II CONFERÊNCIA INTERNACIONAL DE
POLÍTICAS PÚBLICAS E CIÊNCIA DE DADOS

ENSEMBLE OF MACHINE LEARNING APPLIED TO ECONOMIC CYCLES ANALYSIS: A COMPARATIVE STUDY USING ANTECEDENT MACROECONOMIC INDICATORS FOR BRAZILIAN GDP PREDICTION CLASSIFICATION

Eduardo Palhares Júnior[1,2], Antonio M. T. de Araujo[1], Adriano H. de Souza[1], Noam G. da Silva[1], Wenndisson da S. Souza[1]

[1] Federal Institute of Amazonas, Itacoatiara 69101-030, Brazil
[2] Escola Politécnica, University of São Paulo, São Paulo 05424-970, Brazil
eduardo.palharesjr@ifam.edu.br, teixeira2gpt@gmail.com, adriano.honorato@ifam.edu.br, noam.silva@ifam.edu.br, wenndisson.souza@ifam.edu.br

Tipo de apresentação preferencial: presencial

Resumo

This work proposes a comparative study between several machine learning techniques, applied in the analysis of the phases of the Brazilian economic cycle. To this end, several macroeconomic indicators were used to build a model that was able to identify and predict the turning points of the economic cycle, such as the beginning of a recession or a recovery. The discretization of the variables proved to be decisive in the quality of the classification process, due to the diversity of the data and the non-linear nature of the analyzed phenomenon. The different techniques used reinforce a dilemma, because usually the best results come from very abstract methods, making it difficult to interpret the steps and their causes.

Palavras-chave: *Machine Learning, Classification, Economic Cycle, Multiclass-discretization.*

1. Introduction

Expert systems applied to the economy study are used to assist researcher and professionals in decision making regarding the forecast of the stock market, securities and commodities. Understanding the dynamics and interrelation of macroeconomic variables with the future evolution of the economic cycle is a problem that arouses great interest to both academics and market participants. In the literature, we can identify two main currents of research. The theoretical current tries to explain the characteristics that determine the turning points of the cycle, through the behavior of the interest curve such as slope, level and curvature, based on several financial theories (Alexander, 2010; Brown & Dybvig, 1986; Cox et al., 1985; Heath et al., 1992; Hicks, 2001; Ho & Lee, 1986; Nelson & Siegel, 1987; Svensson, 1994; Vasicek, 1977). On the other hand, the empirical current seeks to implement the use of different economic methods and knowledge discovery techniques, with the aim of modeling the structure and dynamics of economic cycles. Recently, the use of knowledge discovery techniques has become increasingly common, mainly because they are capable of capturing and dealing with non -linearities between variables, as well as the complexity involved in seasonality and structural ruptures(Chionis et al., 2010; Enke & Thawornwong, 2005; Gogas et al., 2015; Jacovides, 2008; Ju et al., 1997; Kim & Noh, 1997; Oh & Han, 2000; Vela, 2013; Zimmermann et al., 2002). However, there are many difficulties involved in using this type of tool, as non -linearities considered in each stage of prediction are often extremely complicated, making interpretation unfeasible. Thus, many of these methods, despite having excellent results, are limited to being considered black boxes.

The classification of economic variables has been used for a long time as an instrument to predict relevant movements within economics studies, such as (Burns & Mitchell, 1946) who studied economic slowdown and recessions. Over time, several different variables have been proposed and evaluated as economic indicators, as discussed in (Estrella & Mishkin, 1995), which are notably recognized as important in identifying a point of inflection in economic growth movements. (Kauppi & Saikkonen, 2008; Rudebusch & Williams, 2009).

Thus, this article aims to develop machine learning based models that analyze various economic indicators, seeking a methodology that signals the possibility of a point of inversion of economic growth, in this case, the beginning of a recession phase in several temporal horizons.

2. Methodology overview

The general objective of this methodology is the study characterization of the behavior of economic cycles, based on statistical and machine learning techniques. Specifically, the context chosen for this study was the Brazilian macroeconomic scenario, through a model that describes GDP behavior as a function of other economic variables.

It is important to emphasize that the choice of GDP as a variable of interest occurred due to the world socioeconomic context resulting from the Covid-19 pandemic, given the risk of

widespread economic recession. However, the model is quite flexible and allows broader analysis of other variables of interest, simply changing the research hypotheses.
To facilitate continuous analysis and improvement of the project, the steps were modularized to facilitate the clearance of internships and understanding of the intermediate phenomena, as shown in the **Figure 1**.

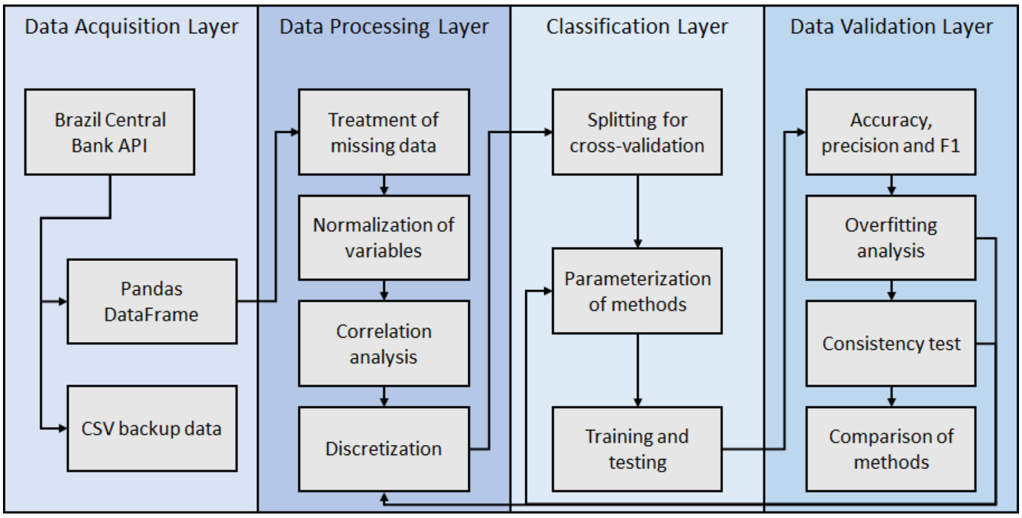

Figure 1: Architeture of methodology

2.1. Data acquisition and understanding

The first step involves understanding the requirements and the search for data available for analysis. In addition, it involves the collection and prior analysis of data consistency, comparing a theoretical perspective with a descriptive statistical analysis.
In this study, the data set used refers to macroeconomic data from the Brazil's economy according **Table 1**, during the period between January 2002 and March 2021. The data were obtained directly from the Central Bank of Brazil, through an API provided by the financial authority from the country.

Table 1: List and description of the economic variables

Economic Variable	Description
PIB	GDP monthly
IPA	Wholesale Price Index-Market
IPEM	Physical Production - Mineral extraction
IPIT	Physical Production - Capital goods
IPBC	Physical Production - Intermediate goods
IPBCD	Physical Production - Durable goods
IVVV	Sales volume index in the retail sector - Vehicles and motorcycles, spare parts - Brazil
VVCCL	Sales of factory authorized vehicle outlets - Light commercial cars sales
VVCC	Sales of factory authorized vehicle outlets - Trucks sales
IEF	Future expectations index

ICC	Consumer confidence index
Spub	Credit operations outstanding of financial institutions under public control - Total
Spriv	Credit operations outstanding of financial institutions under private control - Total
M1	Money supply - M1 (working day balance average)
M2	Broad money supply - M2 (end-of-periodo balance)

Detailed analysis of descriptive statistics is an important component of the data understanding process, allowing us to identify stylized facts of the variables of interest. For example, the average and standard deviation of the variables provide an estimate of the long -term equilibrium value and their medium historical fluctuations that can be used as a reference for future interpretations of models.

2.2. Data preparation

The second layer deals with the preparation of initial gross data, in which the variables are transformed and modified to properly feed the modeling step. This is an iterative process that assists in improving the analysis (modeling step) and may involve the addition or removal of variables or lags, missing data treatment, or normalization of variables to maintain consistent database in all observations and the comparable variables with each other. Several data preparation steps can be applied to ensure that the chosen explanatory variables have coherence with each other.
The discretization step is of fundamental importance, as while it is necessary to ensure that the phenomenon is correctly represented, classification methods often have a lot of numerical sensitivity, causing overfitting problems.

2.3. Modelling

During the modeling step, the techniques are identified and applied to pre-processed variables, and their results are examined in relation to the objective to be achieved. Configuration and calibration parameters may need to be specified for each modeling technique, and these are examined and optimized as much as possible during this step. Each technique has been implemented and parameterized independently to extract maximum efficacy and stability with different dataset subsamples.

2.4. Evaluation

The validation step involves the quantitative evaluation of the proposed models according to the degree in which they meet the objectives listed in the research hypotheses, and to what extent they are useful to correctly predict decision making. Hits/errors count was structured through the confusion matrix, so that the quality of the classification was earned using metric F1. Consistency test is an important tool for identifying overejustment problems or bias.
Comparison between methods transits between a quantitative and qualitative analysis, given that some methods can be poorly robust in some scenarios. Often a less specialized solution can be better generalizable.

3. Modelling and evaluation

3.1. Overview

The stage of choice of indicators is of fundamental importance, however, there is no consensus on the literature on how to make this choice. Comparing several studies conducted in different countries, it is remarkable to realize that each economy is better represented by certain parameters and segments that have correlation with intrinsic characteristics of their economic activities. In other words, there is no specific methodology for this step.
The most common approach is to include the main indicators related to various segments of the economy, and later filter them through statistical correlation studies. Thus, the main indicators widely used in other studies were referenced, and an equivalent sought in the Brazilian economy.

3.2. Data preparation

The variables used are monthly temporal granularity, and were normalized using the percentage variation:

$$\Delta x_i = \frac{x_i - x_{i-1}}{x_{i-1}}$$

To evaluate the series stationarity, KPSS (Kwiatkowski - Phillips - Schmidt - Shin) and ADF (Augmed Dickey Fuller) tests were performed. The pearson coefficients matrix (**Figure 2**) was used for a linear correlation analysis with the GDP variable.

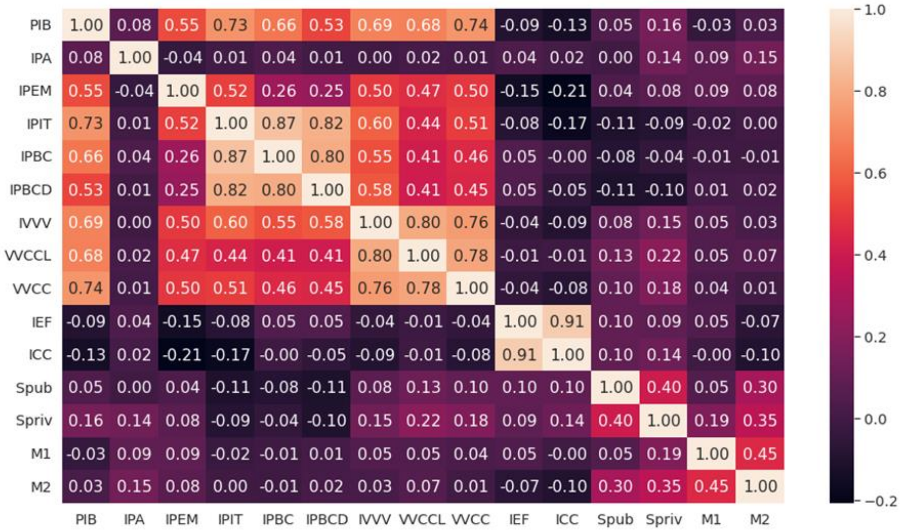

Figure 2: Correlations between the economic variables - complete dataset.

Although the correlation of several variables are apparently low, it must be remembered that the phenomena analyzed can be quite nonlinear. Seeking to consider these subtleties, a new database was proposed considering only the significant correlation variables, named as restricted database as shown as **Figure 3**.

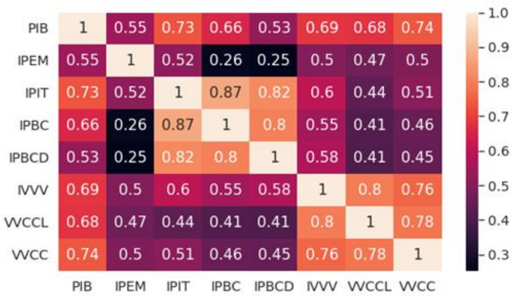

Figure 3: Correlations between the economic variables - restrict dataset.

3.3. Discretization

The discretization step is a very important stage in the construction of the model, because influence on the correct visualization of the phenomenon. In this work several discretizations were proposed, seeking a sufficiently simple representation to avoid numerical problems, but capable of adequately representing decision-making conditions.

3.3.1. Binary discretization

Binary classification is a transformation that considers only the state of high and falling in each indicator:

$$\begin{cases} x \mapsto 0, & \text{if } \Delta x < 0; \\ x \mapsto 1, & \text{if } \Delta x \geq 0. \end{cases}$$

Despite presenting excellent practical results, it is a very simplistic modeling that cannot separate regions of trend and stagnation, being little informative. Still, it was an important model to validate architecture and test the robustness of some techniques used.

3.3.2. 3-categories discretization

Discretization with 3 categories separates significant high and falling movements with a stagnation movement, where the variation is less than a standard deviation in relation to the average:

$$\begin{cases} x \mapsto -1, & \text{if } \Delta x \leq \mu_x - \sigma_x; \\ x \mapsto 0, & \text{if } \mu_x - \sigma_x < \Delta x < \mu_x + \sigma_x; \\ x \mapsto 1, & \text{if } \Delta x \geq \mu_x + \sigma_x. \end{cases}$$

This approach was very promising for a practical application, as it combines a good accuracy in the classification with effective categories in decision making.

3.3.3. 5-categories discretization

The classification considering 5 categories is a refinement from 3 categories, as in addition to separating stagnation cases, it can also identify very rare cases. The first approach proposed for 5 categories considered rare cases to a distance of 2 standard deviations, which caused bias in the model and convergence problems in some classifiers. A new separation interval between the 5 categories was proposed, seeking to better balance these significantly rare events:

$$\begin{cases} x \mapsto -2, & \text{if } \Delta x \leq \mu_x - 1,6745 \cdot \sigma_x; \\ x \mapsto -1, & \text{if } \mu_x - 1,6745 \cdot \sigma_x < \Delta x \leq \mu_x - 0,6745 \cdot \sigma_x; \\ x \mapsto 0, & \text{if } \mu_x - 0,6745 \cdot \sigma_x < \Delta x < \mu_x + 0,6745 \cdot \sigma_x; \\ x \mapsto 1, & \text{if } \mu_x + 0,6745 \cdot \sigma_x \leq \Delta x < \mu_x + 1,6745 \cdot \sigma_x; \\ x \mapsto 2, & \text{if } \Delta x \geq \mu_x + 1,6745 \cdot \sigma_x. \end{cases}$$

With the discretization proposed, the central category (stagnation) includes 50% of the data and the extreme categories correspond to 10% of the data. Thus, classifiers showed a significant improvement in accuracy as well as a very adequate description.

3.4. Classification

3.4.1. Cross-validation

Before starting training, it is necessary to divide the training/test subsets, for application of cross validation that seeks to minimize the effects of overfitting and bias. As the variables are temporal series, it is important that the chronological order is not broken during cross validation. Therefore, the Time Series Split Cross Validation method was adopted, which initially considers small training/testing subsets at the beginning of the historical series, and follows in an iterative manner considering the previous round test subset as the current round training subset. Several settings were tested for the size of the initial test subsets, but due to the large number of scenarios, in this work will be considered only the configuration that considers 30% for the subset of testing and validation.

3.4.2. Classification mathods

Several statistical and data mining techniques have been applied, which are available as part of the Python Scikit-Learn Library set. For all these techniques, the default settings proposed by the tool set were used, as the exhaustive optimization of the model's parameters and architectures was outside the scope of the article. However, seeking to solve some convergence problems, some of these techniques have been explicitly implemented to test some hyper parameters more attention. In detail, the techniques applied were: K-nearest neighbors (KNN), Gaussian naive bayes (NB), decision trees (DT), random forests (RF), logistics regression (LR), support vector machines (SVC), artificial neural networks (NN).

The K-nearest neighbors (KNN) is a method for classifying cases based on their similarity with other cases. The average or median target value of the nearest neighbors was used to obtain the expected value for the new case. Parameter K was automatically defined and the distance calculation metric was Euclidian. The naive bayes classifier (NB) is based on the application of the Bayes theorem with the "naive" assumption of conditional independence between each pair of characteristics given the value of the category variable. In this work was considered a Gaussian kernel whose coefficients were calculated through maximum likelihood. The decision tree (DT) is a map of the possible results of a series of related choices, while the random forest (RF) consists of building various decision trees considering different parts of the training set, seeking to reduce the estimator's variance and increasing accuracy. Logistic regression (LR) divides the dependent variable into two categorical levels, calculating the probability of output as a function of the input. In the proposed problem the input variables were also classified within these same cutting strips. The support vector classification (SVC) is a specific type of support vector machines (SVM) applied to classification problems and can be used to improve the predictive accuracy of a model without overfting training data and can be used to categorize a high space of dimensional appeal. It is noteworthy that this technique has brought excellent results, but it is quite sensitive to discretization, especially when trained with very rare and unbalanced categories. The artificial neural network (NN) can be used to automatically bring nonlinear relationships between inputs and the destination variable. The disadvantage of this flexibility is that the neural network is not easily interpretable. Specifically, the perceptron architecture of multi-charged was used, but the different discretizations brought a lot of variability in relation to hyperparimeters that were parameterized individually through pre-processing. Thus, the activation function, the solver, the number of neurons and hidden layers was parameterized on a case-by-case basis, to avoid problems of numerical convergence.

3.5. Target Variable Forecasting

To compare the various scenarios presented, a comparative analysis is proposed between all methods, in relation to the accuracy achieved in the training and testing steps. In addition, a more detailed analysis of the behavior of F1-score in relation to the different discretizations and the set of explanatory variables used in each scenario.

3.5.1. Accuracy

To analyze accuracy, it is necessary to compare all the methods proposed with all discretization approaches.

Table 2: Comparison of accuracy in the test stage

Methods	2CC	2CR	3CC	3CR	5CC	5CR
kNN	90,00%	90,00%	81,43%	81,43%	55,71%	55,71%
NB	87,14%	87,14%	78,57%	80,00%	50,00%	51,43%
DT	80,00%	88,57%	72,86%	75,71%	48,57%	52,86%
RF	88,57%	90,00%	82,86%	75,71%	65,71%	60,00%
LR	84,29%	87,14%	68,57%	75,71%	57,14%	58,57%
SVC	88,57%	90,00%	81,43%	84,29%	54,29%	57,14%
NN	85,71%	85,71%	84,29%	82,86%	58,57%	60,00%

Column labels refer to the discretization used and the amount of explanatory variables, for example:

- 2CC -> 2 categories with the full base (**Figure 2**)
- 5CR -> 5 categories with the restricted base (**Figure 3**).

Analyzing **Table 2** we can see that there was considerable homogeneity in the results obtained, despite the different discretizations. There was a slight improvement in the scenario considering binary classification, probably due to bias due to overly simple discretization for the analyzed phenomenon.

Comparing the variability between the methods that considers the same discretization, we see that for discretizations with 3 and 5 categories, the reduction in the number of explanatory variables has brought lower precision variability between the methods.

3.5.2. F-score

The quality of each method in relation to each scenario is relative, depending on the criterion. However, it is important to compare the effect that each scenario had on the performance of each percentage variation in the restricted base score F1 (some parameters) compared to the complete base (all parameters). As many scenarios were tested and in each of them there are many categories to evaluate, the number of results is significantly large. Thus, a qualitative analysis will be presented that demonstrates which the best scenario, which will also be discussed quantitatively.

Trying to summarize the total absolute gain achieved in each scenario when considering the restricted base, despite the high variance, **Table 3** combines the absolute liquid gains of each category in relation to each scenario. As 7 methods are considered, the relative gain is precisely the average liquid gain between each of the 7 modeling.

Table 3: Comparison of F-score improvement when some parameters are removed (various scenarios).

Categories	2 categories		3 categories		5 categories	
	absolute	relative	absolute	relative	absolute	relative
-2					46,53%	6,65%
-1			-20,14%	-2,88%	61,26%	8,75%
0	17,85%	2,55%	6,51%	0,93%	7,67%	1,10%
1	11,82%	1,69%	24,55%	3,51%	-42,50%	-6,07%
2					-52,22%	-7,46%
Total	29,67%	4,24%	10,92%	1,56%	20,74%	2,96%

Despite the good performance in the binary case, it is a very simple approach that has no direct practical interest. On the other hand, the approach considering 5 categories, besides presenting a very chaotic performance, also presented convergence problems, showing that the approach is not adequate.

For all that has been exposed, the scenario that considers 3 categories had an effective performance gain when the base was restricted to less significant variables. However, the relative improvement was very small, being significant. Therefore, while removal is a valid option, future studies may consider both cases in an attempt to verify in which cases they are significant.

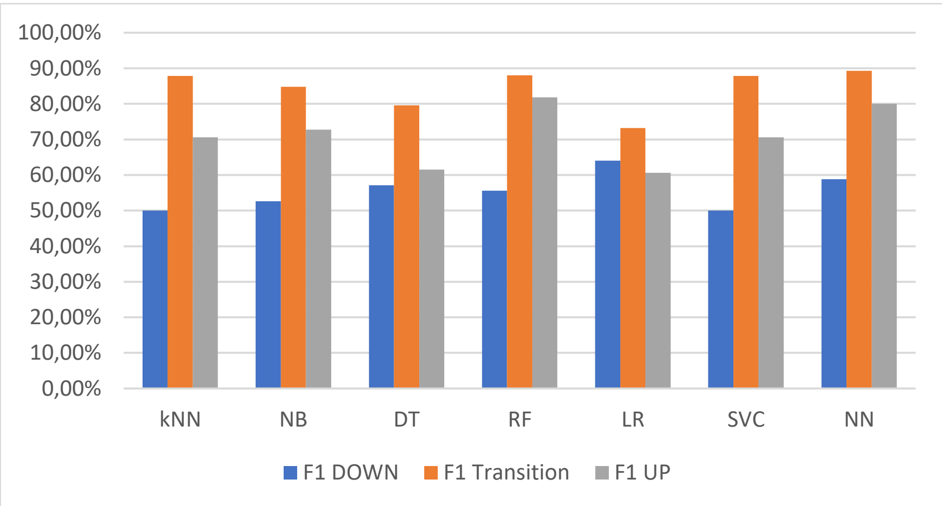

Figure 4: Score-F1 results of the Brazilian GDP forecast step at the best model that considers 3 categories and the complete variable set.

We can see from **Figure 4** that all methods reach comparable accuracy, and virtually everyone is easier to predict the central category (stagnation), as well as a greater difficulty in predicting falls. This corroborates what is empirically observed, as low movements often happen faster and significantly, without leaving such clear signs. Despite this similarity with what is empirically observed, it is nonetheless a limitation of the model. Some methods (SVC and NN) have had problems of convergence and repeatability when the approach with 5 categories is used, as the Extreme Categories (Strong Up and Down) are quite unbalanced, causing model overfitting problems.

Another aspect that could influence overfitting or bias problems is the fact that the number of observations available is relatively small, as indicators have monthly temporal granularity.

Therefore, several rounds of the same tests were performed with the same parameters to ensure the repeatability of the results. Despite consistency in the observed results, it is always desirable to have more data available. Another possible approach to be investigated to improve the quality of the model is the inclusion of backward variables through autoregressive models.

4. Summary and conclusions

This article proposes a methodology to model and predict the value of Brazil's GDP, through the analysis of macroeconomic indicators related to different stages of the economic cycle.

The initial proposal to conduct a comparative study between various machine learning algorithms was exploited, but there was a greater focus on the data preparation stage, to facilitate the phenomenon interpretation. Because it is an unusual approach in the Brazilian economy, besides the general understanding that could be applied in any market, it was necessary to find by attempt and error a configuration that adapted to the local context. Thus, the choice of variables as well as the steps of preparation and discretization were decisive not only to achieve good results, but also to interpret GDP behavior.

The proposed methodology was divided in a modular way to facilitate the understanding and implementation of continuous improvements. The variable categorization stage showed that simpler approaches (using fewer categories) often have better practical results, so it was necessary to find in the middle of simplicity and utility. The use of 3 categories can separate stages of stagnation and trend while ensuring good accuracy. Comparing the influence of low correlation variables, the model trained with fewer variables was slightly superior in general, but looking specifically for each category, it is not possible to establish a consensus since the forecasting methods vary significantly and become more unpredictable (smaller repeatability). Because of this, it is possible to state that despite the low linear correlation, these variables have significant importance in the model, especially when using techniques that can capture these nonlinear relationships.

Regarding the methods employed for forecast, all had similar behavior within the ideal scenario (3 categories with all variables). Some methods such as Support Vector Machines and Neural Networks had problems of convergence and overfitting when many categories were considered, especially because of the side categories (Strong Up and Down) were very unbalanced. Decision trees had a result had a slightly inferior accuracy, but can be preferable as the results can be interpreted more simply. Combined with random forests that always maintain a good result, this can be a set of efficient techniques that ensures good precision without losing the interpretative notion of the phenomenon. Data availability was not a limiting factor, but a larger amount of data available would probably bring better quality to the model. In addition, the thin adjustment of hyper parameters has brought little improvement in the results, showing that most of the methods employed are robust enough for this problem.

Finally, this study does not exhaust the possibilities on the subject, quite the contrary, opens many doors for new investigations. The proposed architecture can be used for other macroeconomic variables with appropriate considerations, can be extended to analyze a broader set of economies from other countries. The use of delay variables through self-regressive models as well as moving averages that include components for seasonal trends are among the many future investigation possibilities.

5. Acknowledgments

The authors would like to thank SAMSUNG ELETRÔNICA DA AMAZÔNIA LTDA. through the Projeto Aranouá and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - Programa de Excelência Acadêmica (PROEX) - Brasil for Financial Support. The present work is the result of the Research and Development (R&D) project 001/2021, signed with Instituto Federal do Amazonas and FAEPI, Brazil, which has funding from Samsung. We would also like to thank collaboration of Afonso de Campos Pinto and Elia Matsumoto (FGV), Flavio Almeida de Magalhães Cipparrone (USP) and James Moraes de Almeida (ILUM-CNPEM), for the review and suggestions for improvement in the steps of the work.

Referências

Alexander, G. J. (2010). Fundamentals of investments/ fundamentos de inversiones. Prentice Hall.

Brown, S. J., & Dybvig, P. H. (1986). The Empirical Implications of the Cox, Ingersoll, Ross Theory of the Term Structure of Interest Rates. The Journal of Finance, 41(3), 617–630. https://doi.org/10.1111/j.1540-6261.1986.tb04523.x

Burns, A. F., & Mitchell, W. C. (1946). Measuring Business Cycles. National Bureau of Economic Research.

Chionis, D., Gogas, P., & Pragidis, I. (2010). Predicting European Union Recessions in the Euro Era: The Yield Curve as a Forecasting Tool of Economic Activity. International Advances in Economic Research, 16(1), 1–10. https://doi.org/10.1007/s11294-009-9247-2

Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985). A Theory of the Term Structure of Interest Rates. Econometrica, 53(2), 385. https://doi.org/10.2307/1911242

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

Estrella, A., & Mishkin, F. (1995). Predicting U.S. Recessions: Financial Variables as Leading Indicators (w5379; p. w5379). National Bureau of Economic Research. https://doi.org/10.3386/w5379

Gogas, P., Papadimitriou, T., Matthaiou, M., & Chrysanthidou, E. (2015). Yield Curve and Recession Forecasting in a Machine Learning Framework. Computational Economics, 45(4), 635–645. https://doi.org/10.1007/s10614-014-9432-0

Heath, D., Jarrow, R., & Morton, A. (1992). Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation. Econometrica, 60(1), 77. https://doi.org/10.2307/2951677

Hicks, J. (2001). Value and capital: An inquiry into some fundamental principles of economic theory (2. ed., reprint). Clarendon Press.

Ho, T. S. Y., & Lee, S. (1986). Term Structure Movements and Pricing Interest Rate Contingent Claims. The Journal of Finance, 41(5), 1011–1029. https://doi.org/10.1111/j.1540-6261.1986.tb02528.x

Jacovides, A. (2008). Forecasting Interest Rates from the Term Structure: Support Vector Machines Vs Neural Networks [University of Nothingham]. http://eprints.nottingham.ac.uk/22097/

Ju, Y. J., Kim, C. E., & Shim, J. C. (1997). Genetic-based fuzzy models: Interest rate forecasting problem. Computers & Industrial Engineering, 33(3–4), 561–564. https://doi.org/10.1016/S0360-8352(97)00193-9

Kauppi, H., & Saikkonen, P. (2008). Predicting U.S. Recessions with Dynamic Binary Response Models. Review of Economics and Statistics, 90(4), 777–791. https://doi.org/10.1162/rest.90.4.777

Kim, S. H., & Noh, H. J. (1997). Predictability of interest rates using data mining tools: A comparative analysis of Korea and the US. Expert Systems with Applications, 13(2), 85–95. https://doi.org/10.1016/S0957-4174(97)00010-9

Nelson, C. R., & Siegel, A. F. (1987). Parsimonious Modeling of Yield Curves. The Journal of Business, 60(4), 473. https://doi.org/10.1086/296409

Oh, K. J., & Han, I. (2000). Using change-point detection to support artificial neural networks for interest rates forecasting. Expert Systems with Applications, 19(2), 105–115. https://doi.org/10.1016/S0957-4174(00)00025-7

Rudebusch, G. D., & Williams, J. C. (2009). Forecasting Recessions: The Puzzle of the Enduring Power of the Yield Curve. Journal of Business & Economic Statistics, 27(4), 492–503. https://doi.org/10.1198/jbes.2009.07213

Vasicek, O. (1977). An equilibrium characterization of the term structure. Journal of Financial Economics, 5(2), 177–188. https://doi.org/10.1016/0304-405X(77)90016-2

Vela, D. (2013). Forecasting Latin-American yield curves: An artificial neural network approach (Borradores de Economia). Banco de la Republica.

Zimmermann, H.-G., Tietz, C., & Grothmann, R. (2002). Yield curve forecasting by error correction neural networks and partial learning. 10th Eurorean Symposium on Artificial Neural Networks, ESANN 2002, Bruges, Belgium, April 24-26, 2002, Proceedings, 407–412. https://www.esann.org/sites/default/files/proceedings/legacy/es2002-205.pdf