Fuzzy Deep Neural Network for Classification of Overlapped Data

Rukshima Dabare[1,2(✉)], Kok Wai Wong[1(✉)], Mohd Fairuz Shiratuddin[1], and Polychronis Koutsakis[1]

[1] Discipline of Information Technology, Mathematics and Statistics, Murdoch University, Perth, WA, Australia
{rukshima.dabare,K.Wong,F.Shiratuddin, p.koutsakis}@murdoch.edu.au
[2] Faculty of Engineering Technology, The Open University of Sri Lanka, Nawala, Nugegoda, Sri Lanka

Abstract. Deep Learning is a popular and promising technique for classification problems. This paper proposes the use of fuzzy deep learning to improve the classification capability when dealing with overlapped data. Most of the research focuses on classification and uses traditional truth and false criteria. However, in reality, a data item may belong to different classes at different degrees. Therefore, the degree of belonging of each data item to a class needs to be considered for classification purposes in some cases. When a data item belongs to different classes with different degrees, then there exists an overlap between the classes. For this reason, this paper proposes a Fuzzy Deep Neural Network based on Fuzzy C-means clustering, fuzzy membership grades and Deep Neural Networks to address the over-lapping issue focused on binary classes and multi-classes. The proposed method converts the original attribute values to relevant cluster centres using the proposed Fuzzy Deep Neural Network. It then trains them with the original output class values. Thereafter, the test data is checked with the Fuzzy Deep Neural Network model for its performance. Using three popular datasets in overlapped and fuzzy data literature, the method presented in this paper outperforms the other methods compared in this study, which are Deep Neural Networks and Fuzzy classification.

Keywords: Classification · Deep learning · Fuzzy · Fuzzy C-means clustering · Overlapped data

1 Introduction

Class overlap occurs when data with similar characteristics appear in the feature space with different degree of belongings [1]. In ideal situations, the feature space should be well separated, and one should be able to identify the class it belongs to. However, this is not the case because real-world data may have overlapping regions which makes it hard for a traditional classifier to classify them into different classes [2] or distinguish between the different classes [3]. Recently, Deep Neural Network (DNN) technique has

become popular in solving classification tasks. However, the work done is minimal on handling overlapped data. Therefore, in this paper, we propose an approach to handle overlapped data, which is different from most of the related literature (see Sect. 2).

When a data item belongs to two or more clusters, it is said to be overlapped, and its degree of belonging can be identified through Fuzzy C-means clustering [4]. This is something that hard clustering techniques are not able to address [5]. Therefore, when overlapped data is clustered, it may belong to two or more clusters [6]. Each data item will have a membership to every cluster it belongs to. In our approach, we identify the features which will have an overlapping behaviour and address them separately. The features which are non-categorical will be fuzzified, and the relevant cluster centre will be identified for the training of the DNN along with the categorical features. DNN was selected as the classifier because of its proven efficiency in recent years [7, 8] due to its multiple processing layers which will allow systems to abstract more and learn a representation of data better than with shallow networks [9]. Therefore, in order to deal with the class overlapping issue, in our study, we propose a Fuzzy Deep Neural Network (FuzzyDNN) based on Fuzzy C-means clustering, fuzzy membership grades, and DNN (Deep Neural Network).

The performance of our proposed method is validated using three datasets, Heart, Wisconsin and Wavform from UCI data repository that is commonly used in overlapped data and fuzzy literature [3, 10] with the use of the accuracy metric. In order to demonstrate the proposed method, we compare our results with two other classifiers, which are a DNN and a fuzzy classifier.

This paper is organized as follows. Section 2 gives an overview of related work carried out by other researchers related to fuzzy deep learning and overlapped data classification. Section 3 presents the novel FuzzyDNN model. In Sect. 4, the experimental setup, datasets and the results of our study are presented and analyzed. The final section concludes the paper.

2   Related Work

Deep Neural Networks (DNN) have shown excellent results across different domains, including but not limited to pharmacological science [11], emotion predictions [12], speech enhancement [13] and medical [14]. The excellent results of DNN across these different domains are the reason that we have selected DNN as the classifier along with the fuzzy approach in our study. However, the DNN has problems handling the overlap uncertainties appears in some datasets [15]. On the other hand, fuzzy sets can be used to represent the uncertainties and impreciseness of data [16]. This is addressed by assigning a membership grade ranging from 0 to 1 for each object (or data item) to the different sets or clusters [16, 17]. The fuzzy concept has been widely used in domains ranging from manufacturing [17] to the medical field [18] and network security [19]. In the following section, we briefly discuss some of the related work on Fuzzy Deep Learning on overlapped data and clustering.

In recent years, few studies have been carried out by combining the fuzzy concept with deep learning. The work of [20–25] and [26] have successfully used the fuzzy concept in different formats with deep learning. Most of the work mentioned have used

fuzzy concept in order to improve the representational ability of the classifiers as well as to ensure the interpretability of the whole system. However, only the work of [15] concentrates on handling overlapped data with the use of fuzzy logic concept and DNN. In [15], the fuzzy C-means is used to handle the overlapping uncertainties by implementing fuzzy inference rules, apply them on the previously clustered data and generate corresponding membership values. Thus their method required the steps of implementing fuzzy inference rules which in some cases could be difficult and complex than our proposed method.

In order to address the overlapping issue in classification, there are various methods adopted by researchers. The work in [27] introduces a novel method based on Fuzzy Adaptive Resonance Theory (ART) which can be used for clustering. However, this work is only applicable to unsupervised learning, whereas our work concentrates on supervised learning. The work of [3, 28, 29] and [30] have concentrated more on the imbalanced data set with overlapped data, and therefore, the focus is different from our work. Therefore, in order to handle the overlapping issue in classification, in our work, the fuzzy and deep learning concepts are combined with the use of Fuzzy C-means clustering, fuzzy membership grades, and a DNN. With the adoption of deep learning, we expect the system to learn a better representation of data [9] and with the use of the fuzzy approach that it will address the overlapping problem [15].

3 Proposed FuzzyDNN Model

The proposed FuzzyDNN comprises of three main phases. The first phase is the fuzzification of each attribute by clustering and identifying the membership grade for each input attribute. The second phase is the identification of the new input attribute values by using different α-cuts and cluster centre identification. The third phase is the learning phase with the use of a DNN with the cluster centres as the input to the DNN. The description of each phase is given below, and the overall proposed method, if the attribute values were only non-categorical values, are shown in Fig. 1. All the phases presented in this proposed model are applicable only for training the model; therefore, only for the training dataset.

3.1 Phase 1:- Fuzzification of Each Attribute

In this phase, the attributes which are not categorical are given a fuzzy value. By doing so, we are trying to solve the overlapping issue of non-categorical features. Here the non-categorical features will be fuzzified in order to check their belongings to each class and thereby identifying the degree of overlapping. This process is explained under step 1 of this phase. In this paper, we refer to this as fuzzification. Each attribute of the dataset is clustered with the use of Fuzzy C-Means clustering. The reason for the selection of this clustering method is its capability of identifying the membership grade for each cluster [4]. The three steps of the first phase are given below in detail.

Step 1. Use Fuzzy C-means to cluster each attribute in the training dataset with continuous values. Here we used three clusters for all three (3) datasets. Experiments were also carried out with different numbers of clusters, but we got the best result from three (3) clusters. The attributes with categorical values were not clustered.
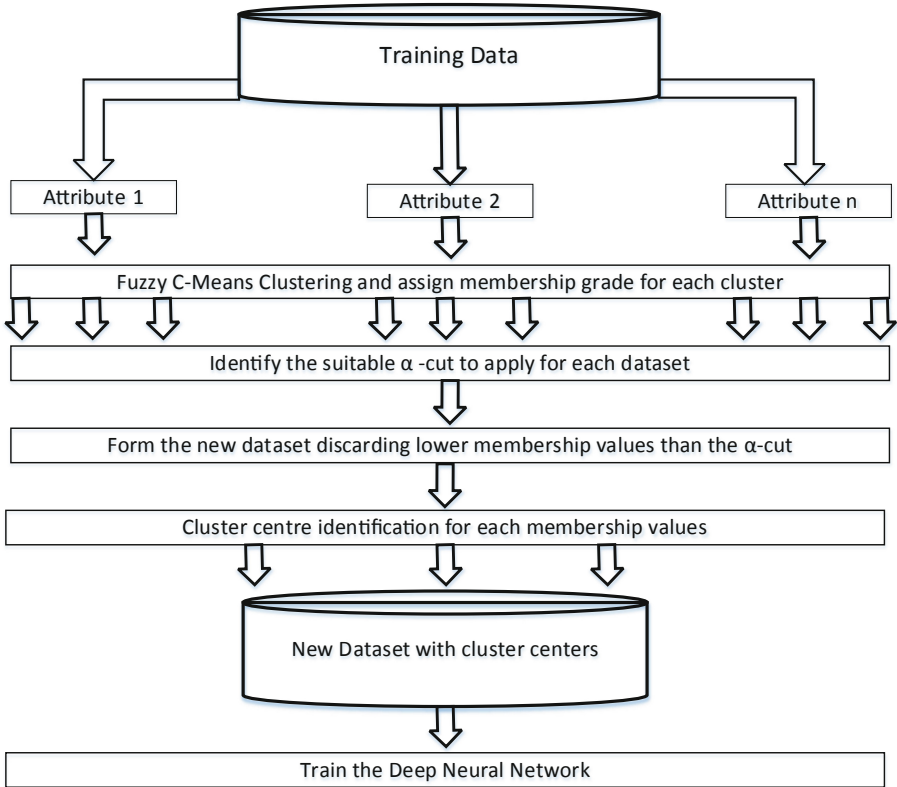
Fig. 1. The proposed FuzzyDNN method

Step 2. Identify the degree of membership of each attribute value to each cluster. By doing so, an attribute will have three (3) membership values relevant to the three (3) clusters it belongs to.

Step 3. When forming the new training dataset, the descending order of the membership grade to each attribute was taken into consideration. Here, for each attribute value considered, the membership grades with the highest values are used to form the new record of data. Then after the first set is formed the second-highest membership grades are considered to form the next set of records. Therefore, after forming the whole dataset, the number of records is increased by the number of clusters. For example, in our experiments for Heart, Wisconsin and Waveform datasets, the number of records increased by three (3) as we used three (3) clusters for each attribute. Thereafter, for all the attributes that were not fuzzified (because of the categorical nature of the attributes), then their original values were used in forming the new dataset. When forming the new input dataset, the original values of the labels will be used as they are not fuzzified.

Phase 1 is explained using the matrix format below: -

Assuming there are n input records (patterns) and k number of attributes for each record (each pattern) in the dataset. The input space and the output space can be

represented in a matrix form as X and Y, respectively.

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \ldots & \ldots & x_{1k} \\ x_{21} & x_{22} & x_{23} & \ldots & \ldots & x_{2k} \\ x_{31} & x_{32} & x_{33} & \ldots & \ldots & x_{3k} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \\ x_{n1} & x_{n2} & x_{n3} & \ldots & \ldots & x_{nk} \end{bmatrix} \tag{1}$$

and,

$$Y = y_1 \; y_2 \; y_3 \; \ldots \; y_n \tag{2}$$

Therefore,

$$y_i = x_{i1} \; x_{i2} \; x_{i3} \ldots \ldots x_{ik} \tag{3}$$

where i is a record of the dataset.

The Fuzzy C-Means clustering will consider each column at a time for clustering of matrix X, which is an attribute value of the dataset. Therefore, Fuzzy C-Means clustering will be performed on each column of the X matrix, and the membership grade relevant to each cluster along with the input attributes will form the new matrix W_i that is:

$$W_n = \begin{bmatrix} \mu_{11}(x_{i1}) & \mu_{11}(x_{i2}) & \mu_{11}(x_{i3}) & \ldots \ldots & \mu_{11}(x_{ik}) \\ \mu_{12}(x_{i1}) & \mu_{12}(x_{i2}) & \mu_{12}(x_{i3}) & \ldots \ldots & \mu_{12}(x_{ik}) \\ \mu_{13}(x_{i1}) & \mu_{13}(x_{i2}) & \mu_{13}(x_{i2}) & \ldots \ldots & \mu_{13}(x_{ik}) \\ \mu_{21}(x_{i1}) & \mu_{21}(x_{i2}) & \mu_{21}(x_{i3}) & \ldots \ldots & \mu_{21}(x_{ik}) \\ \mu_{22}(x_{i1}) & \mu_{22}(x_{i2}) & \mu_{22}(x_{i3}) & \ldots \ldots & \mu_{22}(x_{ik}) \\ \mu_{23}(x_{i1}) & \mu_{23}(x_{i1}) & \mu_{23}(x_{i1}) & \ldots \ldots & \mu_{22}(x_{ik}) \\ \ldots & \ldots & \ldots & \ldots \ldots & \ldots \\ \mu_{n3}(x_{i1}) & \mu_{n3}(x_{i2}) & \mu_{n3}(x_{i4}) & \ldots \ldots & \mu_{n3}(x_{ik}) \end{bmatrix} \tag{4}$$

where μ_{ik} is the fuzzy membership grade. Therefore, for example, μ_{12} is the fuzzy membership grade for the 1st attribute for the 2nd cluster which belongs to the 1st record of the original dataset. Now the X matrix is converted into W which also can be represented as

$$W_n = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \ldots \\ w_n \end{bmatrix} \tag{5}$$

Where W_i is the memberships grade details of the record i.

For example, the W_1 will represent the membership grades of the 1st record when three (3) clusters are used.

$$W_1 = \begin{bmatrix} \mu_{11}(x_{i1}) & \mu_{11}(x_{i2}) & \mu_{11}(x_{i3}) & \ldots \ldots \mu_{11}(x_{ik}) \\ \mu_{12}(x_{i1}) & \mu_{12}(x_{i2}) & \mu_{12}(x_{i3}) & \ldots \ldots \mu_{12}(x_{ik}) \\ \mu_{13}(x_{i1}) & \mu_{13}(x_{i2}) & \mu_{13}(x_{i3}) & \ldots \ldots \mu_{13}(x_{ik}) \end{bmatrix} \tag{6}$$

The final step of this phase is to combine all the W_i to form the inputs of the dataset, which is explained under **Step 3** of **phase 1.**

3.2 Phase 2:- Identification of the New Attribute Values

This phase is only applicable to the input attributes that were given a fuzzy value. Identification of the new input attribute values is achieved by identifying the cluster centre and α-cut assignment to the fuzzified input attributes. Below are the three steps required for phase 2.

Step 1. Identify the best α-cut for each dataset and apply it in order to address the overlapping.
Step 2. If the membership degree is less than the identified α-cut value, then that fuzzified attribute is discarded from the newly formed training dataset.
Step 3. Identify the associated cluster centre for each attribute.

Phase 2 is explained in detail below: -
A threshold for W_n is then identified considering the dataset. Here the threshold value is fuzzy α-cut performed on W_n:

$$W_\alpha = \{u \in U/\mu_{ik} \geq \alpha\} \tag{7}$$

where U is the universe of discourse, μ_{ik} is the membership grade, which takes the values between in interval [0, 1] and α takes values in the interval [0, 1] [31]. After applying the fuzzy α-cut for W_n, the number of rows of W_n will be reduced according to the fuzzy α-cut value. The reason is that if the membership value of any item is less than the α-cut value, then the relevant record is discarded from the dataset. By doing so, we expect our proposed method will respond differently to different fuzziness that is available on the dataset.

Finally, for each element of W_n, the cluster centre for each cluster is assigned instead of the membership grade values. So, if n is the number of records the cluster centre matrix would be,

$$C_n = \begin{bmatrix} C_1 \ C_2 \ C_3 \ \dots \ C_n \end{bmatrix}^T \tag{8}$$

Therefore, for an example in the case where no records were discarded from W_n, meaning $\alpha = 0$, then the cluster centre matrix C_i would be:

$$C_i = \begin{bmatrix} C_{11} & C_{12} & C_{13} & \dots & C_{1k} \\ C_{21} & C_{22} & C_{23} & \dots & C_{2k} \\ C_{31} & C_{32} & C_{33} & \dots & C_{3k} \\ \dots & \dots & \dots & \dots & \dots \\ C_{j1} & C_{j2} & C_{j3} & \dots & C_{jk} \end{bmatrix} \tag{9}$$

Where k is the number of attributes and j is the number of clusters.

Phase 3:- Training of the DNN Using the New Input Attributes
The input for the DNN will be the output of the 2nd Phase. The fuzzified input dataset will then be fed to the DNN, and the DNN will learn according to the new cluster centre values (assigned to the continuous attributes) and the original data for the non-continuous attributes which are also known as categorical attributes.

4 Experimental Procedure and Evaluation of Results

4.1 Experimental Procedure

In our study, we used the Heart, Wisconsin and Waveform datasets from the UCI repository [32]. All three datasets are popular in overlapped and fuzzy data literature and are used for performance evaluation in recent studies such as in [3, 10, 33, 34]. In our study, the dataset was divided into 5-folds for each experiment, and the average of the 5 experiments is used for evaluation. For both datasets, 80% was used for training, and the rest was used for testing. The Heart dataset consists of 13 attributes, and out of those 13 attributes, only 5 attributes were fuzzified by the FuzzyDNN as the others were categorical values. The Wisconsin dataset contains 9 attributes, and for the FuzzyDNN all 9 attributes were fuzzified. Waveform dataset contained 21 attributes as all of them are non-categorical, and we fuzzified all attributes according to the proposed method.

The DNN architecture was decided by trial and error, considering the number of input attributes to the model. The DNN architecture and the relevant details are given in Table 1.

Table 1. DNN architecture and relevant details

Dataset	Number of hidden layers	Type	Number of hidden nodes	Number of epoch	Batch size
Heart	4	Binary	20, 20, 20, 20	500	100
Wisconsin	4	Binary	50, 100, 100, 40	1500	100
Waveform	3	Multi	20, 10, 50	500	1

The activation function used in the nodes of the hidden layers are Rectified Linear Unit (relu) as recommended by many studies [35], and the nodes of the last layer use the sigmoid activation function for the Heart and Wisconsin datasets as they are binary classification problems [36]. The activation function used for the Waveform dataset is relu for the hidden layers and softmax for output layer as the classification task is a multi-class classification problem [36]. The DNN model used the loss function binary cross entropy and used a batch size of 100 for the binary classes and for the Waveform class used a sparse_categorical_crossentropy function while the batch size was 1 for each training phase. Dropout technique [37], which is a popular regularization technique, is used to avoid overfitting when training our model. Table 2 presents the classification performance in terms of accuracy of the trained FuzzyDNN model, DNN and fuzzy classification for the Heart, Wisconsin and Waveform datasets.

4.2   Evaluation of Result

As shown in Table 2, the FuzzyDNN model outperforms the other approaches in terms of accuracy for the Heart, Wisconsin and Waveform datasets. FuzzyDNN improves the classification accuracy by 8.89%, 0.88% and 1.24% for the Heart, Wisconsin and Waveform datasets, respectively, when compared with DNN and by 9.63%, 2.48%, 3.57% when compared with Fuzzy Classification. The FuzzyDNN achieved higher accuracy in all cases as it was able to address the overlapped data issue separately for each attribute and use that knowledge to train the data.

Table 2. The best accuracies of the Heart, Wisconsin, and Waveform datasets

Dataset	Best α-cut	FuzzyDNN	DNN	Fuzzy classification
Heart	0.05	0.8407	0.7518	0.7444
Wisconsin	0.01	0.9739	0.9651	0.9491
Waveform	0.05	0.8332	0.8208	0.7975

It can be observed from the results provided in Table 2 that the use of FuzzyDNN model provides better results for the three datasets considered. Here as mentioned earlier, the threshold value is the α-cut value applied to the fuzzy set identified by the clustering. According to the fuzzy theory, low α-cut values means lesser overlapping between classes. Therefore, the threshold value selected for the α-cut is important for accurate classification.

As we wanted to investigate the most suitable threshold value for each dataset, we carried out experiments for different threshold values, which in our model are the α-cut values. Tables 3, 4 and 5 show the accuracy of the classification task for the Heart, Wisconsin and Waveform datasets, respectively, with the application of different α-cut values. As shown in Tables 3, 4 and 5 by choosing the appropriate α-cut value according to the dataset, we can adjust how much overlap is handled by the proposed model.

Table 3. The accuracy of the Heart dataset for different threshold values

Threshold value	Number of training data	Accuracy
0.05	**229**	**0.8407**
0.1	222	0.8111
0.2	218	0.8148
0.3	216	0.7888

Table 4. The accuracy of the Wisconsin dataset for different threshold values

Threshold value	Number of training data	Fuzzy DNN
0.01	**554**	**0.9739**
0.05	545	0.9710
0.1	545	0.9681
0.2	545	0.9710

Table 5. The accuracy of the Waveform dataset for different threshold values

Threshold value	Number of training data	Fuzzy DNN
0.05	**4001**	**0.8332**
0.1	4000	0.8204
0.2	4000	0.8208
0.3	4000	0.8202

5   Conclusion

In this work, we proposed a novel method to handle overlap between classes using fuzzy concept with the use of a DNN, Fuzzy C-means clustering and fuzzy membership grades. The features or the attributes that had continuous values were fuzzified by first clustering them using Fuzzy C-means clustering and identifying a membership grade for each attribute to the cluster it belongs to. After the identification of the membership values, according to the dataset, fuzzy α-cut was used to select the data which has more belongings or higher membership. Thereafter, this information was used to assign new values to the attribute values, which are the cluster centre of each cluster in order to handle the overlap. This approach was shown to improve deep learning performance in terms of the performance metric, accuracy and thereby improve the accuracy of the classifier. After carrying out experiments on three popular datasets in overlapped and fuzzy literature, our proposed FuzzyDNN approach was shown to perform better than the DNN classifier and a traditional fuzzy classifier.

References

1. Xiong, H., et al.: Classification algorithm based on NB for class overlapping problem. Appl. Math. Inf. Sci. **7**(2L), 409–415 (2013)
2. Tang, W., et al.: Classification for overlapping classes using optimized overlapping region detection and soft decision. In: 13th Conference on Information Fusion (FUSION). IEEE (2010)
3. Lee, H.K., Kim, S.B.: An overlap-sensitive margin classifier for imbalanced and overlapping data. Expert Syst. Appl. **98**, 72–83 (2018)

4. Bezdek, J.C., Ehrlich, R., Full, W.: FCM: the fuzzy c-means clustering algorithm. Comput. Geosci. **10**(2–3), 191–203 (1984)
5. Setyohadi, D.B., Bakar, A.A., Othman, Z.A.: Optimization overlap clustering based on the hybrid rough discernibility concept and rough K-Means. Intell. Data Anal. **19**(4), 795–823 (2015)
6. Banerjee, A., et al.: Model-based overlapping clustering. In: Proceedings of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining. ACM (2005)
7. Faust, O., et al.: Deep learning for healthcare applications based on physiological signals: a review. Comput. Methods Programs Biomed. **161**, 1–13 (2018)
8. Young, T., et al.: Recent trends in deep learning based natural language processing. IEEE Comput. Intell. Mag. **13**(3), 55–75 (2018)
9. LeCun, Y., Bengio, Y., Hinton, G.: Deep. Learn. Nat. **521**(7553), 436 (2015)
10. Patwary, M.J., Wang, X.-Z.: Sensitivity analysis on initial classifier accuracy in fuzziness based semi-supervised learning. Inf. Sci. **490**, 93–112 (2019)
11. Wang, C.-S., et al.: Detecting potential adverse drug reactions using a deep neural network model. J. Med. Internet Res. **21**(2), e11016 (2019)
12. Kim, H.-C., Bandettini, P.A., Lee, J.-H.: Deep neural network predicts emotional responses of the human brain from functional magnetic resonance imaging. NeuroImage **186**, 607–627 (2019)
13. Saleem, N., et al.: deep neural network for supervised single-channel speech enhancement. Arch. Acoust. **44**(1), 3–12 (2019)
14. Katzman, J.L., et al.: DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. BMC Med. Res. Methodol. **18**(1), 24 (2018)
15. Sumit, S.H., Akhter, S.: C-means clustering and deep-neuro-fuzzy classification for road weight measurement in traffic management system. Soft Comput. **23**, 4329–4340 (2019). https://doi.org/10.1007/s00500-018-3086-0
16. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**(3), 338–353 (1965)
17. De Silva, C.W.: Intelligent Control: Fuzzy Logic Applications. CRC Press, Boca Raton (2018)
18. Korenevskiy, N.: Application of fuzzy logic for decision-making in medical expert systems. Biomed. Eng. **49**(1), 46–49 (2015)
19. Dotcenko, S., Vladyko, A., Letenko, I.: A fuzzy logic-based information security management for software-defined networks. In: 16th International Conference on Advanced Communication Technology (ICACT), 2014. IEEE (2014)
20. Park, S., et al.: Intra-and inter-fractional variation prediction of lung tumors using fuzzy deep learning. IEEE J. Transl. Eng. Health Med. **4**, 1–12 (2016)
21. El Hatri, C., Boumhidi, J.: Fuzzy deep learning based urban traffic incident detection. Cogn. Syst. Res. **50**, 206–213 (2018)
22. Davoodi, R., Moradi, M.H.: Mortality prediction in intensive care units (ICUs) using a deep rule-based fuzzy classifier. J. Biomed. Inform. **79**, 48–59 (2018)
23. Zheng, Y.-J., et al.: Airline passenger profiling based on fuzzy deep machine learning. IEEE Trans. Neural Netw. Learn. Syst. **28**(12), 2911–2923 (2017)
24. Deng, Y., et al.: A hierarchical fused fuzzy deep neural network for data classification. IEEE Trans. Fuzzy Syst. **25**(4), 1006–1012 (2017)
25. Chen, C.P., et al.: Fuzzy restricted Boltzmann machine for the enhancement of deep learning. IEEE Trans. Fuzzy Syst. **23**(6), 2163–2173 (2015)
26. Nugaliyadde, A., Pruengkarn, R., Wong, K.W.: The fuzzy misclassification analysis with deep neural network for handling class noise problem. In: Cheng, L., Leung, A.C.S., Ozawa, S. (eds.) ICONIP 2018. LNCS, vol. 11304, pp. 326–335. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-04212-7_28

27. Mak, L.O., et al.: A merging Fuzzy ART clustering algorithm for overlapping data. In: 2011 IEEE Symposium on Foundations of Computational Intelligence (FOCI). IEEE (2011)
28. Xiong, H., Wu, J., Liu, L.: Classification with class overlapping: a systematic study. In: The 2010 International Conference on E-Business Intelligence (2010)
29. Vorraboot, P., et al.: Improving classification rate constrained to imbalanced data between overlapped and non-overlapped regions by hybrid algorithms. Neurocomputing **152**, 429–443 (2015)
30. Das, B., Krishnan, N.C., Cook, D.J.: Handling class overlap and imbalance to detect prompt situations in smart homes. In: 2013 IEEE 13th International Conference on Data Mining Workshops. IEEE (2013)
31. Harris, C.J., Moore, C.G., Brown, M.: Intelligent Control: Aspects of Fuzzy Logic and Neural Nets, vol. 6. World Scientific (1993)
32. Dua, D., Taniskidou, E.K.: UCI Machine Learning Repository (2017). http://archive.ics.uci.edu/ml
33. Singh, H.R., Biswas, S.K., Purkayastha, B.: A neuro-fuzzy classification technique using dynamic clustering and GSS rule generation. J. Comput. Appl. Math. **309**, 683–694 (2017)
34. Liu, X., et al.: A hybrid classification system for heart disease diagnosis based on the RFRS method. Comput. Math. Methods Med. **2017** (2017). 11 pages, Article ID 8272091. https://doi.org/10.1155/2017/8272091
35. Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT Press, Cambridge (2016)
36. Géron, A.: Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media Inc, Sebastopol (2017)
37. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from ovrfitting. J. Mach. Learn. Res. **15**, 1929–1958 (2014)