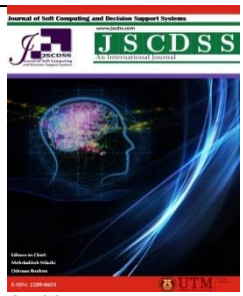
An International Journal
JSCDSS
http://www.jscdss.com
Vol.7 No.1 February 2020: 19-30
Article history:
Accepted 20 February 2020
Published online 12 March 2020

Journal of Soft Computing and Decision Support Systems



E-ISSN: 2289-8603

Disease Diagnosis Using Machine Learning Techniques: A Review and Classification

Mehrbakhsh Nilashi [a,*], Neda Ahmadi [b], Sarminah Samad [c], Leila Shahmoradi [d,e], Hossein Ahmadi [f], Othman Ibrahim [g], Shahla Asadi [h,*], Rusli Abdullah [h], Rabab Ali Abumalloh [i], Elaheh Yadegari [j]

[a] School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
[b] Shahid Chamran Universty of Ahvaz, Faculty of Engineering, Department of Computer Engineering, Ahvaz, Iran
[c] Department of Business Administration, Collage of Business and Administration, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia
[d] Halal Research Center of IRI, FDA, Tehran, Iran
[e] Health Information Management Department, School of Allied Medical Sciences, Tehran University of Medical Sciences, Tehran, Iran
[f] Department of Information Technology, University of Human Development, Sulaymaniyah, Iraq
[g] Department of Software Engineering & Information System, Faculty of Computer Science & Information Technology, University Putra Malaysia, Selangor 43400, Malaysia
[h] Azman Hashim International Business School, Universiti Teknologi Malaysia (UTM), Skudai, Johor 81310, Malaysia
[i] Computer Department, Community College, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia
[j] Center for Software Technology and Management (SOFTAM), Faculty of Information Science and Technology (FTSM), Universiti Kebangsaan Malaysia (UKM), 43600 Bangi, Selangor, Malaysia

*** Corresponding author email address**: nilashidotnet@hotmail.com

Abstract

In this research, we reviewed and classified academic conference and journal papers; which used data mining techniques in disease classification and diagnosis based on public medical datasets published between 2007 and 2019. The results of this review demonstrated that the application of data mining techniques in disease classification has experienced a dramatic rise in recent years. The finding of this paper also revealed that there was minimal focus on developing methods using incremental version of data mining techniques. We hope that this research will provide useful information about various data mining techniques, their application in disease diagnosis, and help researchers in developing medical decision support systems with insights into the state-of-the-art of development methods.

Keywords: Data Mining, Public Medical Datasets, Diseases Diagnosis, Literature Survey, UCI, PRISMA

1. Introduction

Data mining has been applied effectively in developing decision support systems to discover patterns in big datasets (Han et al., 2011). The potential of the data mining as an efficient disease classification tool was identified by World Health Organization (WHO) (Nilashi et al., 2016b). The use of data mining techniques in knowledge discovery for disease classification has been recognized as one of the interesting and important topics addressed by researchers (Ahmadi et al., 2018; Cauchi et al., 2015; Gopalakrishnan et al., 2010; Nilashi et al., 2016a).

There are many of supervised machine learning techniques for prediction and classification tasks (Kotsiantis et al., 2007; Nilashi et al., 2014a; Nilashi et al., 2014b; Nilashi et al., 2015a; Nilashi et al., 2015b). Using these techniques, several studies have developed methods for disease diagnosis and have been evaluated using real-world datasets (Çalişir and Doğantekin, 2011; Nilashi et al., 2019a; Nilashi et al., 2017a; Nilashi et al., 2017c; Polat, 2012). Hence, the classification of these studies based on applied machine learning techniques for investigating different types of disease diagnosis and reporting their outputs are important tasks; which this research aims to perform. Additionally, further researches should be

conducted to demonstrate the usefulness of machine learning techniques for disease classification, compared to other domains and based on their usage, the application of these techniques is less matured and has been rarely investigated. Therefore, the existing studies on disease classification which applied machine learning techniques need to be studied to show their contributions in disease diagnosis research. Accordingly, in this research we reviewed and classified the academic conference and journal papers which used data mining techniques in disease classification and diagnosis based on the public medical datasets of Repository of the University of California, Irvine (UCI) (Asuncion and Newman, 2007), which have been published between 2007 and 2019. We hope that this research will provide useful information about the adopted data mining techniques, their application in disease diagnosis and help researchers in developing medical decision support systems with insight into the state-of-the-art of development methods.

The rest of the paper is organised as follows. In Section 2, we present the research methodology that has been used in this study. In Section 3, we provide the results of the review of the selected research papers of disease diagnosis. Finally, the conclusion and future work is presented in Section 4.

2. Research Methodology

In this work, we have reviewed the selected research papers based on Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Budrionis and Bellika, 2016; Schraagen and Verhoeven, 2013; Zare et al., 2016). The aim of PRISMA is to help researchers in reporting meta-analyses results and conducting systematic reviews. Although PRISMA can't provide high quality assessment of a systematic review studies, this technique has proven to be a useful tool for critical assessment of published systematic reviews. On the other hand, PRISMA has well-defined steps which can help researchers in identifying, interpreting and evaluating the research papers' data effectively. In this review, we followed three main steps of PRISMA (searching the literature, selecting the eligible papers, and extracting and summarizing the papers' data) and accordingly classified the reviewed papers. In our review, in order to conduct PRISMA methodology, three main steps of PRISMA need to be accomplished which are literature search, study selection and data extraction and summarization.

2.1 Literature search

In the first step of PRISMA, we selected the major databases to conduct our review and provide a comprehensive review of the application of data mining techniques in disease classification. We selected the following databases: Scopus and Google Scholar. These databases are selected according to their popularity in health and computer science researches. The papers that have been published in conferences and academic journals on these databases are considered in this research and the literature search is performed based on the following terms: "Diseases Classification and Prediction" combined with each keyword of "Machine Learning", "Data Mining" and "Repository of the University of California, Irvine (UCI)". We considered their combinations for "Diabetes", "Heart", "Parkinson" and "Breast Cancer". As we found from the literature that the application of data mining techniques in disease diagnosis is relatively new, accordingly papers were collected from 2007 to 2019.

2.2. Study selection and eligible papers

In the second step of PRSIMA, full texts of the selected research papers, which have been identified and included from the first stage of PRISMA, have been reviewed carefully by the authors. This step helped us to select the research papers which were suitable and eligible for the inclusion in our review. To do so, papers which had applied data mining technique in disease diagnosis and performed the experiments based on the UCI datasets for the four diseases were chosen. In addition, unpublished papers, papers that provided the abstracts only, non-English papers, and book chapters were excluded. Furthermore, we also excluded the papers that failed to explicitly provide the used data mining technique. Note that the majority of papers which have been excluded in this phase were those papers that did not use public medical datasets in UCI in their experiments.

2.3. Data extraction and summarizing

In the final step, based on PRISMA guideline, the extracted research papers in the previous phase were summarized and the results were presented. In this step of PRISMA, summarizing and categorizing papers enabled the authors to obtain several critical and impressive hints about the use of data mining techniques in disease classification. Based on the results of this stage, we could accordingly find the limitations of previous studies and provide some directions and guidelines for future studies.

3. Results

We first reviewed the research papers which used data mining techniques for disease classification. The reviewed papers are listed in Table 1 (see Appendix A). We focused on four types of diseases which are Heart, Diabetes, Breast Cancer and Parkinson. In addition, all papers which have been listed in this section have used UCI public datasets in their experiments. Next, we classified the papers based on different perspectives. In this review, the academic papers were classified based on different main categories which utilized data mining techniques, year of publication in which they have been published, and types of diseases.

An International Journal
JSCDSS
E-ISSN: 2289-8603

3.1 Classification of Papers

The distribution of data mining techniques that were applied in disease classification is presented by publication year between 2007 and 2019 (see Fig. 1 in Appendix B). In terms of Parkinson, Breast Cancer, Heart and Diabetes diseases, the following methods were used from 2007 to 2019: SVM (36.2%), KNN (20.3%), NN (33.33%), ANFIS (4.3%), Fuzzy Logic (21.7%), K-Means (2.9%), GP (7.2%), EM (10.1%), PCA (10.1%), Random Forest (2.9%), LDA (7.2%), DT (7.2%), Association Rules (7.2%), PSO (5.8%), NB (11.6%), Regression (5.8%) and Deep Learning (10.1%). Specifically, the total number of methods that were utilized for Parkinson, Breast Cancer, Heart and Diabetes diseases are 54 in the time period of 2010 to 2018, 24 from 2007 to 2017, 44 from 2008 to 2018, and 21 from 2008 to 2019, respectively.

Among all of these techniques, the most widely used approaches were SVM, NN and Fuzzy Logic. For example, let pick SVM has the highest percentage as it was applied for Parkinson disease with 11.6%, Breast Cancer disease with 6%, Heart disease with 11.6% and Diabetes disease with 7.2%.

The frequency of data mining applications in disease classification studies has been relatively low between 2007-2011 and relatively high between 2011-2019. In 2018, 14 research papers in different areas of disease classification were published. The application of data mining in disease classification have experienced a dramatic rise, reaching to the peak of 14 papers in 2018 and 13 papers in 2016. Parkinson, Breast Cancer, Diabetes and Heart disease studies have adopted data mining methods with 31, 18, 10 and 19 research papers from 2007 to 2019, respectively. The distribution of academic papers of Diabetes disease was quite low from 2007 to 2019; however, Diabetes disease was widely investigated between 2009 to 2019 and reached the peak (up to 7 research papers) in 2016, after that, Diabetes studies decreased gradually to 5 research papers in 2019. Additionally, both Heart and Breast Cancer diseases had similar trends in 2007, 2014, 2015, 2017, 2018 and 2019 (see Fig. 2 in appendix B). In average, from this chart we can see that, in recent years researchers were more motivated to employ data mining techniques for disease classification. This may be referred to the emerging new data mining techniques and their algorithmic improvements over the last years as they have been adopted for classification tasks in different domains.

Distribution of papers based on the application of data mining techniques is presented in Table 1 in Appendix A. From Table 1, it is clear that SVM, NN and fuzzy logic methods are the most popular data mining techniques for disease classification comparing to other techniques (Adeli and Neshat, 2010; Agrawal and Jayaswal, 2020; Ahmadi and Akbarizadeh, 2018; Ahmadi et al., 2019).

Distribution of research papers conducted for investigating different diseases is shown Table 1 in Appendix A. In general, heart and Parkinson diseases gained researchers' attention as more researchers have focused on these diseases in their studies. These diseases were followed by breast cancer and diabetes diseases. It should be noted that based on the chart, the highest rank of published papers has been in 2016 for Parkinson disease. In Table 1 (see Appendix A), we summarize the employed data mining techniques for exploring different diseases along with the publication year of papers.

4. Discussion

The results of reviewing previous studies presented in this work have several important implications:

- In this review, we found that data mining techniques have contributed significantly in developing methods for disease classification. These methods have been evaluated to classify diseases based on a number of real-world datasets, however, most of the methods developed by supervised methods do not support the incremental learning and only about 1% of methods can support incremental version of data mining techniques.
- Developing efficient frameworks for big data analytic in the healthcare is one of the most important tasks; which should be taken into consideration. Developing a well-designed method for disease diagnosis from big data in healthcare is important. Disease diagnosis system developers need to consider the data sharing mechanism, privacy of data, the security of data, and the growth of data size. Moreover, in data mining it is important to develop methods for disease diagnosis to be modified for big data environment easily.

5. Conclusion and Future Work

This review investigated the applications of data mining techniques for disease classification. We have reviewed previous research papers which used data mining techniques in disease classification which have been published between 2007 and 2019 in the Scopus and Google Scholar databases. These databases were selected according to their popularity in health and computer science research. We focused on the studies that applied data mining techniques to investigate four types of diseases which are Heart, Parkinson, Breast Cancer and Diabetes. All studies reviewed in this work used UCI public datasets for their experiments. We observed that the usage of data mining techniques in disease classification has experienced a substantial growth over the last years. The results from this review showed that a majority of the selected papers has combined data mining techniques in developing methods for investigating a specific disease classification. SVM, NN and Fuzzy Logic techniques were found as the most frequently applied data mining techniques in classification methods (Ahmadi, 2019; Ahmadi and Akbarizadeh, 2017; Wang and Tian, 2019). In addition, most of the studies have investigated Parkinson disease prediction (Parnetti et al., 2019).

The results also demonstrated that a significant number of studies have focused on accuracy improvements of

An International Journal
JSCDSS
E-ISSN: 2289-8603

disease classification, while other issues such as computation time (time complexity or scalability) of methods have been rarely investigated. More specifically, most papers developed data mining techniques were not suitable for incremental learning. In this review, we only surveyed research papers published focusing on four types of diseases. This research can be extended to consider other diseases, which might provide different results. In addition, for this review we only considered the application of data mining techniques in the studies which have used UCI public datasets in their experiments for Parkinson, Heart, Diabetes and Breast Cancer. Therefore, additional databases also need be considered to comprehensively cover the usage of data mining techniques for other types of diseases. Accordingly, we will continue our review in the future to classify research papers on an ongoing basis and based on the limitations of this review. The review in this study provides recommendations for researchers in their future studies. Based on the result of our review, we recommend the researchers to use not only UCI datasets, but also use large datasets in their experiments from other machine learning repository to ensure the effectiveness of the developed method. Focusing on large datasets not only allow them to test the accuracy of the adopted classification, but also can motivate them to develop methods for incremental learning for time complexity issue. We hope that this research will provide useful information about the data mining techniques, their application in diseases diagnosis and help researchers in developing medical decision support systems with insight into the state-of-the-art of development methods.

Appendix A

Table 1
Distribution of academic papers by publication year, disease and data mining techniques

Disease	Author	Methods																
		SVM	KNN	NN	ANFIS	Fuzzy Logic	K-Means	GP	EM	PCA	Random Forest	LDA	DT	Association Rule	PSO	NB	Regression	Deep Learning
Diabetes	(Polat et al., 2008)	*																
	(Aslam et al., 2013)	*	*					*										
	Kahramanli and Allahverdi (2008)			*		*												
	(Erkaymaz and Ozer, 2016)			*														
	(Ganji and Abadeh, 2011)					*												
	(Dogantekin et al., 2010)				*							*						
	(Temurtas et al., 2009)			*														
	(Çalişir and Doğantekin, 2011)	*										*						
	(Nilashi et al., 2018a)	*							*	*								
	(Hayashi and Yukita, 2016)												*					
	(Nilashi et al., 2017c)			*					*									
	(Maniruzzaman et al., 2017)															*		
	(Nilashi et al., 2018a)	*								*								
	(Kannadasan et al., 2019)																	*

An International Journal
JSCDSS
E-ISSN: 2289-8603

Table 1
Distribution of academic papers by publication year, disease and data mining techniques (Cont.)

Disease	Author	Methods																
		SVM	KNN	NN	ANFIS	Fuzzy Logic	K-Means	GP	EM	PCA	Random Forest	LDA	DT	Association Rule	PSO	NB	Regression	Deep Learning
Breast Cancer	(Şahan et al., 2007)					*												
	(Polat and Güneş, 2007)	*																
	(Übeyli, 2007)	*																
	(Marcano-Cedeño et al., 2011)			*														
	(Zheng et al., 2014)	*																
	(Chen, 2014)													*				
	(Bhardwaj and Tiwari, 2015)			*														
	(Onan, 2015)		*			*												
	(Karabatak, 2015)															*		
	(Abdel-Zaher and Eldeib, 2016)			*											*			
	(Sheikhpour et al., 2016)														*			
	(Chaurasia et al., 2018)			*												*		
	(Khuriwal and Mishra, 2018)		*															
	(Sakri et al., 2018)														*			
	(Kadam et al., 2019)																	*
	(Sri et al., 2019)	*	*															
	(Islam et al., 2017)		*															
	(Nilashi et al., 2017b)					*			*			*						

An International Journal
JSCDSS
E-ISSN: 2289-8603

Table 1
Distribution of academic papers by publication year, disease and data mining techniques (Cont.)

Disease	Author	Methods																
		SVM	KNN	NN	ANFIS	Fuzzy Logic	K-Means	GP	EM	PCA	Random Forest	LDA	DT	Association Rule	PSO	NB	Regression	Deep Learning
Parkinson	(Guo et al., 2010)							*	*									
	(Das, 2010)			*														
	(Nilashi et al., 2017a)								*	*			*					
	(Bhattacharya and Bhatia, 2010)	*																
	(Nilashi et al., 2019b)				*				*									
	(Åström and Koker, 2011)					*												
	(Li et al., 2011)					*												
	(Ozcift, 2012)	*																
	(Polat, 2012)		*			*	*											
	(Eskidere et al., 2012)	*		*														
	(Chen et al., 2013)	*	*			*				*								
	(Babu and Suresh, 2013)			*														
	(Peterek et al., 2013)										*							
	(Hariharan et al., 2014)			*					*	*		*						
	(Froelich et al., 2015)												*					
	(Khan et al., 2016)			*				*										
	(Buza and Varga, 2016)			*														
	(Naranjo et al., 2016)															*		
	(Al-Fatlawi et al., 2016)			*														
	(Jain and Shetty, 2016)		*	*										*				
	(Behroozi and Sami, 2016)	*	*													*		
	(Grover et al., 2018)	*																*
	(Gunduz, 2019)																	*
	(Wan et al., 2018)		*									*					*	*
	(Nilashi et al., 2018b)	*																
	(Johri and Tripathi, 2019)																	*
	(Abdar and Zomorodi-Moghadam, 2018)															*		
	(Avci and Dogantekin, 2016)							*										
	(Nilashi et al., 2018a)	*			*				*	*								

An International Journal
JSCDSS
E-ISSN: 2289-8603

Table 1
Distribution of academic papers by publication year, disease and data mining techniques (Cont.)

Disease	Author	Methods																
		SVM	KNN	NN	ANFIS	Fuzzy Logic	K-Means	GP	EM	PCA	Random Forest	LDA	DT	Association Rule	PSO	NB	Regression	Deep Learning
Heart	(Bhatia et al., 2008)	*																
	(Kahramanli and Allahverdi, 2008)			*		*		*										
	(Das et al., 2009)			*														
	(Adeli and Neshat, 2010)					*												
	(Soni et al., 2011)													*				
	(Gudadhe et al., 2010)	*		*														
	(Ghumbre et al., 2011)	*		*														
	(Anooj, 2012)					*												
	(Rout, 2012)					*												
	(Nahar et al., 2013)													*				
	(Shilaskar and Ghatol, 2013)	*																
	(Shao et al., 2014)			*														
	(Long et al., 2015)					*												
	(Nguyen et al., 2015a)					*		*										
	(Nguyen et al., 2015b)					*												
	(Kausar et al., 2016)	*					*			*								
	(Amin et al., 2019)	*	*	*									*			*	*	
	(Bashir et al., 2019)	*									*		*			*	*	
	(Pawlovsky, 2018)		*															
	(Darmawahyuni et al., 2019)																	*
	(Paul et al., 2018)					*												
	(Takci, 2018)	*	*	*										*			*	
	(Khourdifi and Bahaj, 2018)		*												*			

An International Journal
JSCDSS
E-ISSN: 2289-8603

Appendix B

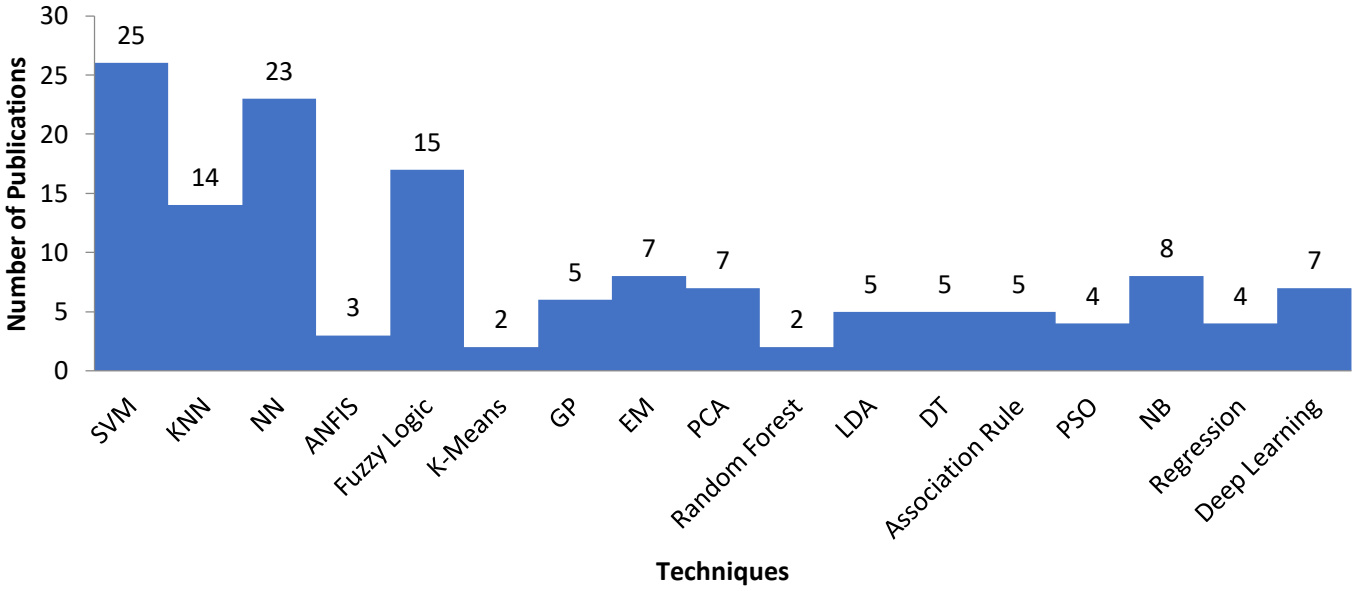


Fig. 1. Techniques and number of publications

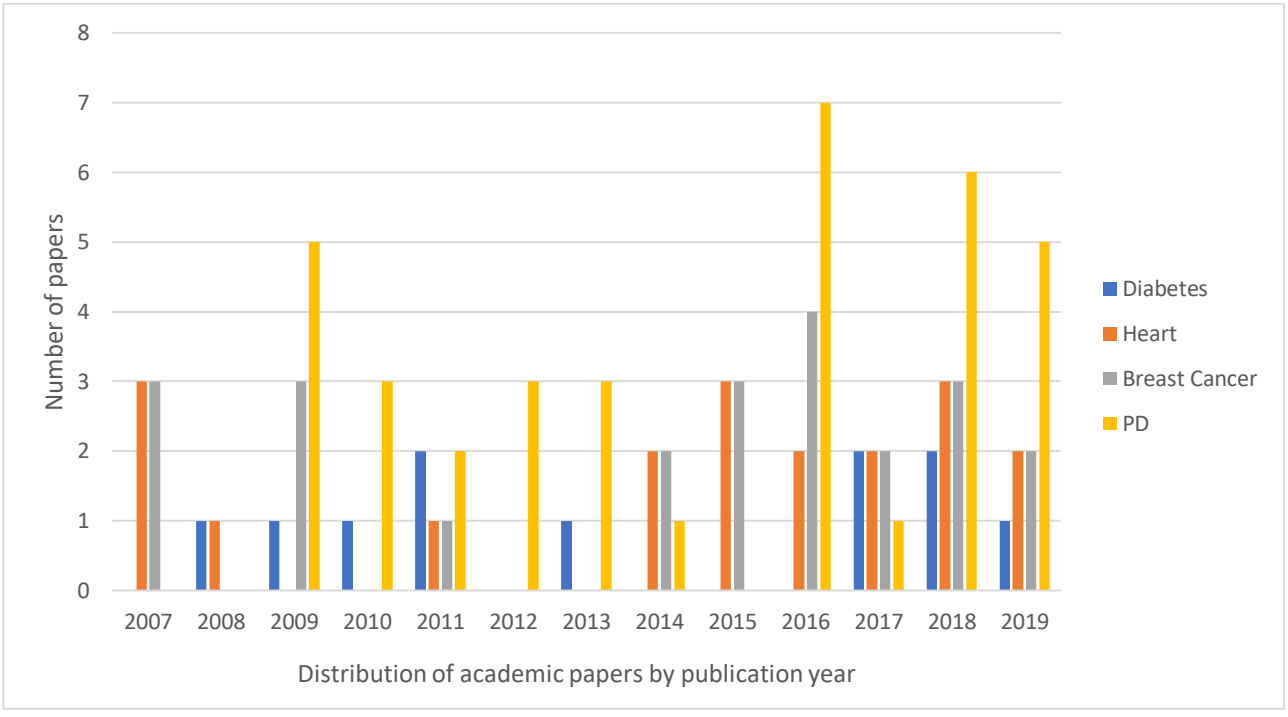


Fig. 2. Number of papers and distribution of academic papers by publication year

An International Journal
JSCDSS
E-ISSN: 2289-8603

References

Abdar, M., Zomorodi-Moghadam, M., 2018. Impact of patients' gender on parkinson's disease using classification algorithms. Journal of AI and Data Mining 6(2), 277-285.

Abdel-Zaher, A.M., Eldeib, A.M., 2016. Breast cancer classification using deep belief networks. Expert Systems with Applications 46, 139-144.

Adeli, A., Neshat, M., 2010. A fuzzy expert system for heart disease diagnosis, Proceedings of International Multi Conference of Engineers and Computer Scientists, Hong Kong. pp. 28-30.

Agrawal, P., Jayaswal, P., 2020. Diagnosis and classifications of bearing faults using artificial neural network and support vector machine. Journal of The Institution of Engineers (India): Series C 101(1), 61-72.

Ahmadi, H., Gholamzadeh, M., Shahmoradi, L., Nilashi, M., Rashvand, P., 2018. Diseases diagnosis using fuzzy logic methods: A systematic and meta-analysis review. Computer Methods and Programs in Biomedicine 161, 145-172.

Ahmadi, N., 2019. Intelligent Approaches towards Fuzzy Segmentation and Fuzzy Edge Detection. Journal of Soft Computing and Decision Support Systems 6(6), 9-13.

Ahmadi, N., Akbarizadeh, G., 2017. Hybrid robust iris recognition approach using iris image pre-processing, two-dimensional gabor features and multi-layer perceptron neural network/PSO. Iet Biometrics 7(2), 153-162.

Ahmadi, N., Akbarizadeh, G., 2018. Iris tissue recognition based on GLDM feature extraction and hybrid MLPNN-ICA classifier. Neural Computing and Applications, 1-15.

Ahmadi, N., Nilashi, M., Samad, S., Rashid, T.A., Ahmadi, H., 2019. An intelligent method for iris recognition using supervised machine learning techniques. Optics & Laser Technology 120, 105701.

Al-Fatlawi, A.H., Jabardi, M.H., Ling, S.H., 2016. Efficient diagnosis system for Parkinson's disease using deep belief network, 2016 IEEE Congress on Evolutionary Computation (CEC). IEEE, pp. 1324-1330.

Amin, M.S., Chiam, Y.K., Varathan, K.D., 2019. Identification of significant features and data mining techniques in predicting heart disease. Telematics and Informatics 36, 82-93.

Anooj, P., 2012. Clinical decision support system: Risk level prediction of heart disease using weighted fuzzy rules. Journal of King Saud University-Computer and Information Sciences 24(1), 27-40.

Aslam, M.W., Zhu, Z., Nandi, A.K., 2013. Feature generation using genetic programming with comparative partner selection for diabetes classification. Expert Systems with Applications 40(13), 5402-5412.

Åström, F., Koker, R., 2011. A parallel neural network approach to prediction of Parkinson's Disease. Expert systems with applications 38(10), 12470-12474.

Asuncion, A., Newman, D., 2007. UCI machine learning repository.

Avci, D., Dogantekin, A., 2016. An expert diagnosis system for parkinson disease based on genetic algorithm-wavelet kernel-extreme learning machine. Parkinson's disease 2016.

Babu, G.S., Suresh, S., 2013. Parkinson's disease prediction using gene expression–A projection based learning meta-cognitive neural classifier approach. Expert Systems with Applications 40(5), 1519-1529.

Bashir, S., Khan, Z.S., Khan, F.H., Anjum, A., Bashir, K., 2019. Improving Heart Disease Prediction Using Feature Selection Approaches, 2019 16th International Bhurban Conference on Applied Sciences and Technology (IBCAST). IEEE, pp. 619-623.

Behroozi, M., Sami, A., 2016. A multiple-classifier framework for Parkinson's disease detection based on various vocal tests. International journal of telemedicine and applications 2016.

Bhardwaj, A., Tiwari, A., 2015. Breast cancer diagnosis using genetically optimized neural network model. Expert Systems with Applications 42(10), 4611-4620.

Bhatia, S., Prakash, P., Pillai, G., 2008. SVM based decision support system for heart disease classification with integer-coded genetic algorithm to select critical features, Proceedings of the world congress on engineering and computer science. pp. 34-38.

Bhattacharya, I., Bhatia, M.P.S., 2010. SVM classification to distinguish Parkinson disease patients, Proceedings of the 1st Amrita ACM-W Celebration on Women in Computing in India. pp. 1-6.

Budrionis, A., Bellika, J.G., 2016. The learning healthcare system: where are we now? A systematic review. Journal of biomedical informatics 64, 87-92.

Buza, K., Varga, N.Á., 2016. Parkinsonet: estimation of updrs score using hubness-aware feedforward neural networks. Applied Artificial Intelligence 30(6), 541-555.

Çalişir, D., Doğantekin, E., 2011. An automatic diabetes diagnosis system based on LDA-Wavelet Support Vector Machine Classifier. Expert Systems with Applications 38(7), 8311-8315.

Cauchi, M., Fowler, D., Walton, C., Turner, C., Waring, R., Ramsden, D., Hunter, J., Teale, P., Cole, J., Bessant, C., 2015. Comparison of GC-MS, HPLC-MS and SIFT-MS in conjunction with multivariate classification for the diagnosis of Crohn's disease in urine. Analytical Methods 7(19), 8379-8385.

Chaurasia, V., Pal, S., Tiwari, B., 2018. Prediction of benign and malignant breast cancer using data mining techniques. Journal of Algorithms & Computational Technology 12(2), 119-126.

Chen, C.-H., 2014. A hybrid intelligent model of analyzing clinical breast cancer data using clustering techniques with feature selection. Applied Soft Computing 20, 4-14.

An International Journal
JSCDSS
E-ISSN: 2289-8603

Chen, H.-L., Huang, C.-C., Yu, X.-G., Xu, X., Sun, X., Wang, G., Wang, S.-J., 2013. An efficient diagnosis system for detection of Parkinson's disease using fuzzy k-nearest neighbor approach. Expert systems with applications 40(1), 263-271.

Darmawahyuni, A., Nurmaini, S., Firdaus, F., 2019. Coronary Heart Disease Interpretation Based on Deep Neural Network. Computer Engineering and Applications Journal 8(1), 1-12.

Das, R., 2010. A comparison of multiple classification methods for diagnosis of Parkinson disease. Expert Systems with Applications 37(2), 1568-1572.

Das, R., Turkoglu, I., Sengur, A., 2009. Effective diagnosis of heart disease through neural networks ensembles. Expert systems with applications 36(4), 7675-7680.

Dogantekin, E., Dogantekin, A., Avci, D., Avci, L., 2010. An intelligent diagnosis system for diabetes on linear discriminant analysis and adaptive network based fuzzy inference system: LDA-ANFIS. Digital Signal Processing 20(4), 1248-1255.

Erkaymaz, O., Ozer, M., 2016. Impact of small-world network topology on the conventional artificial neural network for the diagnosis of diabetes. Chaos, Solitons & Fractals 83, 178-185.

Eskidere, Ö., Ertaş, F., Hanilçi, C., 2012. A comparison of regression methods for remote tracking of Parkinson's disease progression. Expert Systems with Applications 39(5), 5523-5528.

Froelich, W., Wrobel, K., Porwik, P., 2015. Diagnosis of Parkinson's disease using speech samples and threshold-based classification. Journal of Medical Imaging and Health Informatics 5(6), 1358-1363.

Ganji, M.F., Abadeh, M.S., 2011. A fuzzy classification system based on Ant Colony Optimization for diabetes disease diagnosis. Expert Systems with Applications 38(12), 14650-14659.

Ghumbre, S., Patil, C., Ghatol, A., 2011. Heart disease diagnosis using support vector machine, International conference on computer science and information technology (ICCSIT') Pattaya.

Gopalakrishnan, V., Lustgarten, J.L., Visweswaran, S., Cooper, G.F., 2010. Bayesian rule learning for biomedical data mining. Bioinformatics 26(5), 668-675.

Grover, S., Bhartia, S., Yadav, A., Seeja, K., 2018. Predicting severity of Parkinson's disease using deep learning. Procedia computer science 132, 1788-1794.

Gudadhe, M., Wankhade, K., Dongre, S., 2010. Decision support system for heart disease based on support vector machine and artificial neural network, 2010 International Conference on Computer and Communication Technology (ICCCT). IEEE, pp. 741-745.

Gunduz, H., 2019. Deep Learning-Based Parkinson's Disease Classification Using Vocal Feature Sets. IEEE Access 7, 115540-115551.

Guo, P.-F., Bhattacharya, P., Kharma, N., 2010. Advances in detecting Parkinson's disease, International Conference on Medical Biometrics. Springer, pp. 306-314.

Han, J., Kamber, M., Pei, J., 2011. Data mining concepts and techniques third edition. The Morgan Kaufmann Series in Data Management Systems, 83-124.

Hariharan, M., Polat, K., Sindhu, R., 2014. A new hybrid intelligent system for accurate detection of Parkinson's disease. Computer methods and programs in biomedicine 113(3), 904-913.

Hayashi, Y., Yukita, S., 2016. Rule extraction using Recursive-Rule extraction algorithm with J48graft combined with sampling selection techniques for the diagnosis of type 2 diabetes mellitus in the Pima Indian dataset. Informatics in Medicine Unlocked 2, 92-104.

Islam, M.M., Iqbal, H., Haque, M.R., Hasan, M.K., 2017. Prediction of breast cancer using support vector machine and K-Nearest neighbors, 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC). IEEE, pp. 226-229.

Jain, S., Shetty, S., 2016. Improving accuracy in noninvasive telemonitoring of progression of Parkinson'S Disease using two-step predictive model, 2016 Third International Conference on Electrical, Electronics, Computer Engineering and their Applications (EECEA). IEEE, pp. 104-109.

Johri, A., Tripathi, A., 2019. Parkinson Disease Detection Using Deep Neural Networks, 2019 Twelfth International Conference on Contemporary Computing (IC3). IEEE, pp. 1-4.

Kadam, V.J., Jadhav, S.M., Vijayakumar, K., 2019. Breast cancer diagnosis using feature ensemble learning based on stacked sparse autoencoders and softmax regression. Journal of medical systems 43(8), 263.

Kahramanli, H., Allahverdi, N., 2008. Design of a hybrid system for the diabetes and heart diseases. Expert systems with applications 35(1-2), 82-89.

Kannadasan, K., Edla, D.R., Kuppili, V., 2019. Type 2 diabetes data classification using stacked autoencoders in deep neural networks. Clinical Epidemiology and Global Health 7(4), 530-535.

Karabatak, M., 2015. A new classifier for breast cancer detection based on Naïve Bayesian. Measurement 72, 32-36.

Kausar, N., Abdullah, A., Samir, B.B., Palaniappan, S., AlGhamdi, B.S., Dey, N., 2016. Ensemble clustering algorithm with supervised classification of clinical data for early diagnosis of coronary artery disease. Journal of Medical Imaging and Health Informatics 6(1), 78-87.

Khan, M.M., Chalup, S.K., Mendes, A., 2016. Parkinson's disease data classification using evolvable wavelet neural networks, Australasian Conference on Artificial Life and Computational Intelligence. Springer, pp. 113-124.

Khourdifi, Y., Bahaj, M., 2018. K-Nearest Neighbour Model Optimized by Particle Swarm Optimization and Ant Colony Optimization for Heart Disease Classification, International Conference on Big Data and Smart Digital Environment. Springer, pp. 215-224.

An International Journal
JSCDSS
E-ISSN: 2289-8603

Khuriwal, N., Mishra, N., 2018. Breast cancer diagnosis using adaptive voting ensemble machine learning algorithm, 2018 IEEMA Engineer Infinite Conference (eTechNxT). IEEE, pp. 1-5.

Kotsiantis, S.B., Zaharakis, I., Pintelas, P., 2007. Supervised machine learning: A review of classification techniques. Emerging artificial intelligence applications in computer engineering 160, 3-24.

Li, D.-C., Liu, C.-W., Hu, S.C., 2011. A fuzzy-based data transformation for feature extraction to increase classification performance with small medical data sets. Artificial Intelligence in Medicine 52(1), 45-52.

Long, N.C., Meesad, P., Unger, H., 2015. A highly accurate firefly based algorithm for heart disease prediction. Expert Systems with Applications 42(21), 8221-8231.

Maniruzzaman, M., Kumar, N., Abedin, M.M., Islam, M.S., Suri, H.S., El-Baz, A.S., Suri, J.S., 2017. Comparative approaches for classification of diabetes mellitus data: Machine learning paradigm. Computer methods and programs in biomedicine 152, 23-34.

Marcano-Cedeño, A., Quintanilla-Domínguez, J., Andina, D., 2011. WBCD breast cancer database classification applying artificial metaplasticity neural network. Expert Systems with Applications 38(8), 9573-9579.

Nahar, J., Imam, T., Tickle, K.S., Chen, Y.-P.P., 2013. Association rule mining to detect factors which contribute to heart disease in males and females. Expert Systems with Applications 40(4), 1086-1093.

Naranjo, L., Pérez, C.J., Campos-Roca, Y., Martín, J., 2016. Addressing voice recording replications for Parkinson's disease detection. Expert Systems with Applications 46, 286-292.

Nguyen, T., Khosravi, A., Creighton, D., Nahavandi, S., 2015a. Classification of healthcare data using genetic fuzzy logic system and wavelets. Expert Systems with Applications 42(4), 2184-2197.

Nguyen, T., Khosravi, A., Creighton, D., Nahavandi, S., 2015b. Medical data classification using interval type-2 fuzzy logic system and wavelets. Applied Soft Computing 30, 812-822.

Nilashi, M., bin Ibrahim, O., & Ithnin, N. (2014a). Hybrid recommendation approaches for multi-criteria collaborative filtering. Expert Systems with Applications, 41(8), 3879-3900.

Nilashi, M., bin Ibrahim, O., & Ithnin, N. (2014b). Multi-criteria collaborative filtering with high accuracy using higher order singular value decomposition and Neuro-Fuzzy system. Knowledge-Based Systems, 60, 82-101.

Nilashi, M., bin Ibrahim, O., Ithnin, N., & Sarmin, N. H. (2015a). A multi-criteria collaborative filtering recommender system for the tourism domain using Expectation Maximization (EM) and PCA–ANFIS. Electronic Commerce Research and Applications, 14(6), 542-562.

Nilashi, M., Jannach, D., bin Ibrahim, O., & Ithnin, N. (2015b). Clustering-and regression-based multi-criteria collaborative filtering with incremental updates. Information Sciences, 293, 235-250.

Nilashi, M., Ahmadi, H., Shahmoradi, L., Ibrahim, O., Akbari, E., 2019a. A predictive method for hepatitis disease diagnosis using ensembles of neuro-fuzzy technique. Journal of infection and public health 12(1), 13-20.

Nilashi, M., bin Ibrahim, O., Ahmadi, H., Shahmoradi, L., 2017a. An analytical method for diseases prediction using machine learning techniques. Computers & Chemical Engineering 106, 212-223.

Nilashi, M., Bin Ibrahim, O., Mardani, A., Ahani, A., Jusoh, A., 2018a. A soft computing approach for diabetes disease classification. Health Informatics Journal 24(4), 379-393.

Nilashi, M., Ibrahim, O., Ahani, A., 2016a. Accuracy improvement for predicting Parkinson's disease progression. Scientific reports 6, 34181.

Nilashi, M., Ibrahim, O., Ahani, A., 2016b. Accuracy improvement for predicting Parkinson's disease progression. Scientific reports 6(1), 1-18.

Nilashi, M., Ibrahim, O., Ahmadi, H., Shahmoradi, L., 2017b. A knowledge-based system for breast cancer classification using fuzzy logic method. Telematics and Informatics 34(4), 133-144.

Nilashi, M., Ibrahim, O., Ahmadi, H., Shahmoradi, L., Farahmand, M., 2018b. A hybrid intelligent system for the prediction of Parkinson's Disease progression using machine learning techniques. Biocybernetics and Biomedical Engineering 38(1), 1-15.

Nilashi, M., Ibrahim, O., Dalvi, M., Ahmadi, H., Shahmoradi, L., 2017c. Accuracy improvement for diabetes disease classification: a case on a public medical dataset. Fuzzy Information and Engineering 9(3), 345-357.

Nilashi, M., Ibrahim, O., Samad, S., Ahmadi, H., Shahmoradi, L., Akbari, E., 2019b. An analytical method for measuring the Parkinson's disease progression: A case on a Parkinson's telemonitoring dataset. Measurement 136, 545-557.

Onan, A., 2015. A fuzzy-rough nearest neighbor classifier combined with consistency-based subset evaluation and instance selection for automated diagnosis of breast cancer. Expert Systems with Applications 42(20), 6844-6852.

Ozcift, A., 2012. SVM feature selection based rotation forest ensemble classifiers to improve computer-aided diagnosis of Parkinson disease. Journal of medical systems 36(4), 2141-2147.

Parnetti, L., Gaetani, L., Eusebi, P., Paciotti, S., Hansson, O., El-Agnaf, O., Mollenhauer, B., Blennow, K., Calabresi, P., 2019. CSF and blood biomarkers for Parkinson's disease. The Lancet Neurology.

Paul, A.K., Shill, P.C., Rabin, M.R.I., Murase, K., 2018. Adaptive weighted fuzzy rule-based system for the risk level assessment of heart disease. Applied Intelligence 48(7), 1739-1756.

Pawlovsky, A.P., 2018. An ensemble based on distances for a kNN method for heart disease diagnosis, 2018 International Conference on Electronics, Information, and Communication (ICEIC). IEEE, pp. 1-4.

Peterek, T., Dohnálek, P., Gajdoš, P., Šmondrk, M., 2013. Performance evaluation of Random Forest regression model in tracking Parkinson's disease progress, 13th International Conference on Hybrid Intelligent Systems (HIS 2013). IEEE, pp. 83-87.

Polat, K., 2012. Classification of Parkinson's disease using feature weighting method on the basis of fuzzy C-means clustering. International Journal of Systems Science 43(4), 597-609.

Polat, K., Güneş, S., 2007. Breast cancer diagnosis using least square support vector machine. Digital signal processing 17(4), 694-701.

Polat, K., Güneş, S., Arslan, A., 2008. A cascade learning system for classification of diabetes disease: Generalized discriminant analysis and least square support vector machine. Expert systems with applications 34(1), 482-487.

Rout, S., 2012. Fuzzy petri net application: Heart disease diagnosis, Int. Conference on Computing and Control Engineering (ICCCE 2012).

Şahan, S., Polat, K., Kodaz, H., Güneş, S., 2007. A new hybrid method based on fuzzy-artificial immune system and k-nn algorithm for breast cancer diagnosis. Computers in Biology and Medicine 37(3), 415-423.

Sakri, S.B., Rashid, N.B.A., Zain, Z.M., 2018. Particle swarm optimization feature selection for breast cancer recurrence prediction. IEEE Access 6, 29637-29647.

Schraagen, J.M., Verhoeven, F., 2013. Methods for studying medical device technology and practitioner cognition: The case of user-interface issues with infusion pumps. Journal of biomedical informatics 46(1), 181-195.

Shao, Y.E., Hou, C.-D., Chiu, C.-C., 2014. Hybrid intelligent modeling schemes for heart disease classification. Applied Soft Computing 14, 47-52.

Sheikhpour, R., Sarram, M.A., Sheikhpour, R., 2016. Particle swarm optimization for bandwidth determination and feature selection of kernel density estimation based classifiers in diagnosis of breast cancer. Applied Soft Computing 40, 113-131.

Shilaskar, S., Ghatol, A., 2013. Feature selection for medical diagnosis: Evaluation for cardiovascular diseases. Expert Systems with Applications 40(10), 4146-4153.

Soni, J., Ansari, U., Sharma, D., Soni, S., 2011. Intelligent and effective heart disease prediction system using weighted associative classifiers. International Journal on Computer Science and Engineering 3(6), 2385-2392.

Sri, M.N., Sailaja, D., Priyanka, J.H., Chittineni, S., RamaKrishnaMurthy, M., 2019. Performance Evaluation of SVM and Neural Network Classification Methods for Diagnosis of Breast Cancer, International Conference on E-Business and Telecommunications. Springer, pp. 344-349.

Takci, H., 2018. Improvement of heart attack prediction by the feature selection methods. Turkish Journal of Electrical Engineering & Computer Sciences 26(1), 1-10.

Temurtas, H., Yumusak, N., Temurtas, F., 2009. A comparative study on diabetes disease diagnosis using neural networks. Expert Systems with applications 36(4), 8610-8615.

Übeyli, E.D., 2007. Implementing automated diagnostic systems for breast cancer detection. Expert systems with Applications 33(4), 1054-1062.

Wan, S., Liang, Y., Zhang, Y., Guizani, M., 2018. Deep multi-layer perceptron classifier for behavior analysis to estimate parkinson's disease severity using smartphones. IEEE Access 6, 36825-36833.

Wang, B., Tian, R., 2019. Judgement of critical state of water film rupture on corrugated plate wall based on SIFT feature selection algorithm and SVM classification method. Nuclear Engineering and Design 347, 132-139.

Zare, M., Pahl, C., Rahnama, H., Nilashi, M., Mardani, A., Ibrahim, O., Ahmadi, H., 2016. Multi-criteria decision making approach in E-learning: A systematic review and classification. Applied Soft Computing 45, 108-128.

Zheng, B., Yoon, S.W., Lam, S.S., 2014. Breast cancer diagnosis based on feature extraction using a hybrid of K-means and support vector machine algorithms. Expert Systems with Applications 41(4), 1476-1482.

An International Journal
JSCDSS
E-ISSN: 2289-8603