ABSTRACT:

Delirium is a common manifestation of severe acute neuropsychiatric dysfunction prevalent in hospital settings, which due to the complex multi-factorial causes is often under-diagnosed and neglected. Early detection is a critical concern that can be addressed by ML techniques. As such, some methods to improve the accuracy of ML classification models for the detection of delirium are covered in this document. The aim is to develop and validate a tool for use in a hospital setting to accurately identify delirium during the admission of a patient.

A database collected at a Portuguese hospital between 2014 and 2016 was used to conduct this experimental research. Available data comprised 511 records and 124 variables, including patient demographics, medications administered, admission category, urgent admission, hospitalization period, history of alcohol abuse and laboratory results.

The methodologies used included data pre-processing, data imbalance processing, feature selection, train and test model with different ML classifiers, evaluating model performance and development of a Python web-based application.

The model that was achieved consists of 26 predictors assessed during admission to a healthcare facility. Overall, the best outcomes were obtained by combining the SFM method with the LR algorithm, with an AUC-ROC result of 0.833 and an AUC-PR of 0.582. Although the prediction model can be enhanced, this approach could be a useful support tool to identify patients at increased risk for delirium in healthcare settings.

INTRODUCTION

Over the past few years, the average life expectancy has increased. But along with the ageing population, chronic diseases have increased significantly~\cite{ref\_dugoff}. These are characterized by high levels of disability and are often responsible for causing pressure on the healthcare system. The biological ageing which an individual is exposed to has a variable course. In this natural degradation, many alterations can occur in the normal functioning of the body. Emphasizing here the deterioration that occurs at the cognitive level, which is a very common impairment in the elderly. A slight impairment in the person caused by a change in the condition of the disease may lead to disoriented states. Delirium is a frequent manifestation of severe acute neuropsychiatric dysfunction, and is widespread in hospital settings. Due to its clinical presentation and variability, it is often under-diagnosed and neglected.

Delirium can affect people of all ages, but predominantly affects hospitalized older adults. This geriatric syndrome leads not only to an increase in morbidity and mortality, but also to an increase in the period of hospitalization and a deterioration of the physical and mental state of the individual. The ability to assess delirium is an essential component of the patient assessment strategy for the prevention or treatment of delirium. Some studies suggest that early diagnosis and an appropriate approach are associated with decreased rates of morbidity and mortality associated with delirium.

As such, methods to improve the accuracy of ML classification models for the detection of delirium are covered in this document. The aim is to develop and validate a tool that will accurately identify delirium during the admission of patients in a hospital setting.

Delirium can affect people of all ages, but predominantly affects hospitalized older adults. This syndrome leads not only to an increase in morbidity and mortality, but also to an increase in the period of hospitalization and a deterioration of the physical state of the person.

RELATED WORK:

The rapid development of AI has increasingly allowed the application of ML to fulfil the needs of human life. A study conducted in 2020 by Vellido indicates that the current conjecture of technological development has triggered the idea that the use of ML would be the way forward to solve health-related problems, as well as being an asset to improve the quality of health services.

According to Liang et al., the PRE-DELIRIC has a high predictive value and the authors suggested that this model be adopted in ICU to detect delirium in high-risk patients, as it contributes to improving the management of resources, as well as improving patients lives

In 2015, an alternative model for early detection of delirium in intensive care, called E-PRE-DELIRIC, was validated. The E-PRE-DELIRIC model uses the data available at ICU admission to predict the development of delirium during the patient's hospital stay. It is composed of nine predictors: age, history of cognitive impairment, history of alcohol abuse, blood urea nitrogen, admission category, urgent admission, mean arterial blood pressure, use of corticosteroids, and respiratory failure.

This study emerged as a need to fill the gap in the previous model, where predictors are required from the first four hours after admission to the ICU. Data from 29 patients was included. The AUC obtained was 0.7 in the development dataset and 0.6 in the validation dataset. An AUC-ROC of 0.70 (0.67 to 0.74) was obtained in the case of delirium that developed after two days and 0.81 (0.78 to 0.84) for delirium that developed after six days.

In order to understand which of the two models would be better prepared for clinical use, the study "..." was conducted in the year 2017. This study concluded that model 1 predicts delirium better, however, ICU physicians prefer model 2 because is more convenient to use than model 1.

In 2018, a study was conducted to evaluate the prediction of delirium using the RF algorithm.

Materials and METHODS

DATA

This study used data from a Portuguese hospital that was extracted between 2014 and 2016.

Machine learning algorithms

**Random Forest**

The RF classifier consists of randomly selected features or a combination of features at each node to grow a tree.

Imbalanced Data:

In this study, due to the small amount of data, an oversampling technique was considered, namely Adaptive Synthetic Sampling (ADASYN). The main idea behind the ADASYN algorithm is to use a systematic method to create an adaptive way different amounts of synthetic data according to their distributions-\cite{ref\_vluymans}. With this approach, new samples of the minority class were synthetically created and a balance of the proportion of the categories was achieved.

In general, the most appropriate performance measure is dependent on the intended application of the classifier.

RESULTS:

The variables selected by model 8 and the respective coefficients are shown in Table 1. The results show that the coefficient values vary between negative and positive scores. This means coefficients with positive values indicate the variable predicts class delirium, on the other hand, coefficients with negative values determine the absence of delirium.

Based on these statistics, we may consider that this variable may effectively have a significant importance on delirium prediction, since the probability of getting a case of delirium when any of these drugs is present is high.

The variables most contributive to the prediction of delirium are age, antipsychotics, and other medications. The variable related to other drugs present includes drugs such Ranitidine and Throspium.

WEB APPLICATION:

To facilitate the use of this predictive model, a web application was built. To facilitate the use of this predictive model, a web application was built. This type of format allows easy viewing of the prediction results in real time and on any platform. The purpose of this tool is to easily and quickly calculate the probability of occurrence or absence of delirium in individuals admitted to healthcare settings.

The variables asked on the application form are: patient origin, pH, pCo2, etc.

CONCLUSION:

This study demonstrated the importance of data preparation in an ML project to obtain accurate predictions.

The model developed allows the prediction of delirium with an AUC-ROC of 0,833 and an AUC-PR of 0,582. The application developed might be a useful support tool for early delirium detection in healthcare settings. In addition, it is simple to use and adjustable to different types of devices. Although the model obtained has a reasonable predictive capacity, it is postulated that if the set of data collected contained a higher number of records, the predictive model could get a higher hit rate. This assumption is based on the fact that the model was constructed with a reduced number of records and a considerable number of variables. As a consequence, ML algorithms may encounter difficulties to detect patterns.

Finally, we conclude that although there is still room for improvement, this predictive model may be an asset in diagnosis of delirium in hospital settings.

In a future work perspective, another predictive ML model can be built using other classifiers and feature selection methods.

Regarding the application, it would be interesting to integrate this type of predictive model into electronic record of the hospital system and carry out this process automatically in real-time. This approach could reduce the number of cases of delirium that go unnoticed by health professionals, allowing an implementation of a delirium prevention program. Overall, the quality of service could be improved, thus improving the quality of life of hospitalised individuals.

APLICATION:

It is therefore essential that all hospital staff be aware of the possibility of delirium developing, and that prompt assessment and appropriate management are ensured.