Original Article

Analysis And Comparison Of Diabetic Prediction Using Medium KNN Classifier And Cosine KNN Classifier

Nirupama S[1], Jenila Rani D[2*]

[1]Research Scholar,Department of Biomedical Engineering,Saveetha School of Engineering,Saveetha Institute of Medical and Technical Sciences,Saveetha University,Chennai,Tamilnadu,India,Pincode:602105

[2]Project guide, Department of Biomedical Engineering,Saveetha School of Engineering,Saveetha Institute of Medical and Technical Sciences, Saveetha University,Chennai,Tamilnadu,India,Pincode:602105

Abstract

Aim: The goal of the study is to find out the presence of diabetes using the Medium K-NN (K- Nearest Neighbour) and Cosine K-NN (K- Nearest Neighbour) algorithm and comparing the accuracy, specificity and sensitivity.

Materials and Methods: A compilation of information from Kaggle's website was used in this research. The samples were regarded as (N=25) for Medium KNN and (N= 25) Cosine KNN according to clinicalc.com, total sample size calculation was performed by keeping alpha error-threshold value 0.05, enrollment ratio is 0:1, 95% confidence interval and 80% power. The accuracy, specificity, sensitivity was calculated by using Matlab software.

Results: The accuracy (%), specificity (%) and sensitivity (%) is compared using SPSS software using independent sample t tests. There is a statistically insignificant difference, p=0.219, p>0.05 with accuracy (47.2%), p=0.067, p>0.05 with specificity (34.54%) and p=<0.01, p<0.05 with sensitivity (48.36%) and demonstrated a better outcome in comparison to Cosine KNN accuracy (44.12%), specificity (39.28%) and sensitivity (45.64%).

Conclusion: Medium KNN appears to give better accuracy, specificity and sensitivity than Medium KNN to predict diabetes.

Keywords: Diabetic prediction, Novel medium KNN, Cosine KNN, Machine learning, Matlab Programming, Accuracy.

DOI: 10.47750/pnr.2022.13.S04.043

INTRODUCTION

Diabetes is a fatal syndrome happening around the world which has a significant barrier or hitch and it is documented by large amounts of data. (R and Manimaran 2017) and (C. Kumar 2019). Since concerted effort at an early stage of the illness would be beneficial. This research aids in diagnosis of diabetes (Ali et al. 2020). The most significant aspect of the study is determining the existence of diabetes using Matlab Programming, while eliminating the risk of human mistake, making it the easiest and cost- effective method of diabetes forecasting (Ghauri 2016). Hospitals and endocrine centres can conduct this study (Zhang et al. 2018) and (Barakat, Bradley, and Barakat 2010).

This study has been linked to 15 Google scholar publications and 10 ScienceDirect articles that were published in recent years to describe the model and algorithm created using machine learning algorithm such as decision tree, logistic regression and K closest neighbour (KNN) algorithms to access and forecast their performance in terms of accuracy (%), specificity (%) and sensitivity (%). (Barakat, Bradley, and Barakat 2010) with the prediction accuracy of (94%) sensitivity (93%) and specificity (94%) on the real world diabetes dataset, intelligible SVM is a potential technique in diabetes prediction when an understandable dataset has been given. (Fiarni, Sipayung, and Maemunah 2019)the suggested models overall accuracy is (68%) thus it may be utilised as an alternative technique for diabetes prediction complications at an early stage. (Çalişir and Doğantekin 2011) and (Kalaiselvi and Nasira 2014) using the specificity and sensitivity analysis of the accurate diagnostic performance of this automated system based on LDA-MWSVM for the diabetes detection was calculated at third stage. The accuracy is (89.47%) (Jarullah and Al Jarullah 2011) the Pima Indian dataset was used which contained the information of the patient who have and who don't have diabetes. (Ali et al. 2020) and (M. Kumar and Kumar 2019) the results show that KNN suppresses the Cosine and Coarse KNN with the accuracy of (98.8%).Our team has

extensive knowledge and research experience that has translate into high quality publications(Bhansali et al. 2021; Jayanth et al. 2021; Sudhakar, Ravel, and Perumal 2021; Sathiyamoorthi et al. 2021; Deepanraj et al. 2021; Raju et al. 2021; Arun Prakash et al. 2020; Kamath et al. 2020; Shanmugam et al. 2021; Rajasekaran et al. 2020; Adhinarayanan et al. 2020; Rajesh et al. 2020; Aurtherson et al. 2021)

The lack of effective early diabetes detection that eliminates human error rate is the driving force behind this project, which aims to determine the presence of diabetes at an early stage by calculating sensitivity percentage and sensitivity percentage. The authors were experts in the field of machine learning algorithms and were able to conduct biomedical studies comparing Novel Medium KNN and Cosine KNN. The primary goal is to compare and contrast diabetic prediction using Novel Medium KNN and Cosine KNN classifiers.

MATERIALS AND METHODS

This research was conducted at the Microprocessor laboratory, Saveetha School of Engineering, Chennai. The sample size was calculated based on the findings of the previous studies (Zheng et al. 2017; Ali et al. 2020) using clinicalc.com, set the alpha error threshold to 0.05, the enrollment ratio to 0:1, the confidence interval to 95% and the power to 80%. The Medium KNN algorithm (N=25) in Cosine KNN (N=25) was the first group. A total of 50 sample sizes.The dataset for the study came from the Kaggle website. Our source data was first trained using Matlab software, and the data was imported into the classifiers. To train the data, the Medium KNN and Cosine KNN classifiers must be chosen. The data confusion matrix was obtained after the training.

A sample dataset of both Medium KNN and Cosine KNN was exported to Microsoft Excel for importing it to Matlab Programming as an input. Matlab 2021 software has to be installed on the PC for training the source data set. The imported data is trained individually for each algorithm by varying k-fold cross-validation. Cross-validation is a model validation technique used to estimate the performance of the model. A confusion matrix is obtained and true positive, true negative, false positive, and false negative values are noted. Accuracy (%), sensitivity (%) and specificity (%) values are calculated from the confusion matrix.

STATISTICAL ANALYSIS

The accuracy (%), specificity (%) and sensitivity (%) comparison of Fine KNN and Novel Medium KNN was done in IBM-SPSS 27.0.1. Since the variables were independent to each other, an independent sample T-Test was done to compare the accuracy, specificity, sensitivity percentages which are considered as independent variables. There were no dependent variables.

RESULTS

In the study of diabetic prediction, both the techniques appear to produce the same variable results, with the accuracy ranging from (45.64%-48.36%), specificity (34.54%-39.28%) and sensitivity ranging from (45.64%-48.36%). Table 1a represents the accuracy, specificity and sensitivity of diabetes using Medium KNN. Table 1b represents the accuracy, specificity and sensitivity of diabetes using Cosine KNN. Table 2 Reflects the comparison of mean accuracy (45.64%), mean specificity (34.54) and mean sensitivity (45.64) using Cosine KNN and mean accuracy (48.36%), mean specificity (39.28%) and mean sensitivity (48.36%) using Cosine KNN. Table 3 shows the independent sample t-test that appears to be a statistically significant difference of accuracy (P=0.0219%, P<0.05), specificity (P=0.067%, P<0.05) and sensitivity (P=<0.001, P<0.05).

Fig. 1 shows the comparison of accuracy between both the classifiers. Confusion matrix represents the true positive rate is 9, while the false positive rate is 3. False negative accounts for 4 and true negative accounts for 8. Using the Novel Medium KNN classifier the overall accuracy is determined to be 48.36%. Fig. 2a depicts the confusion matrix also represents the true positive accounts for 7 while the false positive rate is 5. False negative accounts for 3 and true negative accounts for 9. Fig. 2b depicts the Cosine KNN classifier's overall accuracy determined to be 45.64 %.

DISCUSSION

In this article of predicting diabetes Cosine KNN and Novel Medium KNN had better accuracy p=0.219, p>0.05 with accuracy (47.2%), p=0.067, p>0.05 with specificity (34.54%) and p=<0.01, p<0.05 with sensitivity (48.36%) and demonstrated a better outcome in comparison to Cosine KNN accuracy (44.12%), specificity (39.28%) and sensitivity (45.64%).

Although not statistically significant, the difference appears to have slightly increased (Table 2).

Cosine KNN is the most straightforward and cost effective method of detecting diabetes. (NirmalaDevi, Appavu, and Swathi 2013) The goal of the study is to determine the value of K for PIDD such as amalgam KNN can

improve the classification accuracy. When K is more than one then the suggested model has a classification accuracy of (97.4%) (Tapak et al. 2013) The support vector machine scores best among all the classifiers evaluated in the prediction of diabetes in terms of accuracy, specificity and sensitivity, the total classification accuracy is (85%), specificity (1.000%) and sensitivity (0.820%) (Jayalakshmi and Santhakumaran 2010) When a classifier was applied to the pima Indian dataset, the results were greatly improved when the mixture of preprocessing approaches was used. The experimental systems obtained the accuracy of (99%) (C. Kumar 2019).

The constant and smaller sample size is the factor that affects the outcomes in the validations. The outcomes will grow proportionally as the sampling size and the proportion of training dataset are both increased. On a local level it is sensitive to data structure and memory constraints because it is supervised learning lazy algorithms, as it runs slowly.

Diabetes technology is used in the medical field and how it can help with more accurate detection of other diseases or illnesses in the near future. As a result this initiative has a bright future because manual forecasting can be converted to computerised output quickly and at a low cost. A larger database of real-time applications, combined with other machine learning and deep learning algorithms such as SVM, Naive Bayes should produce better results.

CONCLUSION

When it comes to diabetic prediction, the Novel Medium KNN with the accuracy of (48.36%), specificity of (39.28%) and sensitivity of (48.36%) that make use of Matlab programming appears to give better results when compared to Cosine KNN with the accuracy of (45.64%), specificity of (34.54%) and sensitivity of (45.64%). Furthermore, the algorithm's performance improved as the amount of data increased, which is not the case with other methods. This model is efficient and has a lot of potential for improving diabetes diagnostic efficiency, so it could be used in hospitals and endocrinology centres.

DECLARATIONS

Conflict of interest

There are no conflicts of interest in this manuscript.

Authors contribution

Author SN was involved in data collection, data analysis, and manuscript writing. Author DJR was involved in conceptualisation, data validation and critical review of manuscript.

Acknowledgement

We would like to express our heartfelt gratitude to Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (formerly known as Saveetha University) management for providing me with necessary infrastructure to complete the project.

Funding We thank the following organisation for providing financial support that enabled us to complete the study.

1.Axis clinics Pvt.Ltd,Hyderabad.
2.Saveetha University.
3.Saveetha Institute of Medical and Technical Sciences.
4.Saveetha School of Engineering.

REFERENCES

1. Adhinarayanan, Rajesh, Aravindh Ramakrishnan, Gopal Kaliyaperumal, Melvinvíctor De Poures, Rajesh Kumar Babu, and Damodharan Dillikannan. 2020. "Comparative Analysis on the Effect of 1-Decanol and Di-N-Butyl Ether as Additive with diesel/LDPE Blends in Compression Ignition Engine." *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, June, 1–18.
2. Ali, Ameer, Mohammed A. T. Alrubei, Laith Falah Mohammed Hassan, Mohannad A. M. Al-Ja'afari, and Saif H. Abdulwahed. 2020. "DIABETES DIAGNOSIS BASED ON KNN." *IIUM Engineering Journal*. https://doi.org/10.31436/iiumej.v21i1.1206.
3. Arun Prakash, V. R., J. Francis Xavier, G. Ramesh, T. Maridurai, K. Siva Kumar, and R. Blessing Sam Raj. 2020. "Mechanical, Thermal and Fatigue Behaviour of Surface-Treated Novel Caryota Urens Fibre–reinforced Epoxy Composite." *Biomass Conversion and Biorefinery*, August. https://doi.org/10.1007/s13399-020-00938-0.

4. Aurtherson, P. Babu, Bhanu Teja Nalla, Karthikeyan Srinivasan, Kulmani Mehar, and Yuvarajan Devarajan. 2021. "Biofuel Production from Novel Prunus Domestica Kernel Oil: Process Optimization Technique." *Biomass Conversion and Biorefinery*, May. https://doi.org/10.1007/s13399-021-01551-5.
5. Barakat, Nahla H., Andrew P. Bradley, and Mohamed Nabil H. Barakat. 2010. "Intelligible Support Vector Machines for Diagnosis of Diabetes Mellitus." *IEEE Transactions on Information Technology in Biomedicine: A Publication of the IEEE Engineering in Medicine and Biology Society* 14 (4): 1114–20.
6. Bhansali, Karan J., Kamlesh R. Balinge, Subodh U. Raut, Shubham A. Deshmukh, M. Senthil Kumar, C. Ramesh Kumar, and Pundlik R. Bhagat. 2021. "Visible Light Assisted Sulfonic Acid-Functionalized Porphyrin Comprising Benzimidazolium Moiety for Photocatalytic Transesterification of Castor Oil." *Fuel* 304 (November): 121490.
7. Çalişir, Duygu, and Esin Doğantekin. 2011. "An Automatic Diabetes Diagnosis System Based on LDA-Wavelet Support Vector Machine Classifier." *Expert Systems with Applications*. https://doi.org/10.1016/j.eswa.2011.01.017.
8. Deepanraj, B., N. Senthilkumar, D. Mala, and A. Sathiamourthy. 2021. "Cashew Nut Shell Liquid as Alternate Fuel for CI Engine—optimization Approach for Performance Improvement." *Biomass Conversion and Biorefinery*, February. https://doi.org/10.1007/s13399-021-01312-4.
9. Fiarni, Cut, Evasaria M. Sipayung, and Siti Maemunah. 2019. "Analysis and Prediction of Diabetes Complication Disease Using Data Mining Algorithm." *Procedia Computer Science*. https://doi.org/10.1016/j.procs.2019.11.144.
10. Ghauri, Sajjad Ahmed. 2016. "KNN BASED CLASSIFICATION OF DIGITAL MODULATED SIGNALS." *IIUM Engineering Journal*. https://doi.org/10.31436/iiumej.v17i2.641.
11. Jarullah, Asma A. Al, and Asma A. Al Jarullah. 2011. "Decision Tree Discovery for the Diagnosis of Type II Diabetes." *2011 International Conference on Innovations in Information Technology*. https://doi.org/10.1109/innovations.2011.5893838.
12. Jayalakshmi, T., and A. Santhakumaran. 2010. "A Novel Classification Method for Diagnosis of Diabetes Mellitus Using Artificial Neural Networks." *2010 International Conference on Data Storage and Data Engineering*. https://doi.org/10.1109/dsde.2010.58.
13. Jayanth, Bellappu Venkat, Melvin Victor Depoures, Gopal Kaliyaperumal, Damodharan Dillikannan, Dilipsingh Jawahar, Kumaran Palani, and Ganesha Prasad Meravanigee Shivappa. 2021. "A Comprehensive Study on the Effects of Multiple Injection Strategies and Exhaust Gas Recirculation on Diesel Engine Characteristics That Utilize Waste High Density Polyethylene Oil." *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, June, 1–18.
14. Kalaiselvi, C., and G. M. Nasira. 2014. "A Novel Approach for the Diagnosis of Diabetes and Liver Cancer Using ANFIS and Improved KNN." *Research Journal of Applied Sciences, Engineering and Technology*. https://doi.org/10.19026/rjaset.8.966.
15. Kamath, Manjunath, Subha Krishna Rao, Jaison, Sridhar, Kasthuri, Gopinath, Sivaperumal, and Shantanu Patil. 2020. "Melatonin Delivery from PCL Scaffold Enhances Glycosaminoglycans Deposition in Human Chondrocytes – Bioactive Scaffold Model for Cartilage Regeneration." *Process Biochemistry* 99 (December): 36–47.
16. Kumar, Chandan. 2019. "Prediction of Diabetes Using Data Mining Algorithm." *International Journal for Research in Applied Science and Engineering Technology*. https://doi.org/10.22214/ijraset.2019.2116.
17. Kumar, Munendra, and Anuj Kumar. 2019. "Ensemble Classification Model for Diabetes Prediction in Data Mining." *International Journal of Computer Sciences and Engineering*. https://doi.org/10.26438/ijcse/v7i5.16431647.
18. NirmalaDevi, M., Subramanian Appavu, and U. V. Swathi. 2013. "An Amalgam KNN to Predict Diabetes Mellitus." *2013 IEEE International Conference ON Emerging Trends in Computing, Communication and Nanotechnology (ICECCN)*. https://doi.org/10.1109/ice-ccn.2013.6528591.
19. Rajasekaran, S., D. Damodharan, K. Gopal, B. Rajesh Kumar, and Melvin Victor De Poures. 2020. "Collective Influence of 1-Decanol Addition, Injection Pressure and EGR on Diesel Engine Characteristics Fueled with diesel/LDPE Oil Blends." *Fuel* 277 (October): 118166.
20. Rajesh, A., K. Gopal, De Poures Melvin Victor, B. Rajesh Kumar, A. P. Sathiyagnanam, and D. Damodharan. 2020. "Effect of Anisole Addition to Waste Cooking Oil Methyl Ester on Combustion, Emission and Performance Characteristics of a DI Diesel Engine without Any Modifications." *Fuel* 278 (October): 118315.
21. Raju, P., K. Raja, K. Lingadurai, T. Maridurai, and S. C. Prasanna. 2021. "Glass/Caryota Urens Hybridized Fibre-Reinforced nanoclay/SiC Toughened Epoxy Hybrid Composite: Mechanical, Drop Load Impact, Hydrophobicity and Fatigue Behaviour." *Biomass Conversion and Biorefinery*, March. https://doi.org/10.1007/s13399-021-01427-8.
22. R, Manimaran, and R. Manimaran. 2017. "Prediction of Diabetes Disease Using Classification Data Mining Techniques." *International Journal of Engineering and Technology*. https://doi.org/10.21817/ijet/2017/v9i5/170905319.
23. Sathiyamoorthi, Ramalingam, Gomathinayakam Sankaranarayanan, Dinesh Babu Munuswamy, and Yuvarajan Devarajan. 2021. "Experimental Study of Spray Analysis for Palmarosa Biodiesel-diesel Blends in a Constant Volume Chamber." *Environmental Progress & Sustainable Energy* 40 (6). https://doi.org/10.1002/ep.13696.
24. Shanmugam, Rajasekaran, Damodharan Dillikannan, Gopal Kaliyaperumal, Melvin Victor De Poures, and Rajesh Kumar Babu. 2021. "A Comprehensive Study on the Effects of 1-Decanol, Compression Ratio and Exhaust Gas Recirculation on Diesel Engine Characteristics Powered with Low Density Polyethylene Oil." *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects* 43 (23): 3064–81.
25. Sudhakar, M. P., Merlyn Ravel, and K. Perumal. 2021. "Pretreatment and Process Optimization of Bioethanol Production from Spent Biomass of Ganoderma Lucidum Using Saccharomyces Cerevisiae." *Fuel* 306 (December): 121680.
26. Tapak, Lily, Hossein Mahjub, Omid Hamidi, and Jalal Poorolajal. 2013. "Real-Data Comparison of Data Mining Methods in Prediction of Diabetes in Iran." *Healthcare Informatics Research* 19 (3): 177–85.
27. Zhang, Shichao, Xuelong Li, Ming Zong, Xiaofeng Zhu, and Ruili Wang. 2018. "Efficient kNN Classification With Different Numbers of Nearest Neighbors." *IEEE Transactions on Neural Networks and Learning Systems* 29 (5): 1774–85.
28. Zheng, Tao, Wei Xie, Liling Xu, Xiaoying He, Ya Zhang, Mingrong You, Gong Yang, and You Chen. 2017. "A Machine Learning-Based Framework to Identify Type 2 Diabetes through Electronic Health Records." *International Journal of Medical Informatics* 97 (January): 120–27.

Tables and Figures

Table 1a. Diabetes sample using Medium KNN classifier

Samples	Accuracy	Sensitivity	Specificity
1	0.5	0.5	0.50
2	0.458333	0.478261	0
3	0.458333	0.473684	0.4
4	0.458333	0.470588	0.428571
5	0.416667	0.054545	0
6	0.458333	0.473684	0.4
7	0.416667	0.45	0.25
8	0.5	0.5	0.5
9	0.458333	0.478261	0
10	0.5	0.5	0.5
11	0.541667	0.52381	0.666667
12	0.5	0.5	0.5
13	0.5	0.5	0.5
14	0.458333	0.478261	0
15	0.458333	0.473684	0.4
16	0.458333	0.470588	0.42871
17	0.416667	0.454545	0
18	0.458333	0.473648	0.4
19	0.416667	0.45	0.25
20	0.5	0.5	0.5
21	0.458333	0.478261	0
22	0.5	0.5	0.5
23	0.458333	0.52381	0.666667
24	0.458333	0.5	0.5
25	0.458333	0.478261	0.5

Table 1b. Diabetes sample using Cosine KNN classifier

Samples	Accuracy	Sensitivity	Specificity
1	0.41667	0.444444	0.33333
2	0.458333	0.47619	0.33333
3	0.375	0.411765	0.285714
4	0.416667	0.428571	0.4
5	0.458333	0.466667	0.44444
6	0.375	0.4	0.33333
7	0.416667	0.44444	0.33333
8	0.461538	0.454545	0.5
9	0.416667	0.45	0.25
10	0.458333	0.473684	0.6
11	0.5	0.5	0.33333
12	0.541667	0.526316	0.33333
13	0.41667	0.444444	0.285714
14	0.458333	0.47619	0.4
15	0.375	0.411765	0.44444
16	0.416667	0.428571	0.33333
17	0.458333	0.466667	0.33333
18	0.375	0.4	0.5
19	0.416667	0.44444	0.25
20	0.461538	0.454545	0.6
21	0.416667	0.45	0.33333
22	0.458333	0.473684	0.33333
23	0.5	0.5	0.5
24	0.541667	0.526316	0.25
25	0.541667	0.526316	0.4

Table 2. Comparison of mean, accuracy, specificity and sensitivity using Medium KNN and Cosine KNN. Group statistic comparison of accuracy, specificity and sensitivity for diabetes prediction by using Medium KNN classifier and Cosine KNN classifier. Medium KNN has better mean compared to Cosine KNN. Medium KNN accuracy

(48.36%) specificity (39.28%) and sensitivity (48.36%) Cosine KNN accuracy (44.12%), specificity (34.54%) and sensitivity (45.64%)

Parameters	Classifiers	N	Mean	Std. Deviation	Std. Error Mean
Accuracy	Medium KNN Classifiers	24	.4722	.03617	.00738
	Cosine KNN Classifier	24	.4412	.04756	.00971
Specificity	Medium KNN Classifiers	24	.4836	.02104	.00429
	Cosine KNN Classifier	24	.4564	.03472	.00709
Sensitivity	Medium KNN Classifiers	24	.3454	.00433	.04579
	Cosine KNN Classifier	24	.3928	.09987	.02038

Table 3. Independent sample t-test in predicting the accuracy, specificity and sensitivity of diabetes using Medium KNN classifier and Cosine KNN classifier. There appear to be statistically insignificant differences (p>0.05) in both the methods.

Parameter	Equal Variance	Levene's Test for Equality of Variances		T-test for Equality of Means							
Accuracy	Assumed	F	Sig	t	df	Significance (one-Sided p)	Significance (two-sided p)	Mean Difference	Std.Error Difference	95% Confidence interval (Lower)	95% Confidence interval (Upper)
		1.553	0.219	2.540	46	.007	.015	.03098	.01220	.00643	.0554
	Not assumed			2.540	42.93	.007	.015	.03098	.01220	.00638	.05558
Sensitivity	Assumed	3.522	0.067	3.281	46	<.001	.002	.02718	.00829	.01050	.04396
	Not assumed			3.281	37.88	<.001	.002	.02718	.00829	.01041	.05354
Precision	Assumed	16.99	<0.00	-.945	46	.175	.350	-.04735	.05012	-.14825	.05477
	Not assumed			-.945	31.77	.176	.350	-.04735	.05012	-.14948	.04396

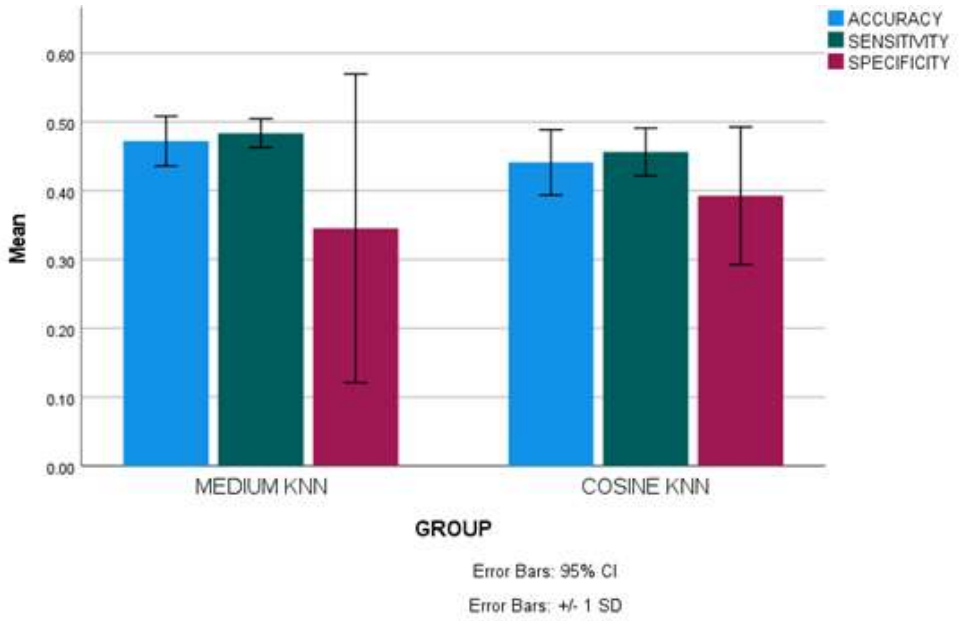

Fig. 1. Simple bar mean of accuracy, mean of specificity and mean of sensitivity using Medium KNN and Cosine KNN. Both techniques appear to produce the same variable results with the accuracy ranging from (44.12%-47.22%), specificity from (45.64%-48.36%) and sensitivity ranging from (34.54%-39.28%). X asis: Medium KNN algorithm vs Cosine KNN, algorithm vs Y axis: Mean accuracy of detection +/- 1 SD

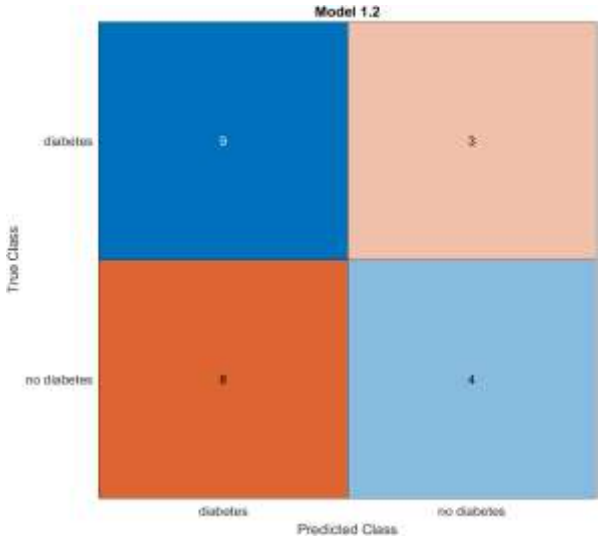

Fig. 2a. Confusion matrix of Medium KNN classifiers. True positive accounts for 9, false positive accounts for 3, false negative accounts for 4 and true negative accounts for 8. The total accuracy was found to be 47.22%.

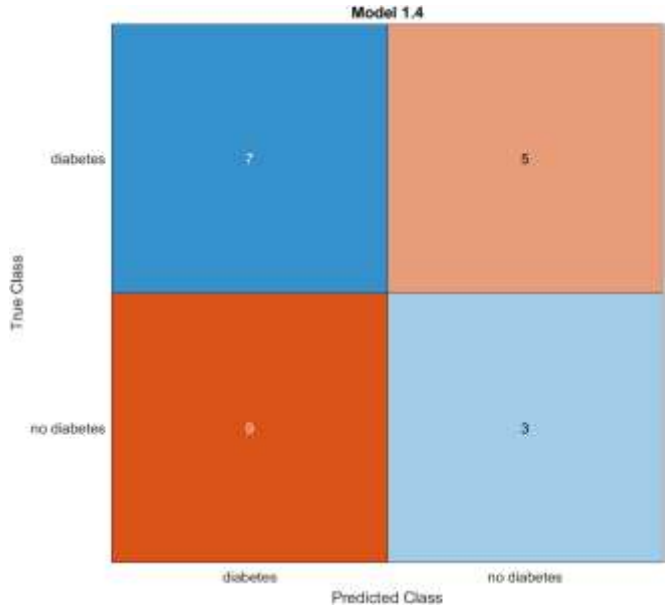

Fig. 2b. Confusion matrix of Medium KNN classifier. True positive accounts for 7, false positive accounts for 5 and false negative accounts for 3 and true negative accounts for 9. The total accuracy was found to be 44.12%.