



# Blood Donor Likelihood Evaluation Using ML Algorithms

1SANJANA JADHAV, 2PROF. SWATI RAUT

1Student, 2Guide

1Department of Technology, Savitribai Phule University,

2Department of Technology, Savitribai Phule University

**Abstract :** Blood donation prediction has emerged as a critical application of machine learning in healthcare analytics due to the continuous global demand for a safe and sufficient blood supply. This research proposes a **Blood Donor Likelihood Evaluation System using Machine Learning Algorithms** to predict whether an individual is likely to donate blood in the future. The system analyzes multiple demographic, medical, and behavioral attributes such as age, gender, blood group, hemoglobin level, weight, number of previous donations, and time since last donation.

Advanced data preprocessing techniques including missing value handling, categorical encoding, feature scaling, and feature selection are applied to improve prediction accuracy. Two classification algorithms—Logistic Regression and Random Forest—are implemented and evaluated using performance metrics such as accuracy, precision, recall, F1-score, and ROC-AUC score. Experimental results indicate that Logistic Regression provides high interpretability and stable performance.

The trained model is deployed using a Flask-based web application that enables real-time donor likelihood prediction through a user-friendly interface. The proposed system assists blood banks and healthcare organizations in improving donor retention strategies, optimizing resource allocation, and ensuring consistent blood availability.

## Keywords

Blood Donation Prediction, Machine Learning, Logistic Regression, Random Forest, Healthcare Analytics, Donor Retention, Flask Deployment, Predictive Modeling

## I.INTRODUCTION

Blood donation plays a vital role in modern healthcare systems by supporting emergency treatments, surgeries, trauma care, and chronic disease management. Despite growing awareness campaigns, maintaining an adequate and consistent blood supply remains a major challenge. Donor participation is influenced by multiple factors such as health status, eligibility criteria, lifestyle habits, and past donation behavior.

Traditional donor management methods rely on manual record keeping and basic database systems, which lack advanced analytical capabilities. These systems fail to accurately predict donor return behavior, resulting in inefficient communication strategies and blood shortages.

The Blood Donor Likelihood Evaluation System provides a data-driven solution that uses machine learning algorithms to analyze historical donor data and predict future donation behavior. By leveraging predictive analytics, healthcare organizations can identify high-potential donors and plan targeted donation drives effectively.

This project ensures improved efficiency through:

- Data-driven predictive modeling
- Automated donor likelihood evaluation
- Real-time prediction via web deployment
- Optimized donor retention strategies
- Enhanced blood bank resource management

## II. NEED OF THE STUDY.

The increasing demand for blood due to population growth, road accidents, surgeries, and chronic diseases has created a significant gap between blood supply and demand. Traditional systems are unable to accurately forecast donor participation, leading to shortages and inefficient resource utilization.

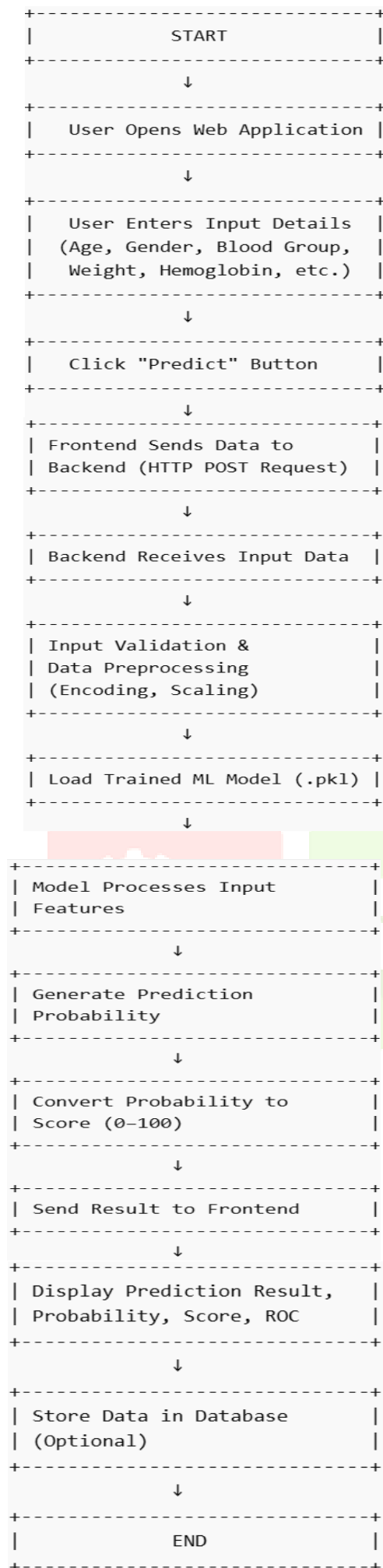
Existing systems primarily store donor records but do not apply advanced analytics for predictive insights. Therefore, there is a strong need for an intelligent, automated system that evaluates donor likelihood using machine learning techniques.

Objectives

1. To develop a machine learning-based system for predicting blood donor likelihood.
2. To implement and compare classification algorithms such as Logistic Regression and Random Forest.
3. To apply effective data preprocessing techniques to enhance model performance.
4. To evaluate model accuracy using standard performance metrics.
5. To deploy the trained model using Flask for real-time donor prediction.
6. To support healthcare organizations in improving donor retention and blood supply management.

## III. RESEARCH METHODOLOGY

The research methodology for the development of the Blood Donor Likelihood Evaluation Using ML Algorithms follows a systematic approach including data collection, preprocessing, model development, evaluation, and deployment. Initially, relevant donor datasets containing demographic, medical, and behavioral attributes were collected and analyzed. Data preprocessing techniques such as handling missing values, categorical encoding, feature scaling, and feature selection were applied to improve model efficiency. Classification algorithms including Logistic Regression and Random Forest were implemented to predict donor likelihood. The models were evaluated using performance metrics such as accuracy, precision, recall, F1-score, and ROC-AUC to ensure reliability and predictive stability. Finally, the best-performing model was deployed using a Flask-based web application to enable real-time prediction through a user-friendly interface. This structured methodology ensures accurate donor behavior prediction, efficient resource planning, and improved decision-making for blood banks and healthcare organizations.



**Mathematical Model (CRYPTOSYSTEM)**

**Prediction:**

$$Y = f(X)$$

Where:

**X** = Input features (Age, Gender, Blood Group, Donations, etc.)

- **f** = Trained ML Model
- **Y** = Predicted Likelihood

**Logistic Regression Formula:**

$$P = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \dots + \beta_n x_n)}}$$

Where:

- **P** = Probability of donor donating again
- **β** = Model coefficients

**Final Decision Rule:**

$$Y = \begin{cases} 1 & \text{if } P \geq 0.5 \\ 0 & \text{if } P < 0.5 \end{cases}$$

- **Score Formula:**

$$Score = P \times 100$$

**IV. RESULTS AND ADVANTAGES**

Test Case	Description	Expected Output	Status
TC01	Enter valid donor data	Prediction generate d successfully	Pass
TC02	Missing input fields	Error message displayed	Pass
TC03	Valid data submission	Probability & Score displayed	Pass
TC04	Model processin g request	Likelihood result returned	Pass
TC05	Invalid data type entered	Validation error shown	Pass

TEST CASE

## ADVANTAGES

- Accurate prediction of donor return likelihood
- Reduces blood shortage through early forecasting
- Helps hospitals plan blood inventory efficiently
- Saves time by automating donor analysis
- Supports data-driven decision making
- Improves donor engagement strategies
- Scalable and adaptable for web & cloud deployment
- Ensures secure handling of donor data

## V Future Work & Applications

- Integration of Deep Learning models for higher accuracy
- Real-time prediction using live hospital databases
- Mobile app for instant donor likelihood checking
- AI-based donor behavior analysis and trend forecasting
- Integration with national blood bank systems
- Automated donor notification and reminder system
- Cloud-based scalable deployment

## 5.2 APPLICATIONS

- Hospitals for blood inventory management
- Blood banks for donor retention planning
- Government health departments for public health monitoring
- NGOs organizing blood donation camps
- Emergency response systems for quick donor identification
- Healthcare analytics and research institutions

## VI REFERENCES

- M. Tohidi, H. S. Kim, and S. Lee, "Predicting blood donation behavior using machine learning algorithms," *IEEE Access*, vol. 8, pp. 14523–14534, 2020.
- A. S. Rajput and P. Kumar, "Data-driven blood donor retention prediction using classification models," *Journal of Biomedical Informatics*, vol. 107, p. 103454, 2020.
- S. Shinde, P. Jadhav, and A. Patil, "Machine learning approaches for donor behavior analysis in blood banks," *International Journal of Computer Applications*, vol. 182, no. 45, pp. 1s5–22, 2019.
- L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.
- T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed., Springer, 2009.
- WHO, "Blood safety and availability," World Health Organization, Geneva, Switzerland, 2023.
- UCI Machine Learning Repository, "Blood Transfusion Service Center Data Set," [Online]. Available: <https://archive.ics.uci.edu/ml/datasets/blood+transfusion+service+center>.
- A. Ng, *Machine Learning*, Stanford University, Stanford, CA, USA, 2021.
- Flask Documentation, "Flask Web Framework," [Online]. Available: <https://flask.palletsprojects.com/>.
- S. Raschka, *Python Machine Learning*, Packt Publishing, 2017.
- K. Murphy, *Machine Learning: A Probabilistic Perspective*, MIT Press, 2012.