Healthcare Transformation with Informatics and Artificial Intelligence
J. Mantas et al. (Eds.)

doi:10.3233/SHTI230417

Prediction of Waiting Times in A&E

Luis F ARIAS-GÓMEZ[a], Thomas LOVEGROVE[b] and Holger KUNZ[a,1]
[a] *Institute of Health Informatics, University College London, London, United Kingdom*
[b] *East Kent Hospitals University NHS Foundation Trust*

Abstract. Predicting waiting times in A&E is a critical tool for controlling the flow of patients in the department. The most used method (rolling average) does not account for the complex context of the A&E. Using retrospective data of patients visiting an A&E service from 2017 to 2019 (pre-pandemic). An AI-enabled method is used to predict waiting times in this study. A random forest and XGBoost regression methods were trained and tested to predict the time to discharge before the patient arrived at the hospital. When applying the final models to the 68,321 observations and using the complete set of features, the random forest algorithm's performance measurements are RMSE=85.31 and MAE=66.71. The XGBoost model obtained a performance of RMSE=82.66 and MAE=64.31. The approach might be a more dynamic method to predict waiting times.

Keywords. Waiting times, A&E, Random Forest, XGBoost

1. Introduction

Recently, increasing waiting times created critical situations in A&E departments [1]. Monitoring the workflow of A&E has been an area of interest in recent years, and the prediction of waiting times and service demands is a critical element in evaluating the services [2–4]. The nature of the A&E department is complex and multifactorial [1]. Several methods have been used to predict the waiting time in emergency departments; however, even though recent research has accomplished better prediction models, the need for more precise results is still present [3–12]. The current project implemented and compared two supervised tree-based ensemble machine learning algorithms, random forest and XGBoost [13,14]. Additionally, this implementation can help the institutions have time to implement reactive action to even the distribution of patients' burden at the A&Es' services based on the workload of the hospital. This estimation is useful for improved capacity planning.

2. Methodology

This study is a cross-sectional secondary data analysis utilizing information from three sources: Firstly, the principal dataset consisted of 68,321 anonymized observations from the A&E department at William Harvey Hospital between 2017 and 2019 (pre-pandemic). Secondly, air temperature and precipitation information were extracted from the southeast regional dataset of the Meteorological Office, the United Kingdom's

[1] Corresponding Author: Dr Holger Kunz; Email: h.kunz@ucl.ac.uk

national weather service. Thirdly, the southeast regional air quality index from the Department for Environment, Food, and Rural Affairs database.

For the predictions, XGBoost and Random Forest algorithms as regression methods were implemented to estimate the expected length of time in minutes. Both algorithms are tree-based ensembles and provide the option to create a feature ranking of the most important features as an additional outcome. Using a random split of the dataset into two components, 70% (47,825 observations) for the training dataset and the remaining 30% (20,496 observations) for the test dataset. A grid search method was used to adjust the hyperparameters in both algorithms. The evaluation of the models and comparison was performed utilizing the mean square errors, the root mean square errors, and the mean absolute errors. Lastly, the final XGBoost regression model was used to plot the features' importance. The project was implemented with Python 3.10, using NumPy, Pandas, Scikit-learn, SciPy, XGBoost, matplotlib and seaborn libraries.

3. Results

Table 1 summarizes the performance measures using different features for both the random forest and the XGBoost models. The best performance in the two cases was achieved using all the feature groups (clinical, sociodemographic, contextual, and workload features).

Table 1. Performance of the models

Set of features	Random forest			XGBoost		
	MAE	**RMSE**	**MSE**	**MAE**	**RMSE**	**MSE**
CL +SO + CO + WL	66.71	85.31	7278.41	64.31	82.66	6833.12
CL	68.76	87.39	7637.96	68.54	87.18	7601.21
CL + SO	67.82	86.40	7465.32	67.26	85.85	7370.11
CL + SO + CO	67.22	85.78	7358.56	65.48	83.93	7044.8

CL = clinical features, SO = Sociodemographic features, CO = Context features, WL = workload features

For the random forest model, the best performance achieved had an RMSE of 85.31, an MAE of 66.71, and an MSE of 7278.41. On the other hand, the best performance achieved by the XGBoost model had an RMSE of 82.66, an MAE of 64.31, and an MSE of 6833.12. The XGBoost model achieved 3.4 % less MAE, 3.11% lower RMSE, and 6.12% less MSE compared to the best random forest algorithm performance. Based on the model with the best performance (XGBoost including all the set of features) the most highly ranked features were identified, including but not limited to mean time to discharge in the previous 3 hours; current number of patients in the department; number of patients that waited more than 4 hours in the previous 24 hours; proportion of patients in the department who arrived by ambulance or helicopter.

4. Discussion

This project predicted the time to discharge in A&E using a random forest and an XGBoost algorithm. Contrary to previous studies, this forecasting model predicts the waiting time before the patient's registration in the hospital and not after the triage evaluation. This implies that the implementation and deployment of this model in the hospital environment could provide the patients with new information on the A&E

waiting time status. The data is based on information from a single health institution. This limits the external validity of the presented results since other institutions may have different protocols, services, and patient characteristics. The study data was fully anonymized and exempted from ethical approval as service evaluation. The model could be implemented in a mobile application to inform patients before they arrive at the hospital and how long they approximately are going to wait. The prediction helps to manage patients' expectations. Patients can inform family members or make other organizational preparations for this often-stressful period. Hospitals might have the advantage of an improved planning to deploy clinical staff and more efficient capacity planning.

5. Conclusions

The objective of this study was to predict waiting time in A&E before the patient arrived at the hospital using random forest and XGBoost regressions. The best performance was obtained with an XGBoost model achieving an RMSE of 82.66, an MAE of 64.31 and an MSE of 6833.12, outperforming around 4.2% of the evaluation measurements acquired by the random forest model. Compared to the 1.3% improved performance using Lasso regression compared to the random forest [6] in similar environments, the XGBoost model achieved slightly better results. Finally, the feature ranking for the best model showed that most of the important features for the model were related to the context of the attention and the workload on the department. A potential application would help patients to better estimate the waiting times in an ambulatory setting and manage expectations. The hospital could use the improved prediction of waiting times to fine-tune and improve the deployment of clinical staff and capacity planning.

References

[1] He J, Hou X, Toloo S, Patrick JR, Gerald GF, Demand for hospital emergency departments: a conceptual understanding, World J Emerg Med. 2 (2011) 253. doi:10.5847/wjem.j.1920- 8642.2011.04.002.

[2] Pianykh OS and Rosenthal DI, Can We Predict Patient Wait Time?, Journal of the American College of Radiology. 12 (2015) 1058–1066. doi:10.1016/j.jacr.2015.04.010.

[3] Sun Y, Teow KL, Heng BH, Ooi CK, Tay SY, Real-time prediction of waiting time in the emergency department, using quantile regression, Ann Emerg Med. 60 (2012) 299–308. doi:10.1016/j.annemergmed.2012.03.011.

[4] Ang E, Kwasnick S, Bayati M, Plambeck EL, Aratow M, Accurate ED Wait Time Prediction, Manufacturing & Service Operations Management. 18 (2016) 141–156. oi:https://doi.org/10.1287/msom.2015.0560.

[5] Hijry H, Olawoyin R, Predicting Patient Waiting Time in the Queue System Using Deep Learning Algorithms in the Emergency Room, International Journal of Industrial Engineering and Operations Management. 03 (2021) 33–45. doi:10.46254/j.ieom.20210103.

[6] Pak A, Gannon B, Staib A, Predicting waiting time to treatment for emergency department patients, International Journal of Medical Informatics. 145 (2021). doi:10.1016/j.ijmedinf.2020.104303.

[7] Walker K, Jiarpakdee J, Loupis A, Tantithamthavorn C, Joe K, Ben-Meir M, Akhlaghi H, Hutton J, Wang W, Stephenson M, Blecher G, Paul B, Sweeny A, Turhan B, Emergency medicine patient wait time multivariable prediction models: a multicentre derivation and validation study, Emergency Medicine Journal. (2021) emermed-2020-211000. doi:10.1136/emermed-2020- 211000.

[8] Suleiman M, Demirhan H, Boyd L, Girosi F, Aksakalli V, Bayesian prediction of emergency department wait time, Health Care Manag Sci. (2022). doi:10.1007/s10729-021-09581-1.

[9] Paling S, Lambert J, Clouting J, González-Esquerré J, Auterson T, Waiting times in emergency departments: Exploring the factors associated with longer patient waits for emergency care in England

using routinely collected daily data, Emergency Medicine Journal. 37 (2020) 781– 786. doi:10.1136/emermed-2019-208849.

[10] Eiset AH, Kirkegaard H, Erlandsen M, Crowding in the emergency department in the absence of boarding - A transition regression model to predict departures and waiting time, BMC Med Res Methodol. 19 (2019). doi:10.1186/s12874-019-0710-3.

[11] Kuo YH, Chan NB, Leung JMY, Meng H, So AMC, Tsoi KKF, and Graham CA, An Integrated Approach of Machine Learning and Systems Thinking for Waiting Time Prediction in an Emergency Department, Int J Med Inform. 139 (2020). doi:10.1016/j.ijmedinf.2020.104143.

[12] Ataman MG, Sarıyer G, Predicting waiting and treatment times in emergency departments using ordinal logistic regression models, American Journal of Emergency Medicine. 46 (2021) 45– 50. doi:10.1016/j.ajem.2021.02.061.

[13] Pak A, Gannon B, and Staib A, Predicting waiting time to treatment for emergency department patients, International Journal of Medical Informatics. 145 (2021). doi:10.1016/j.ijmedinf.2020.104303.

[14] Chen T, Guestrin C, XGBoost: A Scalable Tree Boosting System, (2016). doi:10.1145/2939672.2939785.