43.2 ± 32.1
20-day STDEV	$STDEV_t = \sqrt{\frac{\sum_{k=1}^{20}(P_{t-k} - \overline{P})^2}{4}}$	[0, 81.2]	9.42 ± 7.30
20-day ROC	$ROC_t = \frac{(P_t - P_{t-20})}{P_{t-20}} \times 100$	[−10.9, 15.0]	0.09 ± 2.00
BRN	Brent Crude Oil price	[27.8, 146.1]	82.0 ± 26.8
SI	Silver price	[8.8, 48.6]	20.26 ± 7.34
CU	Copper price	[2770.0, 10205.2]	6773.8 ± 1461.0
USD/CNY	Exchange rate US Dollar/ Chinese Yuan	[6.04, 7.81]	6.62 ± 0.42
DJIA	Dow Jones Industrial Average index	[6547.1, 24838.4]	14236 ± 3589
Target variable	Calculation/description	Number of samples per class (1/0)	
Class $t+1$	1 if $(P_{t+1} - P_t) > 0$, 0 otherwise	1420/2529 (IR = 1.78)	
Class $t+5$	1 if $(P_{t+5} - P_t) > 0$, 0 otherwise	2111/1838 (IR = 1.15)	

Legend: IR – imbalance ratio, RS – ratio of smoothed average of 20-day upward ROC and 20-day downward ROC.

unaltered. The maximum number of 2,000 generations was used as a stopping criterion.

4. Data

For our study, we collected gold prices for the period from 2007 to 2017 (Fig. 2). More precisely, we downloaded daily prices of COMEX Gold futures from the MarketWatch database[1]. The COMEX Gold price is the world's major benchmark for gold.

As discussed above, earlier research has considered various input attributes as the determinants of fluctuations in gold prices. Taken together, this research suggests that previous gold price, oil price, prices of other metals, stock market indexes and exchange rates are the most important factors of gold price. Following this line of research, this study examined several categories of input variables (Table 2).

For the first category, previous gold price was included as an input variable in the form of technical indicators. Indeed, previous studies have found empirical evidence that abnormal returns can be obtained in gold futures using technical trading strategies (Narayan, Ahmed, & Narayan, 2015). Following related work (Ergen & Rizvanoghlu, 2016; Shambora & Rossiter, 2007), we calculated 20-day technical indicators as reported in Table 3, including SMA (Simple Moving Average), EMA (Exponential Moving Average), RSI (Relative Strength Index), STDEV (Standard Deviation) and ROC (Rate of Change). Thus, both trend-type (SMA and EMA) and oscillator-type (RSI) indicators were included, as well as measures of gold price volatility (STDEV) and percentage change (ROC). Oil (Brent Crude Oil price, BRN), silver (COMEX Silver futures, SI) and copper (London Metal Exchange copper price, CU) prices were incorporated to account for the effect of other commodities' prices. The remaining categories of input variables were exchange rate (US Dollar to Chinese Yuan) and Dow Jones Industrial Average (DJIA) index to include the effect of a major exchange rate and stock market. The previous day's closing prices of these input variables were collected from the freely available MarketWatch database.

The output variables were determined as upward/downward (1/0) classes from future gold closing prices, as presented in Table 3. To test the robustness of the proposed prediction system, both one- and five-days-ahead prediction horizons were assumed. Note that the one- and five-days-ahead data sets were imbalanced in favour of downward and upward movement, respectively, as indicated by the values of imbalance ratio IR = 1.78 and IR = 1.15 for the one- and five-days-ahead data, respectively. To provide more insight into the class imbalance problem, Fig. 3 shows the distributions of data samples over the years 2007–2017 for the two data sets.

As discussed above, recent studies have shown that news events have significant information effects on gold prices. In line with earlier research (Maslyuk-Escobedo et al., 2017; Shen et al., 2017; Smales, 2014), we utilized positive and negative news sentiment. In addition to this traditional polarity measure, we evaluated the intensity level, and alternative emotion indicators, such as trust or surprise, were incorporated into the proposed prediction model (because these have been identified as important indicators of online news helpfulness (Rao, Lei, Wenyin, Li, & Chen, 2014)).

The polarity measures were calculated as the sum of positive and negative words matching predefined dictionaries. In this respect, we used two popular dictionaries, Bing Liu's opinion dictionary and the OpinionFinder dictionary (Bravo-Marquez, Mendoza, & Poblete, 2014). These dictionaries are considered reliable in the literature because multiple human judgments are involved in their development (Bravo-Marquez, Frank, Mohammad, & Pfahringer, 2017). To evaluate the intensity of news sentiment, we used four different dictionaries AFINN (Nielsen, 2011), S140 (Kiritchenko, Zhu, & Mohammad, 2014), SentiWordNet (Baccianella, Esuli, & Sebastiani, 2010) and NRC Hashtag (Kiritchenko et al., 2014). High lexical coverage is achieved through their combination due to their uniqueness (Bravo-Marquez et al., 2014). Positive and negative scores calculated using these approaches are based on weights assigned by machine learning methods. The following emotional categories were included based on the Plutchik wheel of emotions (Mohammad & Turney, 2013): anger, anticipation, disgust, fear, joy, sadness, surprise and trust. As with the polarity measures, emotional indicators were calculated based on the numbers of words matching the respective categories in this emotional dictionary.

The news corpus was downloaded from the Thomson Reuters news service for the respective period (2007–2017). Only news items related to precious metals were retained in the sample (266,165 news stories)[2]. The date identifier of each news story enabled calculation of mean values of sentiment indicators for each day. For this calculation, we employed the AffectiveTweets package in the Weka 3.8 environment (Bravo-Marquez, Frank, Pfahringer, & Mohammad, 2019). Table 4 shows the mean values for the news sample, indicating that positive

[1] freely available for download at https://www.marketwatch.com/investing/future/gold

[2] The used news headlines are available at https://www.reuters.com/news/archive/rbssPreciousMetalsMinerals

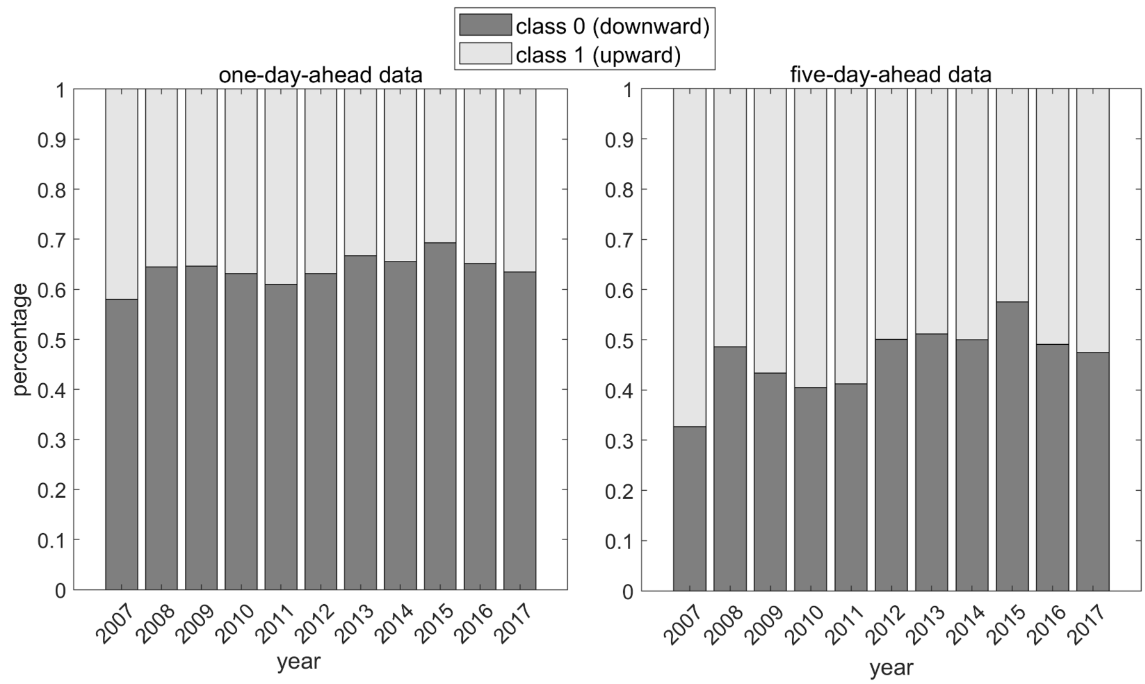

Fig. 3. Distribution of data samples over the years 2007–2017.

Table 4
Descriptive statistics for the input variables from news affect.

Variable	mean	[min, max]	Variable	mean	[min,max]
Bing Liu – pos	47.79	[0,398]	NRC hashtag – pos	2.29	[0,6.0]
Bing Liu – neg	16.68	[0,158]	NRC hashtag – neg	−2.33	[−11.4,0]
OpinionFinder – pos	54.67	[0,337]	Anger	0.28	[0.01,2.08]
OpinionFinder – neg	16.83	[0,163]	Anticipation	0.34	[0.02,0.91]
AFINN – pos	1.50	[0,9]	Disgust	0.12	[0.00,1.16]
AFINN – neg	−0.49	[−6,0]	Fear	0.31	[0.01,2.49]
S140 – pos	1.17	[0.0,4.8]	Joy	0.17	[0.00,0.70]
S140 – neg	−1.19	[−9.4,0]	Sadness	0.17	[0.01,2.02]
SentiWordNet – pos	0.23	[0,1.83]	Surprise	0.13	[0.01,0.49]
SentiWordNet – neg	−0.22	[−1.46,0]	Trust	0.44	[0.02,1.34]

polarity prevailed; trust was the most frequent emotion present in the news corpus. Finally, the obtained data on news affect (20 input variables) were merged with the financial data (10 input variables from Table 3).

5. Experiments

The below-described experiments evaluated the effectiveness of the proposed prediction system. Notably, we extensive compared this system with state-of-the-art methods for predicting gold prices.

5.1. Experimental setting

To evaluate the effectiveness of the proposed system, we used four measures of classification performance Accuracy (Acc), Area under the receiver operating characteristic curve (AUC), sensitivity (true positive rate, TPR) and specificity (true negative rate, TNR). Accuracy is calculated as the percentage of correctly predicted movements in gold price:

$$Acc = \frac{TP + TN}{P + N} \tag{6}$$

where TP and TN are the numbers of correctly predicted positive and negative movements in price, respectively, and P and N denote the numbers of positive and negative movements in the data. As noted above, the two target classes are represented with different frequencies in the data set. Therefore, we used AUC as our primary classification measure, since this measure is robust against class imbalance (Fawcett, 2006). AUC is equivalent to the probability that the prediction system ranks a randomly chosen instance with positive movement higher than a randomly chosen instance with negative movement. To illustrate the classification performance for each of the two target classes, sensitivity (TPR = TP/P) and specificity (TNR = TN/N) values are also reported.

In addition to the measures of classification performance, we also evaluated the interpretability of the fuzzy rule-based system using two complexity measures at the rule base level: the average number of conditions in a rule and the number of rules.

The experiments used data from 2007 to 2014 for training the prediction system and data from 2015 to 2017 for testing the system. Hereinafter, we report the performance measures on the testing (out-of-sample) data. To test the robustness of the results, we examined the performance of the proposed model over two prediction horizons (one- and five-days-ahead), resulting in two data sets.

To determine the learning parameters of FURIA, we experimented with a range of values for: (1) the amount of data used for pruning (#folds = {2, 3, 5}); (2) t-norms for the fuzzy AND operator in Eq. (1) (product t-norm/min t-norm); (3) strategy for handling uncovered instances (rule stretching, voting for the most frequent class or abstain from the decision); and (4) the number of optimization runs from {5, 10, 20, 30}. To demonstrate the superiority of fuzzy rule-based systems, we also compared the results with the original crisp RIPPER algorithm. Experiments were performed in the Keel 3.0 environment. The implementation of the GA tuning method in the Global Optimization Toolbox R2020b was used to further optimize the data base (parameters of membership functions) of the fuzzy rule-based system generated using FURIA.

To demonstrate the effectiveness of the proposed prediction system, we compared its performance with nine methods previous studies used for predicting the prices of precious metals:

- The traditional Random Walk (RW) with drift (Shafiee & Topal, 2010) as a baseline (performed in MATLAB R2013b Econometrics Toolbox).
- Random Forest (RF) (Liu & Li, 2017), a representative of decision trees, was trained with 100 generated trees and $\log_2$(#predictors) + 1 candidate features randomly sampled at each split.
- Adaptive Neuro-Fuzzy Inference System (ANFIS) (Yazdani-Chamzini et al., 2012), initialized using the subtractive clustering algorithm with cluster radius from the range {0.1, 0.2, …, 0.9}. We used the implementation of the ANFIS classifier in the MATLAB 2013b Fuzzy Logic Toolbox.
- Multi-Objective Evolutionary Fuzzy Classifier (MOEFS) (Jiménez, Sánchez, & Juárez, 2014) was chosen as a state-of-the-art, evolutionary fuzzy rule-based system. The classifier constructs its rule base using the ENORA evolutionary algorithm with two objectives: maximizing accuracy and minimizing the number of fuzzy rules. The maximum number of membership functions was set to five (in order to retain good interpretability), the population size was 100 and the number of generations was 20.
- A Multilayer Perceptron Neural Network (MLPNN) (Sánchez Lasheras, de Cos Juez, Suárez Sánchez, Krzemień, & Riesgo Fernández, 2015; Yazdani-Chamzini et al., 2012) was trained using a mini-batch gradient descent algorithm with one hidden layer of #predictors–1 sigmoidal units. Learning rate was set to 0.1 over 1,000 iterations.
- The ARIMA (p, d, q) (Guha & Bandyopadhyay, 2016; Sánchez Lasheras et al., 2015; Sharma, 2016) model was used with different numbers of autoregressive terms p = {0, 1, …, 3}, non-seasonal differences d = {0, 1, …, 3} and lagged forecast errors q = {0, 1, …, 3}.
- GARCH + MLPNN (Kristjanpoller & Hernández, 2017) was trained in two steps. First, a GARCH (p, q) model was estimated by testing p = {0, 1, 2} and q = {0, 1, 2}, where p and q are the orders of the GARCH and ARCH terms, respectively. Second, an MLPNN model (trained as noted above) was fed with both the input variables used for the GARCH model and the GARCH model's predictions.
- An Extreme Learning Machine (ELM) (Sivalingam, Mahedran, & Sivanadam, 2016) model was trained using one hidden layer with rectified linear units in Python. We tested different numbers of hidden units from the range {10, 20, …, 100}.
- Long short-term memory (LSTM) neural networks (Fischer & Krauss, 2018) are state-of-the-art deep, neural network models for predicting financial markets. Following (Fischer & Krauss, 2018), a unidirectional LSTM with one hidden LSTM layer of 25 rectified linear units was trained with the Deeplearning4j library using stochastic gradient descent, with Adam optimizer and dropout.
- Evolutionary fuzzification of RIPPER for regression (EFRiR) (Asadi, 2019) extends RIPPER for regression problems. In agreement with (Asadi, 2019), the target variable (COMEX gold price) was discretized using the equal width method (the lowest mean absolute percentage error was obtained for k = 5 intervals over the tested k= {3, 4, 5}). Then, RIPPER was employed to form the initial Mamdani-type fuzzy rule-based system. Finally, the data base of the system (i.e., the parameters of the triangular membership functions) was tuned using a GA (population size = 100, crossover rate = 0.9, mutation rate = 0.1, and 2000 iterations were used following (Asadi, 2019)) in MATLAB R2020b Fuzzy Logic Toolbox and Global Optimization Toolbox. The mean directional accuracy was used to assess the EFRiR's capacity to forecast the correct upward or downward price movement.

The framework of the research methodology described above is depicted in Fig. 4.

5.2. Experimental results

The first set of experiments examined how learning parameters affect FURIA prediction performance. We first defined baseline settings, following recommendations from the literature (Hühn & Hüllermeier, 2009) and setting the number of folds for pruning to 3. We also used the product t-norm, rule stretching for uncovered instances and 20 optimization runs. Note that rule stretching has been previously identified as a preferable strategy (Hühn & Hüllermeier, 2009), outperforming the default rule (choosing the most frequent class). Rule stretching leads to

Table 5
The effect of FURIA learning parameters on prediction performance.

	one-day-ahead		five-days-ahead	
Method (setting)	Acc [%]	AUC	Acc [%]	AUC
FURIA (5 optimization runs)	57.74	0.564	52.74	0.578
FURIA (10 optimization runs)	57.81	0.570	58.88	0.581
FURIA (20 optimization runs)	57.81	0.570	**61.42**	**0.608**
FURIA (30 optimization runs)	57.81	0.569	57.36	0.558
FURIA (2 folds)	**64.15**	**0.572**	60.91	0.605
FURIA (3 folds)	57.81	0.570	**61.42**	**0.608**
FURIA (5 folds)	59.01	0.556	61.42	0.605
FURIA (min t-norm)	57.81	0.570	**61.42**	**0.608**
FURIA (voting for the most frequent class)	59.99	0.570	**61.42**	**0.608**
FURIA (abstain from the decision)	53.55	0.566	51.90	0.546
RIPPER (3 folds, 20 optimization runs)	61.03	0.531	54.82	0.564

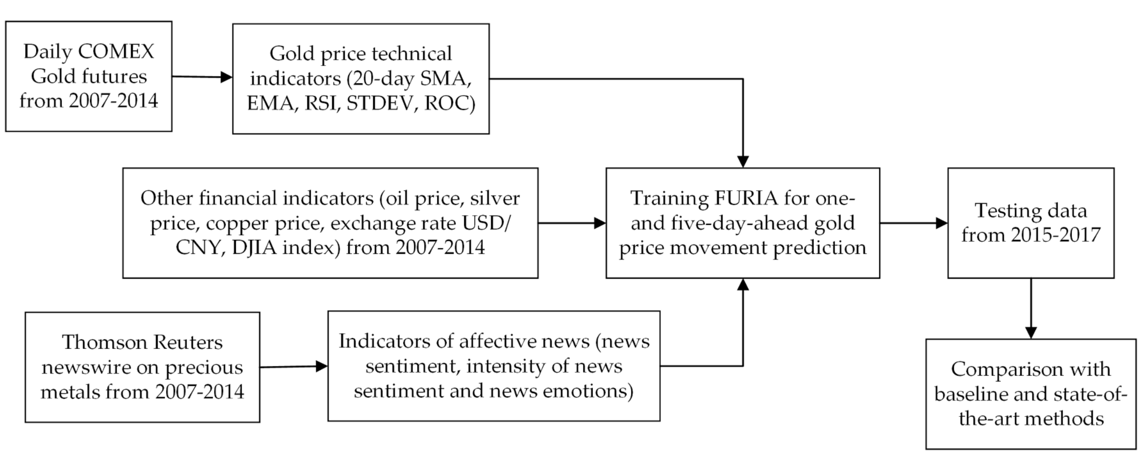

Fig. 4. Framework of research methodology.

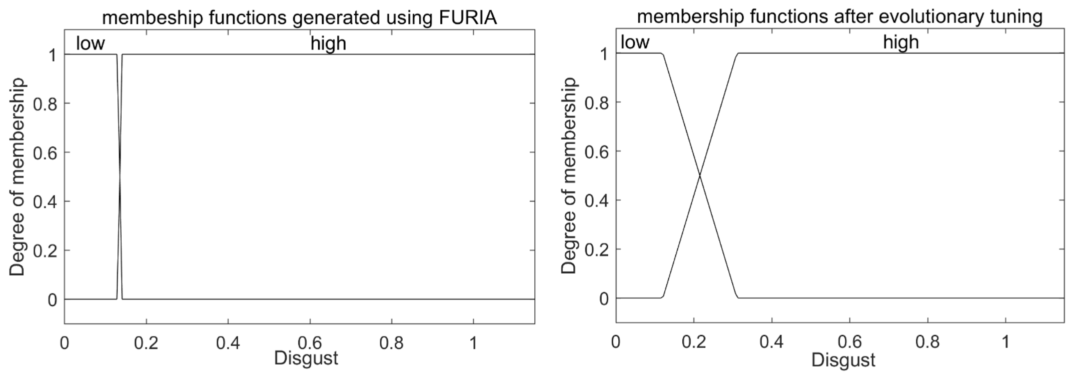

Fig. 5. Trapezoidal membership functions for input variable Disgust.

rule generalizations by deleting rule antecedents which, in turn, increases rule interpretability. Table 5 shows that the FURIA's best performance was obtained with two folds for the one-day-ahead prediction, while three folds were more effective for the five-days-ahead prediction. Changing the number of optimization runs did not improve performance in terms of Acc and AUC; in fact, the results in Table 5 suggest that a lower number of optimization runs was insufficient. Similarly, different strategies for uncovered instances and *t*-norm were ineffective. Importantly, FURIA outperformed its crisp counterpart RIPPER over both prediction horizons.

The rule base of the best FURIA model after evolutionary tuning (FURIA + ET) for the one-day-ahead prediction horizon was as follows:

$R_1^{(\text{one-day})}$: if NRC hashtag – pos is low [-inf, -inf, 2.68, 2.78] and 20-day STDEV is high [8.13, 8.81, inf, inf] then class $c_1 = 0$ with $CF_1 = 0.76$,
$R_2^{(\text{one-day})}$: if 20-day RSI is high [24.1, 24.7, inf, inf] then class $c_2 = 0$ with $CF_2 = 0.68$,
$R_3^{(\text{one-day})}$: if 20-day RSI is low [-inf, -inf, 24.1, 24.7] and S140 – pos is high [1.427, 1.789, inf, inf] and SentiWordNet – neg is low [-inf, -inf, −1.20, −0.68] and Disgust is low [-inf, -inf, 0.12, 0.31] then class $c_3 = 1$ with $CF_3 = 0.65$,
$R_4^{(\text{one-day})}$: if NRC hashtag – pos is high [2.68, 2.78, inf, inf] and NRC hashtag – neg is low [-2.33, −2.26, inf, inf] and AFINN – neg is low [-inf, -inf, −0.42, −0.19] then class $c_4 = 1$ with $CF_4 = 0.55$,

where the parameters of trapezoidal membership functions [-inf, -inf, c_i, d_i] and [a_i, b_i, inf, inf] correspond to low and high linguistic values, respectively. It is worth noting that compared to their FURIA counterparts, the tuned trapezoidal membership functions tended to have larger boundaries (i.e., the core was reduced and the support was enlarged), as presented in Fig. 5.

These rules suggest the polarity and intensity of news affect can predict one-day-ahead movements in gold price. As expected, positive sentiment (or a low level of negative sentiment) indicates positive movement in price, whereas a low level of positive sentiment (NRC hashtag – pos) together with high gold price volatility indicates negative movement. The RSI technical indicator was another important trading rule. Interestingly, the effect of other markets was suppressed in the fuzzy rule generation process, perhaps partly because news stories include information on previous developments in other precious metal markets. For the five-days-ahead prediction model, FURIA + ET produced the following rule base:

$R_1^{(\text{five-day})}$: if 20-day EMA is high [1230.2, 1268.7, inf, inf] and SI is low [-inf, -inf, 16.98, 20.31] then class $c_1 = 0$ with $CF_1 = 0.62$,
$R_2^{(\text{five-day})}$: if USD/CNY is low [-inf, -inf, 6.27, 6.48] and 20-day EMA is high [1230.2, 1268.7, inf, inf] then class $c_2 = 0$ with $CF_2 = 0.71$,
$R_3^{(\text{five-day})}$: if USD/CNY is low [-inf, -inf, 6.27, 6.48] and CU is low [-inf, -inf, 6017.2, 6058.3] then class $c_3 = 0$ with $CF_3 = 0.61$,
$R_4^{(\text{five-day})}$: if SI is high [16.98, 20.31, inf, inf] and USD/CNY is low [-inf, -inf, 6.27, 6.48] and 20-day SMA is high [1332.8, 1421.7, inf, inf] then class $c_4 = 0$ with $CF_4 = 0.61$,
$R_5^{(\text{five-day})}$: if 20-day EMA is high [1230.2, 1268.7, inf, inf] and NRC hashtag – neg is high [-2.78, −1.87, inf, inf] and 20-day STDEV is high [9.17, 11.28, inf, inf] then class $c_5 = 0$ with $CF_5 = 0.56$,
$R_6^{(\text{five-day})}$: if USD/CNY is high [6.27, 6.48, inf, inf] and 20-day EMA is low [-inf, -inf, 1268.7, 1862.6] then class $c_6 = 1$ with $CF_6 = 0.57$.

This rule base suggests that other precious metals (SI and CU) have a more pronounced effect in the five-days-ahead gold price prediction. More precisely, low values for these prices indicate a decrease in gold price. Low values of the exchange rate USD/CNY and 20-day EMA were other indicators of downward movement in gold price, while high values predicted upward movement. The effect of news affect was lower for the five-days-ahead prediction compared with the one-day-ahead

Table 6
Comparison of FURIA + ET with existing methods.

	one-day-ahead				five-days-ahead			
Method	Acc [%]	AUC	TPR	TNR	Acc [%]	AUC	TPR	TNR
Naïve baseline	64.04	0.500	0.000	1.000	53.46	0.500	1.000	0.000
RW	63.71	0.498	0.203	0.875	48.73	0.505	0.761	0.276
RF	62.74	0.541	0.233	0.831	53.81	0.537	0.642	0.451
ANFIS	62.01	0.563	0.448	0.782	54.31	0.579	0.842	0.255
MOEFS	65.12	0.510	0.325	**0.989**	54.31	0.558	**0.989**	0.127
MLPNN	61.03	0.532	0.309	0.774	48.22	0.505	0.821	0.167
ARIMA	58.54	0.549	0.405	0.684	49.75	0.536	0.983	0.039
GARCH + MLPNN	58.54	0.549	0.405	0.684	49.24	0.545	**0.989**	0.029
ELM	**65.99**	0.560	0.101	0.961	49.24	0.527	0.968	0.049
LSTM	62.12	**0.578**	0.364	0.761	50.25	0.539	0.884	0.147
EFRiR	55.06	0.514	0.272	0.700	49.62	0.502	0.289	0.707
FURIA	64.15	0.572	**0.506**	0.715	61.42	0.608	0.432	0.784
FURIA + ET	65.44	0.574	**0.506**	0.719	**62.15**	**0.612**	0.442	**0.808**

prediction, suggesting that investors process and respond to news on the day of announcement, with the effect mitigating after several days. This findings corroborates the results of previous studies (Borovkova & Mahakena, 2015; Dzielinski, 2012; Shen et al., 2017) that also implied news affect has a strong information effect on gold price following the day of announcement.

Regarding the interpretability of the FURIA model, both the number of rules (4 and 6) and the average number of conditions in the rules (2.50 and 2.33) indicate both one- and five-days-ahead predictions offer highly interpretable.

To evaluate the effectiveness of the proposed prediction models, their performance was compared with the nine benchmark methods. The results (Table 6) show that FURIA + ET dominated the compared methods in terms of prediction performance for the five-days-ahead prediction horizon. On the one-day-ahead horizon, FURIA + ET did not perform the best, but was highly competitive and comparable with state-of-the-art ELM and LSTM methods. Note that, in contrast to those methods, FURIA + ET provides better interpretability. In addition, the proposed evolutionary tuning improved the prediction performance of FURIA. Regarding the AUC metric for the one-day-ahead horizon, the LSTM model showed the best performance with AUC = 0.578 (note that the AUC value of 0.5 means that the predictions are no better than random), which is consistent with the results obtained for gold price time series in recent literature (Livieris et al., 2020). FURIA + EP not only performed best for the five-day-ahead horizon, but it also exhibited the most stable performance in terms of the trade-off between TPR and TNR measures. To demonstrate the improvement yielded using the compared methods over a naïve benchmark, the results are shown for the majority class voting method (i.e., 0-R classifier was used). This also explains why the compared methods achieved relatively low accuracies. In view of this, FURIA + EP yielded accuracy improvements of 1.40 % and 8.69 % compared to the naïve baseline for the one- and five-days-ahead predictions, respectively.

Compared with the alternative fuzzy rule-based systems, ANFIS, MOEFS and EFRiR, the proposed prediction system was not only more accurate but also performed comparatively well in terms of interpretability at the fuzzy partition and rule base levels. The number of membership functions is used to control for the granularity at the fuzzy partition level. As noted above, the number of membership functions is fixed to two in the FURIA-based systems. To show the effect of granularity, we therefore present the results of MOEFS, the fuzzy rule-based system that provided the second-best accuracy. The results in Fig. 6 suggest that a low number of membership functions was not sufficient to achieve satisfactory performance in terms of AUC, while the high number indicates system overfitting.

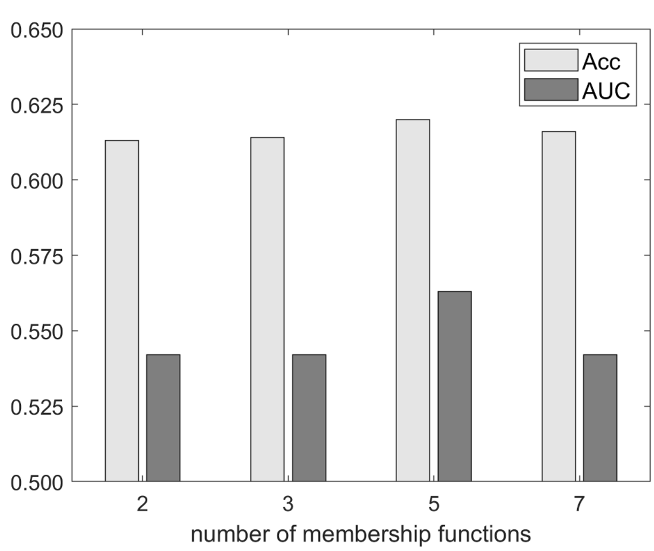

Fig. 6. Effect of the number of membership functions on MOEFS performance.

Table 7
The effect of amount of data used for rule growth on FURIA + ET performance.

	one-day-ahead			five-days-ahead		
# folds	AUC	# rules	# antecedents	AUC	# rules	# antecedents
2	0.574	4	2.50	0.608	5	1.60
3	0.570	10	3.90	0.612	6	2.33
5	0.566	7	4.43	0.605	3	1.33
9	0.508	5	3.20	0.553	8	2.38

Regarding the interpretability at the rule base level, for the two respective prediction horizons, ANFIS produced 5 and 6 rules, MOEFS generated 5 and 7, and EFRiR produced 40 and 61 rules). To examine the effect of rule base complexity on the prediction performance of FURIA, different numbers of folds were tested to determine the amount of data used for growing the rules. The results in Table 7 suggest that increasing the complexity of rule bases leads not only to overfitting but also to an increase in the number of antecedents in the rules. FURIA substantially reduces the input variables required in the rule conditions. Indeed, the embedded feature selection process of FURIA (as in the RIPPER algorithm) significantly enhanced interpretability at the rule base level. Note that neither ANFIS nor MOEFS provide this feature selection component. As a result, they produce more complex rules than does FURIA. It is worth noting that the performance of FURIA + ET and other classification models was also compared with several time-series regression models (RW, ARIMA, and EFRiR) in Table 6. To calculate the classification performance for these models, we adopted the approach of Livieris et al. (2020) so that the models were first trained to predict the gold price on the next day (regression) and, then, the predictions were transformed into the binary classification problem of predicting whether the gold price will increase or decrease on the next day with respect to today's price. Our results further suggest that the FURIA-based classification models outperform their level estimation (regression) counterpart (EFRiR) in terms of predicting the direction of the gold price movement, which is consistent with previous research in stock market forecasting (Enke & Thawornwong, 2005; Leung, Daouk, & Chen, 2000). Indeed, even though more refined trading strategies were developed for the EFRiR regression model (resulting in a larger number of rules) to incorporate the magnitude of the forecasts (Krauss, Do, & Huck, 2017), the classification models more accurate were not only in terms of movement prediction but also in terms of investment trading returns (see results below).

In a further set of experiments, we investigated the effect of news affect on the prediction performance of the compared methods (only those methods in which news affect was used as input attributes were included). The results in Fig. 7 indicate that FURIA and FURIA + ET were among the methods whose performance improved the most after the inclusion of news affect. In contrast, the performance of the regression-based models was not influenced by news affect, suggesting that news affect is crucial for trend prediction of the gold price movement.

In addition to the Acc and AUC measures, we evaluated the prediction performance of the compared methods using the average annual return (AR [%]) obtained by a prediction-based trading strategy. A 'buy' ('hold') signal was generated if upward price movement was predicted, while a downward price prediction indicated a 'sell' signal. We used the closing gold price for trading. The proposed prediction model outperformed the compared methods on both prediction horizons (Table 8). Our model performed about six times better than the baseline RW method. ANFIS, ARIMA + MLPNN, GARCH + MLPNN, ELM and LSTM also performed well in terms of AR, whereas RF, EFRiR and MLPNN performed poorly even when compared with RW. Note that the traditional buy-and-hold (B&H) strategy achieved an AR = 4.85 %.

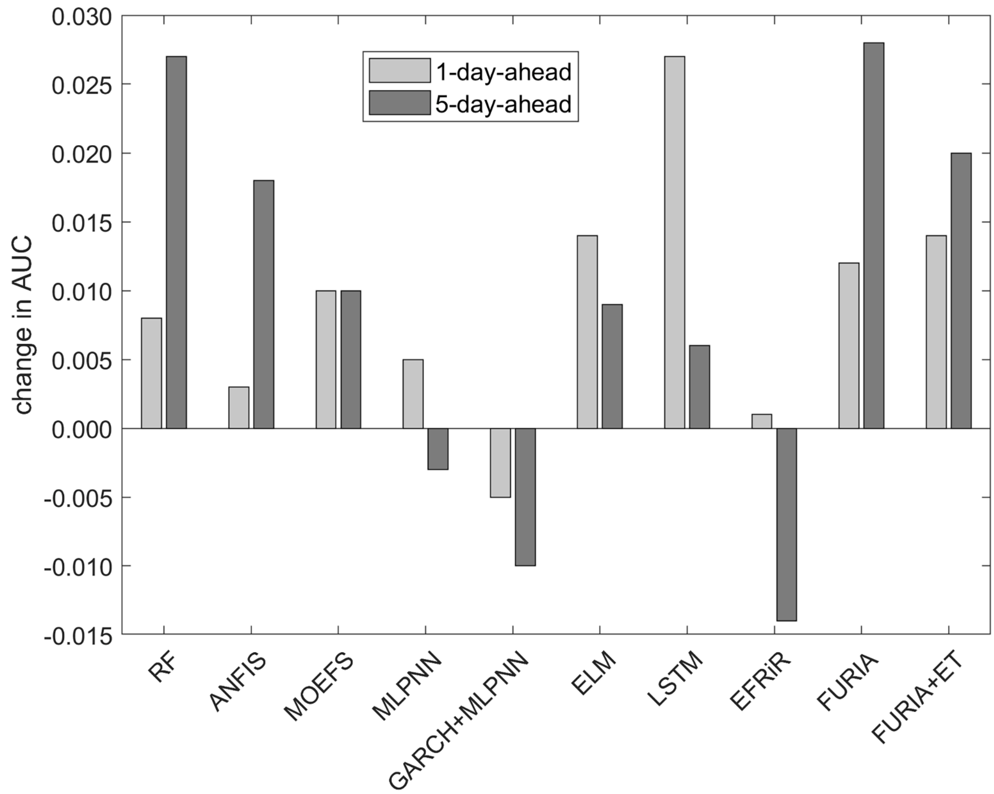

Fig. 7. Sensitivity of AUC to news affect (change in AUC in response to the inclusion of input variables from news affect).

Table 8
Average annual return obtained using the compared methods' predictions.

Method	AR [%] (one-day-ahead)	AR [%] (five-days-ahead)
RW	4.18	2.73
RF	3.26	−18.28
ANFIS	16.49	10.17
MOEFS	1.37	8.75
MLPNN	−2.71	−4.70
ARIMA	6.58	6.71
GARCH + MLPNN	5.21	12.06
ELM	11.29	5.16
LSTM	12.48	7.67
EFRiR	3.67	−6.05
FURIA	23.35	12.31
FURIA + ET	**24.02**	**13.93**

Altogether, a trading strategy based on the predictions of the proposed model would have been highly profitable.

6. Conclusion

This study developed an interpretable, fuzzy rule-based gold price prediction system that uses both historical financial data and linguistic indicators extracted from news stories. The findings suggest that news affect is important for predicting one-day-ahead movement in gold price, while indicators from other financial and precious metal markets are effective in predicting five-days-ahead price movements. These findings clearly support the relevance of news sentiment in the gold market. The most obvious finding that emerges from this study is that an artificial intelligence prediction system can be constructed to perform well in terms of both prediction accuracy and interpretability, two often contradictory objectives. One implication of this is the possibility that fuzzy rule-based trading systems could outperform existing approaches in terms of average annual return while offering investors an interpretable set of trading rules. In addition, investors can also easily modify components of the prediction system (rules and their conditions) in order to improve its accuracy and interpretability in certain contexts. The proposed FURIA + ET classification model can also be easily extended for regression problems by discretizing the target variable in an unsupervised manner, using the method proposed by Asadi (2019).

Several important limitations must also be considered. First, the examined data were limited to COMEX Gold futures. While a major market for trading precious metals, further experimental investigations will be needed to validate the proposed approach for other markets, such as LME. More broadly, research is also recommended to incorporate news sentiment into prediction systems for other precious metals. The proposed prediction system also ignored potentially useful expert knowledge. Indeed, a combination of trading rules obtained from investment experts with those extracted using the proposed system could additionally improve accuracy and interpretability. A related and important problem recommended for future study is the semantic interpretability of the rule bases. Finally, the proposed system utilized only sentiment and sentiment intensity from news stories. Further research should also target richer information hidden in the news, such as topic modelling or word embeddings.

CRediT authorship contribution statement

Petr Hajek: Conceptualization, Funding acquisition, Methodology, Software, Supervision, Visualization, Writing – original draft, Writing – review & editing. **Josef Novotny:** Conceptualization, Data curation, Formal analysis, Investigation, Resources, Validation, Writing – original draft.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

We gratefully acknowledge the help provided by constructive comments of the anonymous referees. This article was supported by the scientific research project of the Czech Sciences Foundation Grant No: 19-15498S.

References

Alameer, Z., Elaziz, M. A., Ewees, A. A., Ye, H., & Jianhua, Z. (2019). Forecasting gold price fluctuations using improved multilayer perceptron neural network and whale optimization algorithm. *Resources Policy, 61*, 250–260. https://doi.org/10.1016/j.resourpol.2019.02.014

Asadi, S. (2019). Evolutionary fuzzification of RIPPER for regression: Case study of stock prediction. *Neurocomputing, 331*, 121–137. https://doi.org/10.1016/J.NEUCOM.2018.11.052

Baccianella, S., Esuli, A., & Sebastiani, F. (2010). Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. *Proc. of the 7th Int. Conf. on Language Resources and Evaluation*.

Balduzzi, P., Elton, E. J., & Green, T. C. (2001). Economic news and bond prices: Evidence from the U.S. treasury market. *The Journal of Financial and Quantitative Analysis, 36*(4), 523–543. https://doi.org/10.2307/2676223

Baur, D. G., Beckmann, J., & Czudaj, R. (2016). A melting pot — Gold price forecasts under model and parameter uncertainty. *International Review of Financial Analysis, 48*, 282–291. https://doi.org/10.1016/j.irfa.2016.10.010

Bhatia, V., Das, D., Tiwari, A. K., Shahbaz, M., & Hasim, H. M. (2018). Do precious metal spot prices influence each other? Evidence from a nonparametric causality-in-quantiles approach. *Resources Policy, 55*, 244–252. https://doi.org/10.1016/j.resourpol.2017.12.008

Bildirici, M. E., & Turkmen, C. (2015). Nonlinear causality between oil and precious metals. *Resources Policy, 46*, 202–211. https://doi.org/10.1016/j.resourpol.2015.09.002

Borovkova, S., & Mahakena, D. (2015). News, volatility and jumps: The case of natural gas futures. *Quantitative Finance, 15*(7), 1217–1242. https://doi.org/10.1080/14697688.2014.986513

Bouri, E., Jain, A., Biswal, P. C., & Roubaud, D. (2017). Cointegration and nonlinear causality amongst gold, oil, and the Indian stock market: Evidence from implied volatility indices. *Resources Policy, 52*, 201–206. https://doi.org/10.1016/j.resourpol.2017.03.003

Brandt, M. W., & Gao, L. (2019). Macro fundamentals or geopolitical events? A textual analysis of news events for crude oil. *Journal of Empirical Finance, 51*, 64–94. https://doi.org/10.1016/j.jempfin.2019.01.007

Bravo-Marquez, F., Frank, E., Mohammad, S. M., & Pfahringer, B. (2017). Determining word-emotion associations from tweets by multi-label classification. In *Proceedings - 2016 IEEE/WIC/ACM International Conference on Web Intelligence*. https://doi.org/10.1109/WI.2016.0091

Bravo-Marquez, F., Frank, E., Pfahringer, B., & Mohammad, S. M. (2019). Affective tweets: A weka package for analyzing affect in tweets. *Journal of Machine Learning Research, 20*(92), 1–6.

Bravo-Marquez, F., Mendoza, M., & Poblete, B. (2014). Meta-level sentiment models for big social data analysis. *Knowledge-Based Systems, 69*, 86–99. https://doi.org/10.1016/j.knosys.2014.05.016

Chai, J., Zhao, C., Hu, Y., & Zhang, Z. G. (2021). Structural analysis and forecast of gold price returns. *Journal of Management Science and Engineering, 6*(2), 135–145. https://doi.org/10.1016/J.JMSE.2021.02.011

Chen, T., Shang, C., Su, P., & Shen, Q. (2018). Induction of accurate and interpretable fuzzy rules from preliminary crisp representation. *Knowledge-Based Systems, 146*, 152–166. https://doi.org/10.1016/j.knosys.2018.02.003

Chen, Y., He, K., & Zhang, C. (2016). A novel grey wave forecasting method for predicting metal prices. *Resources Policy, 49*, 323–331. https://doi.org/10.1016/j.resourpol.2016.06.012

Christie-David, R., Chaudhry, M., & Koch, T. W. (2000). Do macroeconomics news releases affect gold and silver prices? *Journal of Economics and Business, 52*, 405–421. https://doi.org/10.1016/s0148-6195(00)00029-1

Clements, A. E., & Todorova, N. (2016). Information flow, trading activity and commodity futures volatility. *Journal of Futures Markets, 36*(1), 88–104. https://doi.org/10.1002/fut.21724

Cornell, B., & French, K. R. (1986). Commodity own rates, real interest rates, and money supply announcements. *Journal of Monetary Economics, 18*, 3–23. https://doi.org/10.1016/0304-3932(86)90051-6

Creti, A., Joëts, M., & Mignon, V. (2013). On the links between stock and commodity markets' volatility. *Energy Economics, 37*, 16–28. https://doi.org/10.1016/j.eneco.2013.01.005

Dehghani, H., & Bogdanovic, D. (2018). Copper price estimation using bat algorithm. *Resources Policy, 55*, 55–61. https://doi.org/10.1016/j.resourpol.2017.10.015

Dzielinski, M. (2012). Which news resolves asymmetric information? *SSRN Electronic Journal, 800*, 1–36. https://doi.org/10.2139/ssrn.2146698

El Hedi Arouri, M., Lahiani, A., & Nguyen, D. K. (2015). World gold prices and stock returns in China: Insights for hedging and diversification strategies. *Economic Modelling, 44*, 273–282. https://doi.org/10.1016/j.econmod.2014.10.030

Enke, D., & Thawornwong, S. (2005). The use of data mining and neural networks for forecasting stock market returns. *Expert Systems with Applications, 29*(4), 927–940. https://doi.org/10.1016/J.ESWA.2005.06.024

Ergen, I., & Rizvanoghlu, I. (2016). Asymmetric impacts of fundamentals on the natural gas futures volatility: An augmented GARCH approach. *Energy Economics, 56*, 64–74. https://doi.org/10.1016/j.eneco.2016.02.022

Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business, 38*(1), 34–105. https://doi.org/10.1086/294743

Fang, L., Chen, B., Yu, H., & Qian, Y. (2018). The importance of global economic policy uncertainty in predicting gold futures market volatility: A GARCH-MIDAS approach. *Journal of Futures Markets, 38*(3), 413–422. https://doi.org/10.1002/fut.21897

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27*, 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

Fernandez, A., Herrera, F., Cordon, O., Jose Del Jesus, M., & Marcelloni, F. (2019). Evolutionary fuzzy systems for explainable artificial intelligence: Why, when, what for, and where to? *IEEE Computational Intelligence Magazine, 14*(1), 69–81. https://doi.org/10.1109/MCI.2018.2881645

Feuerriegel, S., Heitzmann, S. F., & Neumann, D. (2015). Do investors read too much into news? How news sentiment causes price formation. In *Proceedings of the Annual Hawaii International Conference on System Sciences* (pp. 4803–4812). https://doi.org/10.1109/HICSS.2015.571

Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research, 55*, 654–669. https://doi.org/10.1016/j.ejor.2017.11.054

Frankel, J. A., & Hardouvelis, G. A. (1985). Commodity prices, money surprises and Fed credibilit. *Journal of Money, Credit and Banking, 17*, 425–438. https://doi.org/10.2307/1992439

Gangopadhyay, K., Jangir, A., & Sensarma, R. (2016). Forecasting the price of gold: An error correction approach. *IIMB Management Review, 28*(1), 6–12. https://doi.org/10.1016/j.iimb.2015.11.001

García, D., & Kristjanpoller, W. (2019). An adaptive forecasting approach for copper price volatility through hybrid and non-hybrid models. *Applied Soft Computing Journal, 74*, 466–478. https://doi.org/10.1016/j.asoc.2018.10.007

Guha, B., & Bandyopadhyay, G. (2016). Gold price forecasting using ARIMA model. *Journal of Advanced Management Science, 4*(2), 117–121. https://doi.org/10.12720/joams.4.2.117-121

He, K., Chen, Y., & Tso, G. K. F. (2017). Price forecasting in the precious metal market: A multivariate EMD denoising approach. *Resources Policy, 54*, 9–24. https://doi.org/10.1016/j.resourpol.2017.08.006

Henry, E. (2008). Are investors influenced by how earnings press releases are written? *Journal of Business Communication, 45*(4), 363–407. https://doi.org/10.1177/0021943608319388

Herrera, F. (2008). Genetic fuzzy systems: Taxonomy, current research trends and prospects. *Evolutionary Intelligence, 1*(1), 27–46. https://doi.org/10.1007/s12065-007-0001-5

Hühn, J., & Hüllermeier, E. (2009). FURIA: An algorithm for unordered fuzzy rule induction. *Data Mining and Knowledge Discovery, 19*(3), 293–319. https://doi.org/10.1007/s10618-009-0131-8

Jabeur, S. B., Mefteh-Wali, S., & Viviani, J.-L. (2021). Forecasting gold price with the XGBoost algorithm and SHAP interaction values. *Annals of Operations Research, 2021*, 1–21. https://doi.org/10.1007/S10479-021-04187-W

Jain, A., & Biswal, P. C. (2016). Dynamic linkages among oil price, gold price, exchange rate, and stock market in India. *Resources Policy, 49*, 179–185. https://doi.org/10.1016/j.resourpol.2016.06.001

Jiménez, F., Sánchez, G., & Juárez, J. M. (2014). Multi-objective evolutionary algorithms for fuzzy classification in survival prediction. *Artificial Intelligence in Medicine, 60*(3), 197–219. https://doi.org/10.1016/j.artmed.2013.12.006

Kia, A. N., Haratizadeh, S., & Shouraki, S. B. (2018). A hybrid supervised semi-supervised graph-based model to predict one-day ahead movement of global stock markets and commodity prices. *Expert Systems with Applications, 105*, 159–173. https://doi.org/10.1016/j.eswa.2018.03.037

Kiritchenko, S., Zhu, X., & Mohammad, S. M. (2014). Sentiment analysis of short informal texts. *Journal of Artificial Intelligence Research, 50*, 723–762. https://doi.org/10.1613/jair.4272

Krauss, C., Do, X. A., & Huck, N. (2017). Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. *European Journal of Operational Research, 259*(2), 689–702. https://doi.org/10.1016/J.EJOR.2016.10.031

Kristjanpoller, W., & Hernández, E. (2017). Volatility of main metals forecasted by a hybrid ANN-GARCH model with regressors. *Expert Systems with Applications, 84*, 290–300. https://doi.org/10.1016/j.eswa.2017.05.024

Kristjanpoller, W., & Minutolo, M. C. (2015). Gold price volatility: A forecasting approach using the Artificial Neural Network-GARCH model. *Expert Systems with Applications, 42*, 7245–7251. https://doi.org/10.1016/j.eswa.2015.04.058

Leung, M. T., Daouk, H., & Chen, A. S. (2000). Forecasting stock indices: A comparison of classification and level estimation models. *International Journal of Forecasting, 16*(2), 173–190. https://doi.org/10.1016/S0169-2070(99)00048-5

Li, X., Li, D., Zhang, X., Wei, G., Bai, L., & Wei, Y. (2021). Forecasting regular and extreme gold price volatility: The roles of asymmetry, extreme event, and jump. *Journal of Forecasting*. https://doi.org/10.1002/FOR.2781

Liebmann, M., Hagenau, M., & Neumann, D. (2012). Information processing in electronic markets: Measuring subjective interpretation using sentiment analysis. In J. F. George (Ed.), *International Conference on Information Systems, ICIS 2012* (pp. 1–18). Association for Information Systems.

Liu, C., Hu, Z., Li, Y., & Liu, S. (2017). Forecasting copper prices by decision tree learning. *Resources Policy, 52*, 427–434. https://doi.org/10.1016/j.resourpol.2017.05.007

Liu, D., & Li, Z. (2017). Gold price forecasting and related influence factors analysis based on random forest. *Advances in Intelligent Systems and Computing, 711–723*. https://doi.org/10.1007/978-981-10-1837-4_59

Liu, Y., Yang, C., Huang, K., & Gui, W. (2019). Non-ferrous metals price forecasting based on variational mode decomposition and LSTM network. *Knowledge-Based Systems, 105006*. https://doi.org/10.1016/j.knosys.2019.105006

Livieris, I. E., Pintelas, E., & Pintelas, P. (2020). A CNN–LSTM model for gold price time-series forecasting. *Neural Computing and Applications, 32*(23), 17351–17360. https://doi.org/10.1007/S00521-020-04867-X

Long, W., Song, L., & Tian, Y. (2019). A new graphic kernel method of stock price trend prediction based on financial news semantic and structural similarity. *Expert Systems with Applications, 118*, 411–424. https://doi.org/10.1016/j.eswa.2018.10.008

Maslyuk-Escobedo, S., Rotaru, K., & Dokumentov, A. (2017). News sentiment and jumps in energy spot and futures markets. *Pacific Basin Finance Journal, 45*, 186–210. https://doi.org/10.1016/j.pacfin.2016.07.001

Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word-emotion association lexicon. *Computational Intelligence, 29*(3), 436–465. https://doi.org/10.1111/j.1467-8640.2012.00460.x

Mohtasham Khani, M., Vahidnia, S., & Abbasi, A. (2021). A deep learning-based method for forecasting gold price with respect to pandemics. *SN Computer Science, 2*(4), 1–12. https://doi.org/10.1007/S42979-021-00724-3

Narayan, P. K., Ahmed, H. A., & Narayan, S. (2015). Do momentum-based trading strategies work in the commodity futures markets? *Journal of Futures Markets, 35*(9), 868–891. https://doi.org/10.1002/fut.21685

Nielsen, F. Å. (2011). A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. *ESWC2011 Workshop on "Making Sense of Microposts": Big Things Come in Small Packages*, 93–98.

Ntim, C. G., English, J., Nwachukwu, J., & Wang, Y. (2015). On the efficiency of the global gold markets. *International Review of Financial Analysis, 41*, 218–236. https://doi.org/10.1016/j.irfa.2015.03.013

Palacios, A., Sánchez, L., Couso, I., & Destercke, S. (2016). An extension of the FURIA classification algorithm to low quality data through fuzzy rankings and its application to the early diagnosis of dyslexia. *Neurocomputing, 176*, 60–71. https://doi.org/10.1016/j.neucom.2014.11.088

Piñeiro-Chousa, J., López-Cabarcos, M.Á., Pérez-Pico, A. M., & Ribeiro-Navarrete, B. (2018). Does social network sentiment influence the relationship between the S&P 500 and gold returns? *International Review of Financial Analysis, 57*, 57–64. https://doi.org/10.1016/j.irfa.2018.02.005

Qian, Y., Ralescu, D. A., & Zhang, B. (2019). The analysis of factors affecting global gold price. *Resources Policy, 64*, Article 101478. https://doi.org/10.1016/j.resourpol.2019.101478

Rao, Y., Lei, J., Wenyin, L., Li, Q., & Chen, M. (2014). Building emotional dictionary for sentiment analysis of online news. *World Wide Web, 17*, 723–742. https://doi.org/10.1007/s11280-013-0221-9

Sadik, Z. A., Date, P. M., & Mitra, G. (2019). Forecasting crude oil futures prices using global macroeconomic news sentiment. *IMA Journal of Management Mathematics*. https://doi.org/10.1093/imaman/dpz011

Sadorsky, P. (2021). Predicting gold and silver price direction using tree-based classifiers. *Journal of Risk and Financial Management, 14*(5), 198. https://doi.org/10.3390/JRFM14050198

Sánchez Lasheras, F., de Cos Juez, F. J., Suárez Sánchez, A., Krzemień, A., & Riesgo Fernández, P. (2015). Forecasting the COMEX copper spot price by means of neural networks and ARIMA models. *Resources Policy, 45*, 37–43. https://doi.org/10.1016/j.resourpol.2015.03.004

Sanz, J. A., Bernardo, D., Herrera, F., Bustince, H., & Hagras, H. (2015). A compact evolutionary interval-valued fuzzy rule-based classification system for the modeling and prediction of real-world financial applications with imbalanced data. *IEEE Transactions on Fuzzy Systems, 23*(4), 973–990. https://doi.org/10.1109/TFUZZ.2014.2336263

Shafiee, S., & Topal, E. (2010). An overview of global gold market and gold price forecasting. *Resources Policy, 35*(3), 178–189. https://doi.org/10.1016/j.resourpol.2010.05.004

Shambora, W. E., & Rossiter, R. (2007). Are there exploitable inefficiencies in the futures market for oil? *Energy Economics, 29*(1), 18–27. https://doi.org/10.1016/j.eneco.2005.09.004

Sharma, R. K. (2016). Forecasting gold price with Box Jenkins autoregressive integrated moving average method. *Journal of International Economics, 7*(1), 32–61.

Shen, J., Najand, M., Dong, F., & He, W. (2017). News and social media emotions in the commodity market. *Review of Behavioral Finance, 9*(2), 148–168. https://doi.org/10.1108/RBF-09-2016-0060

Siering, M. (2013). Investigating the impact of media sentiment and investor attention on financial markets. In W. Aalst, J. Mylopoulos, M. Rosemann, M. J. Shaw, C. Szyperski, F. A. Rabhi, & P. Gomber (Eds.), *Lecture Notes in Business Information Processing* (pp. 3–19). Springer. 10.1007/978-3-642-36219-4_1.

Sivalingam, K. C., Mahedran, S., & Sivanadam, S. N. (2016). Forecasting gold prices based on extreme learning machine. *International Journal of Computers, Communications and Control*, *11*(3), 372–380. 10.15837/ijccc.2016.3.2009.

Smales, L. A. (2014). News sentiment in the gold futures market. *Journal of Banking and Finance, 45*, 275–286. https://doi.org/10.1016/j.jbankfin.2014.09.006

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance, 62*(3), 1139–1168. https://doi.org/10.1111/j.1540-6261.2007.01232.x

Wang, C., Zhang, X., Wang, M., Lim, M. K., & Ghadimi, P. (2019). Predictive analytics of the copper spot price by utilizing complex network and artificial neural network techniques. *Resources Policy, 63*, Article 101414. https://doi.org/10.1016/j.resourpol.2019.101414

Wen, F., Yang, X., Gong, X., & Lai, K. K. (2017). Multi-scale volatility feature analysis and prediction of gold price. *International Journal of Information Technology and Decision Making, 16*, 205–223. https://doi.org/10.1142/S0219622016500504

Weng, B., Ahmed, M. A., & Megahed, F. M. (2017). Stock market one-day ahead movement prediction using disparate data sources. *Expert Systems with Applications, 79*, 153–163. https://doi.org/10.1016/j.eswa.2017.02.041

Weng, F., Chen, Y., Wang, Z., Hou, M., Luo, J., & Tian, Z. (2020). Gold price forecasting research based on an improved online extreme learning machine algorithm. *Journal of Ambient Intelligence and Humanized Computing, 11*(10), 4101–4111. https://doi.org/10.1007/s12652-020-01682-z

Yazdani-Chamzini, A., Yakhchali, S. H., Volungevičiene, D., & Zavadskas, E. K. (2012). Forecasting gold price changes by using adaptive network fuzzy inference system. *Journal of Business Economics and Management, 13*(5), 994–1010. https://doi.org/10.3846/16111699.2012.683808

Zhang, P., & Ci, B. (2020). Deep belief network for gold price forecasting. *Resources Policy, 69*, Article 101806. https://doi.org/10.1016/J.RESOURPOL.2020.101806