Using Deep Learning and Interpretable AI to Predict Outcomes in Medical Data

Annaa Osman Jakpa, Musa Taib, Brian Nguyen, Ethan MacDonald, Geoffrey Messier
Department of Electrical and Software Engineering, University of Calgary

Introduction: Administrative medical data is generated at every encounter with the health care centers [1]. By using machine learning (ML) models, prediction of adverse outcomes can be made to avert critical or life-threatening situations. Models are used in the medical field to help experts make the best decisions possible. Therefore, it is very important for predictive models to be interpretable and transparent. This paper seeks to compare the performance of an interpretable model (decision tree) and black box models (neural networks) in predicting adverse outcomes in patients up to a year after their last recorded interaction with the healthcare system.

Methods: A synthetic health care dataset created from Major League Baseball (MLB) data was used for the findings in this abstract. The entries total 452,995 from 1993 to 2011 and contains 915 unique individuals. The synthetic dataset's features were created by mapping MLB data's features to frequently occurring health care events. The target feature that each model predicted was adverse outcomes. Adverse outcomes in medical settings are life threatening and early identification is necessary for support.

Figure 1 represents a decision tree to perform binary classification on a dataset with four features (f1, f2, f3 and f4). To arrive at a decision (A or B), the tree asks sequential questions about the features associated with the examples. Examples are classified by following the path from the root to leaf. Combining simple questions about the data understandably makes decision trees interpretable and transparent.

In this study, two neural networks were used (ANN and RNN). ANNs are universal approximators because of their versatile pattern recognition abilities. As a result, the ANN serves as a great baseline for the comparison of outcomes to other models. The RNN, on the other hand, features a memory cell, which makes it a great model for time series data (medical records are mostly time series data).

Results: The ANN had the best performance out of all three models with precision of 0.68 and recall of 0.92, this was followed by the RNN with precision of 0.60 and recall of 0.85. The decision tree recorded a precision of 0.62 and a recall of 0.65.

The rules produced by the decision tree made it clear that the number of days a patient stays in the hospital was a major factor in determining if they will have an adverse outcome, while no such information could be deduced from the neural networks.

Discussion: Although utilizing the tree model rather than the deep learning model has a 27% reduction in performance in recall, it offers a much more straightforward and interpretative model and is anticipated to gain more acceptance in clinical practice. It should be noted, interpretable AI, and deep learning models both have benefits and limitations.

Future Work: While the results discussed in the abstract are all using the synthetic dataset, the future work will be done with the MIMIC-III dataset. MIMIC-III is a large administrative health database released by MIT in 2016[2], consisting of 46,520 patients admitted to a medical center between 2001 and 2012.

References:

1. Cadarette, S. M., & Wong, L. (2015). An Introduction to Health Care Administrative Data. The Canadian journal of hospital pharmacy, 68(3), 232–237. https://doi.org/10.4212/cjhp.v68i3.1457.
2. Johnson, A., Pollard, T., & Mark, R. (2016). MIMIC-III Clinical Database (version 1.4). PhysioNet. https://doi.org/10.13026/C2XW26.

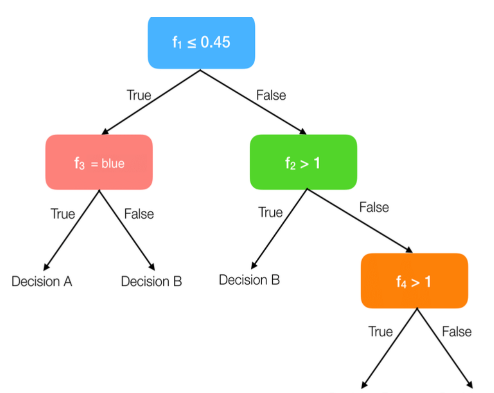

Figure 1: A Simple Decision Tree