PREDICTING READMISSION RISK FOR DIABETIC PATIENTS

Make artificial intelligence work in real life with interpretable machine learning

AUTHORS

Haohui Lu &
Shahadat Uddin

AFFILIATIONS

School of Project Management,
Faculty of Engineering,
The University of Sydney

BACKGROUND

The expenditures of healthcare services associated with unplanned readmissions are enormous. Recognising the reasons that contribute to readmission and identifying at-risk patients are the essential steps to reduce such readmissions. Artificial intelligence is changing the practice of healthcare. It has enabled medical practitioners to provide high-quality treatments to reduce readmissions. While it is essential to employ such solutions, making them transparent to medical experts is more critical.

AIMS

Apparently, healthcare stakeholders do not have a strong data science background. Doctors consider the cause of a prognosis rather than the binary outcome from the machine learning methods. Most of the previous work presented predictions, but they did not explain the models. Treating these models as “black boxes” diminishes confidence in their predictions. This study aims to report high-performing predictive models for readmission along with transparent interpretations.

METHODS

Six machine learning applications are employed in predicting 30-day readmission after hospitalisation for diabetic patients. This study also utilised the white-box machine learning framework by exploring Shapley Additive Explanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME).

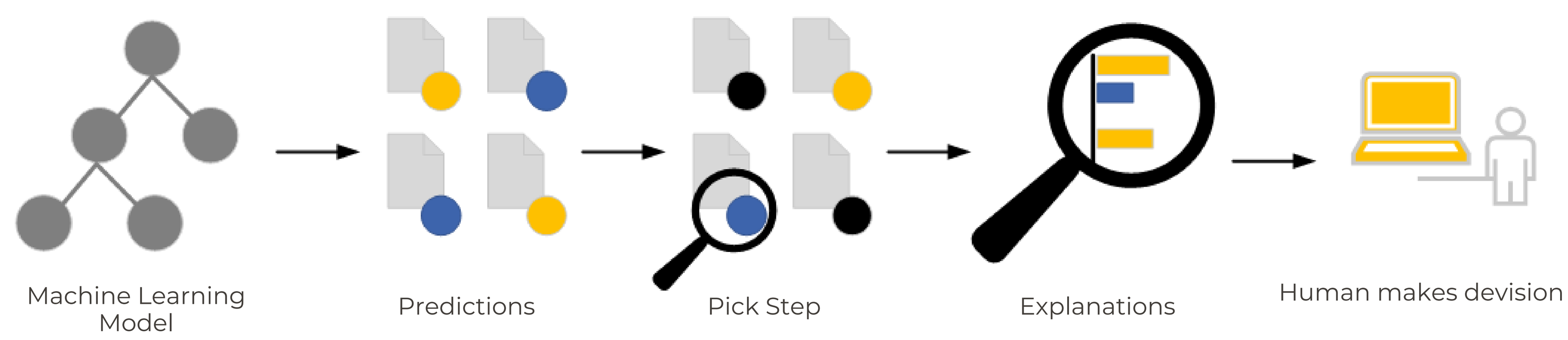
RESULTS

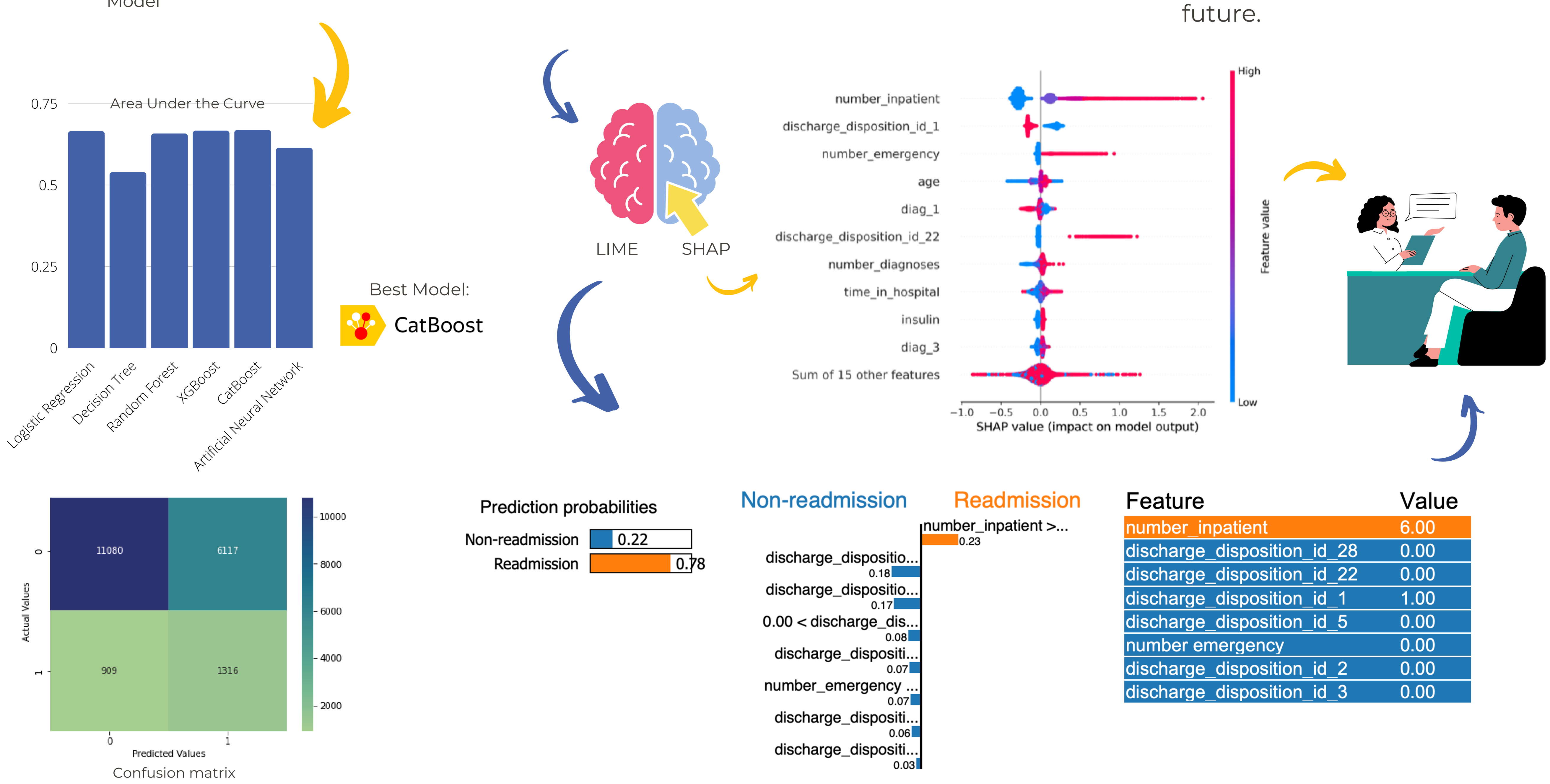
Results show that the CatBoost had the best performance with a higher area under the receiver operating characteristic curve than other models. Prior readmission, discharged at home, the number of emergencies and age were strong predictors. To demonstrated explainability at the individual level, we interpreted the relative variable influence of patient observations.

DISCUSSIONS

According to the findings from SHAP Values, we found the ‘Number of inpatient visits’ variable has the greatest impact on readmission in the CatBoost model.

LIME is applied to the models created and how predictions are made at the individual observation level. For example, a patient had been hospitalised six times before this stay, and this patient was discharged to home and was between the ages of 40 and 50. The model predicted the probability of 30-day readmission is 78%. The model had considered that the six times of inpatient has a positive influence on the prediction with a coefficient of 0.23. However, other most important features such as being discharged to home and the number of emergencies have negative impacts on the outcome of readmission.





CONCLUSION

The findings could be helpful in medical practice and provide valuable recommendations to stakeholders for minimising readmission and reducing public healthcare costs in the future.