

# Accepted Manuscript

Classifier selection and clustering with fuzzy assignment in ensemble model for credit scoring

Haoting Zhang , Hongliang He , Wenyu Zhang

PII: S0925-2312(18)30905-6  
DOI: <https://doi.org/10.1016/j.neucom.2018.07.070>  
Reference: NEUCOM 19823



To appear in: *Neurocomputing*

Received date: 11 September 2017  
Revised date: 9 April 2018  
Accepted date: 28 July 2018

Please cite this article as: Haoting Zhang , Hongliang He , Wenyu Zhang , Classifier selection and clustering with fuzzy assignment in ensemble model for credit scoring, *Neurocomputing* (2018), doi: <https://doi.org/10.1016/j.neucom.2018.07.070>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

# Classifier selection and clustering with fuzzy assignment in ensemble model for credit scoring

<sup>1</sup>Haoting Zhang, <sup>2</sup>Hongliang He and <sup>3</sup>Wenyu Zhang

<sup>1</sup>*St Hilda's College, University of Oxford, Cowley Place, Oxford, UK, OX4 1DY,*

*haoting.zhang@st-hildas.ox.ac.uk*

<sup>2</sup>*School of Information, Zhejiang University of Finance & Economics, China, 310018*

*hhlisme@163.com*

<sup>3</sup>*School of Information, Zhejiang University of Finance & Economics, China, 310018*

*wyzhang@e.ntu.edu.sg*

## Abstract

With the development of statistical methods and machine learning algorithms, credit scoring is no longer a task merely based on experience. From single base classifiers to ensemble classifiers and hybrid models, researches have been focusing on combining classifiers and hybridizing with artificial intelligence algorithms to improve performance of the models. Ensemble classifiers have been proven to have a better predictive accuracy than single classifiers, but the method of ensemble affects performance and is worth studying. This study is based on the ensemble of five of the most widely recognized base classifiers in credit scoring, i.e. logistic regression, support vector machine, neural network, gradient boosting decision tree and random forest. It proposes a new method of selecting classifiers using Genetic Algorithm after they are trained, considering both the accuracy and diversity of the ensemble. Besides, unsupervised clustering is integrated with a fuzzy assignment procedure in the model, to make more use of the data pattern and improve performance. The proposed CF-GA-Ens model is tested on three credit scoring datasets (Australian, German, Japanese) and three performance measures (accuracy, AUC, F-score), and the results show that our classifier selection and clustering procedures have a positive impact on all performance measures.

**Keywords:** Credit scoring; Ensemble learning; Clustering; Genetic Algorithm; Fuzzy assignment.

Download English Version:

<https://daneshyari.com/en/article/8965176>

Download Persian Version:

<https://daneshyari.com/article/8965176>

[Daneshyari.com](https://daneshyari.com)