ISSN:0975-3583,0976-2833 VOL12,ISSUE05,2021

DESIGN INTELLIGENT MACHINE LEARNING CLASSIFIERS IN HEALTHCARE SECTORS FOR DIAGNOSIS ACCURACY

[1]**Channapragada Chinmayi SreeChitra,** Associate Software Engineer,Optum Global Solutions,Hyderabad, ccsreechitra@gmail.com
[2]**C. Rama Seshagiri Rao,** Professor, CSE Department,VignanaBharathi Engineering, College,Hyderabad, crsgrao@gmail.com
[3]**Dr.Yeligeti Raju,** Assoc.Professor in IT, VignanaBharathi Institute of Technology, raju.yeligeti@gmail.com

Abstract: Machine learning techniques are used widely in medical diagnosis to predict and classify tasks. Machine learning techniques are used to diagnose diseases more accurately and more efficiently. The patient's life-care machines and systems experience incremental improvements. This growth leads to an increase in the average lifespan of humans. These health care systems are faced with many challenges, including misinforming patients, privacy, inaccurate data, and lacks of medical information, classifiers to predict, and many other issues. Various systems for diagnosing and predicting diseases have been created, including expert systems, clinical prediction, decision support systems, and personal health records systems. This system is designed to assist doctors in diagnosing diseases accurately. The diagnosis can be described as identifying the symptoms more precisely. It is simple to treat the disease once the symptoms have been identified. These medical systems require a lot of processing power and resources. Medical procedures are also computationally complex.
Researchers have two options when it comes to missing data in medical data: either they detect the problem and remove the relevant data instances from the data set, or they use default methods like mean, median, neighbor, etc., to fill the gap. Both ways do not produce optimal results. Outliers can also be found in the data, which degrades the classifier's performance. Although it is not well explored, few researchers have focused on outlier detection in medical data. This research also addresses two of the most well-known data problems, namely missing value imputation and outlier. KMean++-based data imputation technology addresses the disappeared value imputation problem. This technique validates data via clustering and also calculates missing data values. A hybrid outlier detection method can detect the outlier. LS-SVM classifiers can determine the outcome. This work provides a framework for diagnosing and predicting diabetes using LS-SVM classifiers.

1. Introduction

Over the past decade, patient life care systems and machines have seen incremental improvements. This growth leads to an increase in the average lifespan of humans. These healthcare systems face many challenges, including misinformation, privacy, inaccurate data, lack of medical information, lack of medical information, classifiers to predict, and many other issues. These issues have been addressed by various disease prediction and diagnosis systems, including expert systems, clinical prognosis, decision support systems, personal health record systems, and decision support systems. These systems are designed to assist doctors in diagnosing diseases accurately. Diagnosis of the disease is the ability to identify the symptoms and determine the cause. It is simple to treat the disease once the symptoms have been identified. These medical systems require a lot of processing power and resources. It has also been observed that medical procedures can be computationally complex. The non-communicable disease of diabetes is a major concern in the world. It will affect 300 million people by 2030.

Researchers have two options when it comes to filling in the missing values in medical data: either they detect it and delete the relevant data instances, or they use default methods like mean, median, neighbor, etc. Both ways do not produce the best results. Outliers can also be present in data, which degrades the classifier's performance. Although it isn't well explored, few researchers have focused on outlier detection in medical data. This research also addresses two of the most well-known data problems, namely missing value imputation and outlier. K-Mean++-based data imputation techniques address the missing value issue. This technique validates data using clustering and can also calculate missing data values. Diabetes is when blood sugar levels are too high, and insulin is not enough to regulate them.

These healthcare systems face many challenges, including misinformation, privacy, data loss, inaccuracy, lack of medical information, and classifiers predicting the future. These issues have been addressed by various disease prediction and diagnosis systems, including expert systems, clinical prognosis, decision support systems, and personal health records (1-3). These systems are designed to assist doctors in diagnosing diseases accurately. A diagnosis is the ability to identify the symptoms and determine the cause. It is simple to treat the disease once the symptoms have been identified. These medical systems require a lot of processing power and resources. Medical schemes can also be computationally complex.

ISSN:0975-3583,0976-2833 VOL12,ISSUE05,2021

It is crucial to find new ways and techniques to detect and manage diabetes earlier to reduce diabetic deaths. Late diagnosis can lead to more deaths. Data mining and information technology can be combined to provide a cutting-edge solution for the early detection and diagnosis of diabetes. This includes features selection, meta-heuristic algorithms, diagnostic systems, decision trees, logistic regression, and neural networks. A literature review revealed that diabetes diagnosis has many flaws.

2. Literature survey

Different types of diseases affect people around the globe. Due to the advancement in information and communication technology, many healthcare systems have been developed to diagnose the condition accurately. These systems are designed to detect symptoms and provide early treatment in the event of an emergence. There are many deadly diseases, but diabetes is the most common. It can cause blindness, kidney disease, blindness, and heart attack. These systems are designed to assist doctors in diagnosing diseases accurately. The diagnosis of a disease is the ability to identify the symptoms more precisely. It is simple to treat the disease once the symptoms have been identified. These medical systems require a lot of processing power and resources. Medical procedures can also be computationally complex. Researchers have two options for dealing with missing data in medical data: either delete the data instances that are missing or use default methods like mean, median, neighbor, etc., to fill the gap. Both ways do not produce the best results. Outliers can also be present in data, which degrades the classifier's performance. Although it is understudied, few researchers have focused on outlier detection in medical data. The entire literature can be divided into two parts: * ML Techniques to Predict Disease and Diagnosis // ML Techniques to Predict Diabetes and Diagnosis.

This section explains the various machine learning techniques that can predict and diagnose different diseases, such as stroke, heart disease, cancer, stroke, etc. Lin et al. Lin et al. [22] developed a novel framework for classifying cardiac arrhythmia patients. The proposed framework uses the random forest technique for feature selection and prediction. The MIMIC-III dataset was used to assess the performance of the proposed framework. Results are compared with grid search and genetic algorithm. The proposed framework produces far better results in terms of accuracy rate than grid searching and genetic techniques. Prosed models achieve better accuracy. A computer-assisted diagnosis of lung cancer is possible using an ensemble learning framework [23]. The proposed diagnosis system combines a deep learning model and information about metastasis. Medical Body Area Network provides metastasis information. For accurate diagnosis of thyroid diseases, the CS technique can be applied to various ensemble classifiers [24]. This work examines the use of voting, voting, stacking, and boosting ensemble classifiers. The results are compared using accuracy, sensitivity, and specificity parameters. The stacking ensemble classifier is superior to others.

Abdar et al. proposedtwo methods that were used to predict cerebral strokes based on physiological data. Random forest regression is used to compute missing stroke data values. The prediction of stroke outcomes is also possible using the DL technique. Simulations showed that deep learning techniques achieve a low false-negative rate. The DSS was developed to identify the relevant features and provide effective intervention and treatment for Ischemic stroke. The international stroke data is used in this study. Pearson correlations and Shapiro-Wilk algorithms are used to identify the relevant features. For prediction tasks, several techniques are used, including VC, MLP, and RF.

Bucholc et al. a new classifier were developed using the greedy stepwise and decision trees to improve accuracy. These classifiers use the cheap stepwise method to identify relevant attributes, while the decision tree technique is used for prediction tasks. Based on Japanese stroke patients, the proposed classifier's performance is evaluated. The precision rate of the classifier, as mentioned above, is higher. A smart ensemble method has been developed to treat stroke disease. The authors determine the semantic and syntactic relationships among different attributes of stroke disease. This work covers five hundred seven patients. The patient's medical record contains the symptoms of a stroke. To extract the information from the medical sheet, tagging or entropy methods are used. Artificial neural networks, SVM, and boosting are all used to predict stroke patients better. Results show that ANN produces better results than SVM and other random forest methods, such as bagging, growing, and random forests.

Zhao et al. Different data mining techniques were used to predict Ischemic stroke. This dataset includes medical information for eighty stroke patients as well as one hundred twelve healthy people. It also contains sixteen features that can be used to predict Ischemic stroke. To classify stroke patients as healthily and affected, three machine learning classifiers are used: SVM (SGB), PLR (PLR), and SVM (SGB). SVM methods are more promising than the other two, according to the authors. Nishi et al. [30] used several machine learning techniques to determine ischemic stroke outcomes and provide effective treatment. DNN, RF, and LR are all possible methods. Two thousand forty-three patients were used to evaluate the performance of machine learning techniques. It was found that deep learning predicts a more accurate outcome for Ischemic stroke than random forest or logistic regression techniques. Ali et al. 1 combined historical and real-time data to determine the diagnosis. The author also examines the issues surrounding the quality of health care services.

Vijayakumar et al. Vijayakumar et al. The system was built using IoT sensors and wearables to collect patient data. Fog computing is used to analyze, categorize and share medical information between healthcare service providers and users. The fuzzy k-nearest neighbor approach used the fuzzy k–nearest neighbor method to calculate the similarity coefficients that can be used to distinguish between users and diseases. SNA is also used to indicate the outbreak of disease. To calculate the likelihood that a user will become ill, the Disease Outbreak parameter can be used. The proposed system achieved a 95.9% accuracy in classification, according to the results. Saxena et al. [61] gave an overview of the different treatments based on Ayurveda and complementary therapies and modern homeopathic remedies for the ZIKV viral infection. The author also discussed experimental therapies for ZIKV infections. Ginger et al. Guinier et al. The patient with Zika infection can be recognized by minor skin rashes or edema.

3. HYBRID DISEASE PREDICTION OUTLINE

Diabetes can be prevented by early detection and treatment. Late diagnosis can lead to more deaths. Data mining and information technology can both be used to provide cutting-edge technology solutions that will help diagnose and treat diabetes earlier. Data mining, a branch of computer science, is concerned with exploring hidden data from large databases. This information can be used for medical diagnosis and decision-making. There are many models and decision-making tools that can be used to detect and manage diabetes. [7-11]. Analyzing diabetes data can be difficult because of the complexity of natural medical data and its nonlinearity, normal, structured correlation [7-11]. It has been observed that ML-based systems are dominating the medical healthcare industry [13]. They are used widely for medical imaging, such as stroke, heart disease, and cancer [19]. These systems also allow medical professionals to diagnose and stratify diabetes risk accurately. This chapter addresses missing data imputation and attributes selection issues about the diabetes dataset.

4. PROPOSED METHODOLOGY

This paper discusses the functioning of the proposed framework for diabetes prediction. This system was designed to assist doctors in diagnosing the disease accurately. It is possible to treat the disease once the symptoms have been identified. These medical systems require a lot of processing power and resources. Medical procedures are also computationally complex. Researchers have two options when it comes to missing data in medical data: either they detect the problem and remove the relevant data instances from the data set, or they use default methods like mean, median, neighbor, etc., to fill the gap. Both methods do not produce optimal results. Outliers can also be found in the data, which degrades the performance of the classifier. Although outlier detection in medical data is a common focus of researchers, it has not been extensively explored. K-Mean++ based missing values imputation method, ABC based outlier detection method, and LSVM as a classification method are all part of the proposed framework. The proposed framework for diabetes prediction is illustrated in Figure 1.

4.1 Hybrid outlier Detection Method

Outlier detection is an important part of data analysis. The accuracy of the study can be improved by removing outliers from the dataset. Clustering algorithms are widely used to detect and remove outliers in data sets, according to literature. These algorithms use the distance function to determine outliers. The general approach states that small, low-density clusters can describe an outlier. Normal data, on the other hand, is characterized by large and dense clusters. Many methods can be used to detect outliers within a dataset. These methods can be summarized as statistical, clustering-based, and distance-based.

4.2 K-Mean++ Data Imputation Technique

Sometimes missing data and incomplete data are found in real-world databases. Many factors can cause absent or incomplete data. This could be due to errors in data collection, improper measurement, equipment malfunctioning, or data entry methods. Missing values in databases can cause many problems during knowledge discovery. These problems can be summarized as insufficient efficacy, difficulties with managing data, and difficulty in data analysis. This can lead to biased decisions because of the incomplete data. Therefore, it is impossible to make an exact prediction from these data. Data tolerance techniques and data imputation techniques handle missing values in the literature [136-138]. Data mining algorithms can describe the data tolerance technique, but they cannot compute missing values. Data imputation techniques are used to calculate missing values and then fill the data in the dataset. Table 1 shows the snapshot diabetes dataset. The "*" symbol indicates the missing value.

Table 1: Missing values indication

PR	Gl	BP	ST	IN	BMI	DP	Age
3	116	74	15	105	26.3	0.107	24
*	117	66	31	188	30.8	0.493	22
0	*	65	*	*	24.6	0.66	31
2	122	60	18	106	*	0.717	22

ISSN:0975-3583,0976-2833 VOL12,ISSUE05,2021

*	107	76	*	*	45.3	0.686	24
1	86	66	52	65	*	0.917	29
6	91	*	*	*	29.8	0.501	31
1	*	56	30	56	33.3	1.251	24
4	132	*	*	*	32.9	0.302	23
*	105	90	*	*	*	0.197	46
0	57	60	*	*	21.7	0.735	67

Analyzing the diabetes dataset revealed that 432 instances of data contained missing values. The diabetes dataset contains 763 missing values. Six features out of eight parts have missing values. These missing values include pregnancy, plasma glucose, and diastolic BP. The serum insulin attribute has a 48.6% missing value, while the SFT attribute has 29.5%. It is impossible to be objective due to the number of missing values. The K-Mean $^{++}$ data Imputation method has been proposed to calculate the missing value from the diabetes dataset. Its purpose is to compute the missing values efficiently. Algorithm 1 lists the steps for the K-Mean $^{++}$ data input method.

Algorithm 1: Steps of K-Means++ data assertion method

Input: Dataset (D) and number of clusters (K)

Output: Missing value computation and a complete processed dataset

Step 1: Determine the initial cluster center (c1 ∈ K) fromthe dataset (D) loaded in uniform order.

Step 2: Compute the next cluster center (c) such that (c) = x′ ∈ D using the I probability function mentioned in equation 1.

$$\frac{\boldsymbol{distance}(x`)^2}{\sum_{x\in D}\boldsymbol{distancs}(x)^2} \text{--------------------------------- (1)}$$

Distance (x) denotes the shortest distance between data (x) to the nearest center that is randomly chosen.

Step 3: Continue steps 2 and 3 until you have determined all cluster centers.

Step 4: Compute the Euclidean distance between cluster centers (ci$\in K$) $andeach$ data (x) presented in the dataset (D).

Step 5: Allocate the data(x) to clusters (ci ∈ K) with minimum Euclideandistance.

Step 6: Recomputed the new clusters using equation 2.

$$\left(\frac{1}{C}\right)_j \sum_{K=1}^{c_i} x_i \text{ ---------------------------------(2)}$$

Step 7: Continue repeating steps 4 through 6 until there is no difference in data allocation among clusters. Otherwise, obtain the final centroid.

Step 8: Compute the arrangement of data instances as per the final centroid.

Step 9: Data instances in the cluster were considered to be the nearest neighbor.

Step 10:Calculate the average value of each neighbor (mean value for each cluster), and replace the missing value with the mean value (in the case of numeric attributes).

Step 11:Attain the processed dataset.

Statistics considers the distribution of data to detect outliers. Outliers are data that differ from the standard distribution. These techniques can be used efficiently if the data distribution is known in advance. However, it is difficult to do so with large datasets and high-dimensional data. The clustering-based method uses a clustering model to characterize the data's behavior. To determine outliers, it is possible to use the distance between the data and the respective cluster centroid. If information is located at a significant distance from the individual cluster centroid, it is considered an outlier. Clustering-based methods are also effective in identifying outliers. Distance-based methods can detect outliers by comparing data between them and other data in the data set. This is also called the local outlier factor. The pre-processed data are converted to processed data using the K-Mean++-based missing-data imputation and hybrid outlier detection methods. K-Mean++ effectively addresses the problem of missing data in the diabetes dataset by calculating the optimal values. Outliers in the dataset can be detected using a hybrid outlier detection technique. These outliers are then removed from the data. These techniques convert pre-processed data to processed data, which can then be fed to the classifier to perform prediction tasks.

4.3 Prediction using LS-Support Vector Machine (LS-SVM)

This section describes how to predict the outcome of a task using a support vector model (SVM) on the diabetes data. SVM is a powerful classifier that is widely used for predictive studies. The SVM classifier has an overfitting problem, but this is easily fixed by minimizing structural risk. The LS-SVM variant of the SVM is

ISSN:0975-3583,0976-2833 VOL12,ISSUE05,2021

also presented. Equal constraints can be translated into similar restrictions by minimizing the squared error. LS-SVM is described using linear equations rather than quadratic programming. This allows for the reduction of margin errors and least-squares errors, which can improve prediction accuracy.

5. Experimental Results and Discussion

This section reviews the performance of hybrid diabetes frameworks that are based on the Pima Indian Dataset. Three different methods are used to diagnose diabetes in the proposed framework accurately. These are K-Mean++-based missing value imputation and hybrid outlier detection. To overcome missing values, the K-Mean++ method is used. There are 432 missing data instances in the diabetes dataset. These missing values are present in six of the eight attributes of the diabetic disease dataset. The diabetes dataset contains 763 missing values. Most researchers replace these numbers with "0". The missing value problem in the diabetes dataset was dealt with using the K-Mean++-based missing value imputation technique. This technique calculates the best values. Predictive classifier performance has been affected by the outlier. This work addresses this issue and proposes a hybrid outlier detection method. This technique is designed to increase the accuracy of predictive classifiers. The patterns are also predicted using LS SVM classifiers. Evaluation of the hybrid diabetes framework is done. A confusion matrix is used to derive the parameters discussed above.

Table 2 shows the clustering results from the K-Mean++ data Imputation technique. Cluster_1 has 268 data instances, while Cluster_1 has 768. Cluster_0 has 500 data instances. The imputation technique aims to assign the data instances to Cluster_1 or Cluster_0. Cluster_1 is for diabetes positive, while cluster_0 is for diabetes negative. The previously mentioned imputation technique correctly allocates 207 data cases to Cluster_1 while 394 data instances to Cluster_0. One hundred sixty-seven data instances have been incorrectly classified. Sixty-one data instances belong incorrectly to Cluster_1 while 106 belong to Cluster_0. These data instances are removed from the diabetes dataset. The missing values associated with the remaining data instances are calculated using the average distance between data instances and respective cluster centroids. After applying the K-Mean++ data imputation technique, the diabetes data set now contains 601 instances. The below table shows the simulation results for the K-Mean++-based data imputation technique. It also includes other missing value data techniques. This table is a percentage of instances that were incorrectly classified. These are FKMI and KMI, KNNI and LSSI as well as LSSI, MC and SVDI. The proposed K-Mean++ data input technique has a lower error rate. 21.74 in terms of incorrectly classified data instances, compared to other methods. The SVDI technique has a higher error rate, i.e., 41.95 is the highest error rate of all methods. The study concluded that the K-Mean++ data imputation technique allocates data instances more efficiently to clusters, while SVDI has lower efficacy.

Table 2: Error rate display table

S. No.	Missing Value Imputation Technique	Error Rate (%)
1	K-Mean++	21.45
2	FKMI	30.13
3	KMI	27.26
4	KENNI	25.12
5	LSSI	29.96
6	MC	31.12
7	SVDI	41.34
8	VMI	30.12
9	WKNN	26.34

This study also examines the effectiveness of the Hybrid for outlier detection. These methods include K-NN and Isolation Forest. Simulation results revealed that the proposed Hybrid detects more outliers (i.e., 7.82%) than other techniques. The ABC technique identifies 38 instances of data as outliers. These outliers are then removed from the diabetes dataset. The Percentiles method detects fewer outliers (i.e., 4.33) than other techniques.

ISSN:0975-3583,0976-2833 VOL12,ISSUE05,2021

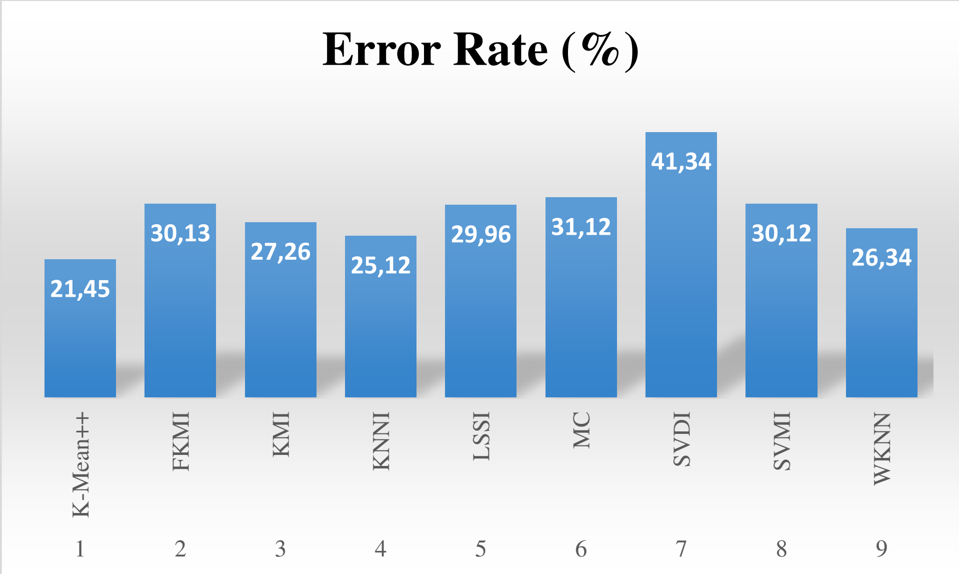
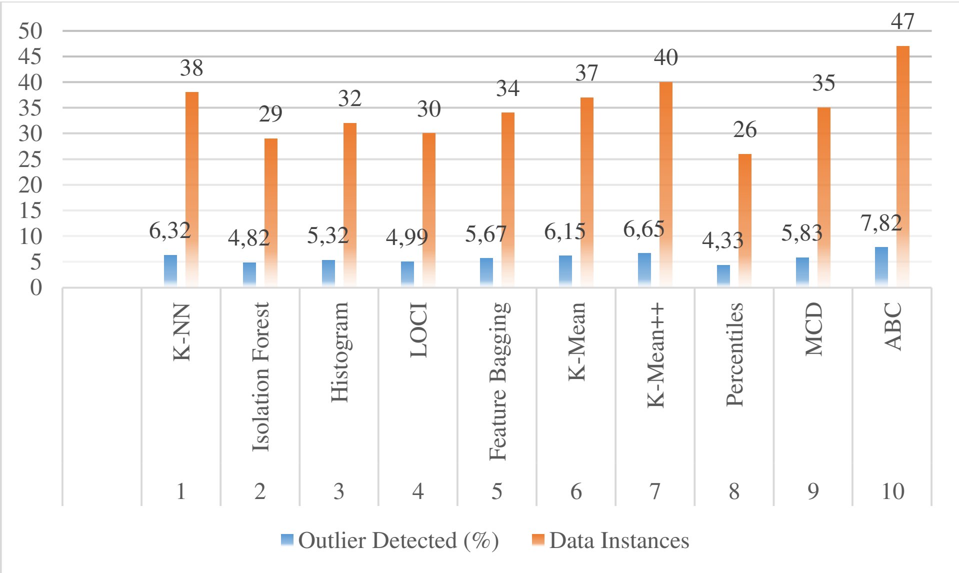


Figure Comparison of accuracy rate of LS-SVM, (K-Mean^{++}+LS-SVM), (Hybrid+SVM) and proposed framework



Comparison of kappa rate of LS-SVM, (K-Mean^{++}+LS-SVM), (ABC+SVM) and proposed framework

6. CONCLUSION

Three different techniques are included in the proposed framework. These are K-Mean++ data imputation and ABC-based outlier detection. Sometimes attributes in a dataset do not contain all the information or have missing values. Missing or incomplete data can cause other problems, such as insufficient efficacy, difficult management of the data, and difficulty in data analysis. You can use the K-Mean++-based method of data imputation to impute missing values. Outliers can have an impact on predictive accuracy. This work uses ahybrid outlier detection method to identify them. These techniques can be used to transform the pre-processed data into processed data. An LS-SVM classification is used to extract the pattern from the processed data. Pima Indian Diabetes dataset is used to evaluate the performance of the proposed framework. This dataset has 763 missing values as well as outliers. These issues are solved by K-Mean++ data Imputation and hybrid Outlier detection techniques. LS-SVM classifications are used to predict diabetes. The proposed framework is compared with several well-known studies. The proposed framework has higher accuracy, sensitivity, and specificity for diabetes prediction and kappa and AUC parameters.

6. REFERENCES

1. Kavakiotis, I., Tsave, O., Salifoglou, A., Maglaveras, N., Vlahavas, I., and Chouvarda, I. (2017). Machine learning and data mining methods in diabetes research. Comput. Struct. Biotechnol. J. 15, 104–116.

ISSN:0975-3583,0976-2833 VOL12,ISSUE05,2021

2. Yadav, G., Kumar, Y. and Sahoo, G., (2012), Predication of Parkinson's disease using data mining methods: A comparative analysis of tree, statistical and support vector machine classifiers. In Computing and Communication Systems (NCCCS), 2012 National Conference on (pp. 1-8). IEEE.
3. Kumar, Y., &Sahoo, G. (2013). Prediction of different types of liver diseases using rule-based classification model. Technology and Health Care, 21(5), 417-432.
4. Yue, C., Xin, L., Kewen, X., and Chang, S. (2008). "An intelligent diagnosis to type 2 diabetes based on QPSO algorithm and WLS-SVM," in Proceedings of the 2008 IEEE International Symposium on Intelligent Information Technology Application Workshops, Washington, DC.
5. Duygu, ç. and Esin, D. (2011). An automatic diabetes diagnosis system based on an LDAwavelet support vector machine classifier. Expert Syst. Appl. 38, 8311–8315.
6. Sahoo, A. J., & Kumar, Y. (2014). Seminal quality prediction using data mining methods. Technology and Health Care, 22(4), 531-545.
7. Gambhir, S., Malik, S. K., & Kumar, Y. (2016). Role of soft computing approaches in healthcare domain: a mini-review—Journal of medical systems, 40(12), 287.
8. Ozcift, A., and Gulten, A. (2011). Classifier ensemble construction with rotation forest to improve medical diagnosis performance of machine learning algorithms. Comput. Methods Programs Biomed. 104, 443–451.
9. Gambhir, S., Malik, S. K., & Kumar, Y. (2017). A PSO-ANN-based diagnostic model for the early detection of dengue disease. New Horizons in Translational Medicine, 4(1-4), 1-8.
10. Kumar, Y., Yadav, G., Singh, P. K., & Arora, P. (2019). A PHR-Based System for Monitoring Diabetes in Mobile Environment. In Mobile Solutions and Their Usefulness in Everyday Life (pp. 129-144). Springer, Cham.
11. Gambhir, S., Malik, S. K., & Kumar, Y. (2018). The Diagnosis of Dengue Disease: An Evaluation of Three Machine Learning Approaches. International Journal of Healthcare Information Systems and Informatics (IJHISI), 13(3), 1-19.
12. Doi, K. (2007). Computer-aided diagnosis in medical imaging: historical review, current status, and future potential. Computerized medical imaging and graphics, 31(4- 5), 198-211. 13. Rangayyan, R. M., Ayres, F. J., & Desautels, J. L. (2007). A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs. Journal of the Franklin Institute, 344(3-4), 312-348.
14. Meystre, S. M., Savova, G. K., Kipper-Schuler, K. C., & Hurdle, J. F. (2008). Extracting information from textual documents in the electronic health record: a review of recent research. Yearbook of medical informatics, 17(01), 128-144.
15. Nilashi, M., Bin Ibrahim, O., Ahmadi, H., & Shahmoradi, L. (2017). An analytical method for diseases prediction using machine learning techniques. Computers & Chemical Engineering, 106, 212-223.
16. Kavakiotis, I., Tsave, O., Salifoglou, A., Maglaveras, N., Vlahavas, I., & Chouvarda, I. (2017). Machine learning and data mining methods in diabetes research. Computational and structural biotechnology journal, 15, 104-116.
17. Shickel, B., Tighe, P. J., Bihorac, A., & Rashidi, P. (2017). Deep EHR: a survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. IEEE Journal of biomedical and health informatics, 22(5), 1589-1604.
18. Nahar, J., Imam, T., Tickle, K. S., & Chen, Y. P. P. (2013). Computational intelligence for heart disease diagnosis: A medical knowledge-driven approach. Expert Systems with Applications, 40(1), 96-104
19. World Health Organization. http://www.who.int 9. Alice, S., & Balachandran, S. (2015), Performance Analysis of Training Algorithms of Multilayer Perceptrons in Diabetes Prediction. Proceedings of the International Conference on Advances in Computer Engineering and Applications (ICACEA), Ghaziabad, India.
20. Zheng, J. W., Zhang, Z. B., Wu, T. H., & Zhang, Y. (2007). A wearable mobihealth care system was supporting real-time diagnosis and alarm. Medical Biological Engineering Computing, 45(9), 877–885.