Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 95 (2016) 264 – 271

Procedia Computer Science

Complex Adaptive Systems, Publication 6
Cihan H. Dagli, Editor in Chief
Conference Organized by Missouri University of Science and Technology
2016 - Los Angeles, CA

Evaluating Forecasting Methods by Considering Different Accuracy Measures

Nijat Mehdiyev[a,b], David Enke[c*], Peter Fettke[a,b], Peter Loos[a,b]

[a]*Institute for Information Systems (IWi). German Research Center for Artificial Intelligence (DFKI), Campus D3 2, 66123 Saarbrücken, Germany*
[b]*Saarland University, Campus D3 2, 66123 Saarbrücken, Germany*
[c]*Department of Engineering Management and System Engineering, Missouri University of Science and Technology, Rolla, MO, 65409-0370, USA*

Abstract

Choosing the appropriate forecasting technique to employ is a challenging issue and requires a comprehensive analysis of empirical results. Recent research findings reveal that the performance evaluation of forecasting models depends on the accuracy measures adopted. Some methods indicate superior performance when error based metrics are used, while others perform better when precision values are adopted as accuracy measures. As scholars tend to use a smaller subset of accuracy metrics to assess the performance of forecasting models, there is a need for a concept of multiple accuracy dimensions to assure the robustness of evaluation. Therefore, the main purpose of this paper is to propose a decision making model that allows researchers to identify the superiority of a forecasting technique over another by considering several accuracy metrics concurrently. A multi-criteria decision analysis approach, namely the preference ranking organization method for enrichment evaluation (PROMETHEE), was adopted to solve this problem. Bayesian Networks, Artificial Neural Networks, SVMs, Logistic Regression, and several Rule and Tree-based forecasting approaches were included in the analysis. After introducing a detailed description of accuracy measures, the performance of the prediction models are evaluated using a chosen dataset from the UCI Machine Learning Repository.

Peer-review under responsibility of scientific committee of Missouri University of Science and Technology

Keywords: Classification, Accuracy Measure, Confusion Matrix, MCDA, PROMETHEE

* Corresponding author. Tel.: +1-573-341-4749
E-mail address: enke@mst.edu

Peer-review under responsibility of scientific committee of Missouri University of Science and Technology
doi:10.1016/j.procs.2016.09.332

1. Introduction

Forecasts serve a crucial need in making rational decisions and planning activities more precisely by handling uncertainty about the future. Efficient prediction is considered as an important perquisite for effective administration and organization in various fields of social, information, human and natural sciences, and related application areas. To deal with the growing variability and complications associated with the domain specific forecasting problems, diverse forecasting methods have been proposed. Decision makers have to consider various aspects of the prediction process, such as the length of forecasting horizon, the goal of forecasting, frequency, structure and nature of the data, etc., when deciding on a forecasting algorithm.

In order to estimate the performance of forecasting methods, in the last three decades various accuracy measures have been adopted by many studies as an evaluation criterion. A number of different forecast accuracy measures for both regression and classification problems have been proposed, and the comments and recommendations on the use of the relevant measures have been intensively discussed in prior studies[1-5]. Such accuracy measures provide necessary and decisive feedback to decision makers for calibrating and refining the model in an effort to improve the preciseness of outcomes[6]. However, research findings suggest that there is no best overall accuracy measure which can be used as a universally accepted single metric for choosing the appropriate forecasting method[2]. Forecasting approaches can realize extremely different performances depending on the chosen metric. Empirical evaluations reveal that some approaches are superior when error based measures are adopted, while others perform better for the same dataset when different metrics are utilized[6].

On these grounds, we can argue that there is a need for a framework to evaluate the forecasting methods considering various accuracy metrics concurrently. In this paper we propose a framework that is assumed to cater to the need for a unique assessment measure that assures a robust comparability of classification methodologies. With this paper we extend the research in forecasting accuracy measurement domain by integrating another important Management Science discipline, namely Multi-Criteria Decision Analysis (MCDA).

The remainder of the paper is structured as follows: Section 2 provides an overview into the steps of proposed MCDA based framework, namely, the related work in the domain of classification accuracy measures and mathematical background of the selected metrics (Section 2.1), a brief description of selected classification methods (Section 2.2), and the details of the selected MCDA approaches, PROMETHEE I and II (Section 2.3). Section 3 presents the performance evaluation of selected approaches and their rankings both in terms of each individual accuracy measure and multidimensional assessment using the MCDA framework. The paper concludes with a discussion of the study implications and future research directions.

2. Proposed Framework

The primary goal of the underlying paper was defined as the evaluation and ranking of competing algorithms for multi-class classification tasks in terms of multiple accuracy criteria. Xu[7] and Ouenniche[8] have proposed MCDA frameworks based on PROMETHEE, ELECTRE, and Data Envelopment Analysis (DEA) methods to assess the performance of forecasting problems for regression problems. Accuracy measures for classification problems with nominal output structure differ significantly from those for regression problems. Peng et al.[9] and Khanmohammadi and Rezaeiahari[10] proposed MCDA frameworks to evaluate the classification algorithms. However, they used single accuracy measure and computational costs as comparison criteria. To the best of our knowledge no prior studies have addressed the multidimensional performance analysis of rival classification algorithms. The next subsections will provide an overview to the selected accuracy measures, alternatives, and the details of PROMETHEE approaches.

2.1. Accuracy Measures

Comparative analysis of classification algorithms is a complicated process since various dimensions of assessment have to be considered. Prior studies suggest that the performance evaluation can be carried out by virtue of statistical tests, performing qualitative analysis by discussing drawbacks and advantages of approaches, or by conducting quantitative analysis using diverse evaluation measures which capture different aspects of classifiers