Technological Forecasting & Social Change 176 (2022) 121466

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

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

The manipulation of Euribor: An analysis with machine learning classification techniques☆

Rubén Herrera [a], Francisco Climent [b], Pedro Carmona [b], Alexandre Momparler [b,*]

[a] Actuarial Consultant, Spain
[b] Universitat de València, Avinguda Tarongers, 46022 Valencia, Valencia, Spain

ARTICLE INFO

Keywords:
Euribor
Rate-fixing
Manipulation
Collusion
Panel bank
Machine learning
Classification

ABSTRACT

The manipulation of the Euro Interbank Offered Rate (Euribor) was an affair which had a great impact on international financial markets. This study tests whether advanced data processing techniques are capable of classifying Euribor panel banks as either manipulating or non-manipulating on the basis of patterns found in quotes submissions. For this purpose, panel banks' daily contributions have been studied and monthly variables obtained that denote different contribution patterns for Euribor panel banks. Thus, in accordance with the court verdict, banks are categorized as manipulating and non-manipulating and Machine Learning classification techniques such as Supervised Learning, Anomaly Detection and Cluster Analysis are applied in order to discriminate between convicted and acquitted banks. The results show that out of seven manipulative banks, five are detected by Machine Learning using Deep Learning algorithms, all five presenting very similar contribution patterns. This is consistent with Anomaly Detection which confirms that several manipulating banks present similar levels of abnormality in their contributions. In addition, the Cluster Analysis facilitates gathering the five most active banks in illicit actions. In conclusion, administrators and supervisors might find these techniques useful to detect potentially illicit actions by banks involved in the Euribor rate-setting process.

1. Introduction

The rigging of benchmark interest rates that took place at the beginning of the 21st century was a major financial scandal that involved both Euribor (Euro Interbank Offered Rate) and Libor (London Interbank Offered Rate) benchmarks rates.

This event caused great social upheaval as both Euribor and Libor are widely known rates used by financial institutions as fundamental benchmarks in their loan operations, especially in the case of mortgages.

In order to grasp the magnitude of the problem, note that according to the European Commission, the nominal value of financial contracts referenced to Euribor is 180 trillion euros (European Central Bank, 2019). In this context, it is worth considering how and why it is possible to carry out the manipulation of such an important financial benchmark supervised by the European Money Markets Institute (hereinafter, EMMI).

To understand why it is possible to manipulate Euribor, it is worth looking at its calculation methodology. According to the EMMI, the Euribor benchmark rate is calculated as a bounded average eliminating 15% of the highest and lowest quotes. These quotes are the rates at which panel banks estimate they can borrow from another panel bank in the interbank market for a specific day (EMMI, 2013).

Thus, one or more panel bank(s) may submit biased quotes seeking to move the benchmark in the direction that benefits them the most. This deceitful behavior can be explained if they are trying to benefit from their positions in financial derivatives referenced to Euribor. This incentive to manipulate quotes has been proved in the transcripts of bank management conversations (Boot et al., 2019). Other incentives for manipulation and their effect on third parties can be found in Gandhi et al. (2020) and Rodríguez-López et al. (2021).

After describing the framework in which manipulation takes place, note that the banks convicted of manipulation of the Euribor involved Barclays Bank, Credit Agricole, HSBC, JP Morgan Chase, Deutsche Bank, Royal Bank of Scotland, and Société Générale (European Commission, 2016, 2013).

There are many state-of-the-art techniques proposed to combat benchmark manipulation. Some of them are qualitative and some are quantitative, the latter with a high statistical component aimed at

☆ Grant PGC2018-093645-B-I00 funded by MCIN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe"

* Corresponding author
E-mail address: alexandre.momparler@uv.es (A. Momparler).

https://doi.org/10.1016/j.techfore.2021.121466
Received 8 June 2021; Received in revised form 10 December 2021; Accepted 29 December 2021
Available online 12 January 2022

detecting anomalies or patterns that help identify manipulation. However, the use of novel techniques to detect patterns in the data is rather unusual.

According to their focus and goals, the most important contributions in recent years can be classified in two groups. One group of studies is aimed simply at detecting manipulation with conventional statistical techniques and focusing on those aspects that make it possible to detect fraudulent behavior. Another large group of researchers have pinpointed the changes required in the governance and methodology of the index to reduce its potential for manipulation, some of them actually proposing the replacement of Euribor with a new benchmark.

The following section cites the most relevant studies from both streams and compares their findings in order to provide a comprehensive theoretical framework to this research.

This study attempts to make a contribution to the existing literature on the manipulation of benchmarks and the classification of manipulators and non-manipulators. Additionally, this paper intends to make a practical contribution to the literature by classifying manipulative and non-manipulative banks with classification techniques based on Machine Learning algorithms.

Therefore, the main research question is whether it is possible to distinguish manipulating banks from non-manipulating banks for the Euribor case. Then, we analyze whether it can be done with the help of Machine Learning algorithms based on both supervised and unsupervised techniques.

Consequently, this study pursues the accurate classifying of banks into manipulating and non-manipulating by applying Machine Learning techniques to quote submission patterns for each panel bank. This is done by creating monthly variables related to the contribution patterns of panel banks. From these variables, different predictive models have been trained in a sample covering the manipulation periods. The subsequent validation is done in a different sample to measure the classification capacity of the Machine Learning models in discriminating manipulating and non-manipulating banks. For this reason, supervised and unsupervised learning techniques have been used and the accuracy of these techniques is tested in different ways.

As a result of this study it is discussed how the Euribor supervisory authority could implement this type of technique and reduce the potential for future manipulations.

The hypothesis we test in this study is whether or not it is feasible to discriminate manipulating and non-manipulating banks on the basis of contribution patterns, accurately identifying those convicted banks that were most active in manipulating activities.

In terms of the contribution to the literature, note that this study is approaching the problem of benchmark index manipulation using for the first time Machine Learning algorithms, both supervised and unsupervised learning. Furthermore, another objective of this research is to propose the use of the tested techniques or any similar techniques by public authorities to improve the administration of the benchmark rates, which could also pave the way for the application of this type of advanced controls to other financial indicators of public interest.

Finally, this work is structured as follows: Section 2 covers the specialized literature on the manipulation of the Euribor and on the application of advanced techniques to financial problems; in Section 3, the methodology is described; in Section 4, the results obtained and their interpretation are presented; and finally, in Section 5, they include the main conclusions of the study.

2. Theoretical framework

The Euribor calculation methodology described in the previous section shows that the problem of manipulation comes from the ease with which participating banks can bias their inputs in the process. This implies that the contributing banks themselves are the relevant variable in any solution to minimize or solve the problem of potential manipulation. Otherwise, the calculation methodology could be changed so that panel banks can no longer make biased contributions.

Although methodological change is always an option, it could lead to costs that are difficult to measure considering all the economic agents involved in financial transactions referenced to Euribor. On the contrary, if current methodology is maintained exercising control over the potential bias of panel banks' contributions, then a wide variety of models are available, both traditional and modern, ranging from the simplest predictive models to complex algorithms, the latter being the subject of this academic paper.

Many studies have focused on the manipulation of benchmark rates since the first paper that highlighted anomalies in Libor was published (Mollenkamp and Whitehouse, 2008). This article argues that the contributions of Libor panel banks were significantly lower than expected based on their behavior in the interbank market. This would support the analysis of contribution patterns for the detection of manipulation.

The relevant literature comprises different types of research and some studies from international organizations stand out. These works are subsequent to the court ruling and contain different recommendations to avoid new manipulations, such as those of IOSCO (2013)[1], ESMA-EBA (2013) and[2] Financial Stability Board (2014). It is noteworthy that recommendations focus on changes in governance and methodology but fail to propose the use of advanced techniques to detect potential manipulations, which is the main contribution of this paper.

Snider and Youle (2012) made a very interesting contribution to the literature showing how convicted banks often submit quotes off the Euribor calculation range, claiming a clear incentive to manipulation due to their positions in derivatives referenced to the benchmark rate. A similar study for the case of Libor is that of Gandhi et al. (2019). Consequently, the purpose of using advanced data processing techniques in this paper is precisely to identify these types of contribution patterns. In line with the results of these studies, our perception is that in the framework of benchmark indices that exclude extreme submissions, providing extreme submissions is shifting the calculation window in the desired direction.

In line with the previous consideration, Herrera et al. (2020) warn of similarities in the contribution patterns of manipulating banks, characterized by a large proportion of contributions excluded from the final calculation. In addition, they point out that the potential for collusive manipulation by several banks of the 3-month and 12-month Euribor rates may represent more than one basis point. This result is consistent with that estimated by Eisl et al. (2017) for the 3-month Euribor and Libor rates. In line with the results of these studies, our understanding is that the ability of a single bank to move the index to its advantage (by providing deceitful quote submissions) is lower than that of a multi-bank cartel with the same goal. This issue is of vital importance in the present investigation and was also pointed out by the authorities investigating benchmark manipulation (European Commission, 2016, 2013).

Other research works focus on methodological changes in Euribor rate-setting or even its replacement by another benchmark. For instance, Youle (2014) supports the substitution of the calculation for the bounded mean by that of the median for the case of Libor as it is a statistic less sensitive to extreme values. Brousseau et al. (2013) propose the replacement of the index by the Overnight Index Swap, while Hou and Skeie (2014) support its replacement by the General Collateral Finance Repurchase Agreement Index, assuming these alternative indexes are harder to manipulate. These solutions may have unintended effects, as changing a benchmark to which a large number of contracts are referenced can lead to legal claims. In addition, there is a potential problem that the resulting new index rate is not even close to that of the current benchmark, generating significant gains or losses to people who are not involved in the manipulation.

[1] International Organization of Securities Commissions

[2] European Securities Market Authority-European Banking Authority

Other authors propose solutions to manipulation without changing the current Euribor methodology. Duffie and Stein (2015) propose to toughen sanctions to discourage manipulation. Instead, Snider and Youle (2012) propose increasing the number of panelists to reduce the possibility of manipulation, a proposal which is in line with that of Herrera et al. (2020). These last research works pursue a solution to manipulation without the potential problems of a methodological change or replacement of the index. In theory, many of the measures proposed in these papers can reduce the potential for manipulation of the benchmark, but they cannot ensure that manipulation will not occur altogether. Therefore, detecting potential manipulations from submission patterns is essential and it is the goal of this study.

Finally, as discussed in the literature review, the cause of manipulation is related to the submission of deceitful quotes by panel banks to move the index to their advantage. The primary objective of this work is to detect manipulation on the basis of submission patterns and Machine Learning techniques are employed to capture complex patterns in the data. The use of Machine Learning methods has become widespread in the field of finance due to their outstanding results in different environments. Machine Learning can be defined as a technique based on computer programming oriented to automatic learning from data (Samuel, 1959). Although this type of technique has not yet been applied to the problem of benchmark manipulation, it is used intensively in the field of finance.

In the literature, we find some examples of the use of Machine Learning techniques to analyze other manipulation or financial distress situations: Krauss et al. (2017) employ a combination of close neighbors and random forests techniques to explore arbitrage of the S&P 500 with precise results; Carmona et al. (2019) use Extreme Gradient Boosting to predict bankruptcy situations in the US banking sector, in what could be considered a classification problem similar to the one proposed in the present work; Ben Jabeur et al. (2021) employ the CatBoost algorithm for corporate failure prediction; Gan et al. (2020) propose a Machine Learning method to price average options accurately; and Palacio (2019) uses Machine Learning for fraud detection in non-life insurance.

3. Methodology

This section is divided into sub-sections that explain the different methodological steps. As an overview, a schematic of the methodological process followed can be found in Fig. 1.

3.1. Data and feature engineering

The data used in this study are the daily quotes submitted by panel banks and final Euribor rates from January 2004 to November 2018 (EMMI, 2019). The choice of the range of years is justified because this is the period for which information on panel banks' contributions is available on the EMMI website.

Before addressing the actual research problem, a data quality analysis is carried out, automating a process in R to replicate the calculation of the final Euribor daily rate, according to the daily contributions and the methodology applied in the Index computation. The result is the detection of some minor errors in the data that, once corrected, have allowed the final rate to be replicated for every single day in the study period.

In a subsequent step, variables representative of contribution patterns are included, which was the foundation of the previous study by Herrera et al. (2020). To study these variables we use data analysis techniques and detect behaviors that allow identification of banks convicted of manipulation. This is carried out for periods of 3 and 12 months, however, due to the similarity of the results and for the sake of simplicity, the explanation of the results focuses only on the 12-month period.

Monthly variables are generated from daily data in a feature engineering process. The main reason for this is the immense number of variables that would involve working with daily data. Another reason is that it is difficult to generate a variable capable of reflecting manipulation or non-manipulation on a daily basis, since an unusual contribution on a certain day can have very different explanations. In this regard, with monthly contributions it is intended to collect the contribution characteristics of each month of the study period for each panel bank.

Note that we are considering only variables related to contribution patterns because the study focuses on this kind of information. However, if benchmark administrators were to use these techniques to detect illegal behavior, it could also take into account other variables that summarize the investment behavior of each panelist for each month or the interest rates applied in operations settled in the interbank market.

In reference to the variable Manipulation, it is formed as a binary variable with a value of 0 for unconvicted banks and a value of 1 for those convicted, that is, Barclays Bank, Credit Agricole, HSBC, JP Morgan Chase, Deutsche Bank, Royal Bank of Scotland, and Société Générale. Thus, 18 additional variables are presented in Table 1.

The major distinctive contribution in this section, as compared to those of other authors, is that a thorough data cleansing and a replication of the historical index are performed to ensure data quality. To the best of our knowledge, this approach is entirely new in the related literature. Additionally, a transformation process of the data has been made to apply advanced algorithms, which is novel due to the innovative techniques applied to deal with the manipulation of the Euribor.

3.2. Application of supervised machine learning techniques

In the application of Machine Learning techniques, we use the H2O platform (H2O.ai 2020a) which includes this type of programmed techniques and allows for using R as interface (r-project.org, 2021). With the use of this platform, we intend to show the potential of these techniques to offer a possible solution to the manipulation of Euribor.

The use of an H2O environment (H2O.ai, 2020b) allows for several algorithms to be tested simultaneously on the data, among which Random Forest, GLM, Gradient Boosting Machine and Deep Neural

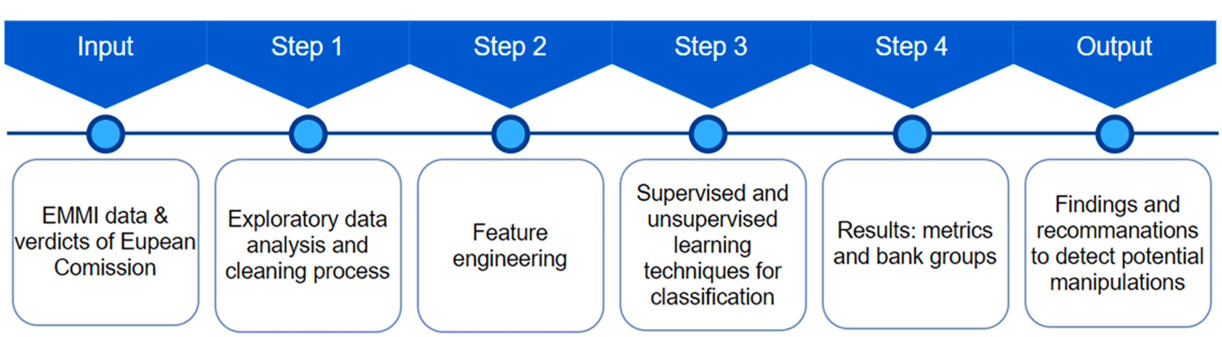

Fig. 1. Methodology scheme.

Table 1
Types of monthly variables identifying contribution patterns.

Number of types	Types of variables
1	Monthly average of the absolute daily variation of contributions
2	Monthly average of the relative daily variation of contributions
3	Percentage of times above the calculation window in the month
4	Percentage of times below the calculation window in the month
5	Percentage of times within the calculation window in the month
6	Percentage of times without contribution in the month
7	Percentage of contributions among the three maximums in the month
8	Percentage of contributions among the minimum three in the month
9	Percentage of non-extreme contributions
10	Average of contributions for the month
11	Standard deviation of contributions for the month
12	Median of contributions for the month
13	Mode of the month contributions
14	Percentage of times contributed mode of the month
15	Maximum of contributions for the month
16	Percentage of times contributed the maximum of the month
17	Minimum of contributions for the month
18	Percentage of times contributed the minimum of the month

Networks stand out. This variety of proven models has been deemed optimal to offer an effective solution to the manipulation problem. The H2O platform is used because it allows the application of many Machine Learning models in a fast and efficient way and allows their use from a free software environment such as R. H2O supports the most widely used statistical and machine learning algorithms including Gradient Boosted Machines, Generalized Linear Models, Deep Learning and other algorithms.

Thus, the Machine Learning methodology is applied as follows. In the first place, the data is processed, the variables being those introduced in the previous subsection and the banks being the observations. Next, missing values are handled because some panel banks are not present during the entire period of study. Consequently, we eliminate the variables with more than 35% of missing values following a methodology similar to that proposed by Momparler et al. (2016), corresponding mainly to the final part of the period due to the fact that in those years there are fewer panel banks. This procedure is applied because the missing values could pose a problem in the use of algorithms. Likewise, for variables with less than 35% missing values, as they add no information, these missing values have been replaced by the mean of the rest of the observations.

After organizing the data, Machine Learning has been implemented using the different algorithms mentioned above, labeling the binary variable Manipulation as the response variable and finding models that, based on the rest of the variables, allow for the classification of banks into two different categories: manipulating and not-manipulating.

Regarding the validation of the fitted models, in order to allow a good generalization of the results, we use K Fold Cross-Validation technique (H2O.ai, 2020c). This is a way to validate the model internally without losing any data for later validation. This technique consists of estimating $K+1$ models, one final model that is trained with all the data and other K models that are validated on 1 / K proportion of the data. This 1 / K proportion of validation sample is different for each of the K times the process is repeated. The results of this validation process are obtained as an average of the predictions of the K models subjected to validation. Furthermore, this technique is explained in depth by Pereira et al. (2009) and used by Carmona et al. (2019).

In addition, the use of K Fold Cross-Validation is justified by the generation of more consistent results. In this regard, we avoid training the model in part of the period and validating it in another, since the form of manipulation may have changed from period to period. Consequently, the use of K-Fold Cross-Validation is one of the best alternatives to assess the effectiveness of the models obtained, especially when both the size of the sample and the number of observations are relatively small.

Regarding the estimation of the models and the K chosen for Cross-Validation, the following criteria have been followed. Testing a multitude of algorithms corresponding to different models implies a high stochasticity component in the results, both in the best model obtained and in the results of the validation samples. That is, the results can have a wide range of variation depending on how the calibration of the models is initially done. Accordingly, the algorithms are run not just once but a total of 30 times because it is a number large enough to obtain reliable results that are less influenced by randomness. Additionally, these 30 times have been run for both 5-Fold Cross-Validation and 10-Fold Cross-Validation. These two modalities are the most used and with this we ensure that the Cross-Validation results are consistent.

Finally, the results discussion will focus both on the best model obtained from all those tested (Random Forest, GLM, Gradient Boosting Machine and Deep Neural Networks) and the average of the performance measures of the 30 executions of each model to better assess the global results. This is the approach used for each type of K-Fold Cross-Validation. The above methodology has been applied throughout the entire period in which the manipulations occurred (from January 2005 to May 2012) and also in six-month terms within the manipulation period so we can compare the results in different terms.

Table 2 shows which magnitudes of the outputs obtained, after the application of different techniques, are chosen to assess the effectiveness of this technique when it comes to correctly classifying manipulating banks and non-manipulating banks. Table 2 also shows which magnitudes of the outputs obtained, after applying the algorithms, are selected to assess the effectiveness of this technique in correctly classifying manipulating and non-manipulating banks.

As an extension to Table 2, note that the AUC (Area Under the Curve) for the problem under study is a measure that determines the ability of the model to distinguish between manipulating and non-manipulating entities. Regarding the interpretation of their possible values, a value close to 1 of the AUC would indicate for this case that the model classifies both manipulating and non-manipulating banks correctly. An AUC value around 0.5 would indicate that the model has no classifying power, and a value close to 0 would imply that the model classifies manipulating banks as non-manipulating and non-manipulating banks as manipulating.

Therefore, according to the results we confirm the hypothesis that it is possible to discriminate between manipulating and non-manipulating banks based on their contribution patterns.

As for the use of unsupervised Machine Learning techniques to detect manipulators and non-manipulators, this is something new in the literature. Even though the problem has been addressed with unsupervised anomaly detection techniques before, it was done from a completely different perspective. Moreover, this approach to the problem is of interest, because it allows us to tie the Euribor problem to other financial problems in which Machine Learning has already been applied.

Table 2
Measures chosen to measure the efficiency of the models obtained.

Measure	Explanation
Best Model	Best model of those tested
AUC	A measure of the model's ability to distinguish between different classes (area under the curve)
% non-manipulative error	Proportion of non-manipulators misclassified over the total of these
% manipulative error	Proportion of poorly classified manipulators over the total of these
% total error	Proportion of classification errors over total banks
Importance of most important variable	Importance percentage of the most important variable
Importance of the least important variable	Importance percentage of the least important variable

3.3. Anomaly detection

Anomaly Detection (also applied through H2O) is a technique used to detect anomalous values in a data sample (H2O.ai, 2020d). It is done through an Autoencoder, which is a type of artificial neural network that treats data in an unsupervised way and seeks to form smaller sets of data by eliminating the noise in the process (Sakurada and Yairi, 2014). Afterwards, it reconstructs the same variables with less anomaly, observing the difference between the original variables and the reconstructed ones.

This technique allows for studying which banks of the panel had the greatest anomalies, and whether or not these are the manipulating banks. It is a complementary analysis to Machine Learning and has been applied using the R software, measuring the anomaly for each panelist using the MSE (mean squared error). Based on this measure, the banks have been ordered from highest to lowest anomaly. Then we identify and analyze the manipulating banks found among the 10 and 15 most anomalous banks. This process has been repeated 30 times for the three periods considered.

Similar techniques for anomaly detection have been applied previously. Yet, what is new in this study is the way in which the results are interpreted. Typically, these techniques measure the level of anomaly of each observation based on the explanatory variables, and then classify them accordingly. The most anomalous observations are labelled as fraud depending on the percentage of anomalies observed empirically.

However, in this paper, we have analyzed whether manipulating banks (convicted banks) presented similar levels of anomaly, even if they were not the most significant ones. We proceed this way because anomalies in contribution patterns may arise not only from potential manipulators, but also from the specific creditworthiness of each financial institution.

3.4. Cluster analysis

This analysis has been implemented to see if it was possible to ratify the results obtained from Machine Learning and Anomaly Detection. Above all, we intend to detect relationships that could indicate possible collusions in a visual manner, as other authors have anticipated (Hartheiser and Spieser, 2010).

Thus, a cluster analysis has been carried out using the K-means method, using the R software and the H2O library, with the purpose of dividing the total number of entities into two groups based on the same variables previously considered. With this, we verify if the banks are classified into groups similar to manipulating and non-manipulating, repeating the process 100 times.

In addition, a hierarchical cluster analysis has been implemented using the single Gower method (Gower,1967). This type of cluster consists of grouping relatively similar banks on the basis of variables, and these groups are grouped hierarchically. In this way we can observe if some of the manipulating banks are grouped. This analysis has been implemented through R software and the stats library.

In order to test the initial hypothesis, we examine whether the clusters coincide with the actual clustering of handlers and non-handlers, both for anomaly detection and cluster analysis. The reason why we use these two unsupervised techniques is to complement supervised learning techniques. Because being related to criminal activities is a sensitive target variable, we deem appropriate the use of techniques that do not rely on the target variable to classify banks. Then, we compare the results with techniques that do use the target variable to calibrate and predict bank membership to one group or the other.

4. Results

4.1. Supervised Machine learning

The manipulation period anticipated in the previous sub-section, from January 2005 to May 2012, has been established according to the information provided by the European Commission (2013) and the European Commission (2016). Likewise, we choose the first six months of 2007 because it is an intermediate period within the period indicated by the authorities as the one with the greatest manipulating activity. In this period, the Royal Bank of Scotland (RBS) was not on the panel, hence the results are shown in relative terms. In addition, with the choice of this six-month period, we test the classification capacity excluding the RBS as the descriptive analysis performed by Herrera et al. (2020) shows different contribution patterns between RBS and the rest of the convicted banks.

Results shown in Table 3 are for the mean of the 30 simulations made for the three periods using Supervised Machine Learning algorithms for the two types of K-fold Cross-Validation implemented.

From Table 3, it should be noted that the only result that cannot be interpreted as an average of the 30 runs of different classification algorithms is the content of the row that indicates "Model". This row shows which model best classifies convicted and non-convicted panel banks for each period and each type of K-fold Cross-Validation, taking into account the total number of runs. In this case, the most accurate is Deep Learning, that is, neural network models. This class of models allow for finding complex relationships implicit in the data and presents better results than other learning algorithms when the amount of data is greater (Ng, 2015). In addition, this result is consistent with the classification problem discussed here, since there are already Deep Leaning applications aimed at predicting or classifying data similar to interest rates, as would be the case for predicting inflation (Yin and Ge, 2012).

The remaining variables are a measure of the aforementioned number of runs, the meaning of which is described in Table 2.

Regarding the rest of the results in Table 3, the high AUC values stand out, situated between 0.78 and 0.88, which in general terms implies a high capacity for classification between manipulating and non-manipulating banks.

Furthermore, the total error shows values around 18% for all K-fold Cross-Validation periods and techniques, which is a relatively low value. However, this is a more common measure in problems with several classes and balanced data, which is not the case with our analysis, as illustrated in Table 4.

Table 4 shows the number of both manipulating and non-manipulating banks, showing that we are trying to detect a very small number out of the total number of panel banks.

In any case, the total error is an informative measure since a high value would indicate that one of the two categories is not well classified. However, it must be analyzed together with the proportion of non-manipulators misclassified over the total of these (false positives) and the proportion of poorly classified manipulators over the total of these (false negatives). Table 3 provides alternative measures to assess the models' performance.

Regarding the proportion of error in the classification of manipulating banks or false negative error, it is over 30% in the total period and in the manipulation period and in the six-month period 13% and 14% for $K = 5$ and $K = 10$ respectively. The interpretation of these values is that, for the two longest periods of time, 30% of error corresponds to two banks of the seven manipulating banks, whereas for the six months the 13%–14% of error represents one bank of the six manipulating banks. Thus, the difference between the two long periods and the shorter six-month period in terms of manipulating banks is the absence of RBS in the shorter period, which seems to be a constant error in the classification of manipulating banks due to RBS.

A justification may be that RBS was on the panel for a short time and manipulated only for eight months, while the other convicted banks manipulated for two to three years according to the European Commission (2013). Thus, RBS role in the manipulation could be residual or different from the rest.

In reference to the error in the non-manipulating or false positive error which is around 15% in the extended periods, being somewhat

Table 3
Average of Supervised Machine Learning results of execution materialized for a period of 12 months.

	Total period		January 2005–May 2012		January 2007–June 2007	
	5 fold CV	10 fold CV	5 fold CV	10 fold CV	5 fold CV	10 fold CV
Model	DeepLearning	DeepLearning	DeepLearning	DeepLearning	DeepLearning	DeepLearning
AUC	0.84	0.8	0.8	0.88	0.78	0.83
% Manipulator Error	30%	29%	32%	27%	13%	14%
% No Manip Error	14%	16%	15%	10%	23%	19%
% Total Error	16%	18%	17%	12%	22%	19%
Importance of most important variable	0.0006	0.0006	0.0008	0.0008	0.0125	0.0126
Importance of the least important variable	0.0004	0.0004	0.0005	0.0004	0.0075	0.0075

Table 4
Distribution of banks between manipulators and non-manipulators.

Period	No Manipulators	Manipulators	Total Banks
Total period	45	7	52
January 2005–May 2012	45	7	52
January 2007–June 2007	41	6	47

higher for the six-month term, about 7 banks, indicating that some non-manipulating banks present contribution patterns similar to those of manipulating banks.

Moreover, in reference to the models' global accuracy on average for each term and K tested, of all the banks there has been a total success that has ranged between 78% and 88%. This success rate we regard as positive in view of the complexity of the problem, and it leads to conclude that it is possible to detect manipulating banks based on the contribution behavior of the panelists with a high success rate. Also, the reduced value of the complementary measure of the total error in comparison with the error in manipulating banks confirms that the magnitude of this last type of error is due to the reduced number of manipulating banks.

Additionally, a tiny difference is observed between the most important and the least important variables, which leads to the conclusion that no variable is singled out in the classification.

When comparing the length of different study periods, it can be concluded that the detection of manipulating banks is easier the longer the period analyzed, despite the fact that for six-month periods it is also possible to identify the manipulating banks (Table 4).

Table 5 shows the model based on the AUC with the highest classification capacity of the 30 runs implemented for each period and each K-fold Cross-Validation modality (Deep Learning). In other words, while Table 3 gives an idea of the results of the Machine Learning technique for the problem under study reducing the randomness of the results, Table 5 shows the maximum potential of these techniques to detect manipulating activity.

Table 5 is analogous to Table 4 but it shows the results for the best models for each of the periods and K-fold used, in order to show the maximum potential in classifying manipulating and non-manipulating banks.

In the best models, the AUC takes values close to 0.9 or even higher, that is, they are values that indicate a high classifying capacity. For the total error in Table 5, more varying values are observed per period and type of Cross-Validation than those in Table 3, even though no big differences are found.

On the basis of classification errors of manipulating and non-manipulating banks, there are two possible outcomes according to the period considered and the type of Cross-Validation: either there is a greater error in classifying manipulating banks and smaller or no error in classifying non-manipulating banks or the opposite.

In this way, a relevant result is the 57% error in the manipulating banks in the total period and $K = 5$, which implies that three manipulating banks are properly classified and four manipulating banks are not, correctly classifying the 45 non-manipulating banks. Something similar happens for the same period and $K = 10$, where two manipulating banks out of seven and one non-manipulating bank out of 45 are classified wrongly. Both results seem to indicate that the obtained models are detecting the most active manipulating banks or those with the longest period of involvement in manipulating activities. This is highly informative for regulators and is consistent with the report by the European Commission (2013).

In any event, the common factor for each period and type of Cross-Validation focuses on total error and global accuracy that mark good results in the classification. Table 5 shows that global accuracy values are close to or greater than 90%. Furthermore, comparing the results for both K-fold Cross-Validation techniques, similar results are found.

In addition, errors in the identification of manipulating banks (two to three banks) may be due to the short-lived role of RBS in illicit activities, and the likely discontinuous participation of Barclays Bank, which is the bank that revealed the cartel behavior according to the European Commission (2013).

Finally, the results support the suitability of the above-mentioned techniques combined with complementary techniques that are introduced in the following subsections for the EMMI to detect manipulating behavior in contribution patterns.

4.2. *Unsupervised machine learning with anomaly detection*

The results of the Anomaly Detection are compiled in Table 6 below.

Table 6 shows the average number of manipulators within the 10 and 15 most anomalous banks in the 30 executions made for the three periods. Also, it shows which banks have been the most anomalous manipulators, in order to see if they are the same in the three periods covered.

Note that results are similar for the three periods. Thus, among the

Table 5
Best supervised models of materialized executions for the 12-month period.

	Total period		January 2005–May 2012		January 2007–June 2007	
	5 fold CV	10 fold CV	5 fold CV	10 fold CV	5 fold CV	10 fold CV
Model	DeepLearning	DeepLearning	DeepLearning	DeepLearning	DeepLearning	DeepLearning
AUC	0.89	0.87	0.89	0.93	0.83	0.89
% Manipulator Error	57%	29%	0%	29%	0%	33%
% No Manip Error	0%	2%	18%	4%	24%	5%
% Total Error	8%	6%	15%	8%	21%	9%
Importance first Var	0.0005	0.0007	0.0007	0.0008	0.0123	0.0137
Importance last Var	0.0004	0.0004	0.0005	0.0004	0.0074	0.0067

Table 6
Anomaly Detection results.

Period	Total period	January 2005–May 2012	January 2007–June 2007
Mean Banks Manip TOP 10	2.1	1.9	1.9
Mean Banks Manip TOP 15	3.3	3.2	2.8
Mode 1st Manip in Anomaly	JP MORGAN CHASE	DEUTSCHE BANK	HSBC
Mode 2nd Manip in Anomaly	SOCIÉTÉ GÉNÉRALE	JP MORGAN CHASE	JP MORGAN CHASE
Mode 3rd Manip in Anomaly	DEUTSCHE BANK	SOCIÉTÉ GÉNÉRALE	SOCIÉTÉ GÉNÉRALE
Mode 4th Manip in Anomaly	CRÉDIT AGRICOLE	CRÉDIT AGRICOLE	–

* The "Model K Manip in Anomaly" rows collect from the manipulator banks which ones have presented the highest level of anomaly.

10 banks with the most anomalous observations, there are two out of 47 banks for the six-month period and two out of 52 for the other two periods. Similarly, for the six-month period, among the first 15 there are around three manipulating banks, while in the other periods there are more than three banks, that is, in some simulations there are three and in other simulations there are four.

The results in Table 6 are consistent with those of Tables 3 and 5 where, on the basis of contribution patterns, correctly classified manipulating banks range 4 to 5. Additionally, Table 6 shows the manipulating banks with the greatest anomalies, which are the first, second, third and fourth most anomalous consistently.

Thus, it is found that for long periods, JP Morgan, Société Générale, Deutsche Bank, and Credit Agricole are the most anomalous manipulating banks, virtually in the same order. In addition, it has been found that these banks are arranged together for most simulations. In short, they present a very similar degree of anomaly in their observations, which suggests the existence of a cartel to carry out sustained manipulating activities over time.

Lastly, the anomalies may be due to different bank situations, different moments of presence in the total period, or even hypothetical undetected manipulations. In any case, the manipulation may have been carried out in many ways, and each manipulating bank may play a different role and have a different period of action. However, based on these results and those of the previous section, it is detected that either four or five banks have played a fundamental role in this entire manipulation process.

4.3. Unsupervised machine learning with cluster analysis

Fig. 2 shows the dendrogram of the applied cluster analysis. The names of the panelists have been modified so that non-manipulating banks are represented with a zero and manipulating banks with one, thus visually identifying the banks involved and their possible relationships.

Also, in Fig. 2 the proximity of a group of four condemned banks can be seen for the entire period, which indicates a possible collusion. Furthermore, those banks are JP Morgan, Société Générale, Deutsche Bank, and Credit Agricole, which coincide with the banks identified in the Anomaly Detection. For the manipulation period, of the 52 banks there are five convicted banks relatively close. As such, these results confirm those obtained in the previous subsections, especially in reference to a group of four to five banks whose contribution patterns allow them to be identified as manipulating banks comprising a cartel.

5. Conclusions

Machine Learning techniques have proven to be an appropriate tool to capture complex data relationships and identify manipulating banks

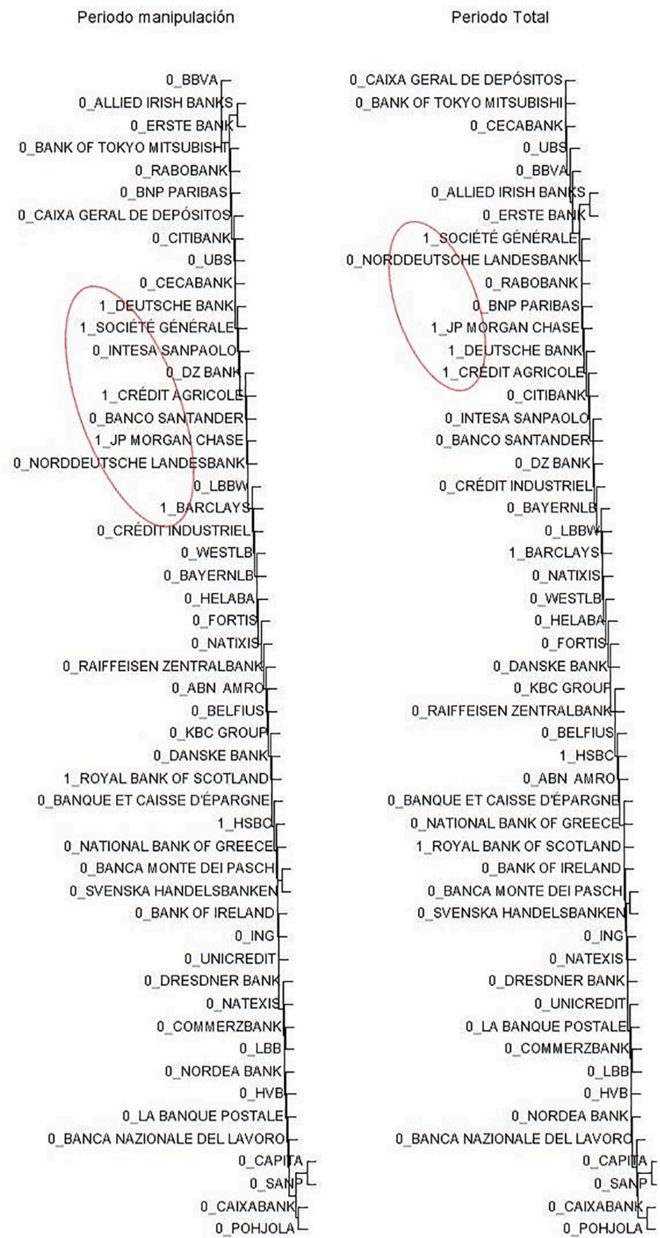

Fig. 2. Dendrogram of hierarchical cluster for total and manipulation periods.

on the basis of their contribution patterns. In this sense, the results of supervised Machine Learning show that it is possible to classify manipulating and non-manipulating banks with an accuracy higher than 90%. We find these techniques appropriate because of their high classifying power.

The unsupervised machine learning with Anomaly Detection has corroborated the possible existence of a cartel formed by JP Morgan, Société Générale, Deutsche Bank, and Credit Agricole that is in agreement with the Cluster Analysis and with the comments by the European Commission (2016).

Consequently, the current methodology should be maintained to avoid potential legal problems and deferrals with the historical rate associated with a methodological change, as proven for the median and mode, but also for real transactions according to EMMI (2017).

Subsidiarily, we suggest the use of Machine Learning techniques to control the existence of possible manipulations as the major contribution of this research. Thus, assuming the incumbent authority does not know which banks are manipulating, it can carry out a descriptive analysis to detect potential relationships between the contributions of panelists, along the lines of those that occurred from 2005 to 2012. On the basis of the relationships detected, the authority could automate some testing procedure for various combinations of manipulating and non-manipulating banks. If any of those groupings of panelists shows signs of abnormal behavior, an investigation should be initiated.

Furthermore, given its privileged position, the authority has the capability to construct other variables beyond those derived from contribution patterns. These variables could be related to the positions in derivatives referenced to the Euribor of the panelists, or else informative of the actual operations closed by panel banks in the interbank market. In this way, the techniques proposed here could be even more decisive in capturing possible manipulations.

With regard to the implementation by the incumbent authority of Machine Learning and other complementary techniques, we suggest the possibility of applying them in a six-month mobile window period, that is, testing the last six-month period every month. This study shows that it is possible to detect criminal patterns in that time frame. Therefore, the continuous application would allow monitoring of the contribution and other possible variables in the short term. Also, we recommend the implementation of the analysis over periods of 4 or 5 years at least once a year because it is proven that for these time periods the classification of convicted and unconvicted is fairly successful. This last recommendation would allow the detection of manipulations from a broader perspective which is complementary to the short-term assessment.

Concerning the theoretical gap in the existing literature, this study shows that the classification of panel banks into manipulating and non-manipulating is possible, at least partially, on the basis of statistical techniques.

Several areas of research are proposed as possible extensions of this work. Accordingly, it is proposed to study the identification capacity of manipulating banks in the central manipulation period the way we recommended to the incumbent authority in this work. That is, to check whether the application of the proposed techniques each month of the period for the previous semester from July 2005 to May 2012 allows the banks that were involved in the manipulation to be correctly identified.

In addition, it is proposed to repeat the study with additional variables related to the contribution process and other interesting aspects. These aspects could be the panelists investment patterns in derivatives referenced to the Euribor or any other interest rate benchmarks applied in the settlement of interbank transactions.

The limitations of the study are to some extent related to the possible extensions stated above and that is because the conclusions are limited to the Euribor benchmark. In addition, the predictive techniques are applied with monthly data, which is not the only possible approach. In relation to the data, the incumbent authority stopped publishing banks' contributions in December 2018, which means that the analysis cannot be updated to include more recent years.

Finally, the objective of this study has been achieved as it shows it is possible to discriminate between manipulating and non-manipulating banks with great accuracy, and most importantly, with the correct classification of all the banks with high involvement in manipulating activities and with very little error in the classification of non-manipulating banks.

Authorship statement

All persons who meet authorship criteria are listed as authors, and all authors certify that they have participated sufficiently in the work to take public responsibility for the content, including participation in the concept, design, analysis, writing, or revision of the manuscript. Furthermore, each author certifies that this material or similar material has not been and will not be submitted to or published in any other publication before its appearance in **Technological Forecasting & Social Change.**

Please indicate the specific contributions made by each author (list the authors' initials followed by their surnames, e.g., Y.L. Cheung). The name of each author must appear at least once in each of the three categories below.

Category 1 Conception and design of study: **R. Herrera, F. Climent;** acquisition of data: **R. Herrera, F. Climent;** analysis and/or interpretation of data: **R. Herrera, F. Climent, P. Carmona, A. Momparler;**

Category 2 Drafting the manuscript: **R. Herrera, F. Climent, P. Carmona, A. Momparler;** revising the manuscript critically for important intellectual content: **R. Herrera, F. Climent, P. Carmona,**

A. Momparler; Category 3 Approval of the version of the manuscript to be published (the names of all authors must be listed): **R. Herrera, F. Climent, P. Carmona, A. Momparler**

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.techfore.2021.121466.

References

Ben Jabeur, S., Gharib, C., Mefteh-Wali, S., Wissal, Ben Arfi, W., 2021. CatBoost model and artificial intelligence techniques for corporate failure prediction. Technol. Forecast. Soc. Change (166), 120650. https://doi.org/10.1016/j.techfore.2021.120658 https://doi.org/.

Boot, N., Klein, T., Schinkel, M.P., 2019. Collusive Benchmark Rates Fixing. Tinbergen Institute Rotterdam.

Brousseau, V., Chailloux, A., Durré, A., 2013. The LIBOR scandal: What's next? A possible way forward. Retrieved June 17, 2019, from VOX CEPR Policy Portal, columns, By Date, 2013, View all 47 columns. https://voxeu.org/article/libor-scandal-and-reform.

Carmona, P., Climent, F., Momparler, A., 2019. Predicting failure in the U.S. banking sector: an extreme gradient boosting approach. Int. Rev. Econ. Finance. 61 (C), 304–323. https://doi.org/10.1016/j.iref.2018.03.008 https://doi.org/.

Duffie, D., Stein, J., 2015. Reforming LIBOR and other financial market benchmarks. J. Econ. Perspect. 29 (2), 191–212. https://doi.org/10.1257/jep.29.2.191 https://doi.org/.

Eisl, A., Jankowitsch, R., Subrahmanyam, M., 2017. The manipulation potential of Libor and Euribor. Eur. Financial Manag. 23 (4), 604–647. https://doi.org/10.1111/eufm.12126 https://doi.org/.

EMMI, 2013. Euribor Code of Conduct. Retrieved June 16, 2019, from EMMI, Euribor, Euribor Reform, Overview of the Euribor Reform: Main Measures. https://www.emmi-benchmarks.eu/assets/files/Euribor_code_conduct.pdf.

EMMI, 2017. Pre-Live Verification Program Outcome and Way Forward. https://www.emmi-benchmarks.eu/benchmarks/euribor/methodology.

EMMI, 2019. Retrieved June 2019, from EMMI, Euribor, Rates, Data. https://www.emmi-benchmarks.eu/euribor-org/euribor-rates.html.

European Central Bank, 2019. Why Are Benchmark Rates So important? Retrieved June 15, 2019, from. https://www.ecb.europa.eu/explainers/tellmemore/html/benchmark_rates_qa.en.html.

European Commission, 2013. AMENDED1 - Antitrust: Commission fines banks €1.49 billion for participating in cartels in the interest rate derivatives industry. Retrieved June 16 2019,, from European Commissio. https://europa.eu/rapid/press-release_IP-13-1208_en.htm.

European Commission, 2016. Antitrust: Commission fines Crédit Agricole, HSBC and JPMorgan Chase €485 million for euro interest rate derivatives cartel. Retrieved June 16 2019,, from European Commission. http://europa.eu/rapid/press-release_IP-16-4304_en.htm.

European Securities Market Authority-European Banking Authority, 2013. ESMA-EBA Principles for Benchmark Setting Processes in the EU. In: https://www.esma.europa.eu/sites/default/files/library/2015/11/2013-659_esma-eba_principles_for_benchmark-setting_processes_in_the_eu.pdf.

Financial Stability Board, 2014. Market Participants Group on Reforming Interest Rate Benchmarks Final Report. https://www.fsb.org/wpcontent/uploads/r_140722b.pdf?page_moved=1.

Gan, L., Wang, H., Yang, Z., 2020. Machine learning solutions to challenges in finance: an application to the pricing of financial products. Technol. Forecast. Soc. Change (153), 119928. https://doi.org/10.1016/j.techfore.2020.119928 https://doi.org/.

Gandhi, P., Golez, B., Jackwerth, J., Plazzi, A., 2019. Financial market misconduct and public enforcement: the case of LIBOR manipulation. Manage. Sci. 65 (11), 5268–5289. https://doi.org/10.1287/mnsc.2018.3065 https://doi.org/.

Gower, J., 1967. A comparison of some methods of cluster analysis. Biometrics 23 (4), 623–637.

Hartheiser, A., Spieser, P., 2010. Libor rate and financial crisis: has the Libor rate been manipulated? Working Paper ESCP EUROPE Business School. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.629.5718&rep=rep1&type=pdf.

H2O.ai, 2020. H2O 3. Retrieved June 2020, from. http://docs.h2o.ai/h2o/latest-stable/h2o-docs/welcome.html.

H2O.ai, 2019b. AutoML: Automatic Machine Learning. Retrieved June 2019, from H2O/Docs. http://docs.h2o.ai/h2o/latest-stable/h2o-docs/automl.html.

H2O.ai, 2020c. Cross-Validation. Retrieved July 2020, from H2O/Docs. http://docs.h2o.ai/h2o/latest-stable/h2o-docs/cross-validation.html.

H2O.ai, 2020d. Anomaly Detection via H2O Deep Learning Model. Retrieved July 2020, from H2o.ai Documentation. http://docs.h2o.ai/h2o/latest-stable/h2o-r/docs/reference/h2o.anomaly.html.

Herrera, R., Climent, F., Carmona, P., Momparler, A., 2020. Can Euribor be fixed? Econ. Res.-Ekon. Istraz. https://doi.org/10.1080/1331677X.2020.1844029 https://doi.org/.

Hou, D., Skeie, D, 2014. LIBOR: origins, economics, crisis, scandal and reform. Staff Report No. 667 Federal Reserve Bank of New York. https://www.newyorkfed.org/medialibrary/media/research/staff_reports/sr667.pdf.

International Organization of Securities Commissions (IOSCO), 2013. Princ. Financ. Benchmarks Final Rep. https://www.iosco.org/library/pubdocs/pdf/IOSCOPD415.pdf.

Krauss, C., Do, X.A., Huck, N., 2017. Deep neural networks, gradient-boosted trees, random forests: statistical arbitrage on the S&P 500. Eur. J. Oper. Res. 259 (2), 689–702. https://doi.org/10.1016/j.ejor.2016.10.031 https://doi.org/.

Mollenkamp, C., Whitehouse, M., May 29, 2008. Study casts doubt on key rate. Wall Street J.

Momparler, A., Carmona, P., Climent, F., 2016. Banking failure prediction: a boosting. Spanish J. Financ. Account. 45 (1), 63–91. https://doi.org/10.1016/j.iref.2018.03.008 https://doi.org/.

Ng, A., 2015. Why Should You Care About Deep Learning? En B. Research (Ed.). Retrieved June 2019, from. https://www.slideshare.net/ExtractConf.

Palacio, S.M., 2019. Abnormal pattern prediction: detecting fraudulent insurance property claims with semi-supervised machine-learning. Data Sci. J. 18 (1), 1–15. https://doi.org/10.5334/dsj-2019-035 https://doi.org/.

Pereira, F., Mitchell, T., Botvinick, M., 2009. Machine learning classifiers and fMRI: a tutorial overview. Neuroimage 45 (1), 199–209. https://doi.org/10.1016/j.neuroimage.2008.11.007 https://doi.org/.

r-project.org, 2021. h2o.pdf. Retrieved June 2021, from cran.r-project.org. https://cran.r-project.org/web/packages/h2o/h2o.pdf.

Rodríguez-López, A., Fernández-Abascal, H., Maté-García, J.J., Rodríguez-Fernández, J.M., Rojo-García, J.L., Sanz-Gómez, J.A., 2021. Evaluating Euribor manipulation: effects on mortgage borrowers. Finance Res. Lett. (40), 101795 https://doi.org/10.1016/j.frl.2020.101795 https://doi.org/.

Sakurada, M., Yairi, T., 2014. Anomaly Detection using Autoencoders with nonlinear dimensionality reduction. MLSDA'14 Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis. https://doi.org/10.1145/2689746.2689747 https://doi.org/.

Samuel, A., 1959. Some studies in machine learning using the game of checkers. IBM J. Res. Dev. 3 (3), 210–229. https://doi.org/10.1007/978-1-4613-8716-9_15 https://doi.org/.

Snider, C., Youle, T., 2012. The Fix Is in: Detecting Portfolio Driven Manipulation of the Libor. Working Paper. UCLA/University of Minnesota. http://ssrn.com/abstract=2189015.

Yin, G., Ge, L., 2012. Application of process neural network on consumer price index prediction. affective computing and intelligent interaction. Adv. Intell. Syst. Comput. 137, 427–432. Publisher: Springer Berlin Heidelberg. https://www.springerprofessional.de/en/application-of-process-neural-network-on-consumer-price-index-pr/3826226.

Youle, T., 2014. How Much Did Manipulation Distort the Libor? Working Paper University of Minnesota.

Francisco Climent Professor in Accounting and Financial Economics at the University of Valencia (Department of Financial and Actuarial Economics). His-research focuses on quantitative methods in Financial Economics; he has published in international journals like Journal of Financial Markets, Energy Policy, Int. Rev. Econ. Finance, Journal of Business Ethics, Quantitative Finance, Financial Analysts Journal and Journal of Business Research.

Pedro Carmona Associate Professor of the Accounting Department at the University of Valencia (Spain). Certified Public Accountant with an experience of four years working in auditing firms. Research interests: machine learning algorithms, fuzzy-set analysis, bank failure prediction, Big-4 transparency reports, business failure prediction, corporate governance risk disclosure, audit quality, consulting services quality, audit risk and co-operatives societies. Author and reviewer of several business and economic journals.

Alexandre Momparler Associate Professor at the Universitat de València (Department of Corporate Finance). PhD in Business Management, Universitat Politècnica de València. MBA in Finance, Fordham University. Working experience in public accounting and banking. Research interests: online banking, bank failure prediction, financial markets, consulting services quality, corporate governance, audit risk, innovation and financial performance, mutual fund performance. Author and reviewer of several business and economic journals.

Rubén Herrera Non-Life Actuarial consultant at Willis Towers Watson (P&C Spain & Portugal) Collegiate actuary with an experience of two and a half years working in actuarial consulting. Master's degree in Actuarial Science from the University of Valencia and Master's degree in Big Data and Data Science from the Complutense University of Madrid. Research interests: Machine Learning and Deep Learning algorithms, non-life actuarial pricing, manipulation and fraud detection in insurance and finance, mortality forecasting.