International Journal of Information Technology Vol. 28 No. 2 2022

Exploring Explainability of Diabetic Diagnosis on Machine Learning Algorithms with the Shapley Value

Yidan Hu[a], Yuan Miao[b], Di Wang[a] and Iain Philip Werry[a]

[a]Nanyang Technological University, Singapore
[b]Victoria University, Australia
{yidan001, wangdi, ipw}@e.ntu.edu.sg, yuan.miao@vu.edu.au

Abstract

Diabetes has affected the lives of millions of patients' and is an economic burden on both patients and society. Machine learning methods are used to detect diabetes in it's early stages to help more people to combat it. However, these methods don't provide explanations for disease diagnosis which can result in patients distrusting the model, and delaying treatment. To address this issue, we aim to implement existing machine-learning methods and explain the predictive outcome of these models. Specifically, we utilize the Shapley value to calculate the impact of individual attributes on the model output. The experimental results show that both medical data and patients' profile information benefit the model prediction.

Keywords: Diabetes, Shapley value.

I. Introduction

Diabetes is a common chronic disease and seriously affects the patient's quality of life. In 1980, there were around 180 million diabetes patients in the world, however this number had increased to 422 million by 2014. Moreover, about 1.5 million people died as a result of diabetes in 2012 [17]. Diabetes also damages the body and vital organs [12, 16]. Early detection is critical for the prevention and treatment of the condition, and can help patients to live a longer, better quality, life. Many machine learning methods are applied to diagnose diabetes and achieve improved performance. However, these methods cannot provide explanations of the predicted results, which can influence a patient's trust in the diagnosis result, and it is often difficult to analyze the reason for the model output. To add insight into the explainability of machine learning methods, we apply the Shapley value [2, 13] to four machine learning algorithms to analyze the contribution of each attribute in the PIDD dataset [15, 5] to the model output. The results show that medical data play the most important role in disease prediction, and that the patients' profile information benefits diabetes diagnosis.

II. Machine Learning Methods for Diabetic Prediction

Machine Learning (ML) methods have been widely used in the healthcare domain [10, 18, 6, 7]. In this work, we utilize four common ML methods (i.e., LightGBM, XGBoost, Decision tree, and Random Forest) to perform diabetic prediction. The details on the ML methods are as follows,

- LightGBM (LGB) [11]: This is a distributed gradient-boosting framework which is based on decision trees.
- XGBoost (XGB) [4]: This is an eXtreme Gradient Boosting package which is a tree learning algorithm. XGB is a gradient-boosting framework library to efficiently use hardware.
- Decision Tree (DT) [14]: This is a tree-based nonparametric supervised learning model which uses greedy search to obtain the best classification results.
- Random Forest (RF) [3]: This is an ensemble learning method which builds up using multiple DTs. The prediction result is decided by all DTs. For the classification task, the output

of the model is the class selected by the most DTs in total.

III. Shapley Value

In this section, we introduce the Shapley value to improve the explainability of ML methods. Shapley value is a concept introduced in game theory to reveal the contribution of individual features on prediction results [8, 9]. Specifically, the Shapley value of a specific feature i is to calculate the mean marginal contribution on all possible combinations of the feature unions. We calculate the Shapley value of feature $i \in N$ as follows,

$$\phi_i(N) = \frac{1}{|N|!} \sum_{S \subseteq N \setminus \{i\}} |S|!(|N| - |S| - 1)! \left(v(S \cup \{i\}) - v(S) \right), \tag{1}$$

where N is the set of all features and S is the subset of N which not includes feature i. $v(\cdot)$ is a characteristic function.

IV. Experiments

A. Dataset

The Pima Indian Diabetes Dataset (PIDD) has been widely used in diabetes prediction tasks, and is collected from the University of California, Irvine (UCI) machine learning repository. The total size of the data sample is 768 records. Each data record contains eight attributes and one ground-truth label. These eight attributes include the number of pregnancies, blood glucose level, skin thickness, insulin level, BMI, blood pressure, diabetes pedigree function, and age.

Model	Accuracy	F1	Recall	Precision	TPR	TNR
LGB	0.7987	0.6804	0.7021	0.6600	0.7021	0.8411
XGB	0.7597	0.6263	0.6596	0.5962	0.6596	0.8037
DT	0.7662	0.6250	0.6383	0.6122	0.6383	0.8224
RF	0.8117	0.6813	0.6596	0.7045	0.6596	0.8785

Table 1: Summary of performance of LGB, XGB, DT, and RF on diabetic prediction.

B. Experimental Setting

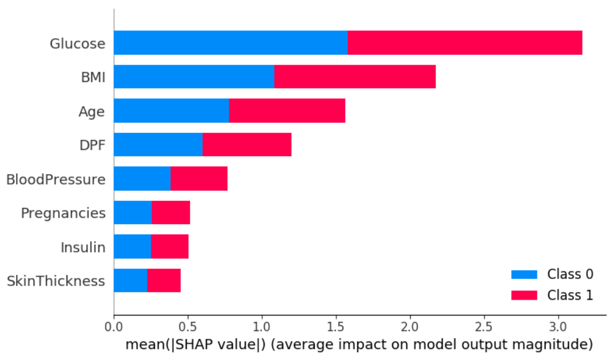
To evaluate the performance of these ML methods, we applied six evaluation metrics, including Accuracy, F1-measure, Recall, Precision, TPR (True Positive Rate), and TNR (False Positive Rate). The dataset was then split into training and testing subsets following an 8:2 proportion.

C. Main Results

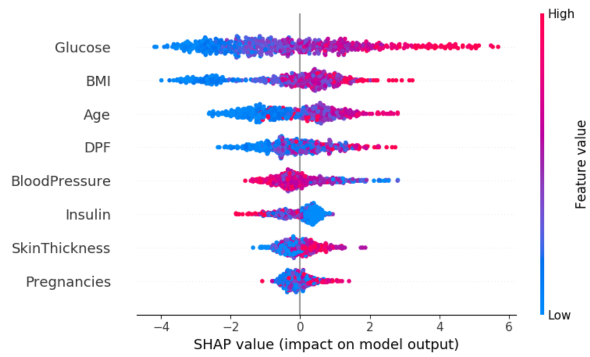
As shown in Table 1, RF achieves the best performance in terms of accuracy, F1, precision, and TNR. DT performs worst among these four models. One possible reason is that the LGB and RF are implemented based on, and extending, decision trees.

D. Analysis of Shapley Value

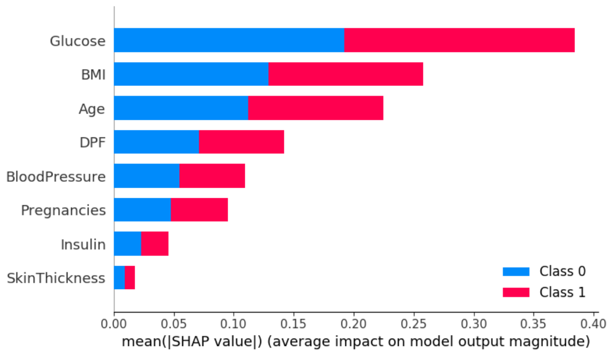
To add further insight into the explainability of the model prediction results, we applied the Shapley value to investigate the impact of individual attributes on the model output. In Figure 1, we can see that blood glucose, BMI, age, and DPF (diabetes pedigree function) play an important role in diabetic diagnosis. Among these attributes, the medical data (i.e., blood glucose) is the most important attribute, which is also to be expected [1], and also the profile information details (i.e., BMI and age) can improve the accuracy of a diabetes diagnosis. In the RF model, the number of pregnancies is also helpful for disease prediction. Among these attributes, skin thickness has a minimal impact on diagnostic results. For LGB, XGB, and DT, the blood pressure rating can



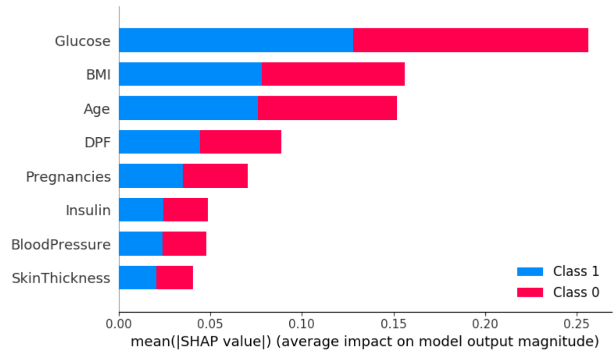
(a) SHAP summary plot of LGB model



(b) SHAP summary plot of XGB model



(c) SHAP summary plot of DT model



(d) SHAP summary plot of RF model

Fig 1: SHAP value summary of four ML methods

benefit the model prediction performance. As shown in Figure 1 (b), we can observe that lower blood glucose, BMI, and age will help participants get eliminate the diabetic risk.

E. Conclusion

In this work, we compared four machine-learning methods using the PIDD dataset. The results show that Random Forest achieves the best performance. To enhance the explainability of the models, we applied the Shapley value to analyze the contribution of individual attributes to the prediction results. The results illustrate that both the medical data and profile data can benefit disease diagnosis.

References

[1] Tsehay Admassu Assegie, Thulasi Karpagam, Radha Mothukuri, Ravulapalli Lakshmi Tulasi, and Minychil Fentahun Engidaye. Extraction of human understandable insight from machine learning model for diabetes prediction. *Bulletin of Electrical Engineering and Informatics*, 11(2):1126–1133, 2022.

[2] Robert Aumann and Sergiu Hart. Handbook of game theory with economic applications. Technical report, Elsevier, 2002.

[3] Leo Breiman. Random forests. *Machine Learning*, 45:5–32, 2001.

[4] Tianqi Chen, Tong He, Michael Benesty, Vadim Khotilovich, Yuan Tang, Hyunsu Cho, Kailong Chen, Rory Mitchell, Ignacio Cano, Tianyi Zhou, et al. XGBoost: extreme gradient boosting. *R Package Version 0.4-2*, 1(4):1–4, 2015.

[5] Dilip Kumar Choubey, Sanchita Paul, Santosh Kumar, and Shankar Kumar. Classification of Pima Indian diabetes dataset using naive Bayes with genetic algorithm as an attribute selection. In *Proceedings of International Conference on Communication and Computing System*, pages 451–455, 2017.

[6] Arianna Dagliati, Simone Marini, Lucia Sacchi, Giulia Cogni, Marsida Teliti, Valentina Tibollo, Pasquale De Cata, Luca Chiovato, and Riccardo Bellazzi. Machine learning methods to predict diabetes complications. *Journal of Diabetes Science and Technology*, 12(2):295–302, 2018.

[7] Murali Krishna Doma, Kayal Padmanandam, Sunil Tambvekar, Keshav Kumar, Bilal Abdualgalil, and RN Thakur. Artificial intelligence-based breast cancer detection using WPSO. *International Journal of Operations Research and Information Systems*, 13(2):1–16, 2022.

[8] Lujain Ibrahim, Munib Mesinovic, Kai-Wen Yang, and Mohamad A Eid. Explainable prediction of acute myocardial infarction using machine learning and shapley values. *IEEE Access*, 8:210410–210417, 2020.

[9] Habib Jafari, Shamarina Shohaimi, Nader Salari, Ali Akbar Kiaei, Farid Najafi, Soleiman Khazaei, Mehrdad Niaparast, Anita Abdollahi, and Masoud Mohammadi. A full pipeline of diagnosis and prognosis the risk of chronic diseases using deep learning and Shapley values: The Ravansar county anthropometric cohort study. *PloS One*, 17(1):e0262701, 2022.

[10] Ioannis Kavakiotis, Olga Tsave, Athanasios Salifoglou, Nicos Maglaveras, Ioannis Vlahavas, and Ioanna Chouvarda. Machine learning and data mining methods in diabetes research. *Computational and Structural Biotechnology Journal*, 15:104–116, 2017.

[11] Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, 30, 2017.

[12] Jamaludin Mohamed, AH Nazratun Nafizah, AH Zariyantey, and SB4868511 Budin. Mechanisms of diabetes-induced liver damage: the role of oxidative stress and inflammation. *Sultan Qaboos University Medical Journal*, 16(2):e132, 2016.

[13] Dov Monderer and Dov Samet. Variations on the Shapley value. *Handbook of Game Theory with Economic Applications*, 3:2055–2076, 2002.

[14] Harsh H Patel and Purvi Prajapati. Study and analysis of decision tree based classification algorithms. *International Journal of Computer Sciences and Engineering*, 6(10):74–78, 2018.

[15] Radhanath Patra et al. Analysis and prediction of Pima Indian Diabetes Dataset using SD-KNN classifier technique. In *Proceedings of IOP Conference Series: Materials Science and Engineering*, volume 1070, page 012059, 2021.

[16] Dahliatul Qosimah, Dhita Evi Aryani, Ma Asuncion Guiang Beltran, et al. Diabetes sepsis on Wistar rat strain (Rattus norvegicus) induced by streptozotocin and bacteria Staphylococcus aureus. *Veterinary World*, 12(6):849, 2019.

[17] Gojka Roglic et al. WHO Global report on diabetes: A summary. *International Journal of Noncommunicable Diseases*, 1(1):3, 2016.

[18] Quan Zou, Kaiyang Qu, Yamei Luo, Dehui Yin, Ying Ju, and Hua Tang. Predicting diabetes mellitus with machine learning techniques. *Frontiers in Genetics*, 9:515, 2018.