|  |  |  |  |
| --- | --- | --- | --- |
| Title | Explainable deep learning for diabetes diagnosis with DeepNetX2 | | |
| Author(s) Name | Sharia Arfin Tanim, Al Rafi Aurnob, Tahmid Enam Shrestha, MD Rokon Islam Emon, M.F. Mridha ∗, Md Saef Ullah Miah | | |
| Contact Email(s) | firoz.mridha@aiub.edu | | |
| Published Journal Name | Biomedical Signal Processing and Control | | |
| Type of Publication | Journal | | |
| Volume | 99 | Issue |  |
| Publisher | Elsevier | | |
| Publication Date | 2024 | | |
| ISSN | 2169-3536 | | |
| DOI | https://doi.org/10.1016/j.bspc.2024.106902 | | |
| URL | https://www.sciencedirect.com/science/article/pii/S1746809424009601 | | |
| Other Related Info. |  | | |
|  | | | |

|  |  |
| --- | --- |
| Abstract |  |
| Diabetes is a leading health global health challenge because of its high blood sugar levels and the risk of extensive damage to other internal organs. Early and accurate identification of diabetes is important because it may cause other diseases including heart diseases and nerve damage. Despite the success of using machine learning, especially deep learning in automated diabetes diagnosis. These models are mostly black boxes which rarely offer comprehensive explanations and interpretations of the results. This study introduces DeepNetX2,  a proposed custom deep neural network designed to overcome these challenges by integrating Explainable Artificial Intelligence (XAI) techniques, specifically Local Interpretable Model-agnostic Explanations (LIME) and Shapley Additive Explanations (SHAP). These techniques make the decision-making process of the model transparent, thereby increasing the credibility of the predictions. The proposed methodology entails a comprehensive data preprocessing technique that involves a customized Spearman’s correlation coefficient feature selection strategy. This preprocessing restricts complexity to only relevant features that promote effectiveness, instead of oversimplification, to the point of decreasing efficiency. DeepNetX2 was rigorously tested on three datasets: the PIMA dataset, local private dataset, and Type-2 diabetes dataset, achieving test  accuracies of 94.81%, 97.87%, and 97.50%, respectively. These results demonstrate not only the superior performance of DeepNetX2 compared with existing models, but also its enhanced interpretability. Based on  the proposed model, an appropriate and rapid strategy for diabetes prediction was developed which is very useful for improving diagnostic integrity and patient health. | |