Chapter 4

Ensemble Learning via Extreme Learning Machines for Imbalanced Data

Adnan Omer Abuassba
Arab Open University, Palestine

Dezheng O. Zhang
School of Computer and Communication Engineering, University of Science and Technology Beijing, China

Xiong Luo
School of Computer and Communication Engineering, University of Science and Technology Beijing, China

ABSTRACT

Ensembles are known to reduce the risk of selecting the wrong model by aggregating all candidate models. Ensembles are known to be more accurate than single models. Accuracy has been identified as an important factor in explaining the success of ensembles. Several techniques have been proposed to improve ensemble accuracy. But, until now, no perfect one has been proposed. The focus of this research is on how to create accurate ensemble learning machine (ELM) in the context of classification to deal with supervised data, noisy data, imbalanced data, and semi-supervised data. To deal with mentioned issues, the authors propose a heterogeneous ELM ensemble. The proposed heterogeneous ensemble of ELMs (AELME) for classification has different ELM algorithms, including regularized ELM (RELM) and kernel ELM (KELM). The authors propose new diverse AdaBoost ensemble-based ELM (AELME) for binary and multiclass data classification to deal with the imbalanced data issue.

DOI: 10.4018/978-1-7998-3038-2.ch004
INTRODUCTION

Among the popular machine learning methods (Abuassba, Zhang, Luo, Zhang, & Aziguli, 2017; Bezdek, 2016; Chen, Li et al., 2018; Luo, Sun et al., 2018; Luo, Jiang et al., 2019; Luo, Xu et al., 2018, Abuassba et al., 2018), extreme learning machine (ELM) is well-known for solving classification and regression problems in real world applications. It is designed for a single hidden layer feed-forward network (SLFN). It is proved theoretically and practically (Huang, Zhu et al., 2006; Huang, Wang et al., 2010; Huang, Zhou et al., 2012; Huang 2014) that ELM is efficient and fast in both classification and regression (Liu, He et al. 2008; Huang, Ding et al., 2010). It eludes parameter tuning on the contrary of traditional gradient based algorithms. Imbalanced data issue appears when negative or majority class dominates another class (positive or minority); which means the number of majority class examples
Related Content

Systematic Literature Review on Empirical Results and Practical Implementations of Healthcare Recommender Systems: Lessons Learned and a Novel Proposal
www.igi-global.com/chapter/systematic-literature-review-on-empirical-results-and-practical-implementations-of-healthcare-recommender-systems/205880?camid=4v1a

Group Process Losses in Agile Software Development Decision Making
www.igi-global.com/article/group-process-losses-agile-software/77873?camid=4v1a

Machine Learning-Based Demand Forecasting in Supply Chains
www.igi-global.com/article/machine-learning-based-demand-forecasting/2426?camid=4v1a

Making the Case for “Architectural Informatics”: A New Research Horizon for Ambient Computing?
Mikael Wiberg (2013). Pervasive and Ubiquitous Technology Innovations for Ambient Intelligence Environments (pp. 128-135).
www.igi-global.com/chapter/making-case-architectural-informatics/68930?camid=4v1a