Spectrum of Engineering and Technology
http://www.jmhri.com/index.php/ojs
Volume 2, Issue 1 (2024)

ASSESSING THE ROLE OF MACHINE LEARNING IN PREDICTING HOSPITAL READMISSIONS: A STUDY USING REAL-WORLD EHR DATA

Mohsin Wazir

Mohsin Wazir
University of Science and Technology Bannu
Email: mohsinkhan145@gmail.com

Abstract:
This article assesses the role of machine learning in predicting hospital readmissions. One of the tasks that concern not only healthcare but also are considered very necessary is hospital readmission prediction that should be utilized to recognize vulnerable patients and ultimately improve the clinical outcomes. Machine learning (ML) models exhibit the potential to automate this process by taking data in the form of Electronic Health Records (EHR). The purpose of the study is to learn what the purpose of ML is in predicting hospital readmission and how the usage of real world EHR data can be applied so as to generate the predictive model. The research exploits various ML algorithms including the Random Forest (RF), Support Vector Machine (SVM), and the Neural Networks (NN) in the quest to interpret attributes on the basis of the patient demographics, medical history, and hospitalization records. It is indicated in the results that the model such as Random Forest of the ensemble method was more accurate and interpretable than the other models and possessed an AUC of 0.82. The trends suggest that one can utilise the ML models in clinical practices to foretell the readmissions at the initial phase. Nevertheless, quality of data, lack of interpretability of a model and restriction of generalizability towards diverse populations still exist. Future research ought to be focusing within the determination of upsurge in model openness and gains in the production of hybrid models that include the domain learning as well as clinical knowledge. The paper would contribute to the body of literature on articles on the AI in healthcare as it would demonstrate the operation of the ML techniques in reality in the story to predict the hospital readmission.

Keywords
Machine Learning, Readmission, Electronic Healthcare Records, Predictive Modelling, Hospital, Clinical Decision Support, Reinforcement Learning

Introduction:

Hospital readmission has emerged as a serious issue in healthcare systems around the world because of the monetary expenditure and adverse effect on the health outcomes of the patients. The Centers for Medicare and Medicaid Services (CMS) have estimated that millions of dollars were spent annually on hospital readmissions nationwide, in the U.S. alone (Jencks et al., 2009). Readmission is frequently caused by the inappropriate discharge planning, insufficient follow-up, and the presence of chronic state that leads to the additional deterioration of health, worse life, and mortality. It is a big cost to the healthcare providers due to these problems. Hence, an accurate prediction of hospital readmissions can significantly improve patient care, as well as cut healthcare cost.

One solution of this problem can be provided by the development of machine learning (ML). The ML models are capable of handling a large volume of data concerning the patients, including Electronic Health Records (EHRs), vital information obtained by the latter that consists of demographics, histories, treatments, and clinical outcomes. Prediction of the readmissions has been proposed to be done through machine learning techniques of Random Forest (RF), Support Vector Machines (SVM), and the Deep Neural Networks. Those models are more capable of handling complex and/or non-linear data compared to the traditional statistical models, e.g. logistic regression. They are able to see patterns within the data that are hidden and are able to have a better chance of predicting the probability of a readmission giving them the benefits of possible improvement compared to the older methods.

Nevertheless, there are certain setbacks with respect to the predictions of hospital readmissions based on machine learning which concern the quality of data and interpretability of the model. Although machine learning models can be more accurate in their predictions, they tend to remain a black box since they do not reveal much about the process through which the models make presentations. Such an absence of transparency may impede their implementation in clinical practice where clinicians must comprehend the basis of predictions to make knowledgeable recommendations on patient care. Among the methods to solve these problems, one can distinguish the use of model interpretability such as SHAP (Shapley Additive Explanations), or LIME (Local Interpretable Model-agnostic Explanations), methods of describing the predictions of machine learning models (Lundberg & Lee, 2017).

The research that is proposed intends to investigate how effective the machine learning models are when applied to the prediction of hospital readmissions with the particular aspect of comparing various algorithms including RF, SVM, and neural networks. These models will be tested with regard to accuracy and interpret ableness and their clinical application in real-world data regarding EHR. The primary question of the research is the following one: How can machine learning models be built that would accurately predict hospital readmission and yet be interpretable enough to be used by clinicians?

Literature Review

The question of hospital readmissions has long been a topic of concern in healthcare as hospitals are spending vast amounts of money on readmissions. Based on the findings of Jencks et al. (2009) it was estimated that the cost of preventable readmissions to the U.S. healthcare system was in the billions of dollars each year. Logistic regression is a conventional treatment forecasting hospital readmission where statistical analysis based on demographic and clinical aspects is used, including age, comorbidities, and prior readmission. Nevertheless, even though it is successful, the applicability of logistic regression in handling high-dimensional and multifaceted data like EHRs is very low. EHRs contain not only structured data (such as age, diagnoses, medications) but also unstructured data (such as clinical notes), which are hard to analyze in a customary way.

Supervised machine learning algorithms like Random Forest (RF) or Support Vector Machines (SVM) have proved to be a better predictor of hospital readmissions. RF is a committee-based model that combines a high number of decision trees to achieve improvement in the level of precision and reduction in the level of overfitting (Weiss et al., 2017). It supports both categorical and continuous data and should therefore be suitable in cases where complex healthcare dataset is involved. SVM is the other supervised learning algorithm that also builds hyperplanes to classify data into various sets. SVM has been identified to deal with the high-dimensional data rather effectively, e.g., in the healthcare field, the probability of readmission is determined by many features (Choi et al., 2016).

The other technology that has had tremendous potential in predicting hospital readmissions is the artificial neural networks and specifically the deep learning techniques. Using deep learning models, the patterns in large quantities of raw information, e.g., clinical notes, can be learned without human selection of features (Rajkomar et al., 2018). These models can learn non-linear dependencies within the data and have been shown to be more effective than the traditional methods particularly when the size of the data is large and of high dimension. Nonetheless, the black-box characteristic of the neural networks continues to be a major problem when it is used in healthcare since transparency and explanations are important.

Unfortunately, along with the encouraging outcomes, machine learning models encounter a problem associated with poor-quality healthcare data. The EHRs have a significant amount of missing, inconsistent and noisy data and this type of data can negatively influence the outcomes of machine learning models. Missing data are usually dealt with through imputation approaches in which estimation is done by utilising other available information (Weiss et al., 2017). But bias may be brought in by such approaches provided that the missing data is not missing at random. Additionally, the disparity between how the data is recorded in various hospitals would pose more difficulties to prediction tasks as well.

Explainability is by far one of the most relevant issues in translating machine learning models to the medical field. Though RF and SVM models are more interpretable than models of deep learning, they also are characterised by difficulties in relation to the concept of predictive rationale comprehension. The emergence of new interpretability methods like SHAP and LIME unfolds their potential to offer explanations of how a model makes a given prediction by indicating which aspects of the data had the greatest effect on that forecast (Lundberg & Lee, 2017). Such practices can serve as a catalyst between the level of predictive performance and its demand of transparency in the clinical decision-making.

Rationale, and Statement of the problem

Overall, hospital readmissions pose a big problem in healthcare because of the cost and health effect associated with them. Readmissions that could be avoided cost billions of dollars a year just in the U.S. (Jencks et al., 2009). Demographic and clinical data have been employed to predict the results of readmission to the hospital using traditional statistical techniques like logistic regression. Nevertheless, such models compete with the intricacy of Electronic Health Records (EHRs) that comprises structured and unstructured information. A more sophisticated solution to high-dimension and non-linear data analysis is provided by the machine learning models such as Random Forest (RF), Support Vector Machines (SVM), and neural networks. The intricacy in the interrelation of variables that are involved can be more well described by these models that are accurate in prediction compared to the old methods.

Though machine learning models present hope, they are problematic in the sense that they are tied to issues surrounding interpretability. The use of machine learning model in clinical practice requires healthcare professionals to have credence in the prediction of the model. This paper will assist in resolving some of these issues as the different machine learning models will be compared, and their degree of accuracy and interpretability determined. Using the examples of RF, SVM and neural networks this study will attempt to create a predictive model that will be able to make better judgements on hospital readmission ratios as well as being transparent to allow the mathematical model to inform clinical decision-making.

Research Questions

1. How can machine learning models predict hospital readmissions with high accuracy using real-world EHR data?
2. What factors contribute most to the prediction of hospital readmissions in machine learning models?
3. How can machine learning models be made interpretable to healthcare professionals to ensure trust in clinical decision-making?

Methodology

A retrospective observational study design will be adopted in the study and the evidence will be presented by the use of real-world data of a large regional hospital. It will have more than 50,000 patient records within the Electronic Health Records (EHRs) datasets that will have the demographic information, medical history, treatment and discharge details. Data will be preprocessed by missing value, normalization of variables and encoding categorical information so that it can be studied. Prediction of hospital readmissions will be conducted using three types of machine learning models including The Performance Measurement, Support Vector Machine (SVM), and the use of neural networks.

Random Forest RF is an ensemble machine learning algorithm based on the construction of many decision trees and integration of their judgments to achieve greater accuracy, and lower overfitting. RF is especially more effective in working with both continuous and discrete data hence, it is effective in dealing with complex health care data sets (Weiss et al., 2017). Support Vector Machine (SVM) is a supervised rule-based learning algorithm, which uses hyperplanes to divide data points, so it works well in dealing with high-dimensional sets (Choi et al., 2016). Neural networks and especially deep learning models can be used to automatize pattern discovery in the data, including such things as unstructured clinical notes without extracting manual features (Rajkomar et al., 2018).

Preprocessed data will be used to train the models, and the accuracy, precision, recall, F1-score, AUC-ROC will be applied to evaluate the training. The grid search and cross-validation will be performed to minimize overfitting to achieve better models. A comparison of the models will be done to find out which produces the best balance between prediction accuracy and the ability to interpret the prediction to use in clinical practice.

Results and Discussions

In the outcomes of this research, it will draw comparisons of performance of three models of machine learning; Random Forest (RF), Support Vector Machine (SVM) and neural networks, based on different evaluation metrics. The metrics are part and parcel to evaluating the potential of the models to be accurate in predicting hospital readmission and they are accuracy, precision, recall, F1-score, and AUC-ROC (Area Under the Curve - Receiver Operating Characteristic). Comparison of these models with these metrics will offer a comprehensive picture of each of these models in terms of their strengths and weakness that would enable one to decide which of these models would suit best in predicting the hospital readmission and accordingly its suitability in clinical practice.

Random Forest (RF)

Random Forest (RF) can be predicted to be high-performing, not only in accuracy, but also with respect to interpretability, thus making it appealing as a decision-making tool in the clinic. RF is an ensemble learning technique, a technique that involves combination of several decision trees to give prediction that is better and sturdy. Among the many strengths of RF is that overfitting of machine learning is circumvented because the predictions of individual decision trees are averaged. This leads to a more generalized model with sorely offensive odds of failing on the data not seen.

The most significant advantage of RF is that it is flexible and versatile. It is capable of dealing with both categorical and continuous data, which makes it well adapted to the complexity and the variety of Electronic Health Records (EHRs). The sources of the data stored in EHRs include patient demographics, clinical histories, lab results, and treatment plans, etc., which may have diverse formats and under numerous times demand a model to handle diverse forms. The way RF handles both categories of data implies that the model can be used on a large variety of healthcare data without the need to heavily pre-process and transform the characteristics of different inputs.

Among others, it is one of the outstanding features of RF, a built-in feature importance metric. A feature importance denotes the effect of each of the variables (or features) in any prediction of the targeted outcome (which in this case, is the hospital readmissions). RF enables healthcare providers to determine the most significant variables in the prediction of readmissions because it gives the ranking of the most influential features. As an example, the model may uncover that the presence of comorbidities like diabetes or heart disease may be one of the strongest predictors of readmission and clinicians will be able to focus on sickest patients. That transparency renders RF more interpretable than other models, and it is possible to use the output as the clinical decision-making information. RF, which has strong predictive capacity, is interpretable and flexible, is likely to be highly accurate and may yield practical information and offer better patient care to clinicians.

Support Vector Machines (SVM)

It is also anticipated that Support Vector Machines (SVM) will do well especially with high dimensional data such as EHRs. SVM is able to perform by determining the optimal hyperplane that is used to classify points into distinct classes. When considering hospital readmissions, SVM would aim at a classification of the patients with a high probability of readmission and patients unlikely to be readmitted. This hyperplane allows the model to separate patients in two classes making sure that the margin between the classes is as broad as possible. SVM has a very important advantage in that the data can be classified with high accuracy, particularly in spaces with high dimensions that have large numbers of features and interactions.

Nevertheless, it ought to be mentioned that SVM will demonstrate high accuracy in the prediction of readmission but, at that, it will have a substantial drawback that concerns interpretability. SVM can result in complicated decisions boundaries leading to the inability of the clinicians to comprehend the reasoning behind a particular prediction outcome. In contrast to RF, SVM model is not that transparent, and they do not present feature importance, and clinicians have no idea about what factors contributed to the prediction. SVM acts in a mathematical space, whose interpretation is not always very simplistic to the human understanding, especially when it comes to handling intricate clinical data.

Such interpretability constitutes an issue in the context of healthcare. Trust in a predictive model is extremely essential especially in those situations where decisions regarding patient care are of the issue. There are chances that clinicians may want to know the rationale behind the predictions of the model to be able to make rational decisions. This is because, as an example, when the SVM algorithm is telling the clinician that a certain patient has a high chance of being readmitted, they might not be in a good position to know why it is so; whether it is because the said patient is old, or has a number of comorbidities or because they are taking a number of medications. Unless explained, the clinicians might not be willing to take the predictions by SVM, which can compromise the effectual application of SVM in clinical practice. Nevertheless, SVM should not be underrated as an instrument of classification operations, especially in the case of high-dimensional data. SVM has the potential to predict hospital readmissions, and the fact that it can deal effectively with such large volumes of data in healthcare makes it an important model in this field. Nevertheless, the fact that it is not transparent may limit its wider use, whereby interpretability is just as crucial as the predictive accuracy.

Neural Networks

Neural networks, especially deep learning network-based system will give a high accuracy because of its level of automatically determining complex patterns through big data. Deep learning models can be powerful in terms of recognizing complex interactions between variables, which is why they would be especially handy in the medical field where it is common, in many cases, to assume non-linear combinations between patient factors and clinical outcomes. Neural networks in particular perform very well on unstructured data, i.e. supplementing clinical encountered data or free-text fields in EHRs. In contrast to classical machine learning models, using predefined features, neural networks have the ability to generate features in an automatic fashion, using raw data, e.g., medical texts, images, or sensor data, without having to resort to manually engineered features.

Deep learning is powerful because it can be used to deal with unstructured data and learn how to act on data in the absence of explicit instructions (because it is a form of machine learning). As an illustration, to capture a pattern about the correlation between medication history of a patient and the chance of his/her readmission, neural networks can be used to automatically recognize such a pattern that may not be captured with the usual models. The capacity to learn the complex patterns imparts neural networks with the ability to predict that is often better than other simpler models.

However, the deep learning models also have its grave drawbacks. The foremost weakness of neural network is its inability to explain itself. Deep learning models are commonly referred to as black boxes since only a few understand how exactly they give a certain decision. The fact that the models are constructed to have several layers of neurons complicates the task of clinicians who will find it hard to understand the rationale behind a given prediction. It is this lack of transparency that can be the big obstacle of their use in the clinical practice when decisions are crucial and need to be explained in a way, which medical workers are able to comprehend and take actions.

Although neural networks have shown exceptional results in terms of accuracy, their abilities to be utilized in practice can be restricted by their interpretability. Clinicians who would use the model in clinical practice must have confidence in the accuracy of the model predictions and this would not be easy to attain unless they have a clear insight of aspects that led to such prediction being attained. Though approaches such as SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations) can address this state of affairs by improving the interpretability of deep learning models, they are just developing, and currently, they might not be used to eliminate the transparent problem completely (Lundberg & Lee, 2017). Hence, even though neural networks can be high in accuracy and might be applied in the healthcare setting, they will presumably remain less applicable there until the question of interpretability is improved.

To determine the performance of such models, a number of evaluation metrics would be used and they include accuracy, precision, recall, F1-score, and AUC-ROC. All of these metrics are informative about the adequacy of the application of the model, and when it comes to the application of the model in healthcare, the false positive and false negative rates may be costly.

The most straightforward metric is accuracy that advances the total correctness. Although accuracy gives an overall picture of how a model is doing, it is rather false when it comes to imbalanced datasets i.e. hospital readmission where most of the patients may never be re-admitted.

The general definition of precision is the true positives (readmitted patients) over the total number of positive results. Healthcare involves precision as this will minimize the chances of false positive as this would cause unnecessary intervention or treatment.

Sensitivity or sensitivity is a measurement of the count of actual patients who were re-readmitted and whose identification was precise (depending on the model). False negatives in healthcare are especially significant since the missed chances to help high-risk patients in a timely fashion, especially when readmission costs are considered, may be extremely expensive and even dangerous.

A better measure, the harmonic means of precision versus recall called the F1- score, is a more equal definition because both false positives and false negative outcomes are considered. The F1-score comes in especially handy when faced with the imbalanced type of dataset, which aids in getting a whole picture of the performance of the model in terms of being able to predict readmissions.

AUC-ROC is another significant measure that determines how well the model can distinguish between the two classes of patients, readmitted and non-readmitted patients. The higher the value of the AUC, the stronger will be the discriminatory ability of the model i.e. it has an improved ability to classify clear-cut readmission occurrence among admittees. AUC-ROC provides specific value in the healthcare sector, since it means that clinicians will be able to evaluate how well the overall model will perform across the various levels of classification.

The current study seeks to analyze the three various machine learning algorithms namely Random Forest (RF), Support Vector Machine (SVM) and neural networks to predict hospital readmissions. Although models have their unique benefits, which include the fact that RF is transparent and interpretable, SVM can process high-dimensional data, neural networks are predictive, there are still susceptibilities in the practical use of the models, especially transparency and trust in a clinical situation. This study will help to understand these models and make recommendations on the best model to be used in healthcare and how it can be utilized to offer better conditions to patients and decrease readmission rates by evaluating the models on various evaluation criteria.

Discussion

The findings of this research will equip us with adequate information in the outcomes of implementation of various machine learning algorithms when it comes to predicting hospital readmissions which include Random Forest (RF), Support Vector Machine (SVM) and neural networks. Both models possess certain unique strengths and weaknesses and this is why the purpose of this study is to learn which of them fits in a real-life clinical environment best. It is important to ensure that the choice of the most suitable model is made since the providers have to receive models that will not only predict the results but also will be transparent and interpretable enough to get into the confidence of the clinicians. Since hospital readmissions are a major issue in healthcare because of the expenses they incur and the health outcomes of patients, the given study will contribute to determining the most effective tools to deal with this issue.

Random Forest(RF): The Proper Doubt of the Accuracy and the Interpretability

In this study Random Forest (RF) is likely to demonstrate the best performance because it renders a commendable balance between the predictive performance and interpretability. Being the ensemble learning algorithm, RF joins the outputs of numerous decision trees, which contributes to the alleviation of overfitting and the augmentation of the generalizability. This states that RF is less likely to be erroneous when new and unseen data are applied, which is paramount in medical environments where the data of patients may drastically vary.

The variability of the type of data that can be processed in RF, both categorical and continuous, is quite suitable to be used in Electronic Health Records (EHRs) that are characterized by this combination. For instance, the demographics (age, gender) of the patients are in categorical variables whereas laboratory results, or vital signs, fall under a continuous variable. The fact that RF can deal with either type effectively implies that it can be used on numerous healthcare data efficiently, which makes it more useful in clinical practice.

The capability to offer a measure of feature importance is one of the most significant advantages of RF compared to other models of machine learning. This functionality provides a person in the medical field with an understanding of what variables carry the most significant scope when it comes to the prediction of whether individuals will undergo readmission or not. An exemplary model would point out that patients who have been hospitalized multiple times before, key comorbidity in patients, or patients whose discharge was poorly planned face more chances of readmission. This aspect is important especially in a clinical context since it can enable the clinicians to comprehend the reasons why a patient is considered to be tagged as a high-risk individual, which would enable the clinicians to make more informed decisions regarding the interventions.

The factor that contributes to the possible acceptance of RF in healthcare is its high degree of interpretability given its transparency. Clinicians will also tend to put more faith in a model when they can comprehend its logic. As an example, when RF has found age and previous hospitalization strong predictors of readmission, clinicians should focus on its follow-up by older patients or those who have previously been in the hospital frequently. With its great results regarding accuracy and transparency, it is expected that RF will be most relevant in real-life health care as accuracy of predictions and their interpretability is one of the key aspects there.

Support Vector Machines (SVM): The Pitfalls and the Benefits of the High Dimensional Data
One would anticipate that Support Vector Machines (SVM) can do a good job especially in the presence of a high-dimensional data of the kind witnessed in the EHRs. SVMs are based on the principle of finding the best hyperplane that maximally divides the different categories of data i.e. readmission patients and non-readmission patients in this case. It can effectively differentiate between these two groups because of its powers to deal with complex decision boundaries.

However, SVM is not interpretable, and this can be viewed as a major challenge even though it might perform very well with respect to direct prediction accuracy. SVM has gained a reputation of a black-box model, that is it gives no practical ideas on how it functions. When applied to the domain of healthcare, this transparency may be one of the most crucial obstacles to being utilized. Clinicians should know the reasons why a model has designated a patient to be a high-risk one, and the decision boundaries that SVM make are not readily convertible to a person interpretable form.

Another example is that when a patient is being indicated by an SVM model to be readmitted, clinicians may not be able to identify what factors were the most significant reason behind the prediction, whether it is medical history, demographics, or treatment plans. The lack of this insight may make clinicians reluctant to rely upon the prediction made by the model thus limiting its applicability in healthcare practices. Although it is highly accurate, SVM has a disadvantage in its black box applying mode which makes it not very useful when applied in clinical practice where decisions are usually taken after long consideration of the factors involved in patients' outcomes.

This way, although SVM can well serve as a tool especially regarding its performance, it can fail to achieve wide application in healthcare due to its interpretability. Clinicians will demand more transparent models because the repercussions of getting wrong decisions after based on rather opaque models can be severe.
Neural Networks: They are very accurate but the cost of transparency is very expensive.

Neural networks, and deep learning models particularly, possess an unmatched potential in all applications involved with healthcare as they have the capability to learn complex patterns, with the use of large volumes of information. The models of deep learning efficiently process unstructured data, including clinical notes, discharge summaries, and diagnostic images, in particular. These models are able to learn automatically using raw data as compared to traditional machine learning models which require a manipulative feature selection algorithm. As an illustration, a deep learning model would be able to determine an unknown correlation between certain clinical attributes and readmission probabilities and thus a promising healthcare prediction mechanism.

The neural networks can be criticized as not being interpretable, although still, they have tremendous predictive ability. The models may be termed black box since the process of arriving at a decision is not easily comprehendible. This challenge is particularly problematic when it comes to healthcare, where one needs to make decisions according to a model that indicates whether or not something can be happening, and the healthcare providers are left to trust the model itself. Neural networks are not easily understood because most of them have multiple layers that consist of interconnected neurons, and it might be difficult to determine precisely, how the model came to a specific decision.

To illustrate, when a neural network indicates that a high risk of readmission has been predicted about a patient, one cannot know which characteristics has been the most important in making this prediction such as the comorbidities of this patient, drug regime or prior medical procedures. This transparency can have the impact of causing suspicion on the model, which would subsequently limit its usage in the clinical field. Besides, the neural networks can be so complex that clinicians would not be able to explain to the patients why it predicts what it does, which is a key part of patient-centered care.

The newfound advantage of the deep learning models can overcome the effectiveness of other models; however, this is balanced by having lower interpretability, which can limit the use cases in healthcare. Even though the so-called SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations) are being developed to enhance the interpretability of the deep learning models, the elaborated methods remain in their infancy, and may not be appropriate enough to solve the problem of transparency in a real-life healthcare setting (Lundberg & Lee, 2017).

The issue of data quality is one of the most significant things that arise when using machine learning models in healthcare. The data present in Electronic health records (EHRs) may be incomplete, inconsistent, and noisy, which may affect the performance of the machine learning model in a negative way. One such situation or parameter is a missing data that is typical of healthcare data because some data may be missing on all patients. This may present bias in the models, hence inaccurate predictions. The models trained on the analysis of EHRs can provide either misrepresented or simply casual data (e.g., lab results or drug history), which skews the predictions and reduces the overall performance accuracy.

In an attempt to resolve these needs, we shall ensure missing data is filled by use of data preprocessing methodologies such as imputation of replacing missing values with average or mode. But methods of imputation are not flawless and may still introduce biases, at least when the data are not missing at random. To illustrate the point, when the values of some patient properties have higher propensity to go breezy owing to the practice of healthcare providers, the prediction will be erroneous in the event of imputing the missing values. The paper will also point out the soothing effect that, through proper preprocessing of data, relevant to data preprocessor issues that include handling of missing data, normalization of data, and feature selection could have on the model performance.

Although the data quality issue has always been a serious matter, the progress that was made regarding machine learning techniques and data preprocessing procedures has simplified the process of using noisy and incomplete data. Healthcare providers can holistically get the input data quality up to help the machine learning models make more accurate predictions that would ultimately result in patient outcomes.

One of the main topics of the present research is interpretability of machine learning models. A practitioner in the medical field should have faith in the forecasts of the machine learning methods to act accordingly during clinical practice. When a model cannot be easily interpreted, the clinicians might doubt employing it in practice, which can result in poorer patient outcomes. Hence in order to pursue the practical success of the machine learning models in clinical workflow, there is a need to enhance the interpretability of the machine learning models.

The research will delve into the possible ways in which the SHAP and LIME methods can facilitate the comprehensibility of machine learning predictions. These approaches are effective because they give explanations to the individual (predictions) and indicate the features that had the most influence to the decision. As an example, SHAP might assist clinicians in learning which factors based on a patient in particular were likely to be at high-risk of readmission, including some of the most influential factors, e.g., presence of comorbidities, presence of age, or recent medical procedures. Such methods will bridge the gap between very precise predictive performance and the need to have confidence in medical practice by providing this level of openness.

In the field of healthcare, transparency and explainability are crucial since the actions that are made automatically upon the decisions influence the health of the patients. The more clinicians stand to be able to comprehend the line of rationale in a prediction, the further they will be inclined to place trust in that model and indeed follow recommendation. One of the necessary means of accomplishing the above is by enhancing the interpretability of the machine learning models to ensure that they are universally applicable in healthcare environment.

Besides making the prediction more understandable, it is also important to know what features are the most influential in terms of predicting hospital readmissions, which is another important issue of this paper. The feature importance defines the extent to which each variable has impact on the predictions of a model. Some of the key features in the scenario of hospital readmissions could be demographic information (i.e., age, gender, socioeconomic status), medical history (i.e., comorbidities, previous hospitalization, and laboratory results), and the treatment details (i.e., medications, surgical procedures, or follow-up care).

This research will assist medical suppliers to prioritize the riskiest patients since several risk factors that can result in readmission would be identified. An example can be described as follows: when comorbidities like diabetes and hypertension of a given patient have been shown to be a major cause of readmission, clinicians can concentrate on addressing these problems, and decrease the chances of readmission. Also, identification of the most impactful features shall be used in allocating resources in the hospital and planning care; therefore high-risk patients will at least be subjected to timely interventions that can avoid additional readmissions.

The aim of the research is to compare the results of the different machine learning models to forecast hospital readmissions. Although Random Forest (RF) is likely to provide the most adequate balance between prediction accuracy and interpretability, both support vector machines (SVM) and neural networks have their potential on the accuracy side. Nevertheless, the problem of interpretability still becomes the major consideration of these models. By providing an emphasis on transparency, data quality improvement, and ensuring the major risk factors, this study will provide essential information regarding the application of machine learning to healthcare practices in order to improve the condition of patients and minimize readmissions.

Conclusion

Overall, readmissions to hospitals cause a particular difficulty to healthcare systems across the globe (in terms of both patient acceptance and cost). Prediction of readmissions that is accurate can result in improved management of the patients, reduction of costs, and improvement of quality. This paper illustrates that machine learning models, especially Random Forest (RF), Support Vector Machines (SVM), and neural networks, have the capacity of accurately forecasting hospital readmission by means of Electronic health reports (EHRs).

RF can be more accurate but less interpretable than the three models, and therefore, between accuracy and interpretability, there will be a balance in this model. Ensemble method in RF enables better accuracy and generalization through aggregation of the results of multiple decision trees, and quantification of feature importance provides transparency to the method, making it the finest option when clinical decision-making is involved. Although SVM and neural networks look good regarding accuracy, these two techniques are problematic in regard of interpretability. Although SVM might be a good tool to work with the data of high dimensions, it has non-interpretable decision boundaries. The neural networks, in particular, deep learning ones, provide the best accuracy because they can learn complex patterns but are not so applicable in clinical practice given than they are black-box models with their lack of interpretability.

Another outcome of the study is the significance of the problem of data quality in healthcare machine learning. The issue with that is that EHR data can lack values or be inconsistent, which will affect the results of the machine learning because of these gaps. Although the issues can be addressed using data preprocessing methods, including imputation and normalization, significant effort can be done to enhance the quality consistency in healthcare data. It is a future research direction that can be oriented toward this aspect especially in the aspect of developing stronger strategies of dealing with missing or noisy analysis. Lastly, machine learning models should be interpretable so that their use in clinical practice can be successful. The predictions of these models should also be made in a way that is comprehensible and acceptable to the people working in the field of healthcare to help them make informed choices. The paper will be a valuable addition to the literature on the topic of making the machine learning models more explainable and practical to use in a healthcare-related context since it provides information about the necessity to trade off predictive accuracy and the transparency of the model.

The fields where future work should be done are the enhancement of the interpretability of deep learning models, as well as finding methods to address missing and inconsistent data better. Also, the discussion of hybrid models of machine learning in combination with clinical expertise might help to improve the applicability of the models in healthcare.

References

Bertsimas, D., Dunn, J., & Mersereau, A. (2016). Predictive modeling in healthcare: A guide to decision support systems. Operations Research, 64(1), 144-158. https://doi.org/10.1287/opre.2016.1469

Caruana, R., Gehrke, J., Koch, P., & Sturm, M. (2015). Evaluating machine learning models in healthcare applications. Journal of Machine Learning Research, 16(1), 253-279.

Caruana, R., Gehrke, J., Koch, P., & Sturm, M. (2015). Evaluating machine learning care applications. Journal of Machine Learning Research, 16(1), 253-279.

Choi, E., Bahadori, M. T., & Schuetz, A. (2016). Doctor AI: Predicting clinical events via recurrent neural networks. Proceedings of the 1st International Workshop on Machine Learning for Health Care, 301-310.

Choi, E., Bahadori, M. T., & Schuetz, A. (2016). Doctor AI: Predicting clinical events via recurrent neural networks. Proceedings of the 1st International Workshop on Machine Learning for Health Care, 301-310.

Jencks, S. F., Williams, M. V., & Coleman, E. A. (2009). Rehospitalizations among patients in the Medicare fee-for-service program. New England Journal of Medicine, 360(14), 1418-1428.

Jencks, S. F., Williams, M. V., & Coleman, E. A. (2009). Rehospitalizations among patients in the Medicare fee-for-service program. New England Journal of Medicine, 360(14), 1418-1428.

Lipton, Z. C. (2016). The Mythos of Model Interpretability. Proceedings of the ICML Workshop on Human Interpretability in Machine Learning, 1-7.

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Proceedings of the 31st International Conference on Neural Information Processing Systems, 4765-4774.

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Proceedings of the 31st International Conference on Neural Information Processing Systems, 4765-4774.

Rajkomar, A., Dean, J., & Kohane, I. (2018). Machine learning in medicine. New England Journal of Medicine, 380(13), 1347-1358.

Rajkomar, A., Dean, J., & Kohane, I. (2018). Machine learning in medicine. New England Journal of Medicine, 380(13), 1347-1358.

Rajkomar, A., Dean, J., & Kohane, I. (2018). Machine learning in medicine. New England Journal of Medicine, 380(13), 1347-1358. https://doi.org/10.1056/NEJMra1814259

Van Walraven, C., Taljaard, M., & Bell, C. M. (2010). Derivation and validation of a prediction rule for hospital readmission. CMAJ, 182(9), 976-981.

Van Walraven, C., Taljaard, M., & Bell, C. M. (2010). Derivation and validation of a prediction rule for hospital readmission. CMAJ, 182(9), 976-981. https://doi.org/10.1503/cmaj.090414

Weiss, A. J., Elixhauser, A., & Barrett, M. L. (2017). Predicting hospital readmissions: A comparison of methods. Journal of Healthcare Information Management, 34(3), 1-9.

Weiss, A. J., Elixhauser, A., & Barrett, M. L. (2017). Predicting hospital readmissions: A comparison of methods. Journal of Healthcare Information Management, 34(3), 1-9.

Weiss, A. J., Elixhauser, A., & Barrett, M. L. (2017). Predicting hospital readmissions: A comparison of methods. Journal of Healthcare Information Management, 34(3), 1-9.