Journal of Engineering and Applied Sciences 14 (19): 7210-7217, 2019
ISSN: 1816-949X

Investigating the Applicability of Several Fuzzy-Based Classifiers on Multi-Label Classification

[1]Mo'ath Al-luwaici, [1]Ahmad Kadri Junoh and [2]Farzana Kabir Ahmad
[1]Institute of Engineering Mathematics, Universiti Malaysia Perlis (UniMAP), Arau, Malaysia
[2]School of Computing, Universiti Utara Malaysia (UUM), Kedah-Sintok, Malaysia

Abstract: In the last few decades, fuzzy logic has been extensively used in several domains such as economy, decision making, logic and classification. In specific, fuzzy logic which is a powerful mathematical representation has shown a superior performance with uncertainty real-life applications comparing with other learning approaches. Many researchers utilized the concept of fuzzy logic in solving the traditional single label classification problems of both types: binary classification and multi-class classification. Unfortunately, very few researches have utilized fuzzy logic in a more general type of classification that is called Multi-Label Classification (MLC). Hence, this study aims to examine the applicability of fuzzy logic to be used with MLC through evaluating several fuzzy-based classifiers on five different multi-label datasets. The results revealed that the utilizing fuzzy-based classifiers on solving the problem of MLC is promising comparing with a wide range of MLC algorithms that belong to several learning approaches and strategies.

Key words: Classification, fuzzy-logic, fuzzy-based classifiers, machine learning, multi-label classification, datasets

INTRODUCTION

A Fuzzy Logic System (FLS) could be defined as a nonlinear mapping from an input dataset into a scalar output dataset (Prabha and Chitra, 2017). A general FLS comprises of four main components: fuzzifier, rules, inference engine and defuzzifier (Elkano *et al*., 2018; Xiang *et al*., 2018). FLS usually deals with fuzzy data, therefore, any input crisp dataset must be converted into a fuzzy dataset using membership functions and fuzzy linguistics variable (Fiderek *et al*., 2017). This process is known as fuzzification.

In general, FLSs have shown a superior performance on modeling complex problems that belong to different classical and modern domains, i.e., decision making, economics, regression and classification (Sundarabalan and Selvi, 2017). In specific, Fuzzy Rule-Based Classification Systems (FRBCSs) have attracted several researchers in the last few decades and have been proven to show a superior performance when compared with other learning approaches and strategies (Xiang *et al*., 2018). The success of FRBCSs could be due to the following two reasons. Firstly, the high interpretability of FRBCSs that could be understand easily by both the ordinary and the domain experts. Secondly, the proper management of the uncertainty achieved through fuzzy-based classifiers.

Unfortunately, most FRBCSs have been applied and tested against single label datasets and very few attempts have utilized the strong capabilities of these FRBCSs in solving the problem of Multi-Label Classification (MLC) (Alazaidah and Ahmad, 2016). MLC is a general type of classification that has attracted several scholars in the last two decades, due to its high applicability to modern real-life applications and domains such as: bioinformatics, Web mining, text classification, scene classification among others (Gibaja and Ventura, 2015).

On the contrary of the conventional single label classification, MLC allows instances in the dataset to be associated with more than one class labels at the same time (Thabtah, 2018). Hence, several class labels in MLC may share the same observation which represents the overlapping aspect of the MLC that makes it a perfect domain for fuzzy logic in general (Prabha and Chitra, 2017).

Consequently, this study aims to examine the applicability of fuzzy logic to handle the problem of MLC, through evaluating several fuzzy-based classifiers on five different multi-label datasets. Also, this study aims to assess fuzzy logic as a learning approach with respect to a several learning approaches that have been represented by several classifiers. Finally, this study aims to determine the most promising fuzzy-based classifier in handling multi-label datasets in order to adapt it to solve the problem of MLC.

Corresponding Author: Mo'ath Al-luwaici, Institute of Engineering Mathematics, Universiti Malaysia Perlis (UniMAP), Arau, Malaysia

Literature review: In this study, an overview of the main concepts related to MLC problem and also a description of the main fuzzy-based classifiers considered in this study. Finally, the few attempts that have utilized fuzzy logic in MLC is discussed.

Overview of MLC: Classification which is one of the main tasks in data mining aims to predict the class label of unseen case as accurately as possible, based on the learning step from the labeled instances (Abdelhamid *et al.*, 2014). In general, classification is divided into three main types. The first type consists only of two class labels and it is called binary classification. The second type is called multi-class classification and consists of more than two class labels (Qabajeh *et al.*, 2015). Both binary and multi-class classification require all instances in the dataset to be associated with only one class label. Hence, class labels in binary and multi-class classification problems are always mutually exclusive (Huang *et al.*, 2017). On the other hand, MLC is more general type of classification that allows instances in the dataset to be associated with any number of class labels at the same time (Gibaja and Ventura, 2015). Thus, class labels in MLC are not mutually exclusive as in the previous types of classification.

As the observations of one instance could be associated with more than one class label at the same time, this makes the learning step of the classifier much more complicated. In fact, the degree of uncertainty in MLC is very high comparing with binary and multi-class classification problems (Corani and Scanagatta, 2016) which makes it the perfect domain for fuzzy logic learning approach.

Generally speaking, two main learning approaches are being used in solving the problem of MLC. The first approach attempts to convert the input multi-label dataset into a single label dataset or more and then, applies any single label classification algorithm (or more) on the converted dataset (s) to construct a classifier. This learning approach has been known as PTM (Alazaidah *et al.*, 2017). The second learning approach attempts to modify or adapts a single label classification algorithm to handle multi-label datasets. This approach has been knows as Algorithm Adaptation Method (AAM) (Markatopoulou *et al.*, 2015). According to the existing literature, very few fuzzy-based single label classification algorithms have been adapted to handle MLC (Alazaidah *et al.*, 2015). Therefore, this study investigates the most promising fuzzy-based classification algorithm to be adapted to handle the problem of MLC with respect to two main criteria: the accuracy of the classification algorithms firstly and the running time of the classification algorithm secondly. This study describes the fuzzy-based single label classification algorithm considered in this study.

MLC has special distinguishable features that could not be applied to binary and multi-class classification. First, the problem search space of any MLC problem is high when compared with the problem search space of both binary and multi-class classification (Gibaja and Ventura, 2015). This is because of the dependency and correlations among the labels that share the same instances (observations). Also, in the contrary of the single label classification (binary classification and multi-class classification), class labels in MLC share some kind of dependencies which should be captured and exploited to reduce the large problem search space of the MLC problem and hence, make the proposed algorithm feasible (Kong *et al.*, 2013). Consequently, it is very significant that the adapted single label classification algorithm has the ability to deal with large number of labels and instances as well in addition to having the capabilities of exploiting, the existing dependencies and correlations among labels.

In fact, several single label classification algorithms have been adapted to handle the problem of MLC. These algorithms belong to different learning strategies and approaches such as decision trees (Comite *et al.*, 2003), neural network (Zhang, 2009), support vector machines (Cesa-Bianchi *et al.*, 2006), associative classification (Thabtah *et al.*, 2004) and several other learning approaches. Nevertheless, based on the existing algorithms, the accuracy of these algorithms still not as expected and much research works should be conducted to determine the best learning approach and the optimal algorithm that has the ability to handle multi-label datasets that are usually suffer from high complexity; due to the large number of features, instances and class labels (Gibaja and Ventura, 2015).

Now a days, MLC is increasingly required; due to its applicability to a wide range of modern domains such as: medical diagnosis (Yan *et al.*, 2010), text classification (Hadi *et al.*, 2018), social network mining (Krohn-Grimberghe *et al.*, 2012), direct marketing (Sousa and Gama, 2018) and drug discovery (Kawai and Takahashi, 2009) among others.

Fuzzy-based classifiers: In this study, a brief description of the considered fuzzy-based classifiers in this study is presented where twelve fuzzy-based classifiers have been considered in the evaluation phase.

Fuzzy Rough Nearest Neighbors (FRNN) (Sarkar, 2007) is an adaptation of the popular algorithm k Nearest Neighbors (kNN) (Aha *et al.*, 1991) using fuzzy logic (Sarkar, 2007). In FRNN, the k neighbors are not with the same importance but they are weighted, according to their distance to a given test example and their contribution to each class membership of each training fuzzy patterns. Even though FRNN showed an improvement over the original KK but it also inherited some of its drawbacks like the high running time, especially with large datasets. Another drawback is the need to store all the training data in the main memory which makes FRNN is not an appropriate choice for datasets with high number of features and instances as well. The FRNN-FRS and FRNN-VQRS algorithms (Jensen and Cornelis, 2011) are an adaptation of the original FRNN algorithm where the nearest neighbors are utilized to the lower and upper approximations of the decision classes. Evaluation results of FRNN-FRS and FRNN-VQRS on eight different datasets showed that both algorithms outperform FRNN algorithm (Jensen and Cornelis, 2011).

By Denoeux (1995), a kNN-based classification algorithm has been proposed. The algorithm dubbed as D-SkNN. D-SkNN addresses the classification problem from the point of view of Dempster Shafer theory. An evaluation of proposed algorithm on several real-life datasets showed an effectiveness comparing with other voting and distance-weighted kNN algorithms.

By Bezdek and Chuah (1986), a general framework to generalize the kNN algorithm to datasets that are not necessarily crisp has been discussed. The researchers have presented their algorithm (FENN) and compared it against other existing algorithms using several datasets and metrics. The FENN algorithm showed a competitive performance with respect to other existing classification algorithms.

The fuzzy-kNN algorithm which is an adaptation of the popular kNN algorithm based on the concepts of the fuzzy sets has been proposed by Keller *et al*.. The fuzzy-kNN utilizes three methods of assigning to the labeled samples. Results on three single label datasets showed that fuzzy-kNN performance is superior comparing with the original kNN algorithm.

By Hu and Xie (2006), an attempt to enhance kNN algorithm using Genetic Algorithm (GA) has been proposed. The enhanced version of kNN dubbed GA fuzzy kNN and depends on a parallel implementation of GA to speed up the process of building the classifier. Another adaptation of kNN could be found by Kuncheva (1995) where Intuitionistic Fuzzy Set (IFS) has been utilized to propose the adapted IFS-KNN algorithm.

By Rhee and Hwang (2003), an extension of type 1 fuzzy kNN has been presented. They dubbed their algorithm IT2F-KNN where the membership values for pattern vectors have been adapted based on type-2 fuzzy membership interval. They proved that their algorithm overcomes other algorithms that belong to crisp and type-1 fuzzy approaches.

The Pruned Fuzzy k Nearest Neighbor (PF-kNN) algorithm has been proposed by Arif *et al*. (2010). Their algorithm aims to classify arrhythmia beat where 103100 beats that belong to six different classes have been classified. Their algorithm is easy to implement but suffer from time complexity when dealing with large number of example and hence, more complicated pruning techniques are needed.

By Han and Kim (1999), another kNN-based algorithm has been proposed and named Variance Weighted Fuzzy kNN. Their main idea was to utilize the standard deviation in giving weights to the neighbors. Their algorithm has been compared against the original kNN algorithm and fuzzy-kNN algorithm and showed a fair enhancement over the previously mentioned two algorithms.

MATERIALS AND METHODS

Utilizing fuzzy logic in MLC: SLC is much easier than MLC in selecting the most related features to a specific class. In SLC, every instance in the training set is associated with only one class label whereas in MLC, an instance could be associated with multiple class labels at the same time. Features in MLC are shared among several class labels with several degrees of correlations. The same feature selection methods that are being used with SLC are being utilized in MLC (Doquire and Verleysen, 2011; Zhang and Wu, 2015). This study briefly explains the presented novel features selection methods which have been designed specifically, for MLC problem. An algorithm named multi-label learning with label-specific features (LIFT) was proposed by Zhang and Wu (2015). The researcher of this algorithm questioned the advantages of using the same feature selection methods that have been designed to suit the problem of SLC with the MLC problem. Hence, LIFT algorithm has been proposed based on a novel feature selection method that is more appropriate to the nature of MLC. The first step in LIFT is groping the instances in the training set using clustering techniques where each cluster have all instances that are associated with a specific class labels. In the second step, the instances in each cluster are used to construct the features specified to every class label in the label set. Finally, the discovered features are selected and (k) classifiers are trained for every label in the label set. To conclude the usefulness of LIFT, it was evaluated using eight different datasets from several domains and compared against several state-of-art algorithms. The researchers conclude the high need of designing new feature selection methods that consider the nature of the MLC problem and utilize the dependencies among labels.

Xu *et al*. (2016) proposed two algorithms of MLC with label specific feature reduction. The goal of these algorithms is to solve the disadvantage of LIFT algorithm which is the redundancy in information. The first algorithm has been called Fuzzy Rough Set (FRS-LIFT) which adopts the theory of FRS as a mathematical tool in feature selection and reduction. In FRS-LIFT, the strategy of forward greedy search is helping the approximation quality in order to assess the significance of certain dimensions. The second algorithm has been called FRS-SS-LIFT which aims to enhance FRS-LIFT by applying Sampling Selection (SS) technique. Five steps are forming this algorithm. First step as in the LIFT algorithm, a label specific features space is built. Second step is SS technique is used to reduce the number of instances. Third step is applying dimensionality reduction for sampled dataset from previous step using principles of FRS. The fourth step is using the output of step 3 to build (k) classification models. Fifth step is the evaluation and the prediction of unseen new instances is done based on (k) classification models. An evaluation of these two algorithms used 10 datasets from several domains and compared against three different algorithms (ML-kNN, MLNB and LIFT). The result of the experimental evaluation showed that the efficiency of the two algorithms and their advantage over the other compared algorithms. The disadvantages of both algorithms are their complexity and ignoring the existing correlations among labels in the dataset.

RESULTS AND DISCUSSION

Empirical analysis: In this study, a complete description of the conducted research is demonstrated. First, the datasets which have been used in this research are described. Then, the evaluation results of the twelve fuzzy-based classifiers are presented. It is worth mentioning that the main evaluation criteria considered in this study are accuracy and the running time. Accuracy metric is calculate the percentage of the correctly predicted labels with respect to the total number of labels and instances. Accuracy metric is computed using the following Eq. 1:

$$\text{Accuracy} = \frac{1}{t}\sum_{i=1}^{t}\frac{|Z_i \cap Y_i|}{|Z_i \cup Y_i|} \quad (1)$$

Datasets characteristics: The five multi-label datasets with different characteristics have been considered in the evaluation of the twelve fuzzy-based classifiers. Table 1 depicts the main characteristics of the five multi-label datasets considered in this study. Datasets could be found in Mulan, a multi-label datasets repository (Tsoumakas *et al*., 2011).

All experiments have been conducted using KEEL (Triguero *et al*., 2017) which is an open source Java Software that is applicable to data mining tasks. KEEL is short for knowledge extraction based on evolutionary learning. Finally, it is worth mentioning that all the twelve fuzzy-based classifiers have been used with their default settings as they have been implemented in KEEL.

Table 1: Datasets main characteristics

Data sets	Instances	Attributes	Labels	LCard	Domain
Yeast	2417	103	14	4.327	Biology
Scene	2712	294	6	1.074	Image
Emotions	593	72	6	1.868	Media
Flags	194	19	7	3.392	Image
Genbase	662	1186	27	1.252	Biology

Table 2: Evaluation results of several fuzzy-based classifiers using accuracy metrics

Classifiers	Emotions	Scene	Flags	Yeast	Genbase	Average
FRNN-FRS	0.997	1.000	0.896	1.000	1.000	0.9786
FRNN	0.439	0.546	0.510	1.000	0.907	0.6804
D-SkNN	0.815	0.561	0.737	0.843	0.987	0.7886
FENN	0.575	0.572	0.597	0.504	0.987	0.6470
FRNN-VQRS	0.994	0.533	0.752	0.621	1.000	0.7800
Fuzzy-kNN	0.997	1.000	0.907	1.000	1.000	0.9808
GAfuzzykNN	0.997	1.000	0.917	0.469	1.000	0.8766
IF-kNN	0.629	0.717	0.639	0.592	0.986	0.7126
IFS-kNN	0.626	0.598	0.608	0.627	0.907	0.6732
IT2F-kNN	0.997	1.000	0.917	1.000	1.000	0.9828
PFKNN	0.501	0.974	0.881	0.996	0.919	0.8542
VWfuzzykNN	0.626	0.591	0.561	0.459	0.980	0.6434

The five multi-label datasets have been transformed into single label datasets using the Most Frequent Label (MFL) transformation method. MFL transformation method starts with counting the frequency of all labels in the training set. Then, all multi-label instances are transformed to be linked with the most frequent label they are associated to (Alazaidah *et al*., 2018).

Evaluation of several fuzzy-based classifiers on multi-label datasets: In general, multi-label datasets have several distinguishable features over the single label datasets such as the high dimensionality, high number of labels and most features in the datasets are usually continuous with big range of value. Thus, the high predictive performance of any single label classifier on the single label datasets does not mean that it will have the same superior performance on multi-label datasets. Therefore, this section is more interested in evaluating the twelve fuzzy-based classifiers on different multi-label datasets in order to identify the best classifier to handle multi-label datasets and consequently, adapts this classifier to handle the problem of MLC. Table 2 to depict the evaluation results of the twelve fuzzy-based classifiers on the five multi-label datasets considered in this study using accuracy as an evaluation metric.

From Table 2, it can be clearly seen that four classifiers have nearly a similar accuracy on most datasets which are they; FRNN-FRS, fuzzy-kNN, GA fuzzy kNN and IT2F-kNN. Therefore, based on accuracy metric, it can be concluded that the previously mentioned

Table 3: Evaluation results of several fuzzy-based classifiers based on running time

Classifiers	Emotions	Scene	Flags	Yeast	Genbase	Average
FRNN-FRS	0.000	0.000	0.000	0	0	0
FRNN	0.000	0.000	0.000	0	0	0
D-SkNN	0.009	0.034	0.015	0,047	0.046	0.026
FENN	0.058	0.265	0.006	0.625	1.594	0.5096
FRNN-VQRS	0.000	0.000	0.000	0	0	0
Fuzzy-kNN	0.000	0.000	0.000	0	0	0
GAfuzzykNN	9.575	42.811	1.292	332.088	944.088	265.9708
IF-kNN	0.029	0.110	0.006	0.313	0.797	0.251
IFS-kNN	0.049	0.006	0.012	0.016	0.031	0.0228
IT2F-kNN	0.127	0.432	0.020	1.64	3.95	1.2338
PFKNN	1.111	5.979	0.134	182.29	3.313	38.5654
VWfuzzykNN	0.029	0.137	0.006	0.312	0.797	0.2562

four fuzzy-based classifiers are the most promising fuzzy-based classifiers to be adapted to handle the problem of MLC.

Another main criterion that could be helpful to break the ties among the best four fizzy-based classifiers is the running time criterion. Table 3 depicts the running time for the twelve fuzzy-based classifiers on the five multi-label datasets in seconds.

Among the twelve fuzzy-based classifiers, it can be clearly noted that FRNN-FRS and fuzzy-kNN have the best running time on all the five multi-label datasets. Therefore, based on Accuracy and running time metrics, it can be concluded that FRNN-FRS and fuzzy-kNN are the best classifiers among others to be adapted to handle MLC problem.

Evaluating fuzzy-based classifiers with respect to other classifiers from different learning approaches: This study aims to assess the applicability of the best two fuzzy-based classifiers determined in this study with respect to other base classifiers from several learning approaches and strategies. Every learning approaches has been represented by two algorithms that are well-known to have a good predictive performance on single label datasets. Decision trees learning approach has been represented by the C4.5 (Quinlan, 1993) algorithm and DT-GA (Carvalho and Freitas, 2004) algorithm where lazy learning approach have been represented by the kNN (Cover and Hart, 1967) and the kNN-adaptive (Wang *et al*., 2007) algorithms. Neural network learning approach has been represented by the GANN (Miller *et al*., 1989) and the NNEP (Martinez-Estudillo *et al*., 2008) algorithms. Finally, statistical learning approach has been represented by the logistic (Le Cessie and Houwelingen, 1992) and the LDA algorithms. All the eight previously mentioned base classifiers have been trained on the transformed version of emotions, scene, flags, yeast and genbase datasets and the accuracy of these base classifiers have been used to assess which learning approach is the best to handle multi-label datasets. Table 4 depicts the accuracy of several base classifiers algorithms that belong to different learning approaches on the five multi-label datasets (Emotions. Scene, Flags, Yeast and Genbase).

Table 4: Accuracy rates for several classifiers on the five multi-label datasets

Learning strategy/Algorithm	Emotions	Scene	Flags	Yeast	Genbase
Fuzzy					
FRNN-FRS	0.997	1.000	0.896	1.000	1.000
Fuzzy-kNN	0.997	1.000	0.907	1.000	1.000
Decision Tree					
C4.5	0.851	0.751	0.510	0.872	0.993
DT-GA	0.630	0.396	0.510	0.454	0.252
Neural network					
GANN	0.553	0.480	0.618	0.368	0.403
NNEP	0.612	0.496	0.530	0.385	
Lazy learning					
kNN	0.546	0.368	0.396	0.352	0.867
kNN-adaptive	0.694	0.468	0.469	0.423	0.953
Statistical					
Logistic	0.581	0.468	0.515	0.524	
LDA	0.597	0.450	0.510	0.492	0.111

Table 5: Running times for several classifiers on the five multi-label datasets

Learning strategy/Algorithm	Emotions	Scene	Flags	Yeast	Genbase
Fuzzy					
FRNN-FRS	0	0	0	0	0
Fuzzy-kNN	0	0	0	0	0
Decision tree					
C4.5	0	0	0	0	5.34
DT-GA	0	1.023	0	0	0
Neural network					
GANN	361	345	321	2545	10983
NNEP	187	196	59	2343	8620
Lazy learning					
kNN	0	0	0	0	0
kNN-adaptive	0.015	0.033	0.005	0.24	0.719
Statistical					
Logistic	0.023	0.47	0.012	2287	3.15
LDA	0	0.34	0	23	36.45

It is obvious clear from Table 4 that FRNN-FRS and fuzzy-kNN algorithms overcome all the other eight algorithms that belong to four learning strategies on the five multi-label datasets. In fact, the accuracy of the two fuzzy-based algorithm is superior comparing with all other algorithms, especially, for the fuzzy-kNN algorithm which highly indicates the high applicability of the fuzzy learning approaches in handling the multi-label datasets that usually suffer from high cardinality, imbalance class distribution and the high dimensionality problems. Table 5 depicts the running time for several classifiers from different learning strategies on the five multi-label datasets considered in this study.

Table 5 clearly shows that the two fuzzy-based algorithms are among the best running time algorithms with the kNN algorithm that follow the lazy learning approach. To summarize this study, the two fuzzy-based algorithms (FRNN-FRS and Fuzzy-kNN) have shown a superior performance over other classification algorithms from several learning approaches and based on two evaluation metrics: accuracy of the output classifier and the total running time of the algorithm. Hence, it can be concluded that utilizing the high capabilities of fuzzy-based algorithms in solving the problem of MLC is fruitful and promising. Therefore, a good future research to suggest is to adapt the FRNN-FRS or the fuzzy-kNN algorithm to handle multi-label datasets with considering the correlations among labels that have been stated by several researchers (Alazaidah *et al*., 2018; Gibaja and Ventura, 2015; Wang *et al*., 2018) is the key solution to reduce the large problem search space of the MLC problem.

CONCLUSION

In this study, an investigate study has been conducted, The study aimed to evaluate and assess the applicability of several fuzzy-based classifiers on solving the problem of MLC. In specific, twelve fuzzy-based classifiers have been trained on five different multi-label datasets. Two metrics have been used in evaluating these classifiers: accuracy and the running time where the results revealed that the FRNN-FRS and the Fuzzy-kNN are the best fuzzy-based classifiers in handling multi-label datasets.

Also, the study examined the performance of the best two fuzzy-based classifiers with respect to eight different algorithms that belong to four learning strategies. The results revealed that the performance of the two fuzzy-based classifiers is superior in term of Accuracy and running time.

RECOMMENDATIONS

As a future work, it is highly recommended to adapt FRNN-FRS and Fuzzy-KNN algorithms to handle the problem of MLC with considering the correlations among labels, especially, the local positive high order correlations. Also, extending this study to consider more learning strategies and other evaluation metrics would be a promising research work.

REFERENCES

Abdelhamid, N., A. Ayesh and W. Hadi, 2014. Multi-label rules algorithm based associative classification. Parallel Process. Lett., 24: 1450001-1-1450001-21.

Aha, D.W., D. Kibler and M.K. Albert, 1991. Instance-based learning algorithms. Mach. Learn., 6: 37-66.

Alazaidah, R. and F.K. Ahmad, 2016. Trending challenges in multi label classification. Intl. J. Adv. Comput. Sci. Appl., 1: 127-131.

Alazaidah, R., F. Thabtah and A.Q. Radaideh, 2015. A multi-label classification approach based on correlations among labels. Intl. J. Adv. Comput. Sci. Appl., 6: 52-59.

Alazaidah, R., F.K. Ahmad and M.F.M. Mohsen, 2017. A comparative analysis between the three main approaches that are being used to solve the problem of multi label classification. Intl. J. Soft Comput., 12: 218-223.

Alazaidah, R., F.K. Ahmad, M.F.M. Mohsen, A.K. Junoh and M. Allwaise, 2018. Evaluating conditional and unconditional correlations capturing strategies in multi label classification. J. Telecommun. Electron. Comput. Eng., 10: 47-51.

Arif, M., M.U. Akram and F.U.A.A. Minhas, 2010. Pruned fuzzy k-nearest neighbor classifier for beat classification. J. Biomed. Sci. Eng., 3: 380-389.

Bezdek, J.C. and S.K. Chuah, 1986. Generalized k-nearest neighbor rules. Fuzzy Sets Syst., 18: 237-256.

Carvalho, D.R. and A.A. Freitas, 2004. A hybrid decision tree/Genetic algorithm method for data mining. Inf. Sci., 163: 13-35.

Cesa-Bianchi, N., C. Gentile and L. Zaniboni, 2006. Hierarchical classification: Combining Bayes with SVM. Proceedings of the 23rd International Conference on Machine Learning, June 25-29, 2006, ACM, Pittsburgh, Pennsylvania, USA., ISBN:1-59593-383-2, pp: 177-184.

Comite, D.F., R. Gilleron and M. Tommasi, 2003. Learning multi-label alternating decision trees from texts and data. Proceedings of the 3rd International Workshop on Machine Learning and Data Mining in Pattern Recognition, July 5-7, 2003, Springer, Berlin, Germany, pp: 35-49.

Corani, G. and M. Scanagatta, 2016. Air pollution prediction via. multi-label classification. Environ. Modell. Software, 80: 259-264.

Cover, T. and P. Hart, 1967. Nearest neighbor pattern classification. IEEE Trans. Inform. Theory, 13: 21-27.

Denoeux, T., 1995. A K-nearest neighbor classification rule based on Dempster-Shafer theory. IEEE. Trans. Syst. Man. Cybern., 25: 804-813.

Doquire, G. and M. Verleysen, 2011. Feature Selection for Multi-Label Classification Problems. In: Advances in Computational Intelligence, Cabestany, J., I. Rojas and G. Joya (Eds.). Springer, Berlin, Heidelberg, Germany, ISBN:978-3-642-21500-1, pp: 9-16.

Elkano, M., M. Galar, J.A. Sanz, P.F. Schiavo and S. Pereira Jr. *et al*., 2018. Consensus via penalty functions for decision making in ensembles in fuzzy rule-based classification systems. Appl. Soft Comput., 67: 728-740.

Fiderek, P., J. Kucharski and R. Wajman, 2017. Fuzzy inference for two-phase gas-liquid flow type evaluation based on raw 3D ECT measurement data. Flow Meas. Instrum., 54: 88-96.

Gibaja, E. and S. Ventura, 2015. A tutorial on multilabel learning. ACM. Comput. Surv., 47: 52-52.

Hadi, W., Q.A. Al-Radaideh and S. Alhawari, 2018. Integrating associative rule-based classification with Naive Bayes for text classification. Appl. Soft Comput., 69: 344-356.

Han, J.H. and Y.K. Kim, 1999. A fuzzy k-NN algorithm using weights from the variance of membership values. Proceedings of the 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149), June 23-25, 1999, IEEE, Fort Collins, Colorado, ISBN:0-7695-0149-4, pp: 394-399.

Hu, X. and C. Xie, 2006. Improving fuzzy k-NN by using genetic algorithm. J. Comput. Inf. Syst., 1: 203-213.

Huang, J., G. Li, S. Wang, Z. Xue and Q. Huang, 2017. Multi-label classification by exploiting local positive and negative pairwise label correlation. Neurocomputing, 257: 164-174.

Jensen, R. and C. Cornelis, 2011. Fuzzy-Rough Nearest Neighbour Classification. In: Transactions on Rough Sets XIII, Peters, J.F., A. Skowron, C.C. Chan, J.W. Grzymala-Busse and W.P. Ziarko (Eds.). Springer, Berlin, Heidelberg, Germany, ISBN:978-3-642-18301-0, pp: 56-72.

Kawai, K. and Y. Takahashi, 2009. Identification of the dual action antihypertensive drugs using TFS-based support vector machines. Chem Bio Inf. J., 9: 41-51.

Kong, X., B. Cao and P.S. Yu, 2013. Multi-label classification by mining label and instance correlations from heterogeneous information networks. Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '13), August 11-14, 2013, ACM, Chicago, Illinois, USA., ISBN:978-1-4503-2174-7, pp: 614-622.

Krohn-Grimberghe, A., L. Drumond, C. Freudenthaler and L. Schmidt-Thieme, 2012. Multi-relational matrix factorization using bayesian personalized ranking for social network data. Proceedings of the 5th ACM International Conference on Web Search and Data Mining (WSDM '12), February 08-12, 2012, ACM, Seattle, Washington, USA., ISBN:978-1-4503-0747-5, pp: 173-182.

Kuncheva, L.I., 1995. An intuitionistic fuzzy K-nearest neighbors rule. Notes Intuitionistic Fuzzy Sets, 1: 56-60.

Le Cessie, S. and J.C. Van Houwelingen, 1992. Ridge estimators in logistic regression. Appl. Stat., 41: 191-201.

Markatopoulou, F., G. Tsoumakas and I. Vlahavas, 2015. Dynamic ensemble pruning based on multi-label classification. Neurocomputing, 150: 501-512.

Martinez-Estudillo, F.J., C. Hervas-Martinez, P.A. Gutierrez and A.C. Martinez-Estudillo, 2008. Evolutionary product-unit neural networks classifiers. Neurocomputing, 72: 548-561.

Miller, G.F., P.M. Todd and S.U. Hegde, 1989. Designing neural networks using genetic algorithms. Proceedings of the 3rd International Conference on Genetic Algorithms, June 4-7, 1989, Morgan Kaufmann Publishers Inc., San Francisco, California, USA., ISBN:1-55860-006-3, pp: 379-384.

Prabha, M.G. and S. Chitra, 2017. Design and development of an efficient hierarchical approach for multi-label protein function prediction. Biomed. Res., 28: S370-S379.

Qabajeh, I., F. Thabtah and F. Chiclana, 2015. A dynamic rule-induction method for classification in data mining. J. Manage. Anal., 2: 233-253.

Quinlan, J.R., 1993. C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, USA., ISBN: 9781558602380, Pages: 302..

Rhee, F.C.H. and C. Hwang, 2003. An interval type-2 fuzzy K-nearest neighbor. Proceedings of the 12th IEEE International Conference on Fuzzy Systems (FUZZ '03), May 25-28, 2003, IEEE, St. Louis, Missouri, ISBN:0-7803-7810-5, pp: 802-807.

Sarkar, M., 2007. Fuzzy-rough nearest neighbor algorithms in classification. Fuzzy Sets Syst., 158: 2134-2152.

Sousa, R. and J. Gama, 2018. Multi-label classification from high-speed data streams with adaptive model rules and random rules. Prog. Artif. Intell., 7: 177-187.

Sundarabalan, C.K. and K. Selvi, 2017. Real coded GA optimized fuzzy logic controlled PEMFC based Dynamic Voltage Restorer for reparation of voltage disturbances in distribution system. Intl. J. Hydrogen Energy, 42: 603-613.

Thabtah, F., 2018. Machine learning in autistic spectrum disorder behavioral research: A review and ways forward. Inf. Health Soc. Care, 13: 1-20.

Thabtah, F., P. Cowling and Y. Peng, 2004. MMAC: A new multi-class, multi-label associative classification approach. Proceedings of the 4th International Conference on Data Mining, November 1-4, 2004, Brighton, UK., pp: 217-224.

Triguero, I., S. Gonzalez, J.M. Moyano, S. Garcia and J. Alcala-Fdez *et al*., 2017. KEEL 3.0: An open source software for multi-stage analysis in data mining. Intl. J. Comput. Intell. Syst., 10: 1238-1249.

Tsoumakas, G., E. Spyromitros-Xioufis, J. Vilcek and I. Vlahavas, 2011. Mulan: A ava library for multi-label learning. J. Mach. Learn. Res., 12: 2411-2414.

Wang, D., J. Wang, F. Hu, L. Li and X. Zhang, 2018. A locally adaptive multi-label k-nearest neighbor algorithm. Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining, June 3-6, 2018, Springer, Cham, Switzerland, ISBN:978-3-319-93033-6, pp: 81-93.

Wang, J., P. Neskovic and L.N. Cooper, 2007. Improving nearest neighbor rule with a simple adaptive distance measure. Pattern Recognition Lett., 28: 207-213.

Xiang, X., C. Yu, L. Lapierre, J. Zhang and Q. Zhang, 2018. Survey on fuzzy-logic-based guidance and control of marine surface vehicles and underwater vehicles. Intl. J. Fuzzy Syst., 20: 572-586.

Xu, S., X. Yang, H. Yu, D.J. Yu and J. Yang *et al*., 2016. Multi-label learning with label-specific feature reduction. Knowl. Based Syst., 104: 52-61.

Yan, Y., G. Fung, J.G. Dy and R. Rosales, 2010. Medical coding classification by leveraging inter-code relationships. Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, July 25-28, 2010, ACM, New York, USA., ISBN:978-1-4503-0055-1, pp: 193-202.

Zhang, M.L. and L. Wu, 2015. Lift: Multi-label learning with label-specific features. IEEE. Trans. Pattern Anal. Mach. Intell., 37: 107-120.

Zhang, M.L., 2009. ML-RBF: RBF neural networks for multi-label learning. Neural Process. Lett., 29: 61-74.