187

Predicting change in target weight of a haemodialysis patient using ensemble machine learning technique

Dr Shoab Saadat[1], Dr Naveen Naqvi[2], Dr Sidra Saadat[3]
[1]Mid & South Essex NHS Trust, Chelmsford, United Kingdom, [2]Shifa International Hospital, Islamabad, Pakistan, [3]Rawalpindi Medical College, Rawalpindi, Pakistan

Introduction:
Target weight estimation is a clinical skill used by clinicians across the world on hemodialysis (HD) patients. The objective of this study was to develop a machine learning algorithm which could be used as a mobile app alongside the conventional clinical estimation to aid the renal physicians.
Methods:
This was a prospective cohort study which was carried out from June 2019 to December 2019, at a tertiary care hospital in Islamabad, Pakistan. All the consenting patients (by non-probability convenience sampling) were included, who had received HD for at least 3 months and who didn't have any disability to communicate. A total of 102 patients were enrolled. An MBBS qualified physician administered a proforma to the patients at the start of a one-month observational period that recorded predictors like age, sex, income, cause of renal failure, HD duration, HD regimen, quality of life score using WHO QOL BREF, dietary compliance, medicinal compliance and electrolytes etc. Target weight as an outcome was estimated clinically by the in-charge renal consultant at the start and end of this observational period. Statistical analysis included descriptive stats and building of four machine learning algorithmic models namely linear regression, gradient boost, random forest, Xgboost and an ensemble model using random forest and Xgboost models. We used R statistical software version 3.5.2 for the above analysis.
Results:
The study population included 60% (62/102) males and a median age of 55 years. The mean duration on HD was 37 months while only 38% (38/102) patients had a HD regime of three times per week. Dietary compliance was observed by 87% (88/102) patients while 70% (71/102) patients observed medication compliance.
We divided the data into training, Validation and testing sets and used predictors as above for estimating the change in target weight over the coming one month. We built five models from the training-Validation data set which included Linear regression (Figure 2, R2=0.22, RMSE=9.1), Gradient Boost model (R2=0.22, RMSE=1.8), Random forest (R2=0.32, RMSE=1.56), Xgboost (R2=0.39, RMSE=1.55) and an ensemble model (R2=0.41, RMSE=1.49) using Xgboost and random forest. The best performing model by a long distance was ensemble model which was able to explain about 41% variance in the dependent variable (Figure 1). We then developed a mobile app based on this model which takes in the predictors from last month and can give an estimation of target weight change expected in a given patient. There are plans to increase the sample size to further improve the accuracy of this model and to perform a cost-benefit analysis in terms of work-hours saved per week down the line.
Conclusion:
We were able to develop a predictive model using ensemble machine learning algorithm which could estimate a change in target weight in a given HD patient one month in future with a fairly good accuracy. This model was then implemented in the form of a mobile/web app which can be used by clinicians around the world to get a better estimate of target weight changes in their HD patients.