Fuzzy Techniques for IPO Underpricing Prediction

David Quintana ^a, Francisco Chávez ^b Rafael M. Luque Baena ^{c,*} and Francisco Luna ^c

^a *Dpto. de Informática, Universidad Carlos III de Madrid, Madrid, Spain. Email: dquintan@inf.uc3m.es*

^b *Dpto. de Ingeniería de Sistemas Informáticos y Telemáticos, Universidad de Extremadura, Mérida, Spain. Email: fchavez@unex.es*

^c *Depto. de Lenguajes y Ciencias de la Computación, Universidad de Málaga, Málaga, Spain. Emails: rmluque,flv@lcc.uma.es*

Abstract. Initial public offerings often show abnormal first-day returns. These, usually referred to as underpricing, are difficult to predict. Among the main obstacles, we could mention challenges like the fact that not all relevant variables have been identified yet; the mix of weak and strong indicators or the prevalence of outliers. In this context, we suggest that adaptive neuro-fuzzy inference systems and fuzzy rule-based system with genetic optimization have a lot to bring to the table. We test the predictive performance of these on a sample of 866 US IPOs and we benchmark them against six fuzzy algorithms and a set of eight classic machine learning alternatives. We conclude that both fuzzy systems, specially the former should be seriously considered in this domain.

Keywords: Fuzzy Rule-based System, ANFIS, Initial Public Offering, Underpricing.

1. Introduction

Initial public offerings (IPOs) are at the core of one of the most researched anomalies in finance. For decades, researchers have been trying to explain the reason behind the abnormal returns that tend to characterize first trading days. This often consists of large returns. Even though the difference between the offering price and the first closing price is not guaranteed to be positive, it is frequent enough to make initial returns often referred to as IPO underpricing.

The average magnitude of the initial returns has changed over time, but during the internet bubble of the nineties it got to a point where a large portion of the offerings ended up more than doubling the offering price in their first trading day. Ritter and Welch [48] reported an average initial return of 18.8% based on their sample of 6,249 companies taken public in the US between 1980 and 2001. Re-

cent changes in regulation and higher prevalence of mergers and acquisitions as exit strategy for initial investors has reduced IPO activity. The number of IPOs has gone down from an annual average of 310 during the 1980-2000 period to a about a third during 2001-2012 [60]. This reduction came together with a significant drop in US IPO average underpricing, which was about half during 2001-2010 [49].

Researchers on IPO underpricing have put forward a large number of theories to explain this phenomenon. These are generally supported with empirical analysis that is often based on linear models, usually multiple linear regressions. There are, however, a few papers regarding IPO analysis based on fuzzy approaches.

Among the efforts on the IPO space in general using fuzzy methods, we could mention the work of Bell et al. [8], who explore US investor perceptions of foreign IPO value using a fuzzy set methodology. The authors of [20] analyze IPO implementation decisions integrating different decision trees

*Corresponding author. E-mail: rmluque@lcc.uma.es

by fuzzy logic using reconciliation, and Cheng and Liu [13] predict the industrial growth rate of a set of companies at the time of their IPO based on rough sets.

If we focus our attention specifically on IPO underpricing prediction, Huang et al. [29] predict how a set of IPOs rank in term first-day returns to pick the best performers using a genetic-fuzzy methodology. Chen et al. [14] discretize the initial returns of a sample of IPOs and frame the problem as a classification task where they predict whether IPOs will have either positive, negative or zero initial returns relying on entropy based rough sets. Two years later, the same authors published an extended version along the same lines [12]. Finally, we will mention [21]. The authors of this work compare radial basis neural networks vs. fuzzy regression as predictors of probability of withdrawal and IPO underpricing on sample of companies taken public on Tehran Stock Exchange to conclude that the former outperforms the latter.

In this paper we argue that two fuzzy systems also have a lot to bring to the table for IPO underpricing prediction as continuous variable. More specifically, we focus on two techniques: adaptive neuro-fuzzy inference systems (ANFIS) and a fuzzy rule-based system with genetic optimization, which uses a 2-tuple linguistic representation model ($FRBS_{l-GA}$). To prove it, we benchmark them against six fuzzy algorithms and a set of eight classic machine learning alternatives that represent different algorithm categories.

ANFIS has previously been successfully used in many different financial contexts like the analysis of market volatility [59], financial prediction [6], trading [57,7], financial institution failure prediction [17] or real estate property price prediction [25].

There are also precedents for uses of 2-tuple linguistic representation models in finance. Given the characteristics of the technique, the nature of some of the applications is closely related to the examples mentioned for ANFIS. They include association rule mining in stock price series [24], financial performance evaluation [2], corporate credit rating [42] or stock evaluation [55], among others.

The rest of the paper will have the following structure: in Section 2, we make an introduction to the prediction algorithms that are part of the study. That will be followed by Section 3, which is devoted to the description of the explanatory

variables and the sample to be used in the analysis. Section 4 will be used to report the details related to the experimental analysis and, finally, Section 5 will cover the summary and conclusions.

2. Predictive Models

This section provides short introductions to the two core fuzzy approaches mentioned in the previous section, followed by brief descriptions of the algorithms used to benchmark the results.

2.1. $FRBS_{l-GA}$

The introduction to the Fuzzy Rule-based System (FRBS) will cover the two main aspects, the so-called linguistic 2-tuples representation, and the evolutionary algorithm used within the technique.

In regards to the first one, in [19] we can see new directions to techniques to design a FRBS applying genetic evolutionary algorithms. This process includes: selection of fuzzy rules, feature selection, learning of Knowledge Base (KB) via genetic derivation of the Data Base (DB), interpretability maintenance via multi-objective genetic processes, learning approaches considering different model structures, etc. These kinds of techniques try to find a better trade off between the efficiency of the learning/post-processing and the dimensionality of the search space.

This is the case of the *learning of KBs via genetic derivation of DBs*, a recent approach involving a simpler search space than the classical learning of KBs. It consists of obtaining the DB and the Rule Base (RB) separately, based on the embedded learning of the DB [10,18,45,56] (see Figure 1). This approach allows us to learn the most adequate context for each fuzzy partition, which is necessary in different application contexts and different fuzzy rule extraction models.

The learning scheme considered in this work belongs to this last group and is composed of two main components:

- A process to learn the DB, which allows to define:
 - * The number of labels for each linguistic variable.
 - * The lateral displacements of such labels.

Fig. 1. Learning scheme of the knowledge base.

Fig. 2. Lateral Displacement of the Linguistic Label M.

- A quick *Ad-hoc data-driven method* to derive the RB [58] considering the DB previously obtained. This method is run from each DB definition generated by a Genetic Algorithm (GA), thus allowing the proposed hybrid learning process to finally obtain the whole definition of the KB (DB and RB) by means of the cooperative action of both methods.

In [4], a new model of tuning of FRBSs, $FRBS_{l-GA}$, was proposed considering the linguistic 2-tuples representation scheme introduced in [27], which allows the lateral displacement of the support of a label and maintains the interpretability associated to the obtained linguistic FRBSs. This proposal also introduces a new model for rule representation based on the symbolic translation concept that is a number within the interval $[-0.5, 0.5]$ expressing the domain of a label when it is moving between its two lateral labels. Formally, we have the pair,

$$(s_i, \alpha_i), \quad s_i \in S, \quad \alpha_i \in [0.5, -0.5] \quad (1)$$

Figure 2 shows the lateral displacement of the label M. The new label “ y_2 ” is located between B and M, being smaller enough than M, but closer to M.

As we mentioned, the second key aspect to be specified is the evolutionary algorithm. In this work, we will consider the use of CHC [22] to de-

sign the proposed learning algorithm. CHC is a GA that presents a good trade-off between exploration and exploitation, being a good choice in problems with complex search spaces. This genetic model makes use of a mechanism of “Selection of Populations”. N parents and their corresponding offsprings are combined to select the best N individuals to take part of the next population.

For the fuzzy inference, we have selected the *minimum t-norm* playing the role of the implication and conjunctive operators, and the *center of gravity weighted by the matching* strategy acting as defuzzification operator.

2.2. ANFIS

The main advantage of fuzzy inference systems is that it is possible to interpret the linguistic rules generated by the model. The combination with neural networks, which incorporate mechanisms for learning from a set of data, forms a hybrid model denoted as Adaptive Neuro Fuzzy Inference System (ANFIS), proposed by [31].

The objective behind this approach, which involves a Sugeno-type fuzzy inference system and neural networks, is to design an architecture that utilizes a fuzzy model to represent knowledge in an interpretable way, with the learning capability derived from a neural network that can adjust the membership functions parameters and linguistic rules directly from data in order to improve the model performance.

Initially, the number and type of the input membership functions should be defined. For that, a substrative clustering is applied to define the initial structure for the ANFIS model [15]. This method is considered as an extension of the mountain clustering method [61].

After this initialization, a training phase is carried out by means of an hybrid strategy based on gradient descent and least squares estimator. Without giving further details about the equations of the gradient descent (see [31] for additional information), the aim of this process is to minimize the error between the real output and the model output, modifying the premise parameters. Additionally, it should be highlighted that there are several parameters related to the learning rate. The main one is the step size k , which can vary the convergence speed of the algorithm. This parameter is adapted dynamically along the train-

ing process using two kind of factors, the step size decrease rate k_d (fewer than 1) and the step size increase rate k_i (higher than 1).

2.3. Benchmark algorithms

The study will enrich the analysis done using ANFIS and the fuzzy rule-based system with genetic optimization and a 2-tuple linguistic representation model benchmarking the results against a number of alternatives. More specifically, we will consider six additional fuzzy algorithms and a set of eight classic machine learning alternatives that represent different algorithm categories. We introduce them very briefly below.

2.3.1. Fuzzy alternatives

The fuzzy alternatives used in the experiments are:

- *SBC* [16]: method based on using a subtractive clustering method to obtain the cluster centers (generating the rules), and then optimizes the membership function by using an efficient gradient descent method.
- *WM method* [58]: model introduced by Wang and Mendel. The WM method is able to generate fuzzy rules from sample data by dividing the input space into fuzzy regions, and then uses a lookup table to extract rules for each fuzzy subspace.
- *HYFIS* [33]: Hybrid neural fuzzy inference system (HYFIS) uses heuristic fuzzy logic rules and input-output fuzzy membership function, which can be further optimized by a hybrid learning scheme composed of two phases, namely, rule generation from data, and rule tuning using error backpropagation for a neural fuzzy system.
- *DENFIS* [32]: the Dynamic Evolving Neural-Fuzzy Inference System (DENFIS) is an on-line model that creates a TSK fuzzy inference system where fuzzy rules are generated using clusters of the input vectors. The system dynamically evolves its rule base to capture changes in the system throughout a continuous updating of the model.
- *FIR.DM* [44]: Fuzzy Inference Rules by Descent Method (FIR.DM) proposed by Nomura *et al.* automatically obtains fuzzy rules expressing the input-output relation of the data

by using a triangular membership functions, which are tuned by a descent method.

- *FS.HGD* [30]: simplified TSK fuzzy rule generation method using heuristics and gradient descent method (FS.HGD). It uses fuzzy IF-THEN rules with nonfuzzy singletons in the consequents of the rules. Space partition techniques generates the rule antecedents, while consequents are computed by a weighted mean value of the input data, which heuristically guided by the user. Finally, a gradient descent method seeks for the optimal value of the consequents.

2.3.2. Classic data mining approaches

The algorithms we will use as benchmarks are: IBK, least median of squared regression, LWL, M5P, M5Rules, multilayer perceptron, radial basis neural networks and SMO-Regression. This selection of algorithms includes different categories such as alternatives based on functions, nearest neighbors, rules and decision trees.

- IBK [1]: K-nearest neighbor algorithm implementation.
- LMSReg [51]: robust linear regression that disregards outliers to improve the model accuracy.
- LWL [5]: learning algorithm that creates classifiers based on weighted instances.
- M5P [47]: implementation of the M5 tree-based classification algorithm that adds linear regression components to handle continuous variables.
- M5Rules [28]: algorithm based on the divide-and-conquer principle. Every iteration, the algorithm relies in M5 to build regression trees, and then makes the best leaf into a rule.
- MLP [52]: multilayer perceptrons are a type artificial neural network that simulates the biological learning processes by means of a number of weight adjustment strategies. In this case, we rely on the most popular one, back-propagation algorithm.
- RBFN [43]: Radial Basis Neural Networks are another type of artificial neural network. It uses radial basis functions to approximate different regions of the input space depending on their characteristics.
- SMO-Reg [54]: implementation of support vector machines trained using sequential minimal optimization.

3. Variables and Data

In this section we introduce the variables considered in the the experimental work. To this end, we provide a formal definition of IPO underpricing together with the set of independent variables used to predict it. In addition to that, we describe the sample used to fit the models and test their performance.

3.1. Variables

We define IPO underpricing as the percentage change from offering price to first trading day close, minus the market performance of the stock market the day the company was taken public, that is:

$$R_i = \left(\frac{P_{c_i} - P_{o_i}}{P_{o_i}} \right) - \left(\frac{M_{c_i} - M_{o_i}}{M_{o_i}} \right) \quad (2)$$

where R_i is the market-adjusted initial return for stock i ; P_{o_i} is the offering price for stock i ; P_{c_i} is the price at market close for stock i ; M_{o_i} is the closing value for the market index of the stock market where i was floated the day before the IPO and M_{c_i} is the closing price on the day of the IPO for the broad index representing the specific market where i was taken public. The market indexes considered in the experimental work are S&P 500 and NASDAQ and AMEX composites.

Regarding the independent variables, as we have already mentioned in the introduction, the number of research papers devoted to discuss potential explanations for IPO initial returns is very large. Even though there is nothing like a consensus variable set accepted and used across publications, there are a few elements related to the offering structure that turn up very often. These variables, which we introduce succinctly below, are: offer price, price range width, price adjustment, retained stock offer size and relation to tech sector.

- *Offering price (PRICE)*: the final offering price has been identified as a relevant variable both as a component of other indicators and in its own right. Studies like [11,3] suggest that offering price might be perceived by potential buyers as a signal of the quality of the stock.
- *Price range width (RANGE)*: in the first stages of the IPO offering process, issuers

gather non-binding purchase offers by investors based on a tentative offering price ranges. Authors like [26,34] claim that the standardized width of these ranges, defined as the difference between the upper and the lower bound of the range divided by the lower one, might be interpreted as a sign of uncertainty and therefore, would be likely to influence underpricing.

- *Price adjustment (P_ADJ)*: researchers like the ones involved in [9,26,36] report that the two previous indicators might be combined to generate a third one that would capture the difference between the initial valuation and the final. According to them, the result of the expression that follows could be perceived by investors as a sign of uncertainty:

$$P_ADJ = \frac{P_f - P_e}{P_e} \quad (3)$$

where P_f is the final offer price and P_e is the expected price defined as the center point of the price range.

- *Offering size (LSIZE)*: the amount of capital involved in the offering is often found to have an statistically significant inverse relationship with underpricing [41,3]. The variable is defined as the offering size (in USD millions), excluding the over-allotment.
- *Retained stock (RETAINED)*: early investors are often perceived to have insider information. This has led researchers to think that the proportion of capital retained at the time of the IPO might be interpreted as a quality signal by new investors [35,36]. Given that we do not have detailed information regarding the exact number of primary and secondary shares, we will proxy this indicator through the number of shares sold at the IPO divided by the difference between the post-IPO number shares and the number of shares sold at the offering.
- *Technology (TECH)*: technology-based companies tend show higher average initial returns than the rest. For this reason, IPO research [53,37,38] often controls for this factor by means of a dummy variable that equals one in case that the core industrial activities of the company are related to technology and zero otherwise. The criterion that we will use consider and IPO as tech-related will be

based on IPO Monitor's classification. This company publishes reports where they label IPO as Tech or not according their US Standard Industry Codes and the activities they describe in their filings.

3.2. Data

The experimental analysis will be based on a sample of 866 US IPOs. The companies represented were taken public between January 1999 and May 2010 in NASDAQ, AMEX and NYSE. Following the standards of IPO research, the sample does not include closed-end funds; American Depositary Receipts; real state investment trusts or unit offerings. The main source of information was the commercial data provider IPO Monitor. In those instances where the information was incomplete, we used Hoovers, another data vendor, as the secondary source. Regarding index information, we obtained the NASDAQ and AMEX composites from the National Association Of Securities Dealers, and the S&P 500 index from Datastream.

Table 1 summarizes the main descriptive statistics, including the mean, median, maximum, minimum and standard deviation for the mentioned sample of 866 companies.

Table 1
Main descriptive statistics.

	Mean	Median	Max	Min	Std. Dev.
RETURN	0.176	0.073	3.718	-0.281	0.389
PRICE	14.641	14.000	85.000	3.250	5.812
RANGE	0.149	0.143	0.500	0.000	0.063
P_ADJ	0.105	0.082	1.509	0.000	0.099
LSIZE	2.054	2.025	3.939	0.061	0.446
RETAINED	0.309	0.262	1.000	0.002	0.198
TECH	0.328	0.000	1.000	0.000	0.470

The basic structure of the distribution of initial returns is consistent with the one reported in most of the referenced studies. As we can see in Fig. 3, among its main features, we could mention a positive mean return; the right-skewness caused by hot IPOs and the high kurtosis.

4. Experimental Analysis

In this section we report the results of the experimental analysis performed to test the suitability

Fig. 3. Return histogram.

ity of the fuzzy approaches in the IPO underpricing prediction domain. We will first introduce the methodology used in the conducted experimentation. The sensitivity of the fuzzy methods with respect to their parameterization is included next. The section ends with a comparison of the output of these methods to the predictions made by eight alternatives that are widely used in data mining.

4.1. Methodology

The models will be assessed by their predictive accuracy in terms of root mean square error (RMSE) on the same data set. In order to make the results as general as possible, we will perform a 10-fold cross validation (the same folds for all the algorithms). Given the stochastic nature of most of the benchmark algorithms, we will run the experiments 15 times using different random seeds and compare the average results. We will, however, report the main descriptive statics to ensure that the image is complete.

As we discussed, we will repeat the cross-validation analysis using a set of fuzzy and machine learning alternatives. For such task, we will rely on the FRBS R package and WEKA (Waikato Environment for Knowledge Analysis) [23].

4.2. Sensitivity analysis

For each parameter of the two core fuzzy methods a valid set of values is considered. Then multiple configurations of the methods are generated by combining the values of the parameters in every possible way. The parameter values for each technique are shown in Table 2. In this section we discuss the sensitivity of ANFIS and $FRBS_{l-GA}$ to these values.

Table 2

Parameter settings explored. ANFIS and $FRBS_{I-GA}$. The combination of all the values of the parameters generate the set of configurations tested for each method.

Algorithm	Parameters Explored
ANFIS	Initial Step Size, $K = \{0.005, 0.01, 0.05\}$
	Step Size Decrease Rate, $K_d = \{0.5, 0.9\}$
	Step Size Increase Rate, $K_i = \{1.1, 1.5\}$
	Neighborhood radius, $r_a = \{0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.7, 0.9\}$
	Quash Factor, $\delta = \{1, 1.25, 1.5\}$
	Accept Ratio, $\epsilon_a = \{0.3, 0.5, 0.7, 0.9\}$
$FRBS_{I-GA}$	Reject Ratio, $\epsilon_r = \{0.05, 0.15, 0.3\}$
	Evaluations = $\{50,000, 100,000, 150,000\}$
	Population = $\{50, 100, 150, 200\}$
	$\alpha = \{1.10, 1.15, 1.20, 1.25\}$
	Max Membership Functions = $\{3, 5, 7, 9\}$

4.2.1. $FRBS_{I-GA}$

$FRBS_{I-GA}$ uses four different parameters (see Table 2). The first parameter is the stopping condition of the algorithm, set up as the number of evaluations it is allowed to perform. We have used three different configurations, 50,000, 100,000, and 150,000 evaluations. The second parameter is the population size, i.e., the number of tentative solutions managed by the inner GA, which ranges between 50 and 200. The parameter α allows $FRBS_{I-GA}$ to determine the trade-off between accuracy, i.e., how good the method predicts the output, and complexity in terms of the number of rules (more rules does also mean less interpretability). The fourth parameter is the maximum number of membership functions (MFs) for each variable; we have used 3, 5, 7 and 9; using a maximum number of 9 MFs per variable, as suggested by the authors of the method [19].

A factorial design of experiments with the values included in Table 2 has been performed. In order to avoid overwhelming the reader with a huge amount of data (192 configurations), only a subset of the results are included in this section, aimed at showing the most relevant findings raised. The first fact that we want to show is the small sensitivity of $FRBS_{I-GA}$ to the number of evaluations (Fig. 4). Even though the best RMSE in testing is reached with 150,000 function evaluations as stopping condition, the differences are very narrow. The same holds with the population size, little-to-no significant differences can be found (not shown).

Fig. 4. Impact of the stopping condition in $FRBS_{I-GA}$.

Fig. 5. Impact of α and the maximum number of MF in $FRBS_{I-GA}$ (150k evaluations).

Once the stopping condition and the population size are fixed, Fig. 5 plots the RMSE in testing for the four different values of α and the maximum number of MF. The names in the axis prefix with p , a , and t , respectively, the population size, the α value, and the maximum number of MF. It is clear that the effect of α in $FRBS_{I-GA}$ is little for this dataset, providing very similar RMSE values both in training and test. The most determining parameter is the maximum number of MF. Regardless of the stopping condition used, the population size, or the α value, using a maximum number of 9 MF has always provided the lowest RMSE (in training and test). This is consistent throughout the entire experimentation and indicates that a higher number of rules are required to increase the accuracy of the method. The point is that this value is limited to 9 as the fuzzy system loses interpretability when a high number of MF is used.

The best configuration found for the $FRBS_{I-GA}$ method, which is used for comparison to other techniques, is detailed below in Table 4.

Fig. 6. Representation of Pearson correlation and RMSE metrics for all ANFIS configurations over test. Each point is a different configuration. The closer the points are to the coordinates (0.0,1.0), the better the RMSE and the correlation, and thus the better is the regression. RMSE in training is represented by color.

4.2.2. ANFIS

A total number of 3,888 different configurations are evaluated in order to analyze the behavior of the algorithm. These configurations are obtained by combining all the test parameters presented in the first row of Table 2. The three first parameters of this row are related to the ANFIS model itself, while the last four correspond to the subtractive clustering method, which is used to initialize the ANFIS structure.

Besides the RMSE, the Pearson correlation [50] is calculated as a measure to evaluate the accuracy of the neuro-fuzzy model. This value not only provides information about the error between the real output values and the estimated ones from the IPO input data, but it provides us with an idea of the tendency of the estimated output function, and its similarity to the real output function (its shape). The Pearson correlation has also been applied in several papers related to the ANFIS model in the literature [46,39].

Fig. 6 shows the mentioned statistic for each ANFIS configuration (represented as a colored point), with respect the RMSE in the testing phase, whereas the color of the points indicates the RMSE reached in the training phase. The first clear conclusion that can be drawn from this figure is that, the more correlated the real and estimated output vectors (i.e., the more similar their shapes), the lower the RMSE in test. These high correlation

Fig. 7. Sensitivity of ANFIS to parameters.

values indicates that the underlying neural network has been trained properly, and thus its prediction capability is enhanced. The zoomed part includes the final configuration used in a section below to compare with other approaches from the literature.

The second remarkable fact has to do with the color of the points. It can be seen that the lowest RMSE values in the training phase (lighter blue ones at the bottom of the figure) can be reached, but with a large penalty in testing due to overfitting. The Pearson correlation is also consistent with these values as it points out that the shapes of the input and output vectors are fairly different, i.e., the ANFIS training resulted in a poorly performing predictor.

The second figure of this section (Fig. 7) shows the sensitivity of seven representative parameters of ANFIS, measured as the average standard deviations of the configurations after setting each parameter of the graph. That is, for the *InitialStepSize* graph, we have averaged the RMSE of the configurations with the same values for all the parameters except for the *InitialStepSize*. For each cluster of configurations, there are as many inputs as different values of the studied parameter. Thus, the only variation in these configurations is the value of the studied parameter. The standard deviation is obtained as the mean of the standard deviations of all the clusters. As stated above, the first three ones have to do with the operation of the ANFIS method itself and, the second, with the initial clustering that initializes the structure. The error bars in the figure clearly shows that ANFIS strongly depends on the initialization method,

Fig. 8. Rule structure of $FRBSI_{l-GA}$.

as the four last parameters show the larger variability. Overall, this analysis points out that ANFIS is highly sensitive to its parameterization (also shown in Fig. 6, where the cloud points is fairly disperse).

The best configuration found for the ANFIS method, which is used for comparison to other techniques, is detailed below in Table 4.

4.3. Interpretation of solutions

In this subsection we discuss the nature of the solutions provided by $FRBSI_{l-GA}$ and ANFIS. For each approach we discuss the structure of the resulting systems and, finally, we extract conclusions regarding the relative importance of the independent variables.

4.3.1. $FRBSI_{l-GA}$

Fig. 8 shows a representative fuzzy system with 25 rules, which allows us to interpret how the $FRBSI_{l-GA}$ system works. It can be observed that all the variables do have 3 MFs, but V6 (TECH variable), which does only have 2 MFs, due to it is a boolean type variable (values 0 and 1). Just as an example, MF1 is the predominant one for V1 (PRICE), as it is present in 18 out of the 25 rules. This means that most of the data used to build the model does have rather low values for this variable. The same holds for V3 (P_ADJ). Analogously, for the LSIZE variable, the data is mostly located MF2, which are the MF that represents the medium values.

If we analyze the percentage of rules that are activated for each input data, displayed in Fig. 9, several interesting facts can be highlighted. First, there are very important rules in the system, e.g.,

Fig. 9. Percentage of activation for the 25 rules in the example for $FRBSI_{l-GA}$.

Rule 4, which covers many input samples, both in training and test. But it is also remarkable that there are several rules that are activated a couple of times (e.g., Rules 7, 8 or 9). These rules are generated by $FRBSI_{l-GA}$ in order to cover outliers. But note that if the outlier is in the training set, but not in the test set, then it is useless. Indeed, these rules might be critical for the final RMSE performance because, if one these rules is activated in test, but the outlier does not exist, the error gets dramatically increased.

4.3.2. ANFIS

The architecture generated by the ANFIS model uses seven rules with different Gaussian MFs for each pair, input variable and rule. Fig. 10 shows an example of the activation of the rules, given a input sample. As we did with $FRBSI_{l-GA}$, Fig. 11 displays the percentage of input samples that activate each of the seven rules, both in training and test. It can be seen that the two first rules are the ones that get activated more times, which means that they represents larger clusters of input data. Contrary to $FRBSI_{l-GA}$, all the rules are eventually activated (no one is useless).

Finally, Fig. 12 shows the importance of the input variables for the system in terms of the weighting coefficients of the output MF. It can be easily seen that PRICE, P_ADJ, and LSIZE are the variables with the largest influence on IPO prediction.

4.4. Benchmarking

As we already mentioned, we extend the basic study with a comparison of the predictive perfor-

Fig. 10. Rule structure of ANFIS

Fig. 11. Percentage of activation for the 7 rules in the example for ANFIS.

Fig. 12. Importance of the variables on the underpricing prediction according to the ANFIS model.

mance of $FRBS_{I-GA}$ and ANFIS vs. the six fuzzy systems to and the eight classic machines learning algorithms introduced in Section 2.3.

Tables 3 and 4 summarize the parameters used to run all the algorithms used for this analysis.

Table 3
Parameterization of fuzzy algorithms.

ANFIS	
Initial Step Size (K)	0.01
Step Size Decrease Rate (K_d)	0.50
Step Size Increase Rate (K_i)	1.10
Neighborhood Radius (r_a)	0.25
Quash Factor (δ)	1.00
Accept Ratio (ϵ_a)	0.50
Reject Ratio (ϵ_r)	0.15
DENFIS	
Threshold value	0.1
Step size	0.01
Max. no. iterations	200
Width of MF	2
FIR.DM	
No. labels	5
Max. no. iterations	50
Step size	0.01
T-norm	Yager
Implication function	Godel
$FRBS_{I-GA}$	
Stopping Condition	150,000 evaluations
Population Size	100
alpha	1.20
Max MF	9
FS.HGD	
No. Labels	9
Max. no. iterations	10
Step Size	0.01
T-norm	Yager
Implication Function	Dienes_Resher
HYFIS	
No. Labels	7
Max. no. Iterations	20
Step Size	0.01
T-norm	Product
Defuzzification	COG
Implication Function	Minim
SBC	
Neighborhood radius	0.2
Upper Threshold	0.3
Lower Threshold	0.15
WM	
No. Labels	5
MF	Bell
T-norm	Hamacher
Defuzzification	Mean,Max
Implication Function	Goguen

The experimental results are reported in Tables 5 and 6. There, we provide the main descriptive statistics for the RMSE obtained in the 15 repetitions of the 10-fold cross validation. For each algorithm we reported the average, median, variance, maximum and minimum RMSE.

As we can observe, ANFIS provides the best average results followed by HYFIS, the fuzzy rule-based system, WM, IBK and M5P. In addition to the relatively low mean and median RMSE, we should highlight the fact that the approach seems

Table 4

Parameterization of benchmark algorithms.

IBK	
<i>Neighbours Method</i>	3 Linear Search
LMSReg	
<i>R. Sample Size</i>	397
LWL	
<i>Weighting Kernel Classifier</i>	Linear Decision Stump
M5P	
<i>Min. Inst/Leaf</i>	5
M5Rules	
<i>Min. Inst/Leaf</i>	4
MLP	
<i>Transfer Function</i>	Sigmoid
<i>Neurons Hidden L.</i>	4
<i>LR</i>	0.1
<i>Epochs (max.)</i>	2500
RBFN	
<i>Clusters</i>	6
<i>Min. Std. Dev.</i>	0.1
<i>Ridge</i>	1.00E-08
SMO	
<i>Complex. Param</i>	1
<i>Epsilon</i>	1.00E-12
<i>Tolerance</i>	0.001
<i>Kernel</i>	Polynomial, exp=1

Table 5

Descriptive statistics of prediction errors of fuzzy methods.

	Mean	Median	Var.	Max	Min
ANFIS	0.3187	0.3187	1.58e-16	0.3187	0.3187
DENFIS	0.4087	0.4085	0.0000	0.4164	0.4017
FIR.DM	0.3732	0.3731	0.0000	0.3732	0.3731
FS.HGD	0.3820	0.3820	0.0000	0.3820	0.3820
<i>FRBS_{l-GA}</i>	0.3468	0.3471	2.13e-05	0.3563	0.3382
HYFIS	0.3456	0.3456	0.0000	0.3456	0.3456
SBC	0.3655	0.3655	0.0000	0.3655	0.3655
WM	0.3484	0.3484	0.0000	0.3484	0.3484

to be the most reliable one among those with a stochastic component. The range of variation for the prediction error is narrow and its variance is very low. *FRBS_{l-GA}* doesn't seem as trustworthy in that regard as it offered a wider range of variation across executions. That is, the impact of the random component of every specific execution seems to be potentially higher. Having said that, it is still in line with the rest of the stochastic methods. The variance is relatively high, though in the same order of magnitude as LMS Regression and SMO.

Table 6

Descriptive statistics of prediction errors of machine learning algorithms.

	Mean	Median	Var.	Max	Min
IBK	0.3570	0.3570	0.00000	0.3570	0.3570
LMSReg	0.3769	0.3745	4.17e-05	0.3945	0.3734
LWL	0.3853	0.3853	0.00000	0.3853	0.3853
M5P	0.3591	0.3591	0.00000	0.3591	0.3591
M5Rules	0.3786	0.3786	0.00000	0.3786	0.3786
MLP	0.3710	0.3681	1.45e-04	0.3941	0.3550
RBFN	0.3756	0.3758	4.45e-06	0.3800	0.3724
SMO	0.3808	0.3817	2.08e-05	0.3894	0.3737

On the opposite side of the spectrum, DENFIS, FS.HGD, LWL and SMO offer the worst average results. Even though support vector machines have been successfully used in many regression tasks, its sensitivity to the presence of outliers in the training set drags its performance in this domain. Conversely, ANFIS seems to handle this difficulty with much more efficacy.

We think that multilayer perceptron deserves to be highlighted despite of its discrete average performance. In one of the executions, it achieved the most competitive result versus ANFIS among the benchmark algorithms with an RMSE of 0.355. Interestingly, this also came together with the second worst prediction error and the highest variance.

The statistical significance of the differences reported in Table 7 was also formally tested. Given the non-normal distribution of the prediction errors, we relied on the Mann-Whitney test [40]. There, we represent the fact that the algorithm in the row has median prediction error that is significantly larger than the one in the column at 1% by ++. A similar difference in the opposite direction is symbolized by -. In case the disparity is such that the first one is significantly smaller at 5%, we use -. Finally, if we cannot discard at 5% the possibility of equal predictive accuracy, we report =.

The difference in median performance of ANFIS vs the rest is always significant at the 1%. This is also the case for FRBS and IBK, the third best algorithm.

As we can see, these results support the idea that fuzzy approaches, ANFIS in particular, might be suitable in this domain.

Table 7
Statistical significance of differences of medians.

	F1	F2	F3	F4	F5	F6	F7	F8	M1	M2	M3	M4	M5	M6	M7	M8
F2	++															
F3	++	-														
F4	++	-	++													
F5	++	-	-	-												
F6	++	-	-	-	=											
F7	++	-	-	-	++	++										
F8	++	-	-	-	=	++	-									
M1	++	-	-	-	++	++	-	++								
M2	++	-	++	-	++	++	++	++	++							
M3	++	-	++	++	++	++	++	++	++	++						
M4	++	-	-	-	++	++	-	++	++	-	-					
M5	++	-	++	-	++	++	++	++	++	++	-	++				
M6	++	-	=	-	++	++	=	++	++	-	-	++	-			
M7	++	-	++	-	++	++	++	++	++	=	-	++	-	=		
M8	++	-	++	=	++	++	++	++	++	++	-	++	=	++	++	

++: significant 1% +: significant at 5%

F1: ANFIS; F2: DENFIS; F3: FIR.DM; F4: FS.HGD; F5: $FRBS_{I-GA}$; F6: HYFIS; F7: SBC; F8: WM

M1: IBK; M2: LMSReg; M3: LWL; M4: M5P; M5: M5Rules; M6: MLP; M7: RBFN; M8: SMO.

5. Summary and Conclusions

IPO underpricing prediction is a domain with some specific characteristics that make it specially challenging. Among them we could highlight the fact that the set of descriptive variables identified by academic research is limited in their predictive power and mixes both weak and strong predictors. In addition to this, the presence of outliers that usually take the form of IPOs with extremely high initial returns add noise and difficulty to the process of training models.

Research on IPO underpricing prediction using fuzzy systems is extremely limited and it frames the problem either as a three-category classification task, or an effort to rank IPOs according to their expected performance. In this paper, however, we suggest that fuzzy systems can also be very powerful tools to handle it as a continuous variable regression task. In order to do that, we test the predictive performance of two different fuzzy approaches: neuro-fuzzy inference systems (ANFIS) and a fuzzy rule-based system with genetic optimization, $FRBS_{I-GA}$, on a sample of

866 US IPOs. To this end, we use six variables identified by financial literature review that are available in public databases.

In order to have a meaningful base of comparison, we benchmarked the results against six fuzzy alternatives and eight popular machine learning algorithms that represent different algorithm categories. In the former group we considered subtractive clustering; Wang and Mendel; Hybrid neural fuzzy inference system; Dynamic Evolving Neural-Fuzzy Inference System; Fuzzy Inference Rules by Descent Method and the simplified TSK fuzzy rule generation method using heuristics and gradient descent method. The latter included several families of predictive algorithms as those based on functions, nearest neighbors, rules and decision trees. These included IBK; least median of squares regression; LWL; multilayer perceptron; M5P; M5Rules; radial basis neural networks and SMO-Regression.

The outcome of this comparison shows that ANFIS outperformed all the alternatives in terms of mean and median predictive accuracy over the 15 repetitions of the 10-fold cross validation analysis.

The technique also provided the smallest variance and error range among the stochastic algorithms. The fuzzy rule-based system $FRBS_{I-GA}$ offered very good average results, beating all the machine learning algorithms and most of the fuzzy benchmarks (the only exception was HYFIS). Compared to ANFIS, it also turned out to be less reliable as it resulted in a higher variance in prediction errors.

Regarding the fuzzy alternatives, even though none beat ANFIS, they tended to offer better average performance than the classic machine learning ones. HYFIS offered very remarkable results. Even though the difference in prediction error vs. $FRBS_{I-GA}$ was not statistically significant, it obtained slightly better mean and median results. Despite of the superiority of fuzzy methods at large, DENFIS provided the worst performance among all the algorithms used in the experimental work.

Among the traditional algorithms used as benchmark, tree-based M5P offered very good performance. We expected that because the algorithm tends to offer good outlier management properties, a key feature in this domain. However, we did not find a clear pattern showing that any broad category performs significantly better than the rest. Among lazy strategies, IBK offered very good results, but LWL did not. Function-based alternatives were much more consistent as a category, but their average performance was in line with the one of lazy approaches.

The fuzzy systems allow a straight forward analysis of the relative importance of the variables considered to forecast IPO underpricing. The ANFIS model suggests that price is, by far, the most important factor among the ones considered. Price adjustment and size are the main distant followers. This finding adds support to the signaling theories put forward in the financial literature. Being a Tech company does not seem to be as relevant as we anticipated. A potential explanation might be that most of the IPOs in the sample are post-internet bubble. It is likely that the crash that followed might have resulted in an structural change, but this would have to be tested.

The above mentioned results suggest that IPO research would benefit greatly from a wider use of fuzzy methods, specially ANFIS. Among the main advantages of this technique would be the strong predictive performance and the possibility to interpret the results, something that cannot be taken

for granted with other black-box machine learning techniques. The possibility of looking into the relative importance of the independent variables allows for better insights from the purely financial perspective.

Possible future lines of work would include replicating the analysis for samples from other countries, or periods of time to detect structural changes; exploring the potential of the mentioned fuzzy approaches to identify a priori hot IPOs, and exploiting this knowledge to devise trading algorithms.

6. Acknowledgements

D. Quintana acknowledges financial support by the Spanish Ministry of Science under grant ENE2014-56126-C2-2-R. F. Chávez acknowledges financial support granted by Seventh Framework Programme of the European Union through the Marie Curie International Research Staff Scheme, FP7-PEOPLE-2013-IRSES, Grant612689 ACoBSEC, Spanish Ministry of Science and Innovation under project TIN2014-56494-C4-2-P (EPHEMECH), regional government Gobierno de Extremadura and FEDER, project GRU10029. Rafael M. Luque and Francisco Luna acknowledge support by Ministry of Economy, Industry and Competitiveness under contract TIN2016-75097-P, and Universidad de Málaga under contract PPIT.UMA.B12017/15.

References

- [1] D.W. Aha, D. Kibler and M.K. Albert, Instance-Based Learning Algorithms, *Machiche Learning* **6**(1) (1991), 37–66, ISSN 0885-6125.
- [2] F.-Y. Ai and J.-Y. Yang, Approaches to dynamic multiple attribute decision making with 2-tuple linguistic information, *Journal of Intelligent and Fuzzy Systems* **27**(6) (2014), 2715–2723.
- [3] S.M. Albring, R.J. Elder and J. Zhou, IPO Underpricing and Audit Quality Differentiation within Non-Big 5 Firms, *International Journal of Auditing* **11**(2) (2007), 115–131.
- [4] R. Alcalá and F. Herrera, Genetic Tuning on Fuzzy Systems Based on the Linguistic 2-Tuples Representation, in: *Proc. of the 2004 IEEE International Conference on Fuzzy Systems*, Vol. 1, 2004, pp. 233–238.
- [5] C.G. Atkeson, A.W. Moore and S. Schaal, Locally Weighted Learning, *Artificial Intelligence Review* **11**(1–5) (1997), 11–73, ISSN 0269-2821.

- [6] A. Bagheri, H.M. Peyhani and M. Akbari, Financial forecasting using {ANFIS} networks with Quantum-behaved Particle Swarm Optimization, *Expert Systems with Applications* **41**(14) (2014), 6235–6250, ISSN 0957-4174.
- [7] S. Barak, J.H. Dahooie and T. Tichý, Wrapper ANFIS-ICA method to do stock market timing and feature selection on the basis of Japanese Candlestick, *Expert Systems with Applications* **42**(23) (2015), 9221–9235, ISSN 0957-4174.
- [8] R.G. Bell, I. Filatotchev and R.V. Aguilera, Corporate governance and investors' perceptions of foreign IPO value: An institutional perspective, *Academy of Management Journal* **57**(1) (2014), 301–320.
- [9] L.M. Benveniste and P.A. Spindt, How investment bankers determine the offer price and allocation of new issues, *Journal of Financial Economics* **24**(2) (1989), 343–361.
- [10] J. Casillas, O. Cerdón, F. Herrera and P. Villar, A hybrid learning process for the knowledge base of a fuzzy rule-based system, in: *Proc. of the 2004 International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, Vol. 3, 2001, pp. 2189–2196.
- [11] A.J. Chalk and J.W.P. III, Initial Public Offerings: Daily Returns, Offering Types and the Price Effect, *Financial Analysts Journal* **43**(5) (1987), 65–69, ISSN 0015198X.
- [12] Y.-S. Chen, J.R.-S. Chen and C.-H. Cheng, An alternate method of examining IPO returns, *Intelligent Automation and Soft Computing* **16**(2) (2010), 151–162.
- [13] C.-H. Cheng and J.-W. Liu, OWA rough set to forecast the industrial growth rate, in: *2007 International Conference on Convergence Information Technology, ICCIT 2007*, 2007, pp. 1862–1867.
- [14] C.-H. Cheng, Y.-S. Chen and J.-S. Chen, Classifying initial returns of electronic firm's IPOs using entropy based rough sets in Taiwan trading systems, in: *Second International Conference on Innovative Computing, Information and Control, ICICIC 2007*, 2008.
- [15] S. Chiu, Fuzzy Model Identification based on cluster estimation, *Journal of Intelligent Fuzzy Systems* **2** (1994), 267–278.
- [16] S. Chiu, Method and software for extracting fuzzy classification rules by subtractive clustering, in: *1996 Biennial Conference of the North American Fuzzy Information Processing Society (NAFIPS)*, IEEE, 1996, pp. 461–465.
- [17] W. Choensawat and P. Polsiri, Financial institution failure prediction using adaptive neuro-fuzzy inference systems: Evidence from the east Asian economic crisis, *Journal of Advanced Computational Intelligence and Intelligent Informatics* **17**(1) (2013), 83–92.
- [18] O. Cerdón, F. Herrera, L. Magdalena and P. Villar, A genetic learning process for the scaling factors, granularity and contexts of the fuzzy rule-based system data base, *Information Sciences* **136** (2001), 85–107.
- [19] O. Cerdón, F.A.C. Gomide, F. Herrera, F. Hoffmann and L. Magdalena, Ten years of genetic fuzzy systems: current framework and new trends, *Fuzzy Sets and Systems* **141**(1) (2004), 5–31.
- [20] K. Doubravský, T. Meluzín and M. Dohnal, Decision-making on implementation of IPO under topological uncertainty, *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis* **63**(1) (2015), 193–200.
- [21] A. Esfahanipour, M. Goodarzi and R. Jahanbin, Analysis and forecasting of IPO underpricing, *Neural Computing and Applications* (2015), 1–8. doi:10.1007/s00521-015-1884-1.
- [22] L.J. Eshelman, The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination, in: *Foundations of Genetic Algorithms*, Vol. 1, 1991, pp. 265–283.
- [23] M.H.E. Frank, G. Holmes, B. Pfahringer, P. Reutemann and I.H. Witten, The WEKA Data Mining Software: An Update, *SIGKDD Explorations Newsletter* **11**(1) (2009), 10–18.
- [24] H.L. Ghazi and M.S. Abadeh, Mining fuzzy association rules with 2-tuple linguistic terms in stock market data by using genetic algorithm, in: *AISP 2012 - 16th CSI International Symposium on Artificial Intelligence and Signal Processing*, 2012, pp. 354–359.
- [25] J. Guan, D. Shi, J.M. Zurada and A.S. Levitan, Analyzing Massive Data Sets: An Adaptive Fuzzy Neural Approach for Prediction, with a Real Estate Illustration, *Journal of Organizational Computing and Electronic Commerce* **24**(1) (2014), 94–112.
- [26] K.W. Hanley, The Underpricing of Initial Public Offerings and the Partial Adjustment Phenomenon, *Journal of Financial Economics* **34**(2) (1993), 231–250.
- [27] F. Herrera, L. Martínez and F. Herrera, A 2-tuple Fuzzy Linguistic Representation Model for Computing with Words, *IEEE Transactions on Fuzzy Systems* **8** (1999), 746–752.
- [28] G. Holmes, M. Hall and E. Prank, Generating Rule Sets from Model Trees, in: *Advanced Topics in Artificial Intelligence*, Lecture Notes in Computer Science, Vol. 1747, 1999, pp. 1–12.
- [29] C.-F. Huang, M.-Y. Tsai, T.-N. Hsieh, L.-M. Kuo and B.R. Chang, A study of hybrid genetic-fuzzy models for IPO stock selection, in: *2012 International Conference on Fuzzy Theory and Its Applications, iFUZZY 2012*, 2012, pp. 357–362.
- [30] H. Ishibuchi, K. Nozaki, H. Tanaka, Y. Hosaka and M. Matsuda, Empirical study on learning in fuzzy systems by rice taste analysis, *Fuzzy Set and Systems* **64**(2) (1994), 129–144.
- [31] J.-S.R. Jang, ANFIS: adaptive-network-based fuzzy inference system, *IEEE Transactions on Systems, Man, and Cybernetics* **23**(3) (1993), 665–685.
- [32] N.K. Kasabov and Q. Song, DENFIS: dynamic evolving neural-fuzzy inference system and its application for time-series prediction, *IEEE Trans. on Fuzzy Systems* **10** (2002), 144–154.
- [33] J. Kim and N. Kasabov, HyFIS: Adaptive neuro-fuzzy inference systems and their application to nonlinear dynamical systems, *Neural Networks* **12**(9) (1999), 1301–1319.
- [34] B. Kirkulak and C. Davis, Underwriter reputation and underpricing: Evidence from the Japanese IPO market, *Pacific-Basin Finance Journal* **13**(4) (2005), 451–

- 470.
- [35] H.E. Leland and D.H. Pyle, Informational Asymmetries, Financial Structure and Financial Intermediation, *The Journal of Finance* **32**(2) (1977), 371–387.
 - [36] A. Ljungqvist and W.J.W. Jr., IPO Pricing in the Dot-Com Bubble, *The Journal of Finance* **58**(2) (2003), 723–752.
 - [37] M. Lowry and K.J. Murphy, Executive stock options and IPO underpricing, *Journal of Financial Economics* **85**(1) (2007), 39–65.
 - [38] M. Lowry, M.S. Officer and G.W. Schwert, The Variability of IPO Initial Returns, *The Journal of Finance* **65**(2) (2010), 425–465.
 - [39] S. Maiti and R.K. Tiwari, A comparative study of artificial neural networks, Bayesian neural networks and adaptive neuro-fuzzy inference system in groundwater level prediction, *Environmental Earth Sciences* **71**(7) (2014), 3147–3160.
 - [40] H.B. Mann and D.R. Whitney, On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other, *The Annals of Mathematical Statistics* **18**(1) (1947), 50–60.
 - [41] W.L. Megginson, and K.A. Weiss, Venture Capitalist Certification in Initial Public Offerings, *The Journal of Finance* **46**(3) (1991), 879–903, ISSN 1540-6261.
 - [42] D. Meng and Y. Xu, Credit rating method with heterogeneous information, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* **5009 LNAI** (2008), 739–746.
 - [43] J. Moody and C.J. Darken, Fast Learning in Networks of Locally-tuned Processing Units, *Neural Computation* **1**(2) (1989), 281–294, ISSN 0899-7667.
 - [44] H. Nomura, I. Hayashi and N. Wakami, A learning method of fuzzy inference rules by descent method, in: *Proc of the 1992 IEEE International Conference on Fuzzy Systems*, 1992, pp. 203–210.
 - [45] W. Pedrycz, Associations and rules in data mining: A link analysis, *International Journal of Intelligent Systems* **19**(7) (2004), 653–670.
 - [46] D. Petkovic, Adaptive neuro-fuzzy approach for estimation of wind speed distribution, *International Journal of Electrical Power & Energy Systems* **73**(0) (2015), 389–392.
 - [47] J.R. Quinlan, Learning with continuous classes, in: *Proceedings of the Australian Joint Conference on Artificial Intelligence*, Singapore, 1992, pp. 343–348.
 - [48] J.R. Ritter and I. Welch, A Review of IPO Activity, Pricing and Allocations, *Journal of Finance* **57**(4) (2002), 1795–1828.
 - [49] J.R. Ritter, Equilibrium in the Initial Public Offerings Market, *Annual Review of Financial Economics* **3**(1) (2011), 347–374.
 - [50] J.L. Rodgers and A.W. Nicewander, Thirteen Ways to Look at the Correlation Coefficient, *The American Statistician* **42**(1) (1988), 59–66.
 - [51] P.J. Rousseeuw and A.M. Leroy, *Robust Regression and Outlier Detection*, John Wiley & Sons, Inc., 1987. ISBN ISBN 0-471-85233-3.
 - [52] D.E. Rumelhart, G.E. Hinton and R.J. Williams, Learning Representations by Back-propagating Errors, *Nature* **323** (1986), 533–536.
 - [53] S.B. Smart and C.J. Zutter, Control as a motivation for underpricing: a comparison of dual and single-class IPOs, *Journal of Financial Economics* **69**(1) (2003), 85–110.
 - [54] A.J. Smola and B. Schölkopf, A Tutorial on Support Vector Regression, *Statistics and Computing* **14**(3) (2004), 199–222, ISSN 0960-3174.
 - [55] L. Teng, L., X. Yang, F. Tang, S. Teng and W. Zhang, 3-dimension evaluation method for stock investment based on 2-tuple linguistic, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* **8944** (2015), 326–339.
 - [56] Y.W. Teng and W.J. Wang, Constructing a user-friendly GA-based fuzzy system directly from numerical data, *IEEE Transactions on Systems, Man, and Cybernetics, Part. B* **34**(5) (2004), 2060–2070.
 - [57] V. Vella and W.L. Ng, Enhancing risk-adjusted performance of stock market intraday trading with Neuro-Fuzzy systems, *Neurocomputing* **141** (2014), 170–187.
 - [58] L.X. Wang and J.M. Mendel, Generating fuzzy rules by learning from examples, *IEEE Transactions on Systems, Man, and Cybernetics* **22**(6) (1992), 1414–1427.
 - [59] L.-Y. Wei, A GA-weighted ANFIS model based on multiple stock market volatility causality for TAIEX forecasting, *Applied Soft Computing Journal* **13**(2) (2013), 911–920.
 - [60] G. -Xiaohui, J.R. -Ritter and Z. -Zhu, Where Have All the IPOs Gone?, *Journal of Financial and Quantitative Analysis* **48** (2013), 1663–1692, ISSN 1756-6916.
 - [61] R.R. Yager and D.P. Filev, Generation of Fuzzy Rules by Mountain Clustering, *Journal of Intelligent and Fuzzy Systems* **2** (1994), 209–219.