Forecasting returns in international markets with fuzzy rule-based classification systems

Silvia Muzzioli [*] Giovanni Campisi[†] Luca Gambarelli[‡] Bernard De Baets[§]

Abstract

The paper aims to investigate the forecasting ability of fuzzy rule-based classification systems (FRBCS) on future direction of the S&P500 index. To this end, we apply four FRBCS methods. Moreover, we compare both the forecasting accuracy and the interpretability of the results of FRBCS with the recently used machine learning techniques. Overall, among the two approaches, we prefer the FRBCS methods, since they allow a good balance between accuracy and interpretability, and provide sharper results than the machine learning techniques.

Keywords: Fuzzy rule based systems Machine learning; Volatility indices; Market risk.

1 Introduction

Stock price prediction has always been a subject of interest for academics from many disciplines. Correctly assessing the probability of upside of downside future price variation of the stock market is a key challenge for most investors and professional analysts. Nevertheless, identifying the best time to buy or sell has remained a challenging task because several factors may influence stock prices (Chang and Liu (2008)). Many studies have dealt with the problem of input selection when it comes to mapping financial indices and stocks. In recent decades, the quantity and quality of available information to predict future market fluctuations have increased dramatically (Campisi *et al.* (2021)).

In particular, global exchange operators such as the Chicago Board Options Exchange (hereafter, CBOE) introduce several option-based indicators that provide crucial information about the perceived level of risk, such as the volatility and the asymmetry computed from the option-implied distribution. These indicators

[*]Marco Biagi Department of Economics, University of Modena and Reggio Emilia, Viale Jacopo Berengario 51, 41121, Modena, Italy, e-mail: silvia.muzzioli@unimore.it

[†]Marco Biagi Department of Economics, University of Modena and Reggio Emilia, Viale Jacopo Berengario 51, 41121, Modena, Italy, e-mail: giovanni.campisi@unimore.it

[‡]Marco Biagi Department of Economics, University of Modena and Reggio Emilia, Viale Jacopo Berengario 51, 41121, Modena, Italy, e-mail: luca.gambarelli@unimore.it

[§]Department of Data Analysis and Mathematical Modelling and KERMIT, Ghent University, Coupure Links 653 B-9000 Gent, Belgium, e-mail: Bernard.DeBaets@UGent.be

are of paramount importance for prediction tasks since the option market embeds information about the forward-looking distribution of asset returns (Barro and Liao (2019)). Moreover, existing studies (see, e.g., Xing *et al.* (2010)) suggest that option prices incorporate additional information content compared to the one embedded in stock prices and embeds critical information about political uncertainty and tail economic risk (see, e.g. Seo and Wachter (2019)).

The use of option-implied indicators as predictors of future market returns is investigated in several studies, including Rubbaniy *et al.* (2014), Gokmenoglu and Fazlollahi (2015), Elyasiani *et al.* (2017), Elyasiani *et al.* (2018), and Mora-Valencia *et al.* (2021). Even if these studies successfully identify relationships between the option-based indicators and market returns, they focus the analysis on using one or at most two indices, thus overlooking the possibility of combining all the available information to improve their return forecast. Moreover, most studies use traditional methods such as OLS regression to investigate the forecasting power of volatility and skewness indices on returns. On the other hand, a few authors adopt machine learning methods (Campisi *et al.* (2021)), that are more flexible than conventional econometric prediction approaches. In addition, recent studies show that there are more effective in predicting stock returns than traditional methods (see, e.g., Gu *et al.* (2020)). While conventional statistical and econometric methods aim to identify the relationships between the variables in the model, machine learning methods aim at maximising the accuracy of predictions (Athey and Imbens (2019)).

Recently, Campisi *et al.* (2021) combined machine learning and several option-based measures to investigate the information content of option-based indicators for predicting the future direction of the stock market. Even if they successfully combine many indicators, and the proposed approach outperforms the classical least-squares linear regression method in forecasting the direction of S&P500 returns, their model lacks interpretability. The higher the machine learning model interpretability, the easier it is for the researcher or the analyst to understand why certain decisions or predictions have been made. Given that these indices are widely adopted by investors and regulators for investment strategies, risk management purposes, and monitoring the stock market health, many players can benefit from a better understanding of the relationship between option-implied indicator and stock market returns. To fill this gap, in this paper we exploit fuzzy rule-based classification systems (FRBCS) to investigate the forecasting ability of several option-implied indicators on the future direction of the S&P500 index. The aim of the study is twofold: to analyse the FRBCS models performance and the knowledge extracted about the risk indices. In this way, we are able to develop a model which could become interpretable by human beings.

Fuzzy rule-based systems (FRBSs) are a well-known method family within soft computing and are based on fuzzy concepts to address complex real-world problems (Riza *et al.* (2015)). They are based on the fuzzy set theory proposed by Zadeh (1965), which aims at representing the knowledge of human experts in a set of fuzzy IF-THEN rules. The FRBSs have been used for different purposes and represent a powerful method to deal with uncertainty, imprecision, and non-linearity. They are commonly used for identification, classification, and regression tasks. FRBSs have been successfully adopted in several application domains, including control engineering (Babuska (1998)), finance (Boyacioglu and Avci (2010)), robotics (Bai *et al.* (2005)), pattern recognition ((Chi *et al.*, 1996)), bioinformatics (Zhu (2013)), and data mining (Ishibuchi *et al.* (2005a)). Compared to traditional techniques used recently, the use of fuzzy rules allows for more robust systems and more readily understandable even for non-expert users. Specifically, fuzzy modelling comes integrated into processes that previously involved the use of regressions. The main problems of these techniques concern the restitution of approximations not always good and difficulties of interpretation. The use of fuzzy rules beyond promoting immediate interpretability for the user reaches a better accuracy in the approximation. Moreover, working with rules at the local level, instead of general, it allows to represent continuous functions with significant robustness.

The advantage that FRBSs offer, even over recent machine learning (ML) techniques, lies in the identification of the preliminary knowledge of experts, which they can then integrate with that obtained from the data. In the context of fuzzy control systems, the use of FRBS compared to traditional control methods allow simplifying the tasks of model design, for which complex mathematical models are not required, and enabling full use of the information available. From the very beginning, using a fuzzy control system has allowed the development of models with an attractive balance between performance and cost. Finally, within the fuzzy classification, they can distinguish two classification methods, the first able to function autonomously, aiming to optimize the percentage of correct system classifications. The second classification model is used as a tool that supports the activity decision-making by users. This system favours the transparency of the entire process classification through a linguistic FRBS.

In this paper, we use a FRBS for classification purposes (FRBCS) based on spatial partition and genetic algorithms. One feature of FRBCS consists of their straightforward interpretation, which makes them preferable for financial analysis. In particular, we propose four FRBCS approaches for designing interpretable-accurate models for predicting the direction of S&P500 index. The methods include: the FRBCS with Chi *et al.* (1995) method (FRBCS.CHI), the Ishibuchi *et al.* (1999) method based on Genetic Cooperative Competitive Learning (GFS.GCCL), the Ishibuchi *et al.* (2005b) method based on hybridization of GCCL and Pittsburgh (FH.GBML), and the Structural Learning Algorithm on Vague Environment (SLAVE) of González and Pérez (2001).

The use of these methods allows us to achieve a threefold goal. First, to compare FRBCSs methods with traditional machine learning classification methods to better understand their balance between performance and cost in predicting the direction of future market returns. Second, to investigate and contrast the four FRBCSs methods (FRBCS.CHI, GFS.GCCL, FH.GBML, and SLAVE) in a financial application to provide evidence and further insights about their interpretability-accuracy trade-off. Third, to offer a wide range of stakeholders precious information about the possibility of using option-implied indicators for financial market forecasting. We find several results. First, the adopted FRBCSs suffer from a slight underperformance in terms of accuracy compared to traditional ML methods adopted in citecampisiforecasting. However, the lower accuracy is rewarded by the possibility of understanding which indicators and how they can be helpful to predict the future direction of the market. Second, the FH.GBML and SLAVE methods are the ones that are able combine high accuracy and interpretability, thus overperforming both the GFS.GCCL and FRBCS.CHI methods. Third, all the FRBCSs agree on the role of the VIX index as an indicator of market fear: a large (small) value of the VIX index indicate a bearish (bullish) market in the next 30 days. The evidence is less straightforward regarding the remaining indicator. However, all the FRBCSs indicate that low values of the PUTCALL index are associated with a bullish market.

The paper proceeds as follows. In Section 2 we illustrate the FRBCSs adopted in our study. In Section

3 we describe our dataset and the estimation methods used to perform our forecasts. In Section 4 we investigate the predictive power of the fuzzy rule-based classification methods to forecast the S&P500 returns. We also compare the four formulation of fuzzy rule-based classifier design with each other. Section 5 is devoted to the analysis of interpretability in detail comparing the FRBCS methods with the most used ML techniques and providing insights into the rules obtained from the four FRBCSs. The last section concludes.

2 Fuzzy rule base for classification problems

Within the fuzzy rule-based systems (FRBSs) family, fuzzy rule-based classification systems (FRBCSs) are specialized FRBSs meant to handle classification tasks. The main characteristic of the classification problem is that the output is a categorical variable. As a consequence, in FRBCSs we preserve the antecedent part of linguistic variables while changing the consequent part to be a class C_j from a prespecified class set $C = C_1, ..., C_M|$ (Riza *et al.* (2015)) Depending on the available information, two main strategies can be exploited to build FRBSs. The first strategy is to obtain information from human experts. In particular, the researcher defined the FRBS knowledge manually by interviewing human experts to extract and represent their knowledge. However, this approach is not feasible in many cases, e.g., experts are not available, lack of enough knowledge on the problem, etc. The second strategy can overcome this issue since it aims to obtain FRBSs by extracting knowledge from data by using learning methods. Given that existing literature does not provide enough guidance on how to combine different option-implied indicators to forecast the direction of future market returns, we follow the second approach. Riza *et al.* (2015) classify the learning methods into different groups. In our study, we will exploit four different FRBSs, based on space partition and genetic algorithms. We discuss the groups and the chosen methods in detail in the following.

2.1 FRBSs based on space partition approaches

FRBS learning methods classified in this first group adopt a strategy of splitting the variable space and then exploit this partition to obtain the parameters of the membership functions. The first technique based on this approach is proposed by Wang and Mendel (1992). Chi *et al.* (1996) extend Wang and Mendel's method for tackling classification problems. Their algorithm is similar to Wang and Mendel's and is based on four steps. The first step is the construction of linguistic labels. In particular, it consists of an equal division of the input and output spaces of the given numerical data into fuzzy regions. Since fuzzy regions refer to intervals for the linguistic term, they are built with the same triangular shape. Thus, their length is related to the number of linguistic terms. In the second step fuzzy rules are generated for each example in the training data. In particular, for each instance and each variable, the linguistic label with the highest membership degree is selected. Third, the antecedent part is determined by aggregating degrees of membership functions in the antecedent and consequent parts, using the intersection of the selected linguistic labels. On the other hand, the consequent is the class label of the example. In the last step, the final rule-base is obtained after deleting redundant rules: those with a lower degree can be eliminated, i.e., in case of duplicated or conflicting rules, only the rule with the highest weight is kept.

2.2 Genetic fuzzy systems for fuzzy rule learning

The approach based on genetic fuzzy systems was originally introduced by Herrera *et al.* (1998), who adopted a genetic algorithm to generate the structure of the fuzzy rules and the membership function parameters simultaneously. Instead of sharing a common definition of linguistic values as in the original Mamdani formulation, the approximate Mamdani type procedure proposed by Herrera *et al.* (1998) may have a different set of linguistic values for each rule. There are two main advantages using this type of procedure (Riza *et al.* (2015)). First, there is an augmented degree of freedom of parameters, and for a given number of rules, the system can better be adapted to the complexity of the problems. Second, the learning processes can simultaneously identify the structure and estimate the model parameters.

Within this group of FRBSs Ishibuchi *et al.* (1999) propose a method based on genetic cooperative competitive learning to handle classification problems. In this method, a chromosome describes each linguistic IF-THEN rule using integers to represent the antecedent part. Then, the heuristic method is exploited to automatically generate the class in the consequent part of the fuzzy rules. In this method, the evaluation is carried out for each rule, meaning that the performance is not based on the entire rule set. The Ishibuchi *et al.* (1999) method is based on the following five steps. The first step consists in the generation of an initial population of fuzzy rules. In the second step, fuzzy rules are evaluated in the current population. The third step involves the use of a genetic operator to generate new fuzzy rules. The fourth consists in replacing part of the current population using the rules generated in the previous step. Finally, the algorithm is terminated if the stopping condition is met, otherwise it returns to the second step. Additionally, the Ishibuchi *et al.* (1999) method introduces a tool to handle high-dimensional data. In particular, if an attribute "don't care" is used in the antecedent fuzzy sets, the corresponding linguistic values are always assumed to have a degree of one.

(Ishibuchi *et al.*, 2005c) propose an alternative method based on the hybridization of genetic cooperative competitive learning (GCCL) and Pittsburgh approach for genetic fuzzy systems. The (Ishibuchi *et al.*, 2005c) algorithm is based on five steps. The first step consists in generating a population where each individual is a fuzzy rule set. In the second step, the fitness value is computed for each rule set in the population. The third step is the generation of new rule sets by selection, crossover, and mutation consistently with the Pittsburgh-style algorithm. Then, iterations of the GCCL-style algorithm are applied to each generated rule set by considering user-defined probabilities of crossover and mutation. In the fourth step, the best rule set are added in the population to newly generated rule sets to generate the next population. In the last step, the algorithm is terminated if the stopping condition is met, otherwise it returns to the second step.

The last method adopted in our analysis of the US stock market is the structural learning algorithm on vague environment (SLAVE) proposed by Gonzalez and Perez (2001). Since SLAVE is based on the iterative rule learning approach, it generates only one fuzzy rule in each execution of the genetic algorithm. In order to eliminate the irrelevant variables in a rule, the algorithm consists of two parts. The first part aims to represent the relevance of variables, while the second one defines the values of the parameters. To obtain the fuzzy rules, the González and Pérez (2001) method consists of four steps. In the first step, a genetic algorithm is used to obtain one rule for the fuzzy rule-based system. In the second step, the obtained rule is collected into the final set of rules. In the third step, the new rule is checked and penalized. Finally, the system returns the set of rules as the solution if the stopping condition is satisfied, otherwise it returns to the first step. The SLAVE method exploits binary codes as representation of the population and conducts the basic genetic operators, i.e., selection, crossover, and mutation on the population ((Riza *et al.*, 2015)). Finally, the best rule is identified as the one characterized by the highest consistency and completeness degree.

Regarding the three methodologies of the FRBCS based on genetic algorithms, it is crucial to emphasize the competition between the various elemetns of the inference system that is created. While being combined with fuzzy inference systems, the genetic algorithms mantain competition among individuals of the population, leading to the creation of single rules or sets of rules. In these learning algorithms, it is possible to define the control parameters, that is, the parameters that outline the operating mode of the algorithm itself. What cannot be changed arbitrarily by the operator are the form of membership functions and, in the case of Ishibuchi's method based on hybridization of GCCL and Pittsburgh Approach, also the number of linguistic values related to the input variables. Such restrictions significantly limit the interaction between the knowledge acquired by experts and the one that it can obtain from the available numerical data.

The four FRBCSs illustrated above will be used to investigate the relationship between option-implied indicators and the future direction of US market returns. In particular, the use of these FRBCSs compared to other classification methods adopted in the literature (see e.g., Campisi *et al.* (2021)) will allow us to extract and better understand the precious information content embedded in these indicators. The option-implied indicators for the US market will be described in the next section.

3 Data and methodology

The data set consists of S&P500 index returns and option-based indicators on the US stock market. The S&P500 is a stock index that accounts for the performance of 500 of the largest public companies in the United States. Formally known as the Standard & Poor's 500 Composite Stock Price Index and commonly referred to as the S&P500, is it the most used benchmark to follow US stock performance and determine the state of the overall economy. Many investors worldwide also exploit the S&P500 as a benchmark for their portfolios.

3.1 The US market data

We obtain daily data for the S&P500 index and option-based indicators covering the period from October 2014 to September 2019 from the Bloomberg database, consisting of 1232 daily observations.

The choice of the option-based indicators as predictors for the S&P500 returns is strongly supported by financial literature, which shows a significant role of option-implied information in predicting stock market returns for many reasons. First, the option market embeds information about the forward-looking distribution of asset returns (Barro and Liao (2019)). Since the option payoff at expiration depends on the future underlying asset value, option prices reflect investor expectations (under the risk-neutral measure) about future fluctuations of the underlying asset ((Gambarelli and Muzzioli, 2019)).

Second, according to the informed investor theory (see, e.g., Xing *et al.* (2010)) option prices incorporate additional information content compared to the one embedded in stock prices, i.e., information that is not already priced into the underlying stock market. Third, several studies (see, e.g. Seo and Wachter (2019)) provide theoretical and empirical evidence that political uncertainty and tail economic risk is priced into the options market. Finally, these indices are widely adopted by investors and regulators for investment strategies, risk management purposes, and to monitor the stock market health. As a result, a large number of players can benefit from a better understanding of the relationship between these indices and stock market returns.

The option-based indicators embedded in the analysis include the following indices: VIX, VVIX, SKEW, GVZ, OVX, PUTCALL, VIX9D, VIX3M, VIX6M, and VXN. The VIX index measures the expected 30-day stock returns volatility of the S&P500 and is the most followed measure of volatility and risk worldwide. Investors widely use exchange-traded VIX futures to hedge their exposures (Tong and Huang (2021)). The VXN (NASDAQ-100 volatility index) is the corresponding of the VIX index for the NASDAQ-100 index, known as the US large-cap growth indices that includes 100 companies at the forefront of innovation such as Alphabet, Amgen, Apple, Facebook, Intel, Microsoft, Starbucks and Tesla.

There is strong evidence that volatility indices provide useful information about current and future stock returns in financial literature. The predictive power of both VIX and VXN indices on the underlying index returns have been investigated in Rubbaniy *et al.* (2014). To capture the option-implied information for shorter and longer horizons beyond the standard 30-day one commonly adopted in financial risk measuring, the CBOE provides further indicators based on the volatility of S&P500 options. These are the VIX9D, VIX3M, and VIX6M indices measuring the expected nine-day, 3-month, and 6-month volatility of S&P500 stock returns, respectively. We also use the CBOE VVIX Index, measuring the volatility of volatility, i.e., the expected volatility of the 30-day forward price of VIX, which is the price of a hypothetical VIX futures contract that expires in 30 days. The VVIX measures how rapidly S&P500 volatility changes and is thus a measure of the volatility of how quickly market sentiment changes. Given the relatively recent introduction of the VVIX, the information content of the index about current and future stock returns is little investigated in the literature.

We also investigate the information embedded in two indices based on commodities: the GVZ, and the OVX index. The GVZ (CBOE/COMEX Gold Volatility Index) measures the expected 30-day volatility of returns on the SPDR Gold Shares ETF. Similarly, the CBOE Crude Oil ETF Volatility Index OVX (Oil VIX) measures market expectations of 30-day volatility of crude oil prices by exploiting the VIX methodology on the United States Oil Fund. The inclusion of the two volatility indices is motivated by the existence of return and volatility spillovers between Gold, Oil, and the stock market. Kang *et al.* (2015) show that oil price shocks drive the stock market return and volatility relationship in the US market. Gokmenoglu and Fazlollahi (2015) exploit GVX and OVX to show that volatility in one market can affect prices in another market.

Finally, the SKEW and PUTCALL provide information about the asymmetry between call and put option prices. The SKEW index of the Chicago Board Options Exchange (CBOE), launched in February 2011, measures the tail risk not fully captured by the VIX index (Elyasiani *et al.* (2021)). While VIX accounts for the overall risk in the 30-day S&P500 log-returns without disentangling the probabilities attached to positive and negative returns, the SKEW index is aimed to measure the tail risk. Regarding the predictability of skewness indices on future market returns, Elyasiani *et al.* (2018) find that asymmetry indices provide higher explanatory power compared to volatility indices. Moreover, Mora-Valencia *et al.* (2021) show that the SKEW index embeds salient information for expected financial downturns. On the other hand, the PUTCALL is the ratio between put and call options on the S&P500 index purchased on a given day. A high put/call ratio should indicate fear in the markets, while a low ratio signals confidence. Existing studies suggest that options trading volume; hence, the PUTCALL ratio is a predecessor to asset price movements (Houlihan and Creamer (2019)), and in particular, it predicts negative future stock returns (Blau *et al.* (2014)).

Unlike previous studies, we will consider all the indices mentioned above in a fuzzy rule-based framework to investigate their forecasting power on future market returns. To this end, we calculate the response variable as the S&P500 return in the next 30 days for each day in our sample, computed at time t referring to a window of t + 30 days.

4 Discussion of the results

In this section we compare and contrast the forecasting performance of the various FRBCSs adopted in this study: the Chi et al. (1996) method based on space partition (FRBCS.CHI), the Ishibuchi *et al.* (1999) method based on Genetic Cooperative Competitive Learning (GFS.GCCL), the Ishibuchi *et al.* (2005b) method based on hybridization of GCCL and Pittsburgh (FH.GBML), and the Structural Learning Algorithm on Vague Environment (SLAVE) of González and Pérez (2001). The metrics used for comparison are accuracy, the area under the curve, and the F-measure. Moreover, in order to highlight the advantage of feature selection in our analysis, for each of the FRBCSs used in the analysis we show the results before and after feature selection, i.e. by considering all the 10 regressors and the 7 option-implied indicators, respectively.

The results about the forecasting performance of the proposed fuzzy rule-based classification models clearly show an improvement after feature selection (Table 1) compared to the case before feature selection (Table 2), according to all the metrics adopted in the analysis. Moreover, two of the three performance metrics (ACC, and F), indicate the FH.GBML and SLAVE methods as the ones with the highest accuracy after feature selection. The only exception is the measure AUC, which attributes a slightly higher score for FRBCS.CHI than for FH.GBML.

Table 1: Accuracy, AUC, and measure F of the proposed fuzzy rule based classification models before feature selection

ML method	ACC	AUC	F
FRBCS.CHI	0.7672	0.7709	0.8125
GFS.GCCL	0.7586	0.7180	0.8205
FH.GBML	0.5991	0.4983	0.7480
SLAVE	0.6595	0.5744	0.7775

Table 2: Accuracy, AUC, and measure F of the proposed fuzzy rule based classification models after feature selection

ML method	ACC	AUC	F
FRBCS.CHI	0.8009	0.8031	0.8403
GFS.GCCL	0.7835	0.7363	0.8428
FH.GBML	0.8528	0.7946	0.9006
SLAVE	0.8528	0.8261	0.8931

Table 3: Accuracy, AUC, and measure F of the proposed fuzzy rule based classification models after feature selection partitioning the dataset in 80% for training and 20% for testing.

ML method	**Training set**			**Test set**		
	ACC	AUC	F	ACC	AUC	F
FRBCS.CHI	0.7820	0.7776	0.8253	0.8009	0.8031	0.8403
GFS.GCCL	0.7710	0.6766	0.8484	0.7835	0.7363	0.8428
FH.GBML	0.7770	0.7077	0.8455	0.8528	0.7946	0.9006
SLAVE	0.7990	0.7696	0.8499	0.8528	0.8261	0.8931

In Table 3, we compare the metrics for the proposed fuzzy rule-based classification models after feature selection and after partitioning the dataset in 80% for training and 80% for testing. Several observation are in order. First, all the methods adopted in the analysis show a pretty good forecasting performance, with the performance indicator for the test set ranging from 0.78 to 0.90. Second, methods based on Genetic fuzzy systems outperforms the (Chi *et al.*, 1996) method based on space partition approach in terms of forecasting accuracy, with the only exception of the GFS.GCCL method that obtain the lowest accuracy on average. Third, the FH.GBML and SLAVE methods show the highest accuracy in the test set according to both the ACC and AUC performance metrics. In particular, the two methods obtain the same performance according to the ACC metrics, while the SLAVE method overperforms (slightly underperforms) the FH.GBML in terms of AUC (F-measure).

5 Interpretability of the four FRBCSs methods

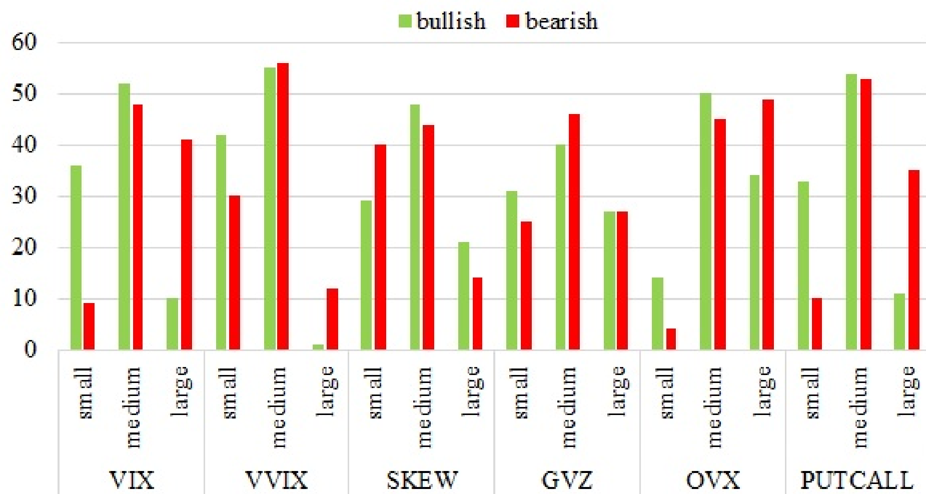
In this section, we evaluate the interpretability of the four FRBCSs methods adopted in our analysis and we provide further insights into the rules generated by the FRBCSs to better understand the relationship between existing option-implied indicators and the direction of future S&P500 returns. Regarding the interpretability-accuracy tradeoff, existing studies (Gacto *et al.* (2011)) argue that commonly used criteria to evaluate the complexity of the rule-based system are the number of rules, and the number of conditions. In particular, the best model is the simplest one fitting the system behavior well according to the principle of Occam's razor. Therefore, the set of fuzzy rules must be as small as possible until the model accuracy is preserved to a satisfactory level. Regarding the number of conditions in the antecedent of a rule must, the number should not exceed the threshold of 7 ± 2 distinct conditions, which corresponds to the number of conceptual entities a human being can handle ((Miller, 1956), Gacto *et al.* (2011)).

In the previous section, we compare and contrast the ability of different FRBCSs to predict the direction of future S&P500 returns. In this section, we propose a comparison between the four FRBCSs methods and traditional ML methods. Moreover, we provide further insights into the rules generated by the FRBCSs to provide investors with more information about the relationship between existing option-implied indicators and the direction of future S&P500 returns.

The FRBCS.CHI, GFS.GCCL, FH.GBML, and SLAVE methods produce 196, 9, 4, and 4 rules, respectively. Based on the discussion above, the FRBCS.CHI method based on space partition, producing a very large number of rules compared to the others, fails to reach a good balance between accuracy and interpretability. Despite the complexity of the output, we will still try to provide more insight about the rules generated by FRBCS.CHI method. While the rules obtained using the methods based on genetic fuzzy systems can be easily arranged in a table to represent the large number of rules obtained using the FRBCS.CHI method, we create a histogram plot (Figure 1). In particular, for each of the six variables exploited by the FRBCS.CHI method, the height of the green (resp. red) bar indicates the number of occurrences in which a small, medium or large value of an option-based indicator is associated with a bullish (resp. bearish) behaviour of the S&P 500 index in the following 30 days. Several observations are noteworthy. First, a small (resp. large) value of the VIX is associated most of the time with a bullish (bearish) direction of the US stock market. This result is consistent with existing studies since a high (resp. low) value of the VIX indicate a high (resp. low) value of fear in the market and thus is commonly associated with negative (positive) market returns. Second, a similar pattern can be observed for the PUTCALL index. More specifically, a small (resp. large) PUTCALL value is generally associated with a bullish (resp. bearish) market. From an economic point of view, this result could be motivated by the fact that investors buy put options when they expect negative outcomes from the stock market. Since the PUTCALL is computed as the ratio between put and call options on the S&P500 index purchased on a given day, if investors buy put options for their hedging purposes, the PUTCALL index is expected to increase. On the contrary, if the market is expected to be bullish, investors sell put options and buy call options, resulting in a decline of the PUTCALL index. Third, the remaining option-based indicators do not provide clear patterns to predict the direction of future S&P500 returns. However, a large value of the VVIX is commonly associated with a bearish stock market, suggesting that the VVIX acts as a fear indicator for the US stock market. On the other hand, small OVX index values are associated with a bullish stock market, suggesting the existence of a link between commodity and stock volatilities. The rules extracted from the remaining FRBCSs are arranged in Table 4. The GFS.GCCL method generates eight rules, but we can see only one clear pattern in the data: a large value of the VIX index is associated with a bearish market. On the other hand, small or medium values of the VIX index are associated to a bullish stock market. The same pattern for the VIX index is also found for the FH.GBML method, with one exception: if the VIX and the PUTCALL are high, but all the remaining indices are low, the future market direction is bullish. A possible interpretation for this result is as follows: if volatility becomes extremely high, then the market has already discounted all the fear, and positive returns can be expected (Elyasiani *et al.* (2018)). Finally, the rules generated by the SLAVE method are arranged in Panel C. A large (resp. medium) value of the VIX index is associated to a bearish (resp. bullish) stock market.

Moreover, a small value of the PUTCALL, independently from the value of the other indices, is associated with a bullish stock market.

Figure 1: number of occurrences in which a small, medium or large value of an option-based indicator is associated with a bullish (in green) or bearish (in red) behaviour of the S&P 500 index



6 Conclusions

In this paper, we investigated the forecasting ability of fuzzy rule-based classification systems (FRBCSs) on the future direction of US market returns. In recent decades, the increasing availability of information about the financial market and new efficient decision-making algorithms pushed financial research towards models characterized by an ever-better performance in forecasting financial markets. In particular, researchers focused on improving the model accuracy without paying particular attention to interpretability (Gacto *et al.*, 2011). Even if many studies successfully combined several variables using machine learning methods to predict future market returns (see e.g. Campisi *et al.* (2021)), most existing models lack interpretability. As a consequence, investors and regulators have been left without a clear indication of which variables, and how, can be helpful to predict the future direction of the stock market. To fill this gap, we investigate the ability of four FRBCSs methods in combining several option-implied indicators to predict the direction of S&P500 future returns. The objective of the study is threefold. First, to compare FRBCSs methods to traditional machine learning classification methods in predicting the direction of future market returns, in order to understand how much does the interpretability cost in terms of accuracy. Second, to compare and contrast different FRBCSs methods in terms of their interpretability-accuracy trade-off. Third, to provide investors and regulators with a better understanding of the relationship between option-implied indicators and the direction of future market returns. The FRBCSs methods adopted in the analysis are the Chi *et al.* (1996) method based on space partition approaches (FRBCS.CHI), the Ishibuchi *et al.* (1999) method based on Genetic Cooperative Competitive Learning (GFS.GCCL), the Ishibuchi *et al.* (2005b) method based on hybridization of GCCL and

Table 4: List of rules obtained through methods based on Genetic fuzzy systems

VIX	VVIX	SKEW	GVZ	OVX	PUTCALL	MARKET
Panel A:						
large	medium	medium	medium	medium	medium	bearish
large	small	don't care	small	small	don't care	bearish
large	don't care	medium	medium	large	medium	bearish
medium	medium	medium	small	medium	medium	bullish
small	small	medium	medium	medium	medium	bullish
medium	medium	medium	medium	medium	medium	bullish
small	medium	medium	don't care	medium	small	bullish
medium	medium	medium	large	large	medium	bullish
small	medium	large	large	don't care	medium	bullish
Panel B:						
high	small	small	small	small	high	bullish
high	don't care	small	medium	medium	don't care	bearish
small	small	don't care	don't care	high	high	bearish
high	high	don't care	small	high	medium	bearish
Panel C:						
large	large	medium	medium	medium	medium	bearish
large	don't care	don't care	don't care	don't care	don't care	bearish
medium	small	medium	large	medium	medium	bullish
don't care	don't care	don't care	don't care	don't care	small	bullish

Pittsburgh (FH.GBML), and the Structural Learning Algorithm on Vague Environment (SLAVE) proposed by Gonzalez and Perez (2001). To predict the direction of future S&P500 returns, we exploit several option-based indicators listed by the CBOE. The choice of these indicators is strongly supported by financial literature that suggests a significant role of option-implied information in predicting stock market returns. Moreover, the relationship between these indicators and the direction of market returns has been investigated in Campisi *et al.* (2021) using traditional machine learning methods, allowing an appropriate comparison. For this reason, we also adopt the same sample period used in Campisi *et al.* (2021) (daily data from October 2014 to September 2019). We found several results. First, when compared with the traditional ML methods such as Logistic Regression, LDA, Random Forest classification, Bagging classification, and Gradient Boosting classification, the FRBCSs adopted in our study suffer from a slight underperformance in terms of accuracy (possiamo dare un ordine / metrica di grandezza?). However, they can provide crucial information about the use of option-based indicators in forecasting future market returns. Second, among the FRBCSs, methods based on Genetic fuzzy systems outperform the standard FRBCS.CHI method based on space partition approach in terms of forecasting accuracy. In particular, the FH.GBML and SLAVE methods show the highest accuracy in the test set according to both the ACC and AUC performance metrics. At the same time, the FH.GBML and SLAVE generate a limited number of rules (4) compared to both GFS.GCCL (8) and FRBCS.CHI (196), highlighting far superior interpretability. Last, the rules obtained from the FRBCSs indicate an important role of the VIX index in predicting the direction of future market returns. More specifically, a large value of the VIX index is associated with a bearish market in the next 30 days. There is only an exception: if the VIX and the PUTCALL are high, but all the remaining indices are low, the future market direction is bullish. On the other hand, small or medium values of the VIX index are associated to a bullish stock market. Regarding the other option-implied indicators, the SLAVE method shows that a small value of the PUTCALL, independently from the value remaining indices, indicate a bullish direction for the stock market.

References

Airola, A., Pahikkala, T., Waegeman, W., Baets, B. D., and Salakoski, T. (2009). A comparison of AUC estimators in small-sample studies. In *Proceedings of the third International Workshop on Machine Learning in Systems Biology*, pages 3–13. PMLR. ISSN: 1938-7228.

Athey, S. and Imbens, G. (2019). Machine learning methods that economists should know about.

Babuska, R. (1998). *Fuzzy Modeling for Control*. Springer Netherlands : Imprint : Springer. OCLC: 840310973.

Bai, Y., Zhuang, H., and Roth, Z. (2005). Fuzzy logic control to suppress noises and coupling effects in a laser tracking system. **13**(1), 113–121.

Barro, R. J. and Liao, G. Y. (2019). Tractable rare disaster probability and options-pricing. **2019**(73).

Blau, B. M., Nguyen, N., and Whitby, R. J. (2014). The information content of option ratios. **43**, 179–187.

Boyacioglu, M. A. and Avci, D. (2010). An adaptive network-based fuzzy inference system (ANFIS) for the prediction of stock market return: The case of the istanbul stock exchange. **37**(12), 7908–7912.

Campisi, G., De Baets, B., and Muzzioli, S. (2021). A comparison of machine learning methods for predicting the direction of the us stock market on the basis of volatility indices. *DEMB WORKINGP PAER*.

Chang, P.-C. and Liu, C.-H. (2008). A TSK type fuzzy rule based system for stock price prediction. **34**(1), 135–144.

Chi, Z., Wu, J., and Yan, H. (1995). Handwritten numeral recognition using self-organizing maps and fuzzy rules. *Pattern Recognition*, **28**(1), 59–66.

Chi, Z., Yan, H., and Pham, T. (1996). *Fuzzy Algorithms: With Applications to Image Processing and Pattern Recognition*, volume 10 of *Advances in Fuzzy Systems — Applications and Theory*. WORLD SCIENTIFIC.

Elyasiani, E., Gambarelli, L., and Muzzioli, S. (2017). The information content of corridor volatility measures during calm and turmoil periods. **1**(4), 454–473.

Elyasiani, E., Gambarelli, L., and Muzzioli, S. (2018). The risk-asymmetry index as a new measure of risk. **22**(3), 173–210. Number: 3/4.

Elyasiani, E., Gambarelli, L., and Muzzioli, S. (2021). The skewness index: uncovering the relationship with volatility and market returns. **0**(0), 1–17. Number: 0.

Gacto, M., Alcalá, R., and Herrera, F. (2011). Interpretability of linguistic fuzzy rule-based systems: An overview of interpretability measures. **181**(20), 4340–4360.

Gambarelli, L. and Muzzioli, S. (2019). Risk-asymmetry indices in europe. Publisher: Dipartimento di Economia Marco Biagi - Università di Modena e Reggio Emilia.

Gokmenoglu, K. K. and Fazlollahi, N. (2015). The interactions among gold, oil, and stock market: Evidence from s&p500. **25**, 478–488.

González, A. and Pérez, R. (2001). Selection of relevant features in a fuzzy genetic learning algorithm. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, **31**(3), 417–425.

Gonzalez, A. and Perez, R. (2001). Selection of relevant features in a fuzzy genetic learning algorithm. **31**(3), 417–425.

Gu, S., Kelly, B., and Xiu, D. (2020). Empirical asset pricing via machine learning. **33**(5), 2223–2273.

Herrera, F., Lozano, M., and Verdegay, J. (1998). A learning process for fuzzy control rules using genetic algorithms. **100**(1), 143–158.

Houlihan, P. and Creamer, G. G. (2019). Leveraging a call-put ratio as a trading signal. **19**(5), 763–777.

Ishibuchi, H., Nakashima, T., and Murata, T. (1999). Performance evaluation of fuzzy classifier systems for multidimensional pattern classification problems. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, **29**(5), 601–618.

Ishibuchi, H., Nakashima, T., and Nii, M. (2005a). *Classification and modeling with linguistic information granules: advanced approaches advanced approaches to linguistic data mining*. Advanced information processing. Springer.

Ishibuchi, H., Yamamoto, T., and Nakashima, T. (2005b). Hybridization of fuzzy gbml approaches for pattern classification problems. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, **35**(2), 359–365.

Ishibuchi, H., Yamamoto, T., and Nakashima, T. (2005c). Hybridization of fuzzy GBML approaches for pattern classification problems. **35**(2), 359–365.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). An introduction to statistical learning: with applications in r. OCLC: ocn828488009.

Kang, W., Ratti, R. A., and Yoon, K. H. (2015). The impact of oil price shocks on the stock market return and volatility relationship. **34**, 41–54.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. **63**(2), 81–97.

Mora-Valencia, A., Rodríguez-Raga, S., and Vanegas, E. (2021). Skew index: Descriptive analysis, predictive power, and short-term forecast. **56**, 101356.

Riza, L. S., Bergmeir, C., Herrera, F., and Benítez, J. M. (2015). **frbs** : Fuzzy rule-based systems for classification and regression in *R*. **65**(6).

Rubbaniy, G., Asmerom, R., Rizvi, S. K. A., and Naqvi, B. (2014). Do fear indices help predict stock returns? **14**(5), 831–847.

Seo, S. B. and Wachter, J. A. (2019). Option prices in a model with stochastic disaster risk. **65**(8), 3449–3469. Number: 8.

Smialowski, P., Frishman, D., and Kramer, S. (2010). Pitfalls of supervised feature selection. **26**(3), 440–443.

Tong, C. and Huang, Z. (2021). Pricing VIX options with realized volatility. **41**(8), 1180–1200.

Vrontos, S. D., Galakis, J., and Vrontos, I. D. (2021). Modeling and predicting u.s. recessions using machine learning techniques. **37**(2), 647–671.

Wang, L.-X. and Mendel, J. (1992). Generating fuzzy rules by learning from examples. **22**(6), 1414–1427.

Xing, Y., Zhang, X., and Zhao, R. (2010). What does the individual option volatility smirk tell us about future equity returns? **45**(3).

Zadeh, L. A. (1965). Fuzzy sets. **8**(3), 338–353.

Zhang, X., Hu, Y., Xie, K., Wang, S., Ngai, E., and Liu, M. (2014). A causal feature selection algorithm for stock prediction modeling. **142**, 48–59.

Zhu, M. (2013). Return distribution predictability and its implications for portfolio selection. **27**, 209–223.