The Importance of Health Informatics in Public Health during a Pandemic
J. Mantas et al. (Eds.)

doi:10.3233/SHTI200491

Comparison of Machine Learning Algorithms for Classifying Adverse-Event Related 30-Day Hospital Readmissions: Potential Implications for Patient Safety

Antoine SAAB [a,b,1], Melody SAIKALI [b] and Jean-Baptiste LAMY [a]
[a] *LIMICS, Université Sorbonne Paris Nord, INSERM, UMR 1142, Bobigny, France.*
[b] *Quality and Patient Safety Department, Lebanese Hospital Geitaoui-UMC, Beirut, Lebanon*

Abstract. Studies in the last decade have focused on identifying patients at risk of readmission using predictive models, in an objective to decrease costs to the healthcare system. However, real-time models specifically identifying readmissions related to hospital adverse-events are still to be elaborated. A supervised learning approach was adopted using different machine learning algorithms based on features available directly from the hospital information system and on a validated dataset elaborated by a multidisciplinary expert consensus panel. Accuracy results upon testing were in line with comparable studies, and variable across algorithms, with the highest prediction given by Artificial Neuron Networks. Features importances relative to the prediction were identified, in order to provide better representation and interpretation of results. Such a model can pave the way to predictive models for readmissions related to patient harm, the establishment of a learning platform for clinical quality measurement and improvement, and in some cases for an improved clinical management of readmitted patients.

Keywords. Readmissions, adverse events, classification, patient safety, machine learning, artificial intelligence

1. Introduction

Nearly one in five patients is re-hospitalized within 30 days of discharge [1], incurring significant costs to the healthcare system [2]. Therefore, minimizing post-discharge adverse events has become a priority for many health care systems around the globe. Many studies in the last decade have focused on identifying patients at risk of readmission using predictive models [3][4]. However, few attempts have been made to identify potentially preventable readmissions [5], and, to the best of our knowledge, no studies have explored models to predict or identify readmissions related to hospital adverse-events. Moreover, study review highlights the need to use more standardized hospital information system (HIS) data related to the readmission (eg. biological, radiological, billing and administrative data) instead of institution-specific or clinical judgement based data, for an earlier and more benchmarkable prediction. The main objective of the study was to construct a model, based on routinely available data from the HIS, that could determine, on a near real time basis, if the patient readmission within

[1] Corresponding Author, Antoine Saab, Quality and Patient Safety Department, Lebanese Hospital Geitaoui-UMC, Beirut, Lebanon; E-mail: qualitedeveloppement@hopital-libanais.com.

30 days was associated with a hospital acquired adverse event that occurred in the previous admission (response variable).

2. Materials and Methods

The dataset used for training and testing of algorithms was built using the gold-standard approach, by a multidisciplinary consensus panel (internal medicine physician, radiologist, nurse, and patient safety professionals), expanded when needed to include physicians from specialized disciplines. The panel analyzed and classified 307 patient readmissions (within 30 days) extracted from the HIS from October 2019 till March 2020 (excluding readmissions of oncology patients and elective readmissions) that occurred in a 250 beds university hospital in Beirut, Lebanon where the study took place. On average, 30 min were required for every case preparation, and 25 min for the panel discussion and classification of each medical case. 46 of these cases were labeled as related to adverse events, to which 47 cases non related to adverse events were randomly chosen from the remaining dataset, to define a final balanced dataset of 93 cases, containing almost equal cases for each of the two classification classes: "Readmission related to adverse event", and "Readmissions not related to adverse events". 23 features (explanatory variables) were identified by the panel based on previous research results accomplished on this subject [6] and were extracted from the HIS. The features consist of both binary and continuous variables corresponding to the readmission and the previous admission cases. A supervised learning approach was adopted using different machine learning algorithms: Random Forests (RF), Decision Tree (DT), Boosting (BT), Artificial Neural Networks (ANN) and Logistic Regression (LR). Cross-validation of the model generated by each algorithm was performed using established methods (K-Folds=5 and Leave-One-Out) to avoid the results being influenced by the partitioning of the original dataset. Results of the different algorithms were reported and compared using model classification accuracy, based on the information entropy. Variable importance was determined by calculating the relative influence of each feature on the classification (except for ANN where this method is complex and not standard). This study was submitted to the hospital's institutional review board (IRB) and was exempt from further review since it does not directly involve human subjects.

3. Results

All algorithms showed good accuracy (>0.85) in the model training phase, which underlines their ability to fit data to a theoretical model using the proposed features. Table 1 shows the accuracy obtained from different algorithms when the predictions of the generated models were compared against the test dataset. Among the different algorithms, ANN showed to be the most predictive (0.88), followed respectively by LR (0.62), RF (0.62), DT (0.60), and BT (0.55). Table 2 shows the features' different weight importance, when this option was possible by the type of algorithm used, and highlights the features with significant weight (>5%).

Table 1. Accuracy result on evaluation for algorithms tested, with chosen algorithm parameters

Algorithm	**Mean accuracy**	**StdDev accuracy**	**Parameters** (algorithm-specific function arguments used for result optimization, as per SKLEARN and KERAS libraries definitions)
ANN	0.88	*0.07*	*24/6/4/1 architecture, optimizer RMSprop, lr=0.0001, epochs=10000, batch size=30*

LR	0.62	*0.10*	*penalty='l2', dual=False, tol=0.0001, C=1, fit_intercept=False, intercept_scaling=1, class_weight='balanced', solver='lbfgs', max_iter=30000*
RF	0.62	*0.06*	*n_estimators = 30, criterion="entropy", max_depth=8, min_samples_leaf=2, min_samples_split=4, max_leaf_nodes=15, bootstrap=False*
DT	0.60	*0.06*	*criterion='entropy', max_depth=8, min_samples_leaf=2, min_samples_split=4*
BT	0.55	*0.07*	*XGBClassifier, max_depth= 8, learning_rate = 0.001, gamma = 1, n_estimators=200*

Table 2. Feature importance by model (highlighted are those with mean importance ≥ 0.05)

Feature	RF	DT	BT	LR	Mean
Days since last discharge	0.15	0.15	0.25	0.01	0.14
White Blood Cells (WBC) value upon readmission	0.08	0.05	0.06	0	0.0475
Microbiological culture ordered in first 48h of readmission	0.03	0.1	0.08	0.07	0.07
C-Reactive Protein (CRP) value upon readmission	0.04	0	0.11	0	0.0375
Creatinine serum differential ratio	0.05	0.04	0.13	0.06	0.07
Potassium level differential	0.05	0	0.05	0	0.025
White Blood Cells (WBC) level differential	0.05	0.08	0.07	0	0.05
C-Reactive Protein (CRP) level differential	0.1	0.11	0	0	0.0525
Combined WBC-CRP levels	0	0	0	0.11	0.0275
Hemoglobin differential ratio	0.05	0.17	0.04	0.01	0.0675
Calcium level differential	0.03	0	0.06	0.01	0.025
Patient age	0.06	0.08	0	0.04	0.045
Patient sex	0.08	0	0.08	0	0.04
Paracetamol use on readmission	0	0	0	0.09	0.0225
Number of biological exams ordered in first 24h of readmission	0.07	0	0.04	0.01	0.03
Number of radiological exams ordered in first 24h of readmission	0.04	0	0.04	0.01	0.0225
Remarks in radiological requests ordered in first 24h of readmission	0.01	0	0	0.12	0.0325
Number of different types of medication during previous admission	0.06	0.06	0.05	0	0.0425
International Normalized Ratio (INR) raise	0.1	0.12	0	0.12	0.085
Number of surgical procedures performed in previous admission	0.04	0	0.11	0.04	0.0475
Cerebral CT-Scan performed on readmission	0	0	0	0.07	0.0175
Abdominal CT-Scan performed on readmission	0	0	0	0.06	0.015
Thoracic CT-Scan performed on readmission	0.03	0.06	0	0.11	0.05
Readmission type (Inpatient/ Outpatient)	0	0	0	0.08	0.02

4. Discussion and Conclusions

Predictive performance of the different algorithms: While accuracy levels across different algorithms seem to be moderate to low, they remain comparable to studies in the literature on 30-days readmissions [3][7]. However, the significance of the features adopted in this study versus other models in the literature is that they permit a near real-time computation of the classification as soon as the patient is readmitted, thus allowing the possibility for immediate proactive administrative and/or clinical interventions to reduce the risk of any preventable adverse event. A few factors could potentially explain and lead to improving this result. First, the sample size can be improved with additional expert time and resources. Another factor is inherent to the nature of the classification outcome variable itself which can only be built on expert opinion and thus contains, despite all methods used to lower the risk of judgement bias, a residual level of uncertainty. The high number of harm categories in the medical field [8], and patient-specific influencing factors that should be taken into consideration, induce a need for a high number of features to encompass this domain's complexity.

Feature importance and possible interpretations: In contrast with other similar studies, the features were chosen in this study to be all directly extractable from a basic HIS, and not needing human intervention for data aggregation or interpretation. This choice is in line with the need to standardize such tools and benchmark results across different healthcare systems. In the majority of tested models, the features impacting most of the results are: Days since last discharge, respective differential of CRP/WBC/INR/Creatinine serum, hemoglobin differential ratio, Thoracic CT-Scanner and Microbiological Cultures performed upon readmission. This result can be intuitively interpreted as a higher sensitivity towards detection of infections, hemorrhages and acute kidney injury cases. Interestingly, from the results obtained, the models insinuate also that readmissions occurring within 12 days of previous discharge are more prone to be associated with patient harm.

Potential practical implications for patient safety: Validating automated models for classifying 30-day readmissions can have some important implications for patient safety efforts in hospitals. Firstly, through correctly estimating the true level of nosocomial harm relative to readmissions, and using this information to analyze and improve clinical practices, hospitals will be able to measure the impact of deployed patient safety efforts over time. Secondly, permitting a proactive management of such cases as soon as they enter the hospital, can help prevent any further harm and address any implications that may arise. Finally, the results can pave the way to more proactive models that can predict risks of preventable readmissions due to adverse-events before patients are physically discharged from the hospital, or identify the patients "at risk" and follow up with them by phone before they actually return to the hospital, thus preventing extra costs for the healthcare system and third-party payers. Given the time and effort that was needed to construct a training and testing dataset on this domain, the validated sample size used in this study was relatively small. Also, some specific data (such as radiology reports, medical notes) could not be extracted from the HIS at this point of the system's integration. These limitations will be taken into consideration in future studies to improve outcomes.

References

[1] Jencks SF, Williams MV, Coleman EA, Rehospitalizations among patients in the medicare fee-for-service program, N. Engl. J. Med. (2009), doi:10.1056/NEJMsa0803563.

[2] Kjellberg J et al., Costs associated with adverse events among acute patients, BMC Health Serv. Res. (2017).

[3] Kansagara D, Englander H, Salanitro A, Kagen D, Theobald C, Freeman M, Kripalani S, Risk prediction models for hospital readmission: A systematic review, JAMA - J. Am. Med. Assoc. (2011).

[4] Artetxe A, Beristain A, Graña M, Predictive models for hospital readmission risk: A systematic review of methods, Comput. Methods Programs Biomed. (2018), doi:10.1016/j.cmpb.2018.06.006.

[5] Garcia-Arce A, Rico F, Zayas-Castro JL, Comparison of Machine Learning Algorithms for the Prediction of Preventable Hospital Readmissions, J. Healthc. Qual. (2018).

[6] Saikali M, A. Tanios, A. Saab, Evaluation of a Broad-Spectrum Partially Automated Adverse Event Surveillance System, J. Patient Saf. 1 (2017), doi:10.1097/PTS.0000000000000442.

[7] Turgeman L, May JH, A mixed-ensemble model for hospital readmission, Artif. Intell. Med. (2016), doi:10.1016/j.artmed.2016.08.005.

[8] Health Quality & Safety Commission, The Global Trigger Tool A Practical Implementation Guide for New Zealand District Health Boards, Health Quality & Safety Commission, Wellington 6146, 2012.