

CSAC224-MachineLearning

Blood Transfusion Prediction Using Machine Learning

**Jesrel Pizzaro**

College of Computer Studies CamarinesSurPolytechnicColleges

\*\*\*@my.cspc.edu.ph

# Abstract

The motivation of this study is describe to face the challenge of solving unsolved the problem and to get the intellectual joy of doing this paper. The importance of transfusion of blood will also be one way to save a person’s life, so it essential to monitor the blood availability based on its type to avoid an inadequate supply of blood. My objective in this study is measure the level of knowledge regarding the blood transfusion. In the methodology it is predictive modeling Techniques investigated the study of Random forest, XG Boost, a Voting Classifier, and individual ML models Logistic Regression, SVM, Decision Tree, KNN, and Gradient Boosting. The model's performance is evaluated using evaluation measures such confusion matrix, precision, recall, F1-score, accuracy, and classification report after the dataset has been preprocessed by scaling features. In order to integrate the advantages of several models, ensemble techniques like Voting Classifier and Stacking Classifier are also used. With the Stacking Classifier obtaining an accuracy of 82.22% on the test set, the results show excellent performance and demonstrate the value of integrating various ML techniques.

# Introduction

The critical issue in Healthcare by predicting an individual is likely to donate the blood. The blood donation plays a crucial role in ensuring a stable and sufficient blood supply for medical treatments, emergencies, and various health conditions. Predicting donor behavior can aid blood centers in developing targeted strategies to encourage and optimize blood donation campaigns. The problem statement revolves around the classification of the individuals based their decency, frequency, monetary contribution, and the time since their last blood donation on March 2007. The binary output represents whether an individual donated blood (1 for donation, 0 for no donation. This classification task is vital for blood centers to identify potential donors and tailor outreach efforts effectively. My motivation for this project stems from significance of maintaining an ample and consistent blood supply, which is crucial for saving lives and providing timely medical interventions. Blood donation patterns are influenced by various factors, and predicting resources efficiently, minimizing shortages, and planning targeted awareness campaigns. The blood donation patterns are influenced by diverse factors, including personal preferences, awareness, and societal consideration. Understanding and predicting these patterns can helps blood centers create proactive strategies to engage donors and ensure a sustainable blood supply. Machine learning algorithms and ensemble methods provide a promising approach to analyze and model these complex relationships. The input to our algorithms consists of historical data related to an individual’s blood donation behavior, while the output is a binary prediction indicating the likelihood of the person donating blood. This explicitly definition of input and output parameters ensures clarity and facilities understanding across different teams and application domains.

# Related work

I gathered some related studies based on given datasets. The several studies have explored the use of machine learning algorithms and ensemble methods in the field of blood transfusion classification. For instance, It is study titled of Machine Learning for Blood Donors Classification Model by using ensemble it is demonstrated the application of machine learning classification algorithms to predict factors contributing to donor return and find homogeneous groups of blood donors. And the another study is Predicting Blood Donors Using Machine Learning Techniques it used regional blood center data to address issues in the US blood supply chain due to innovations in surgical practice, transfusion management, and hospital policy. And next is the research of Machine Learning for Blood Donors Classification Model Using Ensemble Learning was motivated by the constant demand for blood transfusions required for several medical procedures and life-or-death operations. and by review titled Machine Learning in Transfusion Medicine A Scoping Review is described the current trends and key methodological approaches in applying machine learning in transfusion medicine, discussing the challenges and potential solutions to the prospective implementation of machine learning in this field. And lastly, the study Assessment of Machine Learning Methods to Predict Massive Blood was explored the use of modern machine learning methods to develop and validate a model that can accurately predict the need for massive blood transfusion. These studies have made significant strides in using machine learning techniques, including ensemble learning, to predict blood donation patterns and needs. However, there is still room for improvement, particularly in terms of accuracy and real-time prediction capabilities. The integration of more advanced technologies, such as deep learning, could potentially enhance these prediction models.

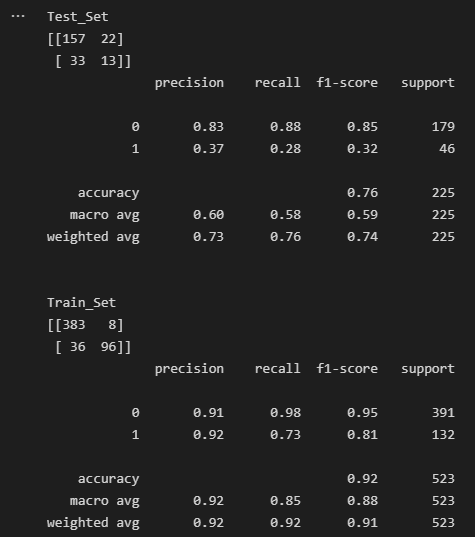
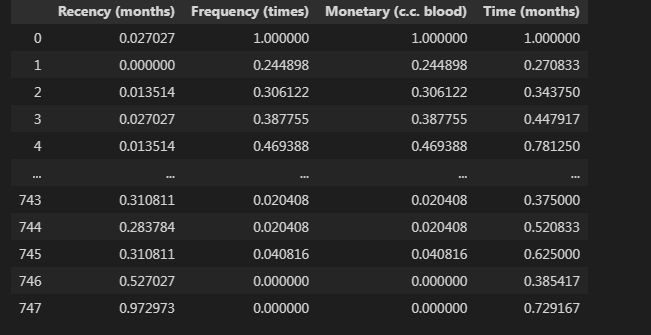


# Methods

My machine learning algorithms and ensemble methods for the classification task related blood transfusion. The primary algorithms employed include random forest, XG Boost, a Voting Classifier combining Logistic Regression, decision Tree, SVM, KNN and Gradient Boosting, as well as individual models of Logistics regression. The datasets containing information such as decency, frequency, monetary contributions, and time since the first donation is meticulously processed and examined for missing values. Descriptive statistics and correlation heat maps provide insight into dataset’s characteristics. The data preprocessing involves scaling the features, enhancing the models' performance. Subsequently, the dataset is split into training and testing sets. Each algorithm is trained on the training set and evaluated on the test set using diverse evaluation metrics such as confusion matrices, precision, recall, F1-score, and accuracy. Additionally, a voting classifier and stacking classifier are employed to leverage the strengths of multiple models. The voting classifier amalgamates combines the predictive capabilities of logistic regressions, Decision tree, SVM, KNN, and Gradient Boosting in a meta- classifier framework. The code conclude with a comparison of the performance of these models, aiding in the identification of the most effective approach for predicting blood donation behavior based on the provided dataset. Overall, the code provides a comprehensive analysis and implementation of various machine learning techniques for the specific blood transfusion classification problem.

Here is the equation provided:

CS230: Deep Learning Winter 2007, Stanford University, CA. (LateXtemplateborrowedfromNIPS2017.)





The result is 0.82222222222222222, is the stacked model the correctly predicted about 82.22%, of the cases of the test data.

# References

1. Sander A, Insisters J, Gumboots H. Patient blood management: the global view. Transfusion. 2016 (supply 1):S94–S102.
2. Ellington KD, Sapient MRP, Harass KA, et al. Continued decline in blood collection and transfusion in the United States‐2015. Transfusion. 2017; 57(supply 2):1588–1598. [PMC free article]
3. Leahy MF, Hofmann A, Towner S, et al. Improve outcomes and reduced costs associated with a health‐system‐wide patient blood management program: a retrospective observational study in four major adult tertiary‐care hospitals. Transfusion. 2017; 57 (6):1347–1358.
4. Hofmann A, Farmer S, Towner SC. Strategies to preempt and reduce the use of blood products: an Australian perspective. Corroding Anesthesia. 2012; 25 (1):66–73.
5. Sander AS, Good enough LT. Blood transfusion as a quality indicator in cardiac surgery. JAMA. 2010; 304 (14):1610–1611.
6. Shih H, Reentrant S. Comparison of time series methods and machine learning algorithms for forecasting Taiwan Blood Services Foundation blood supply. J Health Eng. 2019; 2019:1–6. [PMC free article.
7. Gumboots H, Rehab PH, Sander A, Hofmann A. Blood use in elective surgery: the Austrian benchmark study. Transfusion. 2007; 47(8):1468–1480.
8. Meier J, Iliescu D, Kopeck‐ Lange sneaker S, et al. Intra operative transfusion practices in Europe. Br J A. 2016; 116(2):255–261.