Abstract

In this essay, we use three well known machine learning (ML) models, namely Random Forest (RF), least absolute shrinkage and selection operator (LASSO), and complete subset regression (CSR) to forecast 12-month-ahead Brazilian inflation. Since ML methods started to be used in the context of forecasting economic variables, academic research has been focused on (i) comparing the performance of different models and (ii) analysing which explanatory variables are more relevant. In this paper, we contribute to the literature by analysing if accumulating explanatory variables enhances the forecasts of accumulated inflation.

Keywords: machine learning, time series, forecast, inflation

1 Introduction

Accurately forecasting inflation is important for several reasons. First of all, modern central banks calibrate their economic policies based on expected inflation (Iversen et al, 2016). Bad forecasts would result in ineffective policies with high social costs. Secondly, many long-term contracts are set in nominal terms and therefore bad inflation forecasts would generate undesired uncertainty. Finally, expectations about future prices are a key factor for households' decisions concerning future consumption and investments.

In this paper we use Machine Learning (ML) methods to forecast 12-month-ahead Brazilian inflation. Namely, the models used were: random forest (Breiman, 2001), LASSO (Tibshirani, 1996), Ridge (McDonald, 2009) and CSR (Elliotti; Gargano; Timmermann, 2013). Different authors have already approached this subject. Forni et al. (2003) showed that multivariate models beat univariate models for 12-month forecasts in the main countries of Europe. More recently, Medeiros et al. (2021) showed ML models outperform univariate models specially when inflation is more volatile. In particular, they showed that random forest (RF) provides improvements of almost 25% in terms of the root mean squared error (RMSE) when compared to the random walk (RW) for 12-month forecasts.

This study is innovative because we accumulate explanatory variables in two to twelve months to investigate whether it would enhance our 12-month inflation's forecasts. Recently, Coulombe et al. (2021) showed that transformations in macroeconomic data can enhance forecasts' accuracy. However, the transformations used by the authors were different from the one we are using. Namely, the transformations they used were moving average factors (MAF) and moving average rotation (MARX).

Unlike other papers, our benchmark is not an univariate model. Our benchmark is FOCUS' expected inflation. FOCUS is a report which consists of the projections of many economic variables (such as inflation) and its projections take into consideration the forecasts of more than a hundred professional forecasters.