

Report on Master Thesis

Institute of Economic Studies, Faculty of Social Sciences, Charles University

Student:	Alex Zayat
Advisor:	Ladislav Krištofuk
Title of the thesis:	Machine Learning in Macroeconomic Nowcasting

OVERALL ASSESSMENT (provided in English, Czech, or Slovak):

Short summary

This thesis explores the application of machine learning models for nowcasting, specifically focusing on predicting Argentina's monthly GDP. The author applies various machine learning techniques, such as neural networks, random forests, boosted trees, support vector regression, and K-nearest neighbors, and compares their performance to traditional econometric models like AR(1). The research highlights the predictive advantages of machine learning, offering a framework for integrating these models into central bank prediction tools while also addressing the critical issue of interpretability in "black-box" models.

Contribution

The thesis makes an important contribution to the field by applying machine learning techniques in an area traditionally dominated by econometric models. The use of machine learning in nowcasting, especially for macroeconomic variables such as GDP, is a novel and highly relevant topic. The focus on Argentina provides a unique regional application, adding to its distinctiveness. Additionally, the work delves into interpretability methods, a vital and often overlooked issue in the use of machine learning models for economic forecasting.

Methods

The thesis employs a solid range of machine learning models and justifies their use effectively. The models are well-explained, and the author utilizes proper evaluation metrics such as Mean Absolute Percentage Error (MAPE) and Root Mean Squared Error (RMSE) to assess the performance of these models. The inclusion of interpretability techniques such as SHAP values and feature importance adds depth to the analysis. However, the thesis could have explored more advanced techniques for dealing with data revisions, which could have strengthened the analysis further.

Literature

The literature review is thorough and well-structured, covering the major studies in both machine learning applications and traditional econometrics for nowcasting. The discussion of interpretability in machine learning models is particularly relevant and timely, given the increasing demand for transparency in economic forecasting.

Manuscript form

The manuscript is well-written and logically structured. The technical details of the models and the discussion of results are clearly presented, with well-designed tables and figures. The use of LaTeX ensures that the thesis is visually professional, and the explanations of complex concepts are accessible to readers with varying levels of familiarity with machine learning. Some minor improvements could be made to the flow of the literature review, which occasionally feels disjointed.

Report on Master Thesis

Institute of Economic Studies, Faculty of Social Sciences, Charles University

Student:	Alex Zayat
Advisor:	Ladislav Křišťoufek
Title of the thesis:	Machine Learning in Macroeconomic Nowcasting

Overall evaluation and suggested questions for the discussion during the defense

This is a well-executed and timely thesis, offering a robust application of machine learning techniques in the relatively novel area of macroeconomic nowcasting. The research demonstrates clear potential for machine learning to improve the accuracy of GDP predictions. Additionally, the focus on interpretability methods adds a valuable layer to the research, making it relevant not only for academic purposes but also for practical applications in central banking and policy-making.

Suggested Questions:

- Given the black-box nature of machine learning models, how do you propose central banks could address the challenge of model interpretability when adopting these models?
- What are some potential risks or limitations of relying on machine learning models for nowcasting in an unstable economic environment such as Argentina's?
- Could you discuss how the incorporation of synthetic variables (e.g., lagged EMAE) impacted the predictive accuracy of the machine learning models compared to the baseline models?
- How would you approach improving the performance of machine learning models for nowcasting if provided with a larger, more diverse dataset?

The TurnItIn analysis has not uncovered any suspicious practices (score of 16% with no large portions of overlapping text).

SUMMARY OF POINTS AWARDED (for details, see below):

CATEGORY	POINTS
<i>Contribution (max. 30 points)</i>	27
<i>Methods (max. 30 points)</i>	26
<i>Literature (max. 20 points)</i>	18
<i>Manuscript Form (max. 20 points)</i>	18
TOTAL POINTS (max. 100 points)	89
GRADE (A – B – C – D – E – F)	B

NAME OF THE REFEREE: Ladislav Kristoufek

DATE OF EVALUATION: 10 September 2024

Referee Signature