Review of: "Rules Extraction, Diagnoses and Prognosis of Diabetes and its Comorbidities using Deep Learning Analytics with Semantics on Big Data"

Michael Owusu-Adjei[1]

1 Kwame Nkrumah University of Science and Technology

Potential competing interests: No potential competing interests to declare.

REVIEW REPORT

The research seeks to simplify feature selection use in an area where there is vast feature dimensionality (healthcare). Research on addressing feature dimensionality associated with diseases with comorbidities is a welcome development for diagnostic purpose. This article is well written and structurally commendable.

However, the use of modeling techniques in this domain and its evaluation metrics need clarifications such as the following;

1. Diabetes types (1 & 2) and the impact of non-treatment - types describe disease severity and impact explains consequences (Ndisang JF, Vannacci A, Rastogi S. Insulin Resistance, Type 1 and Type 2 Diabetes, and Related Complications 2017. J Diabetes Res. 2017;2017:1478294. doi: 10.1155/2017/1478294. Epub 2017 Nov 15. PMID: 29279853; PMCID: PMC5723935.) author reference to gestational, neonatal etc as other forms instead of (location or stage occurrence) needs clarifying.
2. Predictive modeling evaluation performance for real-world applications where the possibility of output class imbalance is high using prediction accuracy metric is not enough to determine best model performance. Authors are advised to include other evaluation metrics that take into account the possibility of dataset class imbalance. Some of these include metrics such as model performance on true positive rates, true negative rates, false positives, false negatives, predicted positive values, predicted negative values etc.
3. How did the record of 100 patients impact on prediction outcome especially for the use of deep learning techniques?
4. Predictive modeling for datasets with high dimensionality needs feature engineering. How was feature selection used in determining feature importance for an important modeling subject as disease diagnosis?