



A survey on prediction of diabetes using classification algorithms

A. Khanwalkar ^{a,*}, R. Soni ^b

^a Department of Computer Science Engineering, ITM Group of Institution Gwalior, India

^b Head of the Department of Computer Science Engineering,
ITM Group of Institution Gwalior, India

* Corresponding e-mail address: apeksha.khanwalkar@gmail.com

ORCID identifier:  <https://orcid.org/0000-0003-3464-5408> (A.K.)

ABSTRACT

Purpose: Diabetes is a chronic disease that pays for a large proportion of the nation's healthcare expenses when people with diabetes want medical care continuously. Several complications will occur if the polymer disorder is not treated and unrecognizable. The prescribed condition leads to a diagnostic center and a doctor's intention. One of the real-world subjects essential is to find the first phase of the polytechnic. In this work, basically a survey that has been analyzed in several parameters within the poly-infected disorder diagnosis. It resembles the classification algorithms of data collection that plays an important role in the data collection method. Automation of polygenic disorder analysis, as well as another machine learning algorithm.

Design/methodology/approach: This paper provides extensive surveys of different analogies which have been used for the analysis of medical data, For the purpose of early detection of polygenic disorder. This paper takes into consideration methods such as J48, CART, SVMs and KNN square, this paper also conducts a formal surveying of all the studies, and provides a conclusion at the end.

Findings: This surveying has been analyzed on several parameters within the poly-infected disorder diagnosis. It resembles that the classification algorithms of data collection plays an important role in the data collection method in Automation of polygenic disorder analysis, as well as another machine learning algorithm.

Practical implications: This paper will help future researchers in the field of Healthcare, specifically in the domain of diabetes, to understand differences between classification algorithms.

Originality/value: This paper will help in comparing machine learning algorithms by going through results and selecting the appropriate approach based on requirements.

Keywords: Diabetes, Diabetes prediction, Algorithm, Data mining, Machine learning

Reference to this paper should be given in the following way:

A. Khanwalkar, R. Soni, A survey on prediction of diabetes using classification algorithms, Journal of Achievements in Materials and Manufacturing Engineering 104/2 (2021) 77-84. DOI: <https://doi.org/10.5604/01.3001.0014.8490>

BIOMEDICAL AND DENTAL ENGINEERING AND MATERIALS

1. Introduction

Data mining will be a subdirectory within software package engineering [1]. This is a collective process to discover information on the knowledge of technologies, insights, information systems, etc. People's sweets contain a great amount of sugar and fats. Because of these reasons, there has been a multilevel disease among individuals worldwide. As a result, many people visit health centers for a blood test [2]. However, many of these may not have the least possible risk of getting polymyxin. The Polygon Disease Maladie Prediction System, which analyzes the multiple cancer diagnoses of the patient's use of polygon's disease.

Because the polytechnic disease is normal, the complexity of the polygenic disorder, pollinic infection, for diabetes has a proportional increase. It is, therefore, the worst disease in the diseases. There is nothing but a handful [3]. Diseases like diabetes, such as chronic illness, and additional complications have symptoms. Nevertheless, early detection of disease and proper care management will make a difference [4]. Signs of blood are like a heart attack, blindness, morbidity, nervous breakdown, and death. Improves polycystic disease area units. Thirst, improper exhaustion, weight loss, impulsive stomach or inexpensive

reflex – blurred vision, slow-healing infections, weakness of human beings.

2. Data mining

Data Mining Techniques (Fig. 1) are researchers at bioinformatics. Bioinformatics is to collect, extract, organize, shape, and use information from biological sequences, molecules [5]. In recent years, data detection and data processing technologies are used for extracting patterns from extensive biological databases. The amount of biological knowledge is increasing rapidly. Analyzing this knowledge cell creates a source of information by influencing structural or generalization from the information. Communication between data processing and bioinformatics plays an important role in the name of many diseases.

Strong awareness of obtaining helpful information receives large-scale medical knowledge on the need for analysis tools. Researchers have long been involved in applying mathematics and data processing tools to expand knowledge analysis in advanced knowledge sets. The unpleasant phrase about diabetes is one of the appendices that prove that in recent years the instrument of data processing is ensured [6].

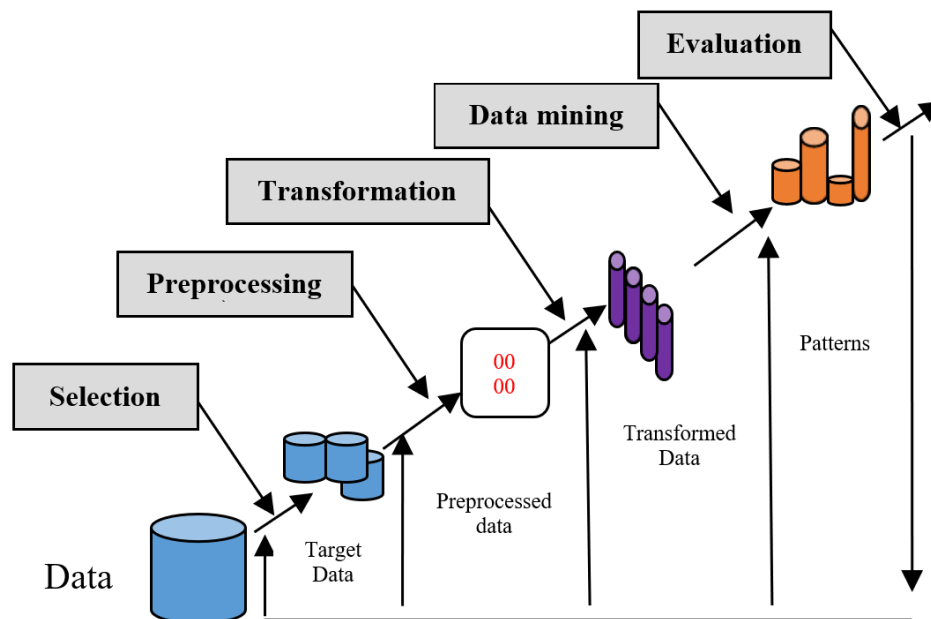


Fig. 1. Steps of data mining process

3. Machine learning

Machine Learning (ML) is a subset of computer science. It solves major world issues "without programming" by providing the mental capacity to the laptop. Machine learning has evolved from the pursuit of research on whether computers should collect data or not emulate the human brain. Arthur Samuel was one of the semi-final semi-semis of 1952 to get enough skills to win the game as the primary game to be a World Semifinal Champion. Later in 1957, Frank Rosenblatt created the Associate in Nursing. Imitating the way in the human brain is a way to solve extensive problems. The development of cubic centimeters has helped to use a computer in drugs [7].

4. Machine learning techniques

There are many technologies for better diagnostic performance in machine learning.

4.1. J48 rule

Each section of each section is to divide the partition into smaller subclasses in a call. Examining the J48 of Basic Knowledge the result of dividing the information by selecting one component [8]. To prepare a conclusion, the intellectual knowledge of the product is growing. If a set related to the same kind of classes is in full context, intensive techniques surround a hold. J48 represents the value of a result. J48, specific attributes, is prepared to select lost attribute values of data, but rather inverse structures.

First of all, if the equal portion of pressure is acceptable, the leaves node takes into account the specified set.

All attributes follow the growth of data from testing on activity.

The best factor is derived from this identification limit.

4.2. Classification And Regression Tree (CART)

Leo Brien, St Economist, Richard Olsen along with associations with nursing Charles Stonn, formed a law known as the Classification and Regression Tree (CART) and developed a daily structure for developing arithmetic models that lack simpler knowledge. It is linked to features that are based on cardboard and completely unfulfilled, based and contribute [9]. This technology can be used to study the number of information, retain information exchange, and continue until few abstracts can be reached.

There is a double chance of canon attributes at this time. This is provided by an associate of non-standard general calls.

Step 1. Choosing to draw close to the canon only attribute is selected.

Step 2. To determine what to stop, you need rules.

Step 3: But nodes divisions have been subordinated.

4.3. Support Vector Machine (SVM) support

Support the vector machine may be a supervised study law, and each correction and regression downside may be resolved, and only think that this focus is distinctive. To illustrate briefly, this is again appearing for a deep hyperplane as shown in a pair (Fig. 2). If such a hyperplane exists, the action is complete. If you do not have such a hyperplane, SVM changes it to the next dataset [10].

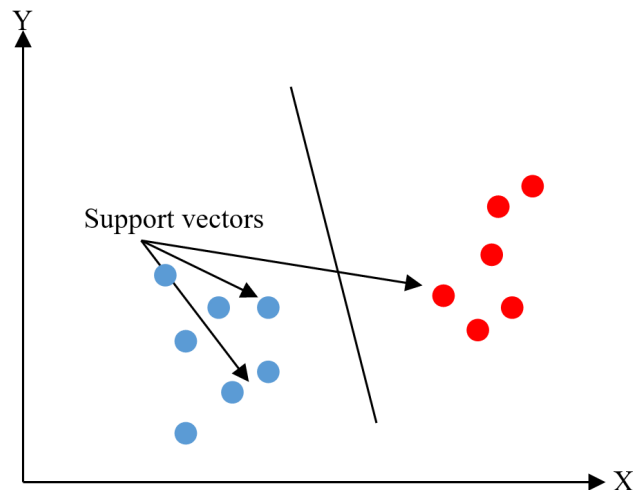


Fig. 2. Support vector machines

As a training law, SVM may not be at a time when compared to other classification methods; however, this is suitable for linear boundaries. SVM has high accuracy. SVM overflow is less likely. SVM successfully applied for letter recognition, text verification, talker identification, etc. SVM may be a characteristic feature. Therefore, the tagged coaching data is provided, which Produces productivity for the formation of unknown causes [11].

4.4. Naïve Bayes classifier

A theory classifier may be mathematical litigation. This classifier helps in overcoming the possibility of a given set. Say the given set or section N [12]. This is often known as

the membership in the category. Bayes classification derives from base theory.

An Assistant degree may be information collected from an unknown category.

Let me assume that there is an assumption of classical issues in the M category.

Calculation $P(H/A)$: Sample A is provided for the find information.

$P(H)$: The former theory H (i.e. the tendency to wait for any information, i.e., before reflecting the background knowledge).

$P(A)$: An opportunity to find sample information.

$P(A|H)$: As long as financing persists, a sample can be understood.

The idea of A is the successor to the universe H, $P(H|A)$ followed by the Bass theory:

$$P(H|A) = (P(A|H) \cdot P(H)) / P(A)$$

4.5. Decision Tree Classifier

A deciding tree is a critical classification algorithm in data mining. The most important benefit of the algorithm of decision making is that they are. It's easy to construct. This is commonly used in most areas [13]. Scientists have developed a range of algorithms to limit your time to improve performance and skills to handle various kinds of knowledge. Decision tree algorithms that are well-liked with ID3, CART, C4.5, C5.0, J48

4.6. k-Nearest Neighbor Algorithm

The neighboring AI-Algorithm program may be a simple method, where it stores all the places and classifies new cases in similar situations. Performance is very simple, and there is a delay in the whole manifestation. One of the most effective parts of neighbors is an entity. k is often a positive number. The exact classification is considered to be the result of the square movement chosen by a group of neighbors.

Step 1. Check k cap number of neighboring neighbors.

Step 2. Calculate the space between the query and all training models.

4.7. K-means cluster

K-meaning K object cells start K in the algorithmic program, which represents the initial cluster center. Cross Browser Center Rights Gate should link all the components

to the nearest hub to support the interconnection of an article with a distance. Once all materials were distributed to the square feet, it was time to find new K clusters. This will continue until the K cluster centers are modified [14].

5. Literature survey

X. Sun et al. [15]. In this paper, the kernel-based adaptive filtering algorithm program is applied to generate predictions. Then, we will look for important things to choose the correct kernel performance. Finally, the accessibility of planned kernel-based caution methods is evaluated in the light glucose prediction. The planned results analyze and evaluate the performance of the IMS algorithm.

Purushottam et al. [16]. A system of diagnosis designed to determine the patient's potential is well-designed. We tend to compare the C45 laws that violate this technology and compare the partial nature of the tree. Systemic activity, generated laws, categorization accuracy, classification error, world-class error, and experimental results all prove that 85 percent of polygenic disorder predict the disease.

S.A. Saji and K. Balachandran [17]. The emphasis is on the performance of caching algorithms which are sorted out of multiple peripherality in polytechnic disorder prediction. During this study, we utilized the Pima Indian Polytechnic Disorder information from the UCI Machine Learning Repository, such as Input Data System is implemented in 2013. Pima has been associated with 768 incidents of polyenoic disorder. The computer file is that the history and goal of the patient's patient are tested as positive or experimental adverse. From performance analysis, all courses of algorithms have found that the Levenberg-Marquardt algorithmic program has given the best training results.

A. Anand and D. Shakti [18]. Personal activity from personal activities of daily activities such as BPI (Body Mass Index), physical activity along with the alternatives like left limits, checking C-squared freedom, classification and regression trees, machine learning algorithm program, crime-validation, etc. Applied and the basis.

S. Bae and T. Park [19]. The goal is to match the performance of predictive models by victimization in every common and rare form. We have a step 1 rectilinear regression, at least for a complete reduction of the choice operator, a pre-risk pricing model, such as elastic-net support and vector machines. We tend to forecast accuracy compared to conniving the world below. Our results show

that manifesting variations in general and unusual variations are generally different or rarely rare.

A. Negi and V. Jaiswal [20]. Experts have developed a double policy, the victimization system. This technology is very reliable due to training, testing, and valid mixing of data. It earned seventy-two precautions.

N. Douali et al. [21]. Failure and Exercise Behavioral Polygenic Disorder Prognosis more and More Basic psychological feature Maps decision network They achieved ninety accuracies in predicting the physical condition polygenic disorder.

K. Yan et al. [22]. In this essay, we tend to plan a unique breathing system entirely. The system is used beautifully to choose chemical sensors to monitor biodiversity. General interaction factors and tenderness, the relation between the number of airflow levels in the lungs, and the square in the square or algorithm. Considering the in-house variation of the lungs, we tend to make models of depressing agents to increase the accuracy of BGL prediction. 295 triggers from health problems and 279 samples collected from diabetic patients. The sensitivity and reproduction of 91.51% and 90.77% polytechnic disorder square measurement score.

S. Perveen et al. [23]. DMD and patients in the discussion of Adopest and Cloth Assembly Machine Learning AIDS Jaws are treated as a diabetic or neurotropic polytechnic disorder risk factor. Results obtained from the test proved, AdaBoost machine Learning umbrella techniques are relatively fabricable compared to a J48 decision tree.

K.M. Orabi et al. polygenic disorder prediction the polygenic disorder of a candidate predict that at a particular stage, [24] designed for a system. The concept of machine learning is supported by the planned system, applies the decision tree. The mechanism designed to predict polygenic disorder events for a specific period has been achieved, resulting in the correct result.

And no one else applied for the Associate Degree algorithm program [25] to classify the possibilities of the DMD. The target Author was found to be using four main machine learning classifications to decision the target, artificial neural networks, logistical regression, and the nick Thomas base. The design of the model is the beauty of the square of the box of boosting techniques. Experimental results of the random forest algorithm offer the best results among all algorithms utilized by the algorithm.

A. Sharma and P.C. Gupta [26] indicates that the processing of data is necessary for the blood protection unit. The J48 method and tree hen code are used for analytical

purposes. Types of classification of the blood group were correct. Its accuracy ranged from eight nine to 94%. This analysis is capable of analyzing the presence of polygenic disorder and the ability of women and women who are plagued with polygenic disorders.

A.G. Karegowda et al. [27] used the K-Mean, KNN Algorithmic Program for Diabetic Patients. They divide diabetes by recommending the results of KNN and k-mean of diabetic patients. 82 accuracy is a precautionary measure.

M. Hardik et al. [28] carried out relative studies of Thomas Beas Classifier and KNN. Besides, to increase the effectiveness of KNN, which effectively increases the number of notification mechanisms, weave degrees can be categorized by selecting options or by increasing the difference.

Y. Wang and Z. Wang [29]. The results of the investigation, indicate that the ancient algorithmic program, which determines the cost of conventional analysis, is the fifth place in the fifth edition of the Ancient Writer.

Y. Angeline Christobel and P. Sivaprakasam [30]. Hyperbolic can be significantly computed in the context of the CNN protocol, which is planned for classification, specialization, and accuracy.

C. Estébanez et al. [31] used analysis in journaling programming for classification programs. The error rate for SVM is 20 seconds, easy to get by twenty-two 14% and mallyerse peruspone 33.31%. In another word, the algorithm of the species is 11 square feet, compared to the use of matrix and clarification. The 10-set cross-validation method uses three different types of cardinal databases and calculations.

J. Han et al. [32] had used a pair of polyenoic disorder information, and the woods called Wood Corn could be turned into a predetermined model. Most of his analysis of the disease prediction is the model of plasma hypoglycemic agents.

A.A. Al Jarullah [33], using the J48 decision tree classifier for her analytical performance. The Southern Infotermist Information Set was enough to implement the Association Act because of the prevalence.

E.G. Yildirim et al. [34]. In his analytic discovery, a pair of pollen disorder information sets were used for data processing in practical measurement in unlimited measurement designing. The adaptation is Neuro Fuss's frocked thinking theory and his filled square foot.

G. Parthiban et al. [35]. In their analysis, the heart-related prophecies and changes are associated with an anemic patient. In their analysis, weaving was used to classify Thomas Bees.

B.M. Patil et al. [36]. Differences from the exact number of distinct classifications in polygenic disorders are quite different. The intensive statistical process support vector machine, quantum connective swinging support for the algorithm program.

6. Conclusions

In this analogy, J48, CART, SVMs and KNN square metabolites were analyzed in medical data for the best answer to the polygenic disorder. Performance indicators are the accuracy, specificity, sensitivity, accuracy, error rate, square meter calculation, and the given data. Get the accuracy of the allegory classifier beyond the accurate information preprocessing techniques. Several authors analyzed it here. Research on polygenic disorder has provided a thorough review of the literature.

References

- [1] D. Shetty, K. Rit, S. Shaikh, N. Patil, Diabetes disease prediction using data mining, Proceedings of the International Conference on Innovations in Information, Embedded and Communication Systems "ICIIECS", Coimbatore, India, 2017, 1-5. DOI: <https://doi.org/10.1109/ICIIECS.2017.8276012>
- [2] R. Mirshahvalad, N.A. Zanjani, Diabetes prediction using ensemble perceptron algorithm, Proceedings of the 9th International Conference on Computational Intelligence and Communication Networks "CICN", Girne, Northern Cyprus, 2017, 190-194. DOI: <https://doi.org/10.1109/CICN.2017.8319383>
- [3] A. Iyer, S. Jeyalatha, R. Sumbaly, Diagnosis of diabetes using classification mining techniques, International Journal of Data Mining and Knowledge Management Process 5/1 (2015) 1-14. DOI: <https://doi.org/10.5121/ijdkp.2015.5101>
- [4] V.P. Kumar, V. Lakshmi, A Data Mining Approach for Prediction and Treatment of diabetes Disease, International Journal of Science Inventions Today 3/1 (2014) 73-79.
- [5] R. Sanakal, T. Jayakumari, Prognosis of Diabetes Using Data mining Approach-Fuzzy C Means Clustering and Support Vector Machine, International Journal of Computer Trends and Technology 11/2 (2014) 94-98. DOI: <https://doi.org/10.14445/22312803/IJCTT-V11P120>
- [6] B.L. Shivakumar, S. Alby, A Survey on Data-Mining Technologies for Prediction and Diagnosis of Diabetes, Proceedings of the International Conference on Intelligent Computing Applications, Coimbatore, India, 2014, 167-173. DOI: <https://doi.org/10.1109/ICICA.2014.44>
- [7] P. Yasodha, M. Kannan, Analysis of a Population of Diabetic Patients Databases in WEKA Tool, International Journal of Scientific and Engineering Research 2/5 (2011) 1-5.
- [8] B. Alić, L. Gurbeta, A. Badnjević, Machine learning techniques for classification of diabetes and cardiovascular diseases, Proceedings of the 6th Mediterranean Conference on Embedded Computing "MECO", Bar, 2017, 1-4. DOI: <https://doi.org/10.1109/MECO.2017.7977152>
- [9] R.M. Khalil, A. Al-Jumaily, Machine learning based prediction of depression among type 2 diabetic patients, Proceedings of the 12th International Conference on Intelligent Systems and Knowledge Engineering "ISKE", Nanjing, China, 2017, 1-5. DOI: <https://doi.org/10.1109/ISKE.2017.8258766>
- [10] H. Walia, A. Rana, V. Kansal, A Naïve Bayes Approach for working on Gurmukhi Word Sense Disambiguation, Proceedings of the 6th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) "ICRITO", Noida, India, 2017, 432-435. DOI: <https://doi.org/10.1109/ICRITO.2017.8342465>
- [11] S. Vijayarani S. Sudha, Disease Prediction in Data Mining Technique-A Survey, International Journal of Computer Applications and Information Technology II/I (2013) 17-21.
- [12] P. Radha, B. Srinivasan, Predicting Diabetes by sequencing the various Data Mining Classification Techniques, International Journal of Innovative Science, Engineering and Technology 1/6 (2014) 334-339.
- [13] N.J. Vispute, D.K. Sahu, A. Rajput, A survey on naive Bayes Algorithm for Diabetes Data Set Problems, International journal for research in Applied Science and Engineering Technology 3/XII (2015) 487-494.
- [14] W. Chen, S. Chen, H. Zhang, T. Wu, A hybrid prediction model for type 2 diabetes using K-means and decision tree, proceedings of the 8th IEEE International Conference on Software Engineering and Service Science "ICSESS", Beijing, China, 2017, 386-390. DOI: <https://doi.org/10.1109/ICSESS.2017.8342938>
- [15] X. Sun, X. Yu, J. Liu, H. Wang, Glucose prediction for type 1 diabetes using KLMS algorithm, Proceedings of

- the 36th Chinese Control Conference “CCC”, Dalian, China, 2017, 1124-1128.
DOI: <https://doi.org/10.23919/ChiCC.2017.8027498>
- [16] Purushottam, K. Saxena, R. Sharma, Diabetes mellitus prediction system evaluation using C4.5 rules and partial tree, Proceedings of the 4th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) “ICRITO”, Noida, India, 2015, 1-6. DOI: <https://doi.org/10.1109/ICRITO.2015.7359272>
- [17] S.A. Saji, K. Balachandran, Performance analysis of training algorithms of multilayer perceptrons in diabetes prediction, Proceedings of the International Conference on Advances in Computer Engineering and Applications, Ghaziabad, India, 2015, 201-206. DOI: <https://doi.org/10.1109/ICACEA.2015.7164695>
- [18] A. Anand, D. Shakti, Prediction of diabetes based on personal lifestyle indicators, Proceedings of the 1st International Conference on Next Generation Computing Technologies “NGCT”, Dehradun, India, 2015, 673-676.
DOI: <https://doi.org/10.1109/NGCT.2015.7375206>
- [19] S. Bae, T. Park, Risk prediction using common and rare genetic variants: Application to Type 2 diabetes, Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine “BIBM”, Kansas City, MO, USA, 2017, 1757-1760. DOI: <https://doi.org/10.1109/BIBM.2017.8217926>
- [20] A. Negi, V. Jaiswal, A first attempt to develop a diabetes prediction method based on different global datasets, Proceedings of the 4th International Conference on Parallel, Distributed and Grid Computing “PDGC”, Wagnaghat, India, 2016, 237-241.
DOI: <https://doi.org/10.1109/PDGC.2016.7913152>
- [21] N. Douali, J. Dollon, M. Jaulent, Personalized prediction of gestational Diabetes using a clinical decision support system, Proceedings of the IEEE International Conference on Fuzzy Systems “FUZZ-IEEE”, Istanbul, Turkey, 2015, 1-5. DOI: <https://doi.org/10.1109/FUZZ-IEEE.2015.7337813>
- [22] K. Yan, D. Zhang, D. Wu, H. Wei, G. Lu, Design of a Breath Analysis System for Diabetes Screening and Blood Glucose Level Prediction, IEEE Transactions on Biomedical Engineering 61/11 (2014) 2787-2795. DOI: <https://doi.org/10.1109/TBME.2014.2329753>
- [23] S. Perveen, M. Shahbaz, A. Guergachi, K. Keshavjee, Performance Analysis of Data Mining Classification Techniques to Predict Diabetes, Procedia Computer Science 82 (2016) 115-121.
DOI: <https://doi.org/10.1016/j.procs.2016.04.016>
- [24] K.M. Orabi, Y.M. Kamal, T.M. Rabah, Early Predictive System for Diabetes Mellitus Disease, in: P. Perner (ed.) Advances in Data Mining. Applications and Theoretical Aspects, Lecture Notes in Computer Science, vol. 9728, Springer, Cham, 2016, 420-427. DOI: https://doi.org/10.1007/978-3-319-41561-1_31
- [25] N. Nai-Arun, R. Mounngmai, Comparison of Classifiers for the Risk of Diabetes Prediction, Procedia Computer Science 69 (2015) 132-142.
DOI: <https://doi.org/10.1016/j.procs.2015.10.014>
- [26] A. Sharma, P.C. Gupta, Predicting the Number of Blood Donors through their Age and Blood Group by using Data Mining Tool, International Journal of Communication and Computer Technologies 1/6 (2012) 6-10.
- [27] A.G. Karegowda, M.A. Jayaram, A.S. Manjunath, Cascading K-means Clustering and k Nearest Neighbor Classifier for Categorization of Diabetic Patients, International Journal of Engineering Advanced Technology 1/3 (2012) 147-151.
- [28] M. Hardik, M.I. Hasan, K.P. Patel, Comparative study of Naive Bayes Classifier and kNN for Tuberculosis, International Journal of Computer Applications 1 (2011) 22-26.
- [29] Y. Wang, Z. Wang, A Fast KNN Algorithm for Text Categorization," Proceedings of the International Conference on Machine Learning and Cybernetics, Hong Kong, China, 2007, 3436-3441. DOI: <https://doi.org/10.1109/ICMLC.2007.4370742>
- [30] Y. Angeline Christobel, P. Sivaprakasam, A New Class wise k Nearest Neighbor (CkNN) Method for the Classification of Diabetes Dataset, International Journal of Engineering and Advanced Technology 2/3 (2013) 396-400.
- [31] C. Estébanez, R. Aler, M. Valls, Genetic Programming Base Data Projections for Classification Tasks, International Journal of Computer and Information Engineering 1/7 (2007) 2195-2200. DOI: <https://doi.org/10.5281/zenodo.1084684>
- [32] J. Han, J.C. Rodriguez, M. Beheshti, Diabetes Data Analysis and Prediction Model Discovery Using RapidMiner, Proceedings of the 2nd International Conference on Future Generation Communication and Networking, Hainan, China, 2008, 96-99. DOI: <https://doi.org/10.1109/FGCN.2008.226>

- [33] A.A. Al Jarullah, Decision tree discovery for the diagnosis of type II diabetes, Proceedings of the International Conference on Innovations in Information Technology, Abu Dhabi, United Arab Emirates, 2011, 303-307. DOI: <https://doi.org/10.1109/INNOVATIONS.2011.5893838>
- [34] E.G. Yildirim, A. Karahoca, T. Ucar, Dosage Planning for diabetes patients using data mining methods, Procedia Computer Science 3 (2011) 1374-1380. DOI: <https://doi.org/10.1016/j.procs.2011.01.018>
- [35] G. Parthiban, A. Rajesh, S.K. Srivatsa, Diagnosis of Heart Disease for Diabetic Patients using Naive Bayes Method, International Journal of Computer Applications 24/3 (2011) 7-11.
- [36] B.M. Patil, R.C. Joshi, D. Toshniwal, Hybrid prediction model for type-II diabetic patients, Expert Systems with Applications 37/12 (2012) 8102-8108. DOI: <https://doi.org/10.1016/j.eswa.2010.05.078>



© 2021 by the authors. Licensee International OCSCO World Press, Gliwice, Poland. This paper is an open access paper distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) license (<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en>).