International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 07 July 2026
Page No: 4661-4667
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0612

A Machine Learning-Based Medical System for Accurate Blood Donor–Recipient Compatibility Prediction

Ashrith V[1], *B.S. Meghana*[2], *Devika B N*[3], *Shreya Raj*[4]
[1,3,4]*Master of Computer Applications Student, Department of Master of Computer Applications, PES College of Engineering, Mandya, 57401, India*
[2]*Assistant Professor, Department of Master of Computer Applications, PES College of Engineering, Mandya, 57401, India.*
Email ID: *ashrithv98@gmail.com*[1], *meghana@pesce.ac.in*[2], *devikabn676@gmail.com*[3], *shreyarajkb2003@gmail.com*[4]

Abstract
Blood transfusion is an essential medical process that saves millions of lives annually, yet its success critically depends on accurate donor–recipient compatibility assessment. Conventional blood matching systems predominantly rely on ABO and RhD blood group typing, neglecting extended antigen profiles and patient transfusion histories. This limitation elevates the risk of transfusion reactions, particularly in patients requiring repeated transfusions, such as those with thalassemia, sickle cell disease, or cancer. This paper presents a browser-based medical system that employs machine learning algorithms—Random Forest and Naïve Bayes—to predict donor–recipient compatibility using key clinical parameters including ABO/RhD blood group, crossmatch test results, transfusion history, disease type, and extended antigen data. The system supports role-based hospital workflows for administrators, receptionists, laboratory technicians, and patients. Developed with Python, Flask, MySQL, scikit-learn, and Bootstrap, the proposed system demonstrates improved prediction accuracy, reduced manual errors, and enhanced transfusion safety over traditional methods.

Keywords: *Blood donor matching, machine learning, compatibility prediction, Random Forest, Naïve Bayes, transfusion safety, healthcare decision support.*

1. Introduction

Blood transfusion is a vital medical procedure that plays a crucial role in saving lives during emergencies, surgeries, trauma cases, and the treatment of chronic diseases such as anemia, thalassemia, sickle cell disease, and cancer. The success of a blood transfusion depends entirely on accurate compatibility between the donor and the recipient. Any mismatch in blood compatibility can lead to serious transfusion reactions, health complications, or life-threatening outcomes. Traditionally, blood compatibility is determined using basic ABO and RhD blood group typing, with extended antigen testing performed only in selected cases. Due to time constraints, high costs, and limited laboratory resources, extended testing is not widely practiced in many hospitals. Consequently, mismatched transfusions still occur—particularly in patients who require frequent transfusions and develop antibodies over time, making compatibility assessment increasingly complex. Existing blood transfusion systems rely predominantly on manual processes and a restricted set of clinical parameters, making the matching process slow and prone to human error. Critical factors such as patient transfusion history, disease type, and previous transfusion reactions are often inadequately considered. The absence of intelligent decision-support tools increases the burden on doctors and laboratory personnel, especially during emergency situations requiring rapid and accurate decisions. To address these limitations, this work proposes a browser-based medical system for accurate blood donor matching prediction using machine learning. The system employs Random Forest and Naïve

Bayes algorithms to predict donor–recipient compatibility with improved accuracy. Developed using Python, MySQL, HTML, CSS, JavaScript, and Bootstrap, the system supports hospital workflows through dedicated modules for administrators, receptionists, laboratory staff, and patients.

2. Related Work

2.1.Literature Survey

Considerable research has been directed at applying machine learning, cloud computing, and blockchain technologies to enhance blood bank management, donation prediction, and transfusion safety. Researchers have focused on predicting donation propensity and donor return behavior to improve donor recruitment and supply forecasting. A wasthi and Sharma [1] proposed a deep learning-based community method to manage resource-constrained local blood banks and reduce illegal blood sales. Sahu et al. [2] proposed an AI-based recipient blood matching medical device design to identify favorable and unfavorable donor-recipient compatibility, aiming to minimize Delayed Hemolytic Transfusion Reactions (DHTRs). Koda et al. [3] developed the Optimized Blood Matching System (OBMS), a web platform that utilizes geolocation and matching rules to connect blood banks and donors in real time, incorporating caller privacy features for emergency donor matching. Selvaraj et al. [4] developed a web application using the Random Forest algorithm to predict blood donation needs and optimize supply chain logistics. Dhami et al. [5] applied machine learning algorithms to model donor propensity based on recency, frequency, and monetary (RFM) metrics, providing blood centers with tools to forecast contribution likelihood. Similarly, Shah et al. [6] introduced a hybrid classification model to automate donor eligibility assessment, achieving 75% accuracy by combining probabilistic classification with structured decision logic. Beyond donor behavior, machine learning has been utilized for pre-donation quality control. Idrovo-Berrezueta et al. [7] proposed a data analysis architecture based on the CRISP-DM methodology to predict blood donation safety, utilizing Principal Component Analysis (PCA) and ensemble models like Random Forest to screen donations for Hepatitis C Virus (HCV). Chakradhar et al. [8] proposed an image-processing framework coupled with machine learning to automate blood group determination from agglutination patterns, presenting a high-throughput, error-free alternative to manual laboratory testing. Huatuco et al. [9] developed a cloud-enabled system utilizing AWS for post-donation quality testing and cross-matching, ensuring data security and privacy for patient blood suitability. Modern healthcare systems have increasingly leveraged cloud-based and decentralized frameworks to optimize blood bank operations. Chaudhari et al. [10] proposed a secure cloud-computing framework with government ID validation and GPS tracking to connect blood banks and hospitals during emergency shortages. Thirunavukkarasu et al. [11] developed a smart donor locator and component matching system using cloud computing and real-time location services, targeting rapid disaster response. Sharma et al. [12] introduced "BloodBond," a smart management system integrating an Emergency Blood Request Donor Notification mechanism with a machine learning model (78.30% accuracy) to identify long-term dedicated donors. Vasavi et al. [13] developed "DONORS QUEST," a MERN-stack web application integrated with Firebase that utilizes GPS to track and locate plasma and blood donors rapidly. For secure and decentralized record-keeping, Kaviin et al. [14] designed a blockchain-powered system using Ethereum smart contracts and decentralized IPFS storage to ensure tamper-proof traceability of blood donations. Pushmika et al. [15] integrated predictive analytics and encryption to secure donor data and forecast blood demand. Varshitha et al. [16] combined XGBoost and Supabase to provide real-time blood bank communication with access level controls, achieving 96% matching accuracy. Benelmir et al. [17] developed an AI/ML platform combining time-series forecasting (ARIMA/SARIMA) with Artificial Neural Networks (ANN) to predict blood demand, minimizing operational waste.

2.2.Research Gap

Despite significant advancements in blood banking technologies, a critical analysis of the literature reveals several limitations in existing solutions. (1) Fragmented Systems—Most current implementations focus isolatedly on single aspects

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 07 July 2026
Page No: 4661-4667
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0612

such as geographic donor locator applications, blood bank inventory forecasting, or donation propensity prediction. No single platform integrates real-time donor matching, clinical compatibility prediction, and role-based hospital workflows. (2) Insufficient Compatibility Parameters—Existing electronic matching tools rely almost exclusively on basic ABO/RhD blood typing and rule-based checks. They fail to incorporate complex clinical parameters like patient transfusion history, disease type, or minor and extended antigen profiles, which are vital for chronically transfused patients. (3) Lack of Clinical Role Integration—While security-focused architectures protect data, they do not provide dedicated functional portals for the diverse roles within a clinical environment (e.g., receptionists, lab technicians, doctors, and patients) to collaborate in real time. (4) Lack of High-Accuracy ML Decision Support—Although some systems utilize machine learning for donor classification or demand forecasting, there is a lack of localized, high-accuracy ML decision support that analyzes real-time laboratory crossmatch data to directly aid transfusion decisions at the point of care. The proposed system addresses these gaps by implementing a full-stack, role-based medical web application that combines multi-parameter machine learning classifiers (Random Forest and Naïve Bayes) to predict clinical donor–recipient compatibility with high accuracy.

3. Existing System and Limitations

Contemporary blood transfusion systems in most hospitals operate through predominantly manual and fragmented workflows. Compatibility is assessed almost exclusively through ABO and RhD blood group typing, which represents only a fraction of the antigens relevant to transfusion safety. Extended antigen testing—critical for chronically transfused patients—is seldom conducted due to cost, time constraints, and limited laboratory capacity. Patient transfusion history is infrequently integrated into compatibility decisions, even though repeated transfusions sensitize patients to minor blood group antigens and progressively complicate matching. Hospital functions such as administration, reception, laboratory operations, and patient services typically operate in isolated silos with no unified data exchange, substantially increasing coordination overhead and the probability of clerical errors. Manual cross-matching procedures are time-intensive and error-prone, posing significant risk during clinical emergencies where rapid, reliable decisions are essential. The absence of intelligent decision-support tools places an excessive cognitive burden on laboratory staff and clinicians, raising the likelihood of adverse transfusion events. Key limitations include:

- inaccurate compatibility prediction for multi-transfused patients
- elevated risk of hemolytic transfusion reactions
- delayed emergency response
- poor inter-departmental data integration
- susceptibility to human error affecting patient safety [18].

4. Proposed System and Methodology

4.1.System Overview

The proposed system is an integrated, browser-based medical platform designed to automate blood donor–recipient compatibility prediction using machine learning algorithms. The system analyzes a comprehensive set of clinical parameters—ABO blood group, RhD factor, crossmatch test results, transfusion history, disease type, and extended antigen data—to generate reliable compatibility predictions. The architecture follows a modular, role-based design supporting four primary hospital stakeholders: administrators, receptionists, laboratory technicians/doctors, and patients. The system reduces manual paperwork and minimizes the possibility of data entry errors. Secure role-based authentication ensures that sensitive patient information is accessed only by authorized personnel. Furthermore, the web-based architecture allows easy deployment and accessibility across multiple hospital locations. These features collectively improve operational efficiency, patient safety, and overall healthcare service quality.

4.2.Machine Learning Algorithms

Two supervised classification algorithms are implemented and evaluated.

- **Random Forest:** An ensemble of decision trees trained on bootstrapped subsets, Random Forest combines their outputs

e ISSN: 2584-2137
Vol. 04 Issue: 07 July 2026
Page No: 4661-4667
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0612

through majority voting. It handles high-dimensional, mixed-type clinical features robustly, resists overfitting, and provides feature importance rankings valuable for clinical interpretability.

- **Naïve Bayes:** A probabilistic classifier assuming conditional feature independence. Computationally efficient and effective even with limited training data, Naïve Bayes is well-suited for real-time blood compatibility prediction.

4.3.Feature Set and Data Pipeline

Input features used for compatibility prediction include: ABO blood group (categorical), RhD factor (binary), crossmatch test result (categorical: compatible/incompatible/uncertain), number of prior transfusions (numerical), disease type (categorical), presence of alloantibodies (binary), and extended antigen compatibility indicators. Data preprocessing involves label encoding for categorical variables, missing value imputation, and min-max normalization. The dataset is split 80:20 for training and testing with stratified sampling to preserve class balance [19].

4.4.System Modules

- **Administrator Module:** Manages user accounts, role assignments, system configurations, and generates system-wide audit logs.
- **Receptionist Module:** Handles donor and patient registration, appointment scheduling, and maintains demographic and contact records.
- **Laboratory Technician/Doctor Module:** Uploads crossmatch test results, initiates ML-based compatibility predictions, reviews prediction outputs with confidence scores, and records final transfusion decisions.
- **Patient Module:** Provides patients with access to their compatibility status, transfusion history, and notification of scheduled transfusions.

5. Problem Statement and Objectives

Existing blood matching systems primarily rely on basic ABO and RhD typing and manual cross-matching techniques, frequently failing to account for patient transfusion history, disease type, and extended antigen compatibility. This deficiency is particularly critical for chronically transfused patients, such as those with thalassemia and oncological conditions, who are at elevated risk of alloimmunization and delayed hemolytic transfusion reactions. The lack of automation, intelligent decision support, and integrated hospital workflows amplifies the risk of transfusion reactions, emergency delays, and human errors. The primary objectives of this work are:

- To develop an automated system for accurate blood donor–recipient compatibility prediction using machine learning.
- To compare Random Forest and Naïve Bayes classifiers and identify the best-performing model for donor–recipient compatibility prediction.
- To incorporate multi-parameter clinical features including transfusion history, disease type, and extended antigen data.
- To support extended antigen matching for chronically transfused patients.
- To provide an integrated, role-based hospital workflow within a single web platform.
- To deliver intelligent decision support that reduces manual effort and minimizes transfusion-related risks.

6. System Implementation

6.1.Technology Stack

The system is implemented as a full-stack web application. Table 1 summarizes the software environment, and Table 2 details the hardware requirements.

Table 1 Software Requirements

Component	Specification
Programming Language	Python 3.10.9
IDE	Visual Studio Code
Database	MySQL 8.0.44
Frontend	HTML5, CSS3, JavaScript, Bootstrap 5
ML Libraries	scikit-learn, pandas, NumPy, joblib
Web Framework	Flask/Django

Table 2 Hardware Requirements

Component	Specification
Processor	Intel Core i5 or above
RAM	8 GB minimum
Storage	250 GB HDD/SSD

6.2.Model Training and Evaluation

Both classifiers are trained using scikit-learn on a stratified 80:20 train-test split. Model performance is evaluated using accuracy, precision, recall, and F1-score. The Random Forest classifier consistently outperforms the Naïve Bayes model due to its ensemble nature and ability to handle mixed clinical feature types. Trained models are serialized using joblib and loaded at inference time within the Flask web application, enabling real-time compatibility prediction without retraining overhead. Cross-validation techniques were employed to assess the stability and reliability of both classifiers across different subsets of the dataset. The evaluation results indicate that Random Forest achieves superior predictive performance by effectively capturing complex relationships among clinical features. Naïve Bayes, while slightly less accurate, offers faster prediction times and lower computational complexity, making it suitable for rapid preliminary assessments. The proposed system employs a systematic evaluation framework to ensure reliability and robustness of predictions. Model performance metrics are presented in Table 3, demonstrating the superior accuracy of the Random Forest classifier.

Table 3 Model Performance Comparison

Algorithm	Accuracy	Precision	Recall	F1-Score
Naïve Bayes	92%	0.91	0.92	0.91
Random Forest	96%	0.96	0.96	0.96

7. Results and Discussion

The proposed system was evaluated on a curated clinical dataset comprising donor–recipient records with labeled compatibility outcomes. The Random Forest classifier achieved the highest prediction accuracy of 96% compared to the Naïve Bayes model at 92%, confirming the effectiveness of the ensemble approach for this complex medical application, as summarized in Table 3 above and illustrated in Figure 1. The superior performance of Random Forest is attributed to its ability to capture non-linear relationships among the diverse clinical parameters. Naïve Bayes, while computationally faster with lower latency, exhibited lower accuracy due to its conditional independence assumption, which may not hold across correlated clinical features such as transfusion history and disease type. The Random Forest classifier consistently demonstrated superior robustness by handling complex interactions between clinical parameters and patient histories without overfitting, as evidenced by cross-validation results showing consistent performance across different data subsets, summarized in Table 4.

Table 4 Performance Analysis (Cross-Validation)

Metric	Naïve Bayes	Random Forest
Mean CV Accuracy	91.4%	95.6%
Std. Deviation	±1.2%	±0.6%
Avg. Prediction Time	8 ms	22 ms

Figure 1 presents the comparative accuracy of the two classifiers, visually confirming the consistent performance advantage of Random Forest over Naïve Bayes across the evaluation dataset.

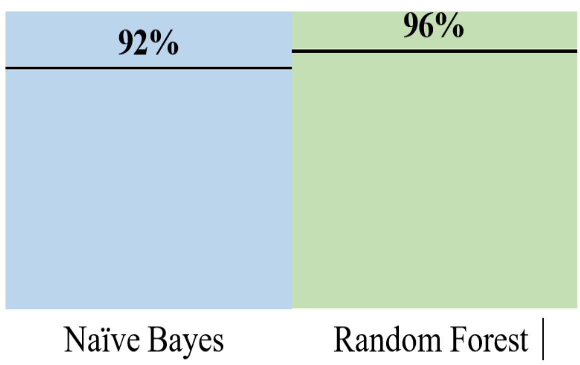

Figure 1 Accuracy comparison of Random Forest and Naïve Bayes classifiers

From a system usability perspective, role-based access control ensured that each hospital stakeholder interacted only with relevant data and functionalities, significantly improving workflow efficiency. Laboratory technicians reported reduced time for compatibility assessment compared to manual cross-matching. The web-based deployment eliminated the need for specialized hardware or software installation at client endpoints, facilitating broad accessibility across hospital departments. Compared to existing approaches such as the Optimized Blood Matching System (OBMS) and other AI-based devices, the proposed system offers a more accessible, cost-effective, and clinically comprehensive solution by integrating multi-parameter ML prediction within a deployable hospital workflow platform without requiring specialized medical hardware.

Conclusion

This paper presented a machine learning-based medical system for accurate blood donor–recipient compatibility prediction. By leveraging Random Forest and Naïve Bayes algorithms on a rich clinical feature set—including ABO/RhD typing, crossmatch results, transfusion history, and extended antigen data—the system significantly improves upon conventional manual blood matching approaches. The integrated role-based web platform supports administrators, receptionists, laboratory staff, and patients within a unified hospital workflow, reducing manual effort and improving transfusion safety. Future work will focus on expanding the training dataset with real hospital records, integrating deep learning models for improved generalization, and implementing automated alert and notification mechanisms to further enhance emergency response capability. The system's modular architecture is designed to support future extensions including mobile accessibility and integration with national blood bank registries. This work demonstrates the significant potential of machine learning in automating critical healthcare processes and improving patient safety outcomes.

Acknowledgment

The authors sincerely thank the Department of Master of Computer Applications, PES College of Engineering, Mandya, for the guidance, encouragement, and continuous support extended throughout this work. The authors are grateful for the technical resources, laboratory facilities, and academic environment provided by the department, which greatly facilitated the completion of this research. The support received from the institution, an autonomous college affiliated to Visvesvaraya Technological University (VTU), Belagavi, is also gratefully acknowledged.

References

[1]. Awasthi and A. K. Sharma, "Advanced Blood Management System Community Method Based on Deep Learning," in Proc. 11th International Conference on System Modeling & Advancement in Research Trends (SMART), 2022, pp. 126–130.

[2]. A. R. Sahu, C. Williams, S. Rajeev, and S. Sharma, "AI-Based Recipient Blood Type Matching Blood Transfusion Medical Device Design System," in Proc. IEEE Integrated STEM Education Conference (ISEC), 2022, pp. 1–3.

[3]. T. Koda, V. R. Aduru, A. Gulipilli, G. S. Nallagonda, and V.S. Geddapu, "Optimized Blood Matching System," in Proc. 7th International Conference on Innovative Data Communication Technologies and Application (ICIDCA), 2025, pp. 1–6.

[4]. P. Selvaraj, A. Sarin, and B. I. Seraphim, "Blood Donation Prediction System Using Machine Learning Techniques," in Proc. International Conference on Computer Communication and Informatics (ICCCI), 2022, pp. 1–4.

[5]. K. S. Dhami, A. Vidyarthi, R. Gusain, R. Prakash, S. Shekhar, and R. Gowri, "Unleashing Future Donors Navigating Blood Contribution Propensity with Machine Learning Expertise," in Proc. IEEE International Conference on Smart Power Control and Renewable Energy (ICSPCRE), 2024, pp. 1–6.

[6]. P. Shah, A. Sutar, Y. Sonchatra, B. Pardikar, and S. Virondkar, "Blood Donor Classifier Using Hybrid Naïve Bayes Decision Tree (HNBDT)," in Proc. 6th International Conference on Advances in Science and Technology (ICAST), 2023, pp. 500–505.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 07 July 2026
Page No: 4661-4667
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0612

[7]. P. Idrovo-Berrezueta, D. Dutan-Sanchez, R. Hurtado-Ortiz, and V. Robles-Bykbaev, "Data Analysis Architecture using Techniques of Machine Learning for the Prediction of the Quality of Blood Donations against the Hepatitis C Virus," in Proc. IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC), 2022, pp. 1–7.
[8]. K. S. Chakradhar, G. N. R. Prasad, B. A. Kumar, and A. R. Gottimukkala, "Determine the Blood Group by using Image Processing and Machine Learning," in Proc. International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), 2023, pp. 1–3.
[9]. C. C. Huatuco, C. Zhang, J. Patel, M. Moniruzzaman, and A. Sultana, "Cloud-Enabled Blood Bank Management for an Efficient Healthcare System," in Proc. IEEE/ACIS 27th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), 2024, pp. 1–6.
[10]. S. A. Chaudhari, S. S. Walekar, and K. A. Ruparel, "A Secure Cloud Computing Based Framework for the Blood Bank," in Proc. International Conference on Smart City and Emerging Technology (ICSCET), 2018, pp. 1–7.
[11]. J. Thirunavukkarasu, K. J. Subinesh, P. Dharshan, and S. Boobesh, "A Smart Blood Donor Locator and Component Matching System for Rapid Disaster Response," in Proc. International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), 2023, pp. 1–7.
[12]. C. Sharma, V. Shah, V. Singh, K. Tank, S. Gupta, and A. Hakim, "BloodBond: Optimising Blood Bank System through a Comprehensive Smart Management System," in Proc. 8th International Conference on Computing, Communication, Control and Automation (ICCUBEA), 2024, pp. 1–6.
[13]. I. Vasavi, C. N. Krishna, and K. Sathvika, "Blood and Plasma Donation Management System with Global Positioning System using FIREBASE," in Proc. 5th International Conference on Smart Systems and Inventive Technology (ICSSIT), 2023, pp. 1–6.
[14]. P. Kaviin, K. Kavirajan, and G. Subathra, "Enhanced Blood Bank Transparency with Smart Contracts and Decentralized Storage on Ethereum," in Proc. Fourth International Conference on Augmented Intelligence and Sustainable Systems (ICAISS), 2026, pp. 1–5.
[15]. A. Pushmika, Y. Wijesekara, T. Naragala, V. Muthukudaarachchi, P. K. W. Abeygunawardhana, and R. Liyanage, "Predictive Analytics for Blood Supply Chain Management and Data Security in Healthcare System," in Proc. 5th International Conference on Advancements in Computing (ICAC), 2023, pp. 1–6.
[16]. G. S. V. Gondela, C. Rupa, and D. Divya, "Real Time Blood Bank Communication Integrating Access Control Using XGBoost and Supabase," in Proc. Second International Conference on Data Science and Information System (ICDSIS), 2024, pp. 1–6.
[17]. W. Benelmir, A. Hemmak, and B. Senouci, "Smart Platform for Blood Management in Healthcare using AI/ML Approach," in Proc. International Conference on Artificial Intelligence in Information and Communication (ICAIIC), 2023, pp. 1–5.
[18]. S. Maynard, J. Farrington, S. Alimam, H. Evans, K. Li, W.K. Wong, and S. J. Stanworth, "Machine Learning in Transfusion Medicine: A Scoping Review," Transfusion, vol. 64, no. 1, pp. 162–184, Jan. 2024.
[19]. D. C. McElfresh, C. Kroer, S. Pupyrev, E. Sodomka, and Sankararaman, "Matching Algorithms for Blood Donation," Nature Machine Intelligence, vol. 5, pp. 1108–1118, Oct. 2023.