Fuzzy classification based on Combinative Algorithms

N A M Saffie¹, K A Rasmani¹ and N H Sulaiman²

¹Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Seremban Campus, 70300 Seremban, Negeri Sembilan, Malaysia
²Faculty of Computer and Mathematical Sciences Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia

E-mail: khairulanwar@ns.uitm.edu.my

Abstract. Classification models are used in order to predict the class label of objects for which the class label is unknown. The performance of a single classification model can be determined on the basis of the classification accuracy. However, it is difficult to determine which single-model is the best classification technique in a specific application domain since a single learning algorithm may not uniformly outperform other algorithms over various datasets. Furthermore, most classification algorithms, using either fuzzy or non-fuzzy approaches, produce results in the form of crisp or categorical classification outcomes. Unfortunately, in certain applications, the classification outcomes that represent the class labels may involve categorisations which are fuzzy in nature. In this study, a new technique is developed to handle the classification task based on combinative algorithms by adopting multiple-model concepts. The aggregation process is based on a modified fuzzy Ordered Weighted Averaging (OWA) operator. The performance of the proposed technique has been analysed using the Domestic Water Consumption (DWC) dataset obtained from a previous study. The findings show that the proposed method consistently outperforms all selected single-models and performs better than a statistical multiple-model approach, known as the Pool Method. This result indicates that the proposed technique has the potential to assist decision-making, especially in the following situations: i) when it is difficult to decide which model is better for a specific classification problem and, ii) to decide selected unconfirmed cases.

1. Introduction

The classification model, also known as a classifier, is used in order to predict the class label of objects for which the class label is unknown. The performance of a single-model (classifier) can be determined on the basis of the classification accuracy. However, it is difficult to determine which single-model is the best classification technique in a specific application domain since single learning algorithms may not uniformly outperform other algorithms over various datasets.

The multiple-model concept that combines different single-models has been introduced by many researchers with the aim of improving prediction performance. In general, there are three mechanisms used to combine different single-models, which are: i) models created using different subsets of training data based on a single learning method, ii) models created using different training parameters based on a single training method, and iii) models created by combining different learning methods [1]. From the literature, it can be observed that little work has been focused on the classification technique based on the combination of different learning methods (termed as combinative algorithms in this study).
Moreover, most classification algorithms using either fuzzy or non-fuzzy approaches produce results in the form of crisp or categorical classification outcomes. Unfortunately, in certain application domains such as business administration and management science, the classification outcomes that represent class labels may involve categorisations which are fuzzy in nature. For example, the classification outcomes from the single algorithm can be in the form of several classes such as ‘low’, ‘moderate’ and ‘high’, which are obviously more suited to being represented in fuzzy sets. Presently, few techniques have been developed to combine the outcomes from several different algorithms which take into consideration the elements of fuzziness and uncertainties in the classification outcomes. Therefore, this study aimed to develop a new technique to handle the classification task based on combinative algorithms by adopting the multiple-model concepts and a well-known fuzzy aggregation for multiple criteria, namely fuzzy Ordered Weighted Averaging (OWA) operator [2].

The rest of the paper is planned as follows: Section 2 reviews the multiple-model concepts and fuzzy OWA operator. Section 3 describes the proposed fuzzy classification technique based on combinative algorithms. Section 4 presents the implementation of the proposed technique using the Domestic Water Consumption dataset and provides comparison of classification performance of the proposed technique with the performance of every participating single-model as well as with the performance of a multiple model based on the traditional statistical approach. Finally, the work is concluded in Section 5.

2. Background
This section describes the multiple-model concepts and fuzzy Ordered Weighted Averaging (OWA) operator.

2.1. Multiple-model concepts
Various studies show that multiple-models generally have similar purposes, which are to improve prediction over a single-model or to make the resulting ensemble more representative of the observed data, but they differ in their technique of combination. The simplest multiple-model is the classification model constructed on the basis of a statistical approach with the assumption that each participating method is equal in terms of weight. This approach is known as the Pool Method [3]. The multiple-model approach has been claimed to be better than the single-model in many aspects such as accuracy, consistency and reliability [3].

A simple probabilistic multiple-model, namely the Pool Method [3], [4], can be generated by simply pooling together the participating single-models, with all ensemble members having equal weight. The probabilistic multiple-model prediction under kth category, $y_k^{pool}$ is constructed by:

$$y_k^{pool} = \frac{\sum_{\eta=1}^{M} \tilde{m}_{\eta,k}}{\sum_{\eta=1}^{M} M_{\eta}}$$  \hspace{1cm} (1)

where

$\tilde{m}_{\eta,k}$ = number of single - model under category kth,

$M_{\eta}$ = number of participating single - models.

2.2. Fuzzy Ordered Weighted Averaging (OWA) operator
Two extreme cases arise for aggregating criteria functions that have been considered important in order to perform an overall decision in many disciplines. The first case is when ‘all the criteria’ need to be satisfied and the other case is when ‘at least one of the criteria’ is satisfied. Fuzzy Ordered Weighted Averaging (OWA) operator [2] has been introduced as a technique to overcome the problem of multiple-criteria aggregation in decision making. The fuzzy OWA operators provide connections between the logical ‘or’ and ‘and’ for the aggregation process [5].
The fuzzy OWA operator of \( j \) dimension is defined as a mapping \( \hat{F} : I^j \rightarrow I \) where \( I = [0,1] \) that has associated weighting vector, \( \hat{W} \) of dimension \( j \)

\[
\hat{W}_j = [\hat{w}_1, \hat{w}_2, \ldots, \hat{w}_j]
\]

with the following properties

i. \( \hat{W}_j \in [0,1] \)

ii. \( \sum_{j=1}^{j} \hat{W}_j = 1 \)

such that

\[
\hat{F}(a_1, a_2, \ldots, a_j) = \sum_{j=1}^{j} \hat{W}_j R_j = \hat{w}_1 r_1 + \ldots + \hat{w}_j r_j
\]  

(2)

where \( R \) is an ordered argument vector consisting of the arguments of \( \hat{F} \), where \( r_j \) is ranked accordingly to a general criterion \( \hat{F}(a_j) \).

3. Proposed fuzzy classification based on combinative algorithms

This section describes the proposed fuzzy classification based on combinative algorithms. The two main concepts adopted to develop the proposed fuzzy classification technique are the multiple-model concepts [3] and fuzzy Ordered Weighted Averaging (OWA) operator [2]. The multiple-models concept is used to combine different learning methods, whereas the concept of fuzzy OWA operator is used to aggregate the fuzzy outcomes produced by each participating single-model. The proposed method will be used to perform classification tasks. The framework of the proposed fuzzy inference method for combinative algorithms is illustrated in Figure 1 and the process can be described as follows:

**Step 1.** Algorithm selection based on classification performance.

Let \( \hat{A}_\delta = \{\hat{A}_1, \hat{A}_2, \ldots, \hat{A}_m\} \) be classification algorithms where \( m \) is the number of algorithms under study. The percentage of classification accuracy of each algorithm, \( \Phi(\hat{A}_\delta) \) defined as

\[
\Phi(\hat{A}_\delta) = \frac{\text{correctly classified instances}}{\text{total number of instances present in the dataset}} \times 100\%
\]  

(3)

Then, \( \hat{A}_\lambda \) where \( \lambda = 1, 2, \ldots, n \) such that \( n \) is the number of selected algorithms under study with \( n \leq m \), representing the selected algorithm if its performance or its classification accuracy is greater than a predefined degree of confidence, \( \varepsilon \) (determined by expert of the study) such that \( \Phi(\hat{A}_\delta) \geq \varepsilon \). Therefore, algorithms with \( \Phi(\hat{A}_\delta) < \varepsilon \) will be omitted.

**Step 2.** Creation of fuzzy weight vector.

Fuzzy weight vector is based on the classification accuracy of the selected algorithms. The weight vector, \( W_\lambda \) of dimension \( \lambda \) has the same properties as the weighting vector in OWA operator [2] where

\[
W_\lambda = [w_1, w_2, \ldots, w_\lambda, \ldots, w_n]
\]

(4)

such that

\[
W_\lambda \in (0,1) \text{ and } \sum_{\lambda=1}^{\lambda} W_\lambda = 1.
\]
Figure 1. Methodological framework of fuzzy inference method for combinative algorithms

**Step 3.** Creation of fuzzy ordered vector.
A fuzzy ordered vector is created based on the relationship between dataset classes and classification outcomes of each selected algorithm, $\hat{A}_\lambda$. The fuzzy ordered vector, $\sigma$ is defined as

$$
\sigma = \begin{bmatrix}
c(\hat{A}_1) \\
c(\hat{A}_2) \\
\vdots \\
c(\hat{A}_\lambda) \\
\vdots \\
c(\hat{A}_n)
\end{bmatrix},
$$

where the fuzzy classification outcomes of each selected algorithm

$$
c(\hat{A}_l) \in \{c_i\}, \quad t=1,2,...,L,$$

are such that $l$ is the number of dataset classes.

This process will transform the crisp categorical classes into fuzzy classes using linguistic terms in
the form of normal Generalized Trapezoidal Fuzzy Number (GTFN). Each class can be represented as
\[ C_t = [c^l_t, c^u_t, c^m_t, c^r_t], \]
where
\[ C_1 \prec C_2 \prec \cdots \prec C_t, \]
with \( \prec \) representing the ordering of fuzzy classes in ascending order.

**Step 4.** Aggregation of the fuzzy classification outcomes of selected algorithms.

The aggregated fuzzy classification outcomes, \( F(\hat{A}_t) \) are given as
\[ F(\hat{A}_t) = W_2, \sigma = (f_1, f_2, f_3, f_4) \]
where \( W_2 \) is the fuzzy weight vector and \( \sigma \) is the fuzzy ordered vector.

**Step 5.** Determination of fuzzy class.

A fuzzy similarity measure will be used to determine the final classification outcomes. For the purpose of this research, the fuzzy similarity measure proposed in [6] will be employed. Other fuzzy similarity measures may be employed depending on the suitability of the technique with regards to the application domain. The similarity degrees will be calculated between the aggregated fuzzy classification outcomes, \( F(\hat{A}_t) \) and the fuzzy classes, \( C_t \). The degree of similarity \( S(F, C_t) \) is given by
\[ S(F, C_t) = \left\{ \begin{array}{ll}
1 - \frac{\sum_{i=1}^{4}|f_i - c^l_t|}{4} & \left( 1 - \frac{x_F - x_{C_t}}{h_{F_t}} \right) b(S_F, S_{C_t}) \\
\frac{y_F - y_{C_t}}{\max(y_F, y_{C_t})} + \frac{|Ar(F) - Ar(C_t)| + \frac{|Pe(F) - Pe(C_t)|}{\max(Pe(F), Pe(C_t))}}{4} & \end{array} \right. \]
where
\[ y_F = \left\{ \begin{array}{ll}
\frac{w_F (f_3 - f_2 + 2)}{6} & \text{if } f_1 \neq f_4, \\
\frac{w_F}{2} & \text{if } f_1 = f_4 \\
\end{array} \right. \]
\[ x_F = \left\{ \begin{array}{ll}
\frac{y_F (f_3 + f_2) + (f_4 + f_1) w_F - y_F}{2 w_F} & \text{if } w_F \neq 0, \\
\frac{f_4 + f_1}{2} & \text{if } w_F = 0 \\
\end{array} \right. \]
\[ b(S_F, S_{C_t}) = \left\{ \begin{array}{ll}
1 & \text{if } S_F + S_{C_t} > 0 \\
0 & \text{if } S_F + S_{C_t} = 0 \\
\end{array} \right. \]
such that \( S_F = f_4 - f_1 \) and \( S_{C_t} = c^r_t - c^l_t \). The heights of normal GTFN is \( h_F = 1 \) and \( h_{C_t} = 1 \).
\[ Ar(F) = \frac{1}{2} w_F (f_4 + f_3 - f_2 - f_1) \] 
\[ Pe(F) = \sqrt{(f_1 - f_2)^2 + w^2_F} + \sqrt{(f_3 - f_4)^2 + w^2_F} + (f_3 - f_2) + (f_4 - f_1). \]

4. Experiments and findings

A series of experiments was conducted to investigate the performance of the proposed method using the existing Domestic Water Consumption (DWC) dataset obtained from a previous study [7]. This dataset consists of 718 instances classified into two classes which are ‘unlikely excessive’ and ‘likely excessive’. For the purpose of this research, the dataset has been reclassified into three classes which are ‘unlikely excessive’, ‘likely excessive’ and ‘most likely excessive’.

Seven fuzzy classification algorithms were used for the experimentations. The performance of each single-model for the 3-class DWC dataset is presented in Table 1. The experiments were carried out with two different degrees of confidence: sets as \( \varepsilon = 65\% \) and \( \varepsilon = 70\% \), respectively for both Combination-1 and Combination-2.

| Algorithm | Performance (%) |
|-----------|-----------------|
| \( \hat{A}_1 \) | 73.69 |
| \( \hat{A}_2 \) | 75.63 |
| \( \hat{A}_3 \) | 77.18 |
| \( \hat{A}_4 \) | 79.27 |
| \( \hat{A}_5 \) | 79.27 |
| \( \hat{A}_6 \) | 81.22 |
| \( \hat{A}_7 \) | 67.85 |

The 3-class dataset is labelled as \( C_i \), \( C_2 \) and \( C_3 \), respectively to represent ‘unlikely excessive’, ‘likely excessive’ and ‘most likely excessive’. Three sets of selected predefined fuzzy numbers for the 3-class dataset were used for the experimentation (Table 2).

| Linguistic term for fuzzy class | Predefined fuzzy number | Set 1 | Set 2 | Set 3 |
|---------------------------------|-------------------------|------|------|------|
| Unlikely excessive              | \( C_1 = (0.000,0.000,0.045,0.256) \) | \( C_1 = (0.000,0.000,0.210,0.316) \) | \( C_1 = (0.000,0.000,0.045,0.231) \) |
| Likely excessive                 | \( C_2 = (0.045,0.210,0.256,0.843) \) | \( C_2 = (0.135,0.210,0.256,0.316) \) | \( C_2 = (0.151,0.231,0.248,0.271) \) |
| Most likely excessive            | \( C_3 = (0.210,0.843,1.000,1.000) \) | \( C_3 = (0.136,0.248,1.000,1.000) \) | \( C_3 = (0.241,0.323,1.000,1.000) \) |

The performance of single-models, the Pool Method and the proposed method using three sets of predefined fuzzy numbers were investigated under Combination-1 and Combination-2. The results of the experiments for Combination-1 and Combination-2 are presented in Table 3 and Table 4, respectively. Note that the results were obtained on the basis of 10-fold cross-validation methods.
Table 3. The performance of single-models, the Pool Method and the proposed method under Combination-1

| Algorithm | Single-model | Pool Method | Proposed method Set 1 | Proposed method Set 2 | Proposed method Set 3 |
|-----------|--------------|-------------|-----------------------|-----------------------|----------------------|
| $\hat{A}_1$ | 0.7369       |             |                       |                       |                      |
| $\hat{A}_2$ | 0.7563       |             |                       |                       |                      |
| $\hat{A}_3$ | 0.7718       |             |                       |                       |                      |
| $\hat{A}_4$ | 0.7927       |             | 0.8760                | 0.9053                | 0.8928               |
| $\hat{A}_5$ | 0.7927       |             |                       |                       | 0.8886               |
| $\hat{A}_6$ | 0.8122       |             |                       |                       |                      |
| $\hat{A}_7$ | 0.6785       |             |                       |                       |                      |

Table 4. The performance of single-models, the Pool Method and the proposed method under Combination-2

| Algorithm | Single-model | Pool Method | Proposed method Set 1 | Proposed method Set 2 | Proposed method Set 3 |
|-----------|--------------|-------------|-----------------------|-----------------------|----------------------|
| $\hat{A}_1$ | 0.7369       |             |                       |                       |                      |
| $\hat{A}_2$ | 0.7563       |             |                       |                       |                      |
| $\hat{A}_3$ | 0.7718       |             |                       |                       |                      |
| $\hat{A}_4$ | 0.7927       |             | 0.8649                | 0.8719                | 0.8747               |
| $\hat{A}_5$ | 0.7927       |             |                       |                       | 0.8788               |
| $\hat{A}_6$ | 0.8122       |             |                       |                       |                      |
| $\hat{A}_7$ |             |             |                       |                       |                      |

The first objective of this research is to compare the performance of the single-model with the performance of the proposed technique. Results from the analysis of classification performance show that the highest performance of a single-model under Combination-1 and Combination-2 is produced by $\hat{A}_6$ with classification performance of 0.8122. However, the performance of the combinative algorithms for both the proposed method and the Pool Method showed that the multiple-model concept, using either the fuzzy approach or the non-fuzzy approach, is better than the performance of each participating single-model. Hence, it can be concluded that the performance of the proposed technique is better than each participating single-model. The second objective is to compare the performance of the proposed technique with the probabilistic Pool Method. The results show that the Pool Method produced higher performance under Combination-1 with 0.8760 compared to Combination-2 with 0.8649. For Combination-1, the performance of the proposed method using Set 1, Set 2 and Set 3 is 0.9053, 0.8928 and 0.8886, respectively, whereas for Combination-2, the performance of the proposed method is 0.8719, 0.8747 and 0.8788, respectively. The findings clearly show how the performance of the proposed method for both Combination-1 and Combination-2 is higher than the performance of the Pool Method. These results indicate that the proposed method produced better performance compared with that of all selected single-models and the Pool Method.

5. Conclusion

This paper has presented a fuzzy classification for combinative algorithms by adopting the concepts of the multiple-model and the fuzzy Ordered Weighted Averaging (OWA) operator. It has been shown that the proposed combinative algorithms have been successfully developed and implemented. The findings of this research showed that the proposed method is better in terms of classification accuracy compared with the best single-model as well as the probabilistic Pool Method and hence can be used to improve prediction. Additionally, the finding also showed that the proposed method has the potential to be used for datasets with fuzzy classes such as the Domestic Water Consumption (DWC)
dataset. These results indicate that the proposed technique has the potential to assist decision-making, for example, in the situation when it is difficult to decide which model is better for a specific classification problem or to decide selected unconfirmed cases decided by several single algorithms.

Acknowledgements
This research work is funded by the Ministry of Education, Malaysia under the Fundamental Research Grant Scheme (FRGS) with reference number 600-RMI/FRGS TD 5/3 (1/2015). The authors would like to thank the Institute of Research Management and Innovation (IRMI), Universiti Teknologi MARA, Malaysia.

References
[1] Kotsiantis S B, Zaharakis I and Pintelas P 2007 Supervised machine learning: A review of classification techniques, Emerging Artificial Intelligence Applications in Computer Engineering 160 3-24
[2] Yager R R 1988 On ordered weighted averaging aggregation operators in multicriteria decisionmaking, IEEE Transactions on Systems, Man, and Cybernetics 18 183-190
[3] Weigel A P, Liniger M A and Appenzeller C 2008 Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts?, Quarterly Journal of the Royal Meteorological Society 134 241-260
[4] Hagedorn R, Doblas-Reyes F J and Palmer T N 2005 The rationale behind the success of multi-model ensembles in seasonal forecasting – I. Basic concept, Tellus 57 219–233
[5] Yager R R 1993 Families of OWA operators, Fuzzy Sets and Systems 59 125-148
[6] Mat Saffie N A, Rasmani K A and Sulaiman N H 2017 Similarity measure for fuzzy number based on distances and geometric shape characteristics, presented at The 3rd International Conference on Computing, Mathematics and Statistics 2017 (iCMS 2017), Langkawi Malaysia, 7-8 November
[7] Jalaluddin A S, Shahari N A and Rasmani K A 2016 Excessive water usage estimation based on domestic per capita water consumption and consumer data, Advanced Science Letters 22 2959-2962