International Journal of Engineering Research and Science & Technology

IJERST

ISSN : 2319-5991



www.ijerst.com

Email: editor@ijerst.com or editor.ijerst@gmail.com

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

https://doi.org/10.62643/ijerst.2025.v21.i2.pp1125-1138 Vol. 21, Issue 2, 2025

A HYBRID CLASSIFICATION AND GAME-THEORETIC STRATEGY FOR OPTIMIZING PATIENT RECRUITMENT IN CLINICAL TRIALS

[1]K. Venkata Prasad Reddy, MCA Student, Department of MCA

[2] Shaik Haseena, M.Tech, (Ph.D), Assistant Professor, Department of MCA

[12]Dr KV Subba Reddy Institute of Technology, Dupadu, Kurnool

ABSTRACT

The urgent need for better patient recruitment in clinical trials, addressing issues like delays and excessive costs, makes this study essential. We want to improve patient outcomes, progress healthcare research, and increase trial efficiency by using a categorisation model and a game theoretic method for clinical trial settings. In the past, manual or semi-automated techniques including ads, referrals, and eligibility screening using databases have been used to recruit patients for clinical trials, despite the fact that this process is essential to the advancement of medical research. Trial completion is often delayed by these methods' inefficiencies, which include exorbitant expenses, drawn-out schedules, and inadequate participant matching. These difficulties show how conventional approaches fall short of the increasing need for effective and efficient hiring practices. In order to overcome these drawbacks, this paper suggests a sophisticated AI-driven solution that uses a Random Forest Classifier in a hybrid architecture. Through a stacking classifier, the suggested model combines Random Forests, Decision Trees, Support Vector Machines, and Logistic Regression. To improve prediction accuracy, autoencoders are used for feature extraction. Furthermore, via strategic interaction, a game-theoretic strategy involving patients, clinical investigators, and research firms optimises recruiting techniques, offering equilibrium-seeking solutions for stakeholder payoffs that are balanced. By simplifying the hiring process, this AI-based strategy provides a game-changing answer to the inefficiencies of conventional hiring practices and greatly advances the area of healthcare research.

I. INTRODUCTION

1.1 Background

Although patient recruitment is still a major obstacle, especially in India, clinical trials are essential to the advancement of medical knowledge. Reports indicate that recruiting problems cause more than 40% of clinical trials worldwide to be delayed, with India particularly affected by concerns including low accessibility and knowledge. The clinical trial industry in India is expected to reach $3.15 billion by 2025 at a compound annual growth rate (CAGR) of 8.7%; however, patient recruiting inefficiencies impede this expansion. Traditional approaches often overlook India's heterogeneous population, cultural obstacles, and poor level of healthcare digitisation, which leads to an average recruiting delay of 4-6 months per trial, increasing expenses and delaying medical advancements.

Overview with Uses: Clinical trial patient recruiting done well improves healthcare outcomes, lowers costs, and advances medical research. Using game theory for strategic decision-making and machine learning for prediction accuracy, this project uses AI to expedite the hiring process. By ensuring optimal enrolment and expedited trial completion, the hybrid strategy advances medication discovery and healthcare delivery.

1.2 Problem Definition

Prior to machine learning, patient recruitment was done manually using labour-intensive and prone to mistakes methods such as doctor referrals, marketing, and database screening.

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1125-1138 Vol. 21, Issue 2, 2025

Limited outreach, mismatched eligibility, and data mismanagement were among the difficulties, which resulted in substantial delays and expensive expenses. Healthcare innovations were stalled by recruitment inefficiencies, which resulted in 40% of trials failing to fulfil enrolment targets and trials being delayed by an average of 6 months. These antiquated techniques were not flexible enough to meet the demands of contemporary medicine.

1.3 Research Motivation

The necessity for sophisticated, scalable solutions is highlighted by the inefficiencies of conventional hiring practices. Data-driven solutions are required by the expanding clinical trial business to handle high costs and recruiting delays. AI-powered methods may use a variety of patient data to match patients precisely, guaranteeing inclusion for all demographics. The potential to drastically shorten trial durations, save expenses, and improve trial results—all of which would eventually benefit patients and hasten medical advancements—is what motivates the effort.

1.4 Existing System and Drawbacks

The process is dominated by manual recruiting techniques including ads, recommendations, and simple database screening, which often leave out or include unqualified applicants. These approaches lack a framework for strategic optimisation among stakeholders and are expensive, sluggish, and prone to mistakes. Recruitment difficulties and trial delays result from their inability to manage different patient pools and massive amounts of data due to a lack of reliable forecasting tools.

1.5 Proposed System

A hybrid AI-driven architecture for patient recruiting is introduced by the suggested system. Within a stacking classifier, it incorporates many machine learning models, such as Random Forests, Decision Trees, Support Vector Machines (SVM), and Logistic Regression. Autoencoders are used to optimise feature extraction, increasing the accuracy of predictions. In order to provide strategic solutions for balanced recruitment, a game-theoretic method is used to simulate interactions between patients, clinical investigators, and research companies. Studies like "AI-Based Clinical Trial Recruitment: A Hybrid Framework" and "Game Theory in Healthcare Optimisation" provide insight into this system, demonstrating its capacity to lower expenses and expedite hiring procedures.

1.6 Real-Time Need of the Project

Clinical trials are crucial for improving medical treatments and drug discoveries, yet ineffective recruiting practices lead to major delays and higher expenses. India's expanding healthcare needs need solutions that tackle logistical issues and a variety of patient demographics. A hybrid recruiting approach powered by AI guarantees quicker enrolment, better patient-physician matching, and increased trial accuracy. Achieving quick and economical trial completion depends on real-time apps that can identify qualified individuals across databases and adjust to trial-specific requirements.

1.7 Applications of the Project

1. Clinical Research: Simplifies recruiting, speeds up drug discovery, and improves trial efficiency.
2. Healthcare Providers: Enhances patient matching for therapeutic advances and trial accessibility.
3. Pharmaceutical Companies: Shortens trial durations and expenses, allowing new medications to reach the market more quickly.
4. Regulatory Agencies: Promote adherence by guaranteeing accurate and varied representation of participants.
5. Patient-Centric Care: Improves results by matching and involving patients in tailored trials.

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1125-1138 Vol. 21, Issue 2, 2025

6. Academic Research: Offers a strong foundation for researching healthcare recruiting optimisation.
7. AI and Data Science: Promotes the use of strategic interaction and predictive modelling in medical research.
8. Impact on Global Healthcare: Encourages fair access to trials, especially in underprivileged areas.

1.8 Module Split

Upload Dataset: Gather and prepare trial data and patient profiles.

1. Data Cleaning: Deal with missing values and standardise data to ensure consistency.
2. Feature Extraction: To extract significant features, use autoencoders.
3. Model Training: Develop Random Forest, SVM, Decision Trees, and Logistic Regression classifiers.
4. Stacking Classifier: Combine models to improve the accuracy of predictions.
5. Game-Theoretic Analysis: Model stakeholder strategic interactions.
6. Model Validation: Use cross-validation methods to test and verify the model.
7. Test Prediction: Apply the learnt technique to assess actual hiring situations.

II. LITERATURE SURVEY

The creation of novel therapies and illness cures depends on clinical trials. Patient recruitment, however, is often a significant obstacle for clinical trials, particularly for uncommon illnesses or those with intricate qualifying requirements. Delays in starting and finishing the experiment may result from this, which may eventually lower the likelihood of success. The difficulties in recruiting patients for clinical trials, especially in uncommon disorders, have been brought to attention by a number of publications [1]. Clinical trial patient engagement and retention may also be enhanced by the adoption of convenience-enhancing solutions, such innovative technology [2]. Additionally, the complexity and subtleties of trial results are not well reflected by the categorisation of clinical trials as "negative" based just on statistical significance. Patients and healthcare professionals may find a more accurate evaluation of trial characteristics and more utility if trials are reclassified into groups that take into account variables including power, termination, and start [3]. Clinical trial recruitment is still difficult, and the costs and complexity are rising. The recruiting time of industry-sponsored phase III clinical trials has significantly increased over the last 12 years, suggesting that they are less successful than they were before [4], [5]. Clinical trial patient recruitment may be enhanced by the use of game theory and ensemble learning. By modelling interactions between sponsors, researchers, and patients, game theory may help find more effective recruiting tactics [6]. The accuracy of patient recruitment forecasts may be increased by combining predictions from many machine learning models using ensemble learning [7]. Clinical trial patient recruiting remains a significant concern. In clinical trials, virtual or decentralised approaches have shown promise in enhancing diversity, recruitment, and retention; some studies have even reported better outcomes than conventional approaches [9]. Furthermore, better participant retention has also been linked to decentralised approaches [10]. Natural language processing is one example of an informatics intervention that has been used to increase the accuracy and efficiency of trial recruitment and eligibility determination [11]. These approaches have shown the ability to simplify the procedure and lessen the time-consuming nature of determining whether individuals qualify. In general, decentralised approaches and informatics interventions, such game theory and machine learning techniques, may assist in overcoming the obstacles to patient

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1125-1138 Vol. 21, Issue 2, 2025

recruitment in clinical trials. The usefulness of game theory as a tool in healthcare practice is becoming more widely acknowledged [12]. Even in situations when there are competing interests, it enables the simultaneous optimisation of many goals. By taking patient-provider interactions into account, game theory also provides a framework for clinical decision-making, such as clinical decision support systems [13]. One such methodology was put out by Bilal et al. [14] and optimises illness categorisation in healthcare systems using a Bayesian game-theoretic approach. The optimal treatment of patients has also been the focus of game theoretic frameworks. The impact that is patient-centric was shown by Udok et al. [15]. The best course of action and the game's worth were determined by evaluating the patients' preferences for medical facilities based on variables including service fees and provider attitudes. In the healthcare industry, game theory may be used to foster transformative leadership, fair competition, trust, and adaptability. In the healthcare industry, a worthy purpose, a culture of trust and no blame, recognising opponents, existential adaptability, and brave leadership are the five guiding principles required to lead an endless game [16]. Game theory may improve the use of resources, the precision of illness diagnosis, and the general effectiveness of healthcare delivery [17]. Using machine learning techniques to optimise patient recruitment for clinical trials has advanced significantly in recent years [18]. Response-adaptive randomisation in clinical trials has also been carried out using machine learning techniques, leading to more individualised optimum treatment assignments and increased trial participant response rates overall [19]. Although recruiting participants for clinical trials is a persistent problem, attempts to enhance study recruitment have been guided by theories of human behaviour change and decision-making, including the Theoretical Domains Framework (TDF) and shared decision-making (SDM) [20]. High accuracy in identifying patient eligibility for clinical trials has been achieved via the exploration of automating the patient recruitment job through the use of machine learning and natural language processing methods [21]. Additionally, a broad framework for the useful use of XAI in medical research has been presented, which may verify and explain cancer biomarkers and provide doctors with information [22]. A potential method in healthcare is ensemble learning, which mixes many models to enhance diagnostic and prediction capabilities. It has been used in a number of healthcare areas, including the diagnosis of hepatitis C [26], the categorisation of diabetic illness [24], the prediction of postpartum haemorrhage [25], and the prediction of medication adherence [23]. These works used memory-based and gradient boosting-based approaches, as well as ensemble learning techniques including deep ensemble learning (DEL) and gradient deep learning boosting. The findings demonstrate that ensemble models perform better than both deep learning and conventional machine learning methods, obtaining greater values for accuracy, recall, F1-score, and area under the curve. The significance of models in ensemble games has been measured using the Shapley value, a notion from game theory. Additionally, base learners in an ensemble have been dynamically assigned coefficients of impact via the use of learning automata, a reinforcement learning approach. These methods have shown encouraging outcomes in enhancing ensemble learning models' performance [27], [28]. There are still issues that need to be resolved even if game theory and ensemble learning seem to have potential for patient recruitment in clinical trials.

III. EXISTING SYSTEM

3.1 Traditional System Overview

The process of finding patients for clinical trials has historically required a lot of work and resources. Advertisements in print, radio,

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1125-1138
Vol. 21, Issue 2, 2025

and internet media are the main ways to contact potential participants. Furthermore, direct outreach at clinics or hospitals and referrals from medical professionals are standard procedures. To find qualified applicants based on trial criteria, eligibility screening is carried either manually or by semi-automated techniques using patient databases, medical records, or health registries. These conventional approaches often include hiring firms or specialised personnel who screen, get in touch with, and interact with participants.

These systems are limited by their manual nature, even if they depend on interpersonal communication and professional networks. Spreadsheets and simple data organisation software are examples of tools that don't provide much assistance when it comes to expanding recruiting efforts. In order to identify qualified volunteers, investigators must go through a large number of applications, which often prolongs the recruiting process.

Additionally, patient enrolment monitoring usually entails human follow-ups and sporadic updates, both of which are laborious and prone to mistakes. Geographical restrictions further limit recruitment tactics since outreach initiatives are often localised, which leaves out a substantial pool of possible participants. Even with small advancements in semi-automated databases, current systems are unable to effectively manage extensive hiring.

3.2 Limitations of Traditional Systems

There are many drawbacks to the conventional approaches of clinical trial patient recruiting that reduce their efficacy and efficiency. Among these restrictions are:

1. Expensive: Manual hiring procedures need a large amount of human capital, which raises operating costs.
2. Extended Timelines: Because manual screening and outreach are time-consuming, participant enrolment is delayed, which prolongs trial completion dates.
3. Limited Reach: By not reaching out to larger, more varied patient pools, localised advertising and referrals may leave out qualified applicants.
4. Low Precision in Matching: Because of human error and insufficient tools for evaluating eligibility criteria, manual techniques often provide participant matching that is less than ideal.
5. High Dropout Rates: Participants often leave a trial in the middle of it due to a lack of engagement tactics.
6. Data Handling Challenges: Large datasets are difficult for manual and semi-automated methods to organise and analyse, which results in ineffective participant selection.
7. Inconsistent Monitoring: The integrity of the hiring process is impacted by the error-proneness of follow-ups and data updates.
8. Geographic Restrictions: Diversity and scalability are limited since recruitment efforts are usually limited to certain locations.
9. Regulatory bottlenecks: Because regulatory regulations need manual paperwork and certification, they often hold down the hiring process.
10. Barriers to Patient Engagement: Conventional approaches lack dynamic engagement tools, which results in inadequate communication and lower patient motivation.

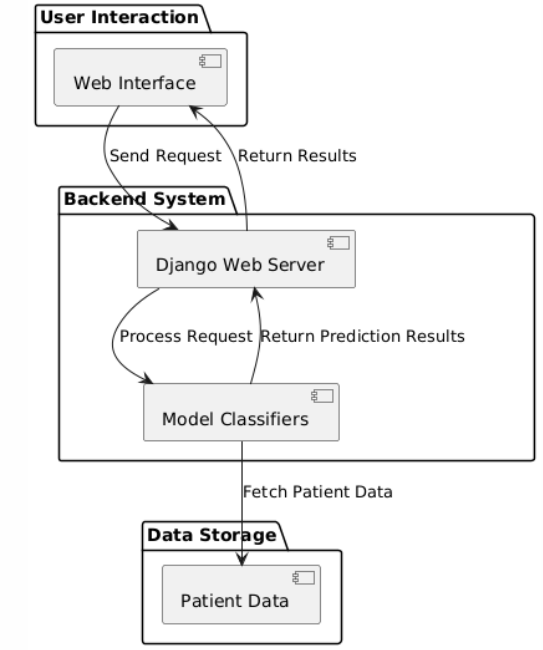
SYSTEM ARCHITECTURE

International Journal of Engineering Research and Science & Technology

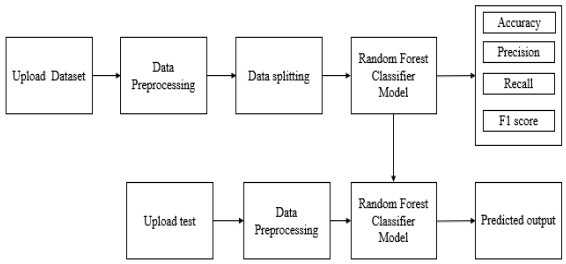
ISSN 2319-5991 www.ijerst.com

https://doi.org/10.62643/ijerst.2025.v21.i2.pp1125-1138

Vol. 21, Issue 2, 2025



IV. PROPOSED SYSTEM



4.1 Existing Algorithms

1. Decision Tree Classifier (DTC)

What is DTC? A supervised machine learning approach called the Decision Tree Classifier (DTC) divides the data into subgroups according to feature values. The end product is a structure that resembles a tree, with each internal node standing for an attribute, each branch for a decision rule, and each leaf node for the conclusion.

How it Operates: Recursively dividing the dataset into subsets according to feature values that provide the greatest possible separation of the target variable is how DTC constructs a model. Information Gain and Gini Impurity are two examples of criteria used in the splitting process. Until a stopping condition is satisfied, such a certain tree depth or data purity, the algorithm keeps splitting.

Architecture:

1. Start with the entire dataset as the root node.
2. At each node, select the best feature to split based on a criterion.
3. Recursively split the dataset into subgroups based on feature values.
4. Continue the process until a predefined stopping condition is met.
5. The final model is represented as a tree structure.

Disadvantages:

1. Prone to overfitting, especially with deeper trees.
2. Sensitive to noisy data and outliers.
3. Can become computationally expensive with large datasets.

2. Logistic Regression

Logistic regression: what is it? A statistical model for binary classification is called logistic regression. By mapping every real-valued integer into a value between 0 and 1, it represents the likelihood that a given input belongs to a certain class using a logistic (sigmoid) function.

How it Operates: The approach uses the maximum likelihood estimation method to estimate the model's parameters, or weights. After applying the logistic function to the weighted sum of the input characteristics, it generates a probability. The input is categorised into one of the two groups using the decision threshold, which is typically 0.5.

Architecture:

1. Input features are weighted linearly.
2. A logistic function (sigmoid) is applied to the weighted sum of inputs.
3. The output is a probability between 0 and 1.
4. A threshold value (usually 0.5) is used for classification.

Disadvantages:

1. Assumes linearity between dependent and independent variables.
2. Not effective with highly complex relationships.
3. Sensitive to outliers in the data.

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1125-1138 Vol. 21, Issue 2, 2025

3. Support Vector Machine (SVM)

What is SVM? A supervised learning approach for regression and classification is called Support Vector Machine (SVM). In the feature space, it locates the ideal hyperplane that maximises the difference between two classes.

How it Operates: Finding the hyperplane that best divides the data into distinct groups is how SVM operates. It seeks to optimise the distance between each class's nearest points, often known as support vectors. When the data is not linearly separable, the technique transforms the feature space using kernel functions.

Architecture:

1. The dataset is mapped into a higher-dimensional space using a kernel.
2. An optimal hyperplane is calculated that maximizes the margin between classes.
3. The hyperplane separates the data into different classes.
4. Data points closest to the hyperplane are the support vectors.

Disadvantages:

1. Can be computationally expensive, especially with large datasets.
2. Choosing the right kernel can be difficult.
3. Not suitable for noisy data or large-scale datasets.

4.2 Proposed Algorithm

Random Forest Classifier (RFC)

What is RFC? To increase classification accuracy, the Random Forest Classifier (RFC), an ensemble learning technique, mixes many decision trees. A portion of the data and a subset of the characteristics are used to construct each tree in the forest. The average of the outcomes from each individual tree is used to get the final forecast.

How it Operates: Building several decision trees, each trained on a different portion of the training data, is how RFC works. Each tree generates a class prediction for classification problems, and the ultimate outcome is determined by the majority vote of all trees. Compared to individual decision trees, this method enhances accuracy and decreases overfitting.

Architecture:

1. The dataset is divided into multiple random subsets.
2. Each subset is used to train an individual decision tree.
3. Features are randomly selected at each split to build diverse trees.
4. Each tree independently makes predictions.
5. The final prediction is based on the majority vote of all trees.

Advantages:

1. Reduces overfitting compared to individual decision trees.
2. Can handle both classification and regression tasks.
3. Effective for large datasets and can handle high-dimensional feature spaces.
4. Robust to noise and outliers in the data.
5. Parallelizable, making it scalable and faster on large datasets.

V. RESULTS AND DISCUSSION

Implementation and Description

1. User Management: o Verification of duplicate emails and usernames is part of user registration.
   - o Secure login is ensured using authentication.
   - o Feedback and redirection messages improve user experience.

2. Dataset Handling: Pandas is used to handle uploaded CSV files.
   - o Data is separated into training and testing sets after being label-encoded.

3. Model Training & Saving: o If the models aren't already stored in the static/model2/directory, they are trained.

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1125-1138 Vol. 21, Issue 2, 2025

- To make the models reusable, they are serialised using joblib (or h5 for ANN).

4. Metric Evaluation: F1-score, recall, accuracy, and precision are among the metrics that are computed.
   - Seaborn is used to visualise a confusion matrix.
5. Machine Learning Models: o Offers support for a number of methods, such as ANN, SVM, Decision Tree, Logistic Regression, and Naive Bayes.

settings.py

This file serves as a Django project's main configuration file. It includes options that specify the behaviour of the project.

Key Sections:

1. Installed Applications:
   - Provides a list of every Django and external application currently used by the project. This covers both custom applications you make and built-in apps like django.contrib.admin.
2. Middleware:
   - A collection of middleware elements that handle requests and answers on a worldwide scale. Among many other things, middleware may enforce security settings and change headers.
3. Database Configuration:
   - Indicates the database name, user, connection settings, and database engine (such as SQLite, PostgreSQL, or MySQL).
4. Templates:
   - Sets up the Django template engine's folders and backend. It contains the locations of your HTML template files.
5. Static and Media Files:
   - Configure the storage and serving locations for user-uploaded media files and static files (such as CSS, JavaScript, and pictures).
6. Authentication Settings:
   - Specifies parameters for user authentication, including authentication backends, password validators, and login URLs.
7. Internationalisation:
   - Manages the project's time zone and language settings.
8. Logging:
   - Sets up the recording and archiving of log messages. This is helpful for monitoring and debugging.
9. Third-Party and Custom Settings:
   - You may provide extra parameters needed for your unique use cases or by third-party libraries.

urls.py

The URL patterns for your Django project are specified in this file. It serves as a router, mapping views to URLs.

Key Components:

1. Import Statements:
   - Contains the imports required to create and connect routes, such as path and include.
2. URLpatterns List:
   - All of the URL patterns are specified in this Python list. Every pattern connects a particular URL route with a class-based view or view function.
3. Root URLs:
   - Specifies the application's primary entry points. For instance, /login for a login page or /admin for the Django admin interface.
4. Adding App-Specific URLs:
   - For improved organisation, you may divide routes into distinct files for every app and add them to the main urls.py file.
5. Static Files in Development:
   - You may add static() configurations to include settings to serve static files locally while developing.

Project Settings:

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1125-1138
Vol. 21, Issue 2, 2025

- To determine where to search for things like database configurations, middleware, and installed applications, the manage.py file leverages the settings from the project's settings.py file (usually via the DJANGO_SETTINGS_MODULE environment variable).

Execute_from_command_line():

- The instructions submitted in the terminal are processed by this function. After processing the command-line parameters, it invokes the relevant Django command.

Dataset Description

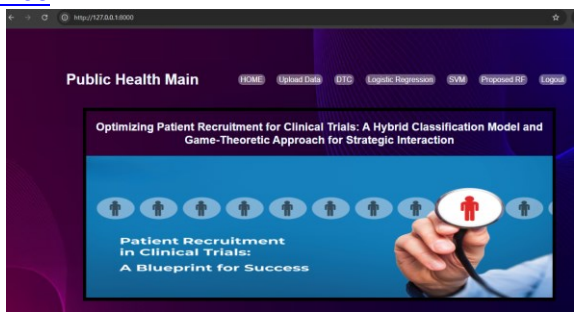
1. **HAEMATOCRIT (HCT):**
   - **Type:** Continuous
   - **Description:** The proportion of blood volume that is occupied by red blood cells. It is typically measured as a percentage and provides insights into the oxygen-carrying capacity of the blood.
   - **Unit:** Percentage (%)
2. **HAEMOGLOBINS (HGB):**
   - **Type:** Continuous
   - **Description:** The concentration of hemoglobin in the blood. Hemoglobin is the protein in red blood cells responsible for carrying oxygen.
   - **Unit:** Grams per deciliter (g/dL)
3. **ERYTHROCYTE (RBC count):**
   - **Type:** Continuous
   - **Description:** The count of red blood cells in a given volume of blood. Red blood cells are responsible for transporting oxygen throughout the body.
   - **Unit:** Millions of cells per microliter (million/µL)
4. **LEUCOCYTE (WBC count):**
   - **Type:** Continuous
   - **Description:** The count of white blood cells in a given volume of blood. White blood cells are part of the immune system, defending the body against infections.
   - **Unit:** Thousands of cells per microliter (K/µL)
5. **THROMBOCYTE (Platelet count):**
   - **Type:** Continuous
   - **Description:** The count of platelets in the blood. Platelets are crucial for blood clotting and wound healing.
   - **Unit:** Thousands of cells per microliter (K/µL)
6. **MCH (Mean Corpuscular Hemoglobin):**
   - **Type:** Continuous
   - **Description:** A measure of the average amount of hemoglobin in a single red blood cell. It provides insight into the oxygen-carrying capacity of individual red blood cells.
   - **Unit:** Picograms (pg)
7. **MCHC (Mean Corpuscular Hemoglobin Concentration):**
   - **Type:** Continuous
   - **Description:** A measure of the average concentration of hemoglobin in a given volume of red blood cells. It is a useful parameter to assess the hemoglobin content relative to the size of the red blood cell.
   - **Unit:** Grams per deciliter (g/dL)
8. **MCV (Mean Corpuscular Volume):**
   - **Type:** Continuous
   - **Description:** The average volume of a red blood cell. This measurement helps diagnose various types of anemia.
   - **Unit:** Femtoliters (fL)
9. **AGE:**
   - **Type:** Continuous
   - **Description:** The age of the individual providing the blood sample, typically in years.
   - **Unit:** Years

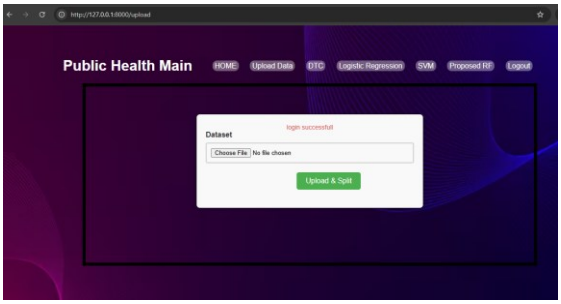
International Journal of Engineering Research and Science & Technology

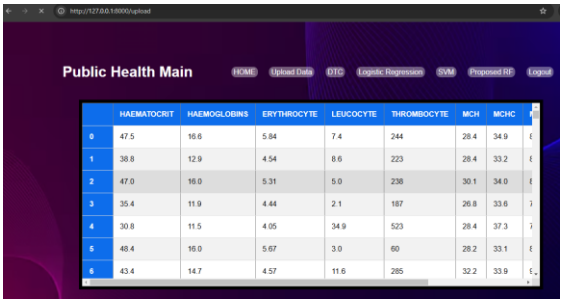
ISSN 2319-5991 www.ijerst.com

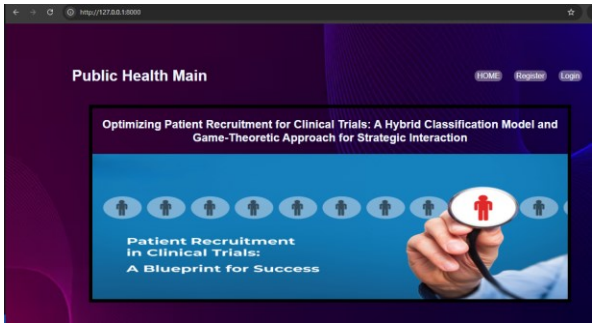
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1125-1138 Vol. 21, Issue 2, 2025

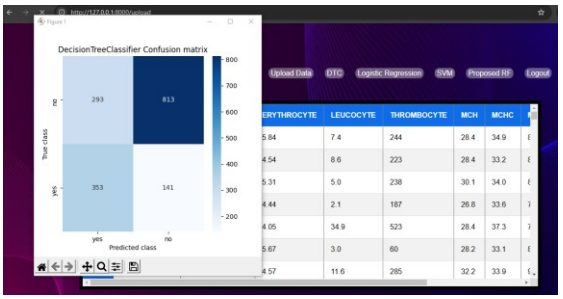
10. **SEX:**
    - **Type:** Categorical
    - **Description:** The biological sex of the individual providing the blood sample. This is typically recorded as 'Male' or 'Female'.
    - **Values:** 'Male', 'Female'
11. **SOURCE:**
    - **Type:** Categorical
    - **Description:** The origin or source of the blood sample. This may refer to different categories or conditions under which the sample was collected (e.g., hospital, clinical lab, research).
    - **Values:** Varies depending on the dataset context (e.g., 'Hospital', 'Clinic', 'Research').

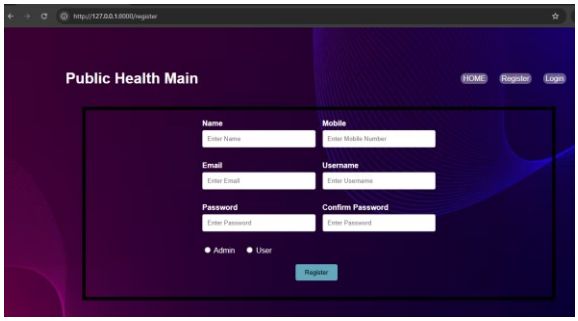


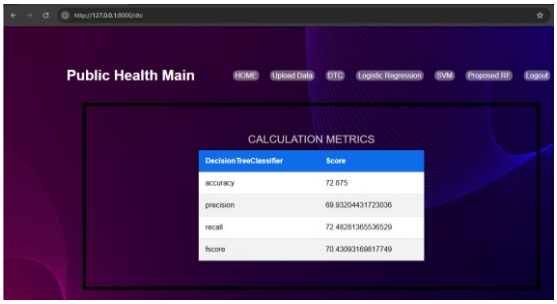


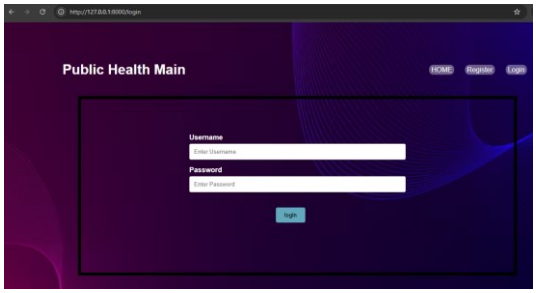








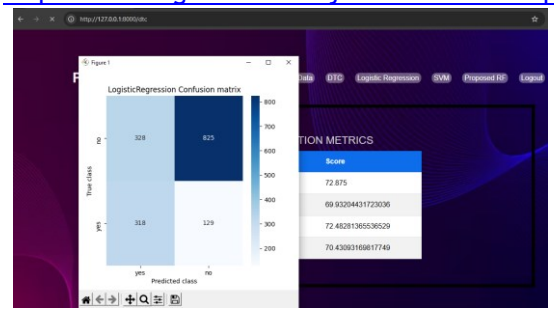


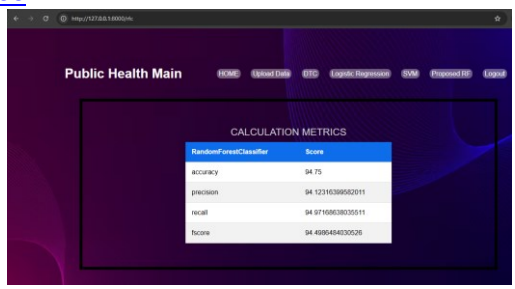


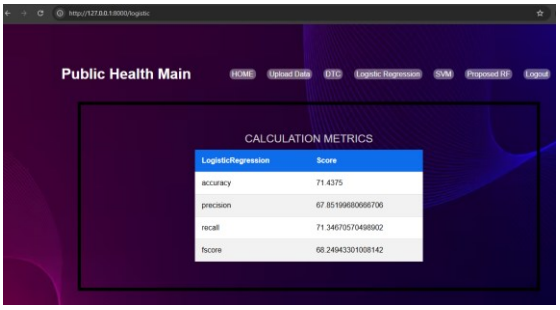
International Journal of Engineering Research and Science & Technology

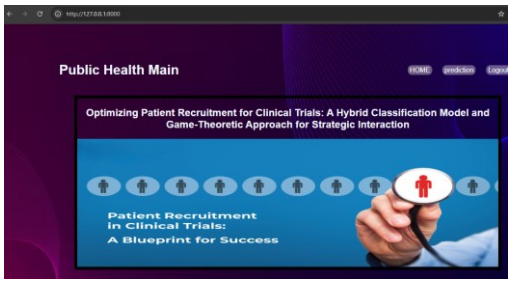
ISSN 2319-5991 www.ijerst.com

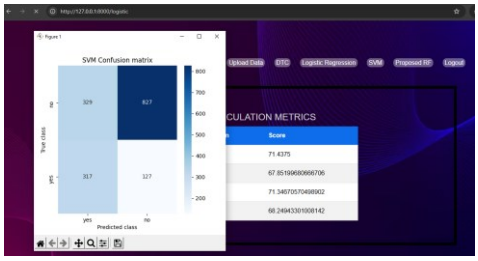
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1125-1138 Vol. 21, Issue 2, 2025

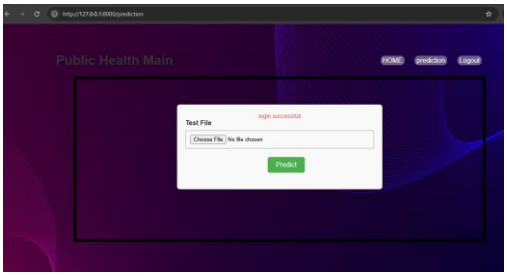


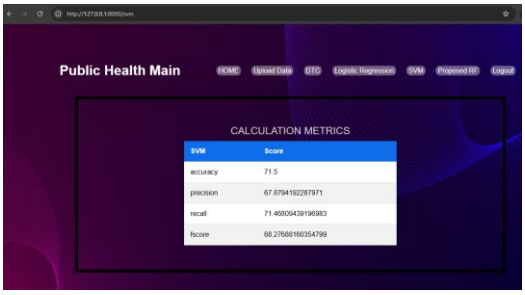


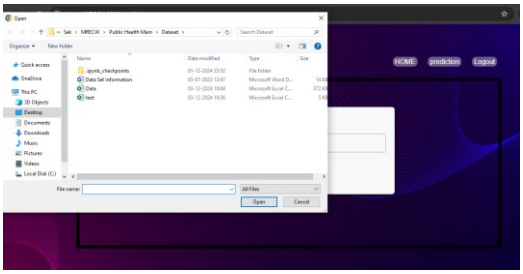


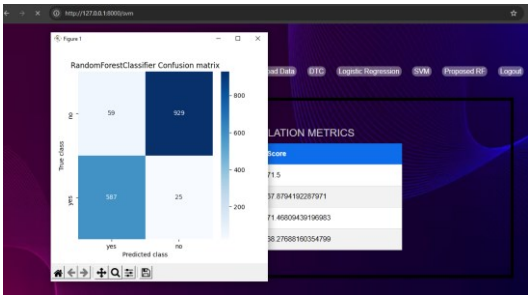


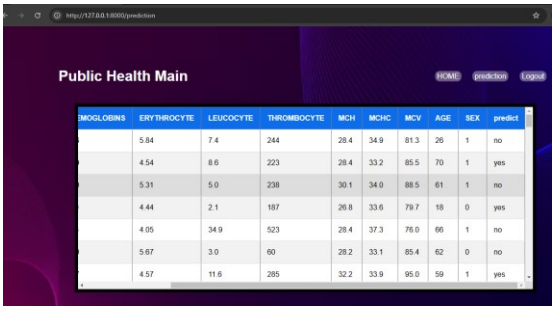












VI. CONCLUSION AND FUTURE SCOPE

This work shows how using cutting-edge AI techniques, especially a hybrid classification

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1125-1138
Vol. 21, Issue 2, 2025

model with game-theoretic tactics, may significantly transform clinical trial patient recruitment. Despite being fundamental, traditional recruiting techniques are tainted by inefficiencies, exorbitant expenses, and protracted deadlines. The suggested approach improves participant matching prediction accuracy by using a hybrid framework that combines machine learning methods such as Random Forest Classifier, Logistic Regression, Support Vector Machines, and Decision Trees. By adding autoencoders for feature extraction, the hiring process is further optimised, increasing the possibility that the trial will be finished within the allotted time.

The use of game theory offers a fresh viewpoint on the strategic interactions between patients, clinical researchers, and research companies, promoting fair hiring practices and balanced rewards. This all-encompassing strategy promotes cooperation and mutual gain by increasing efficiency and balancing the interests of all stakeholders. This study advances healthcare innovation and decreases delays in life-saving medical research by tackling the problems of conventional systems.

The results highlight how AI is revolutionising healthcare and provide a scalable and effective way to address hiring issues. In the end, this hybrid approach opens the door for clinical trials that are more reliable, efficient, and economical, improving patient outcomes and hastening the creation of medical innovations.

Future Scope

The foundation for a number of upcoming developments in clinical trial recruiting is laid by the suggested hybrid AI-driven approach.

1. Scalability: By using large and varied patient datasets, the system may be expanded internationally to accommodate multi-center studies with further improvement.
2. Real-Time Monitoring: Throughout the course of the study, participant engagement and retention may be improved by integrating real-time analytics and patient monitoring technologies.
3. Personalised Recruitment Strategies: By using AI to create outreach campaigns that are tailored to demographic, regional, and behavioural data, better focused recruitment may be achieved.
4. Integration with Electronic Health Records (EHRs): By automating data extraction and analysis, connecting the system to EHRs helps expedite eligibility screening.
5. Automation of Regulatory Compliance: AI may help with compliance by automating reporting, paperwork, and adherence to legal and ethical requirements.
6. Better Stakeholder Interaction: By adding insurance companies and pharmaceutical sponsors to the game-theoretic framework, resource allocation and financial incentives may be optimised.
7. Dynamic Feedback Loops: Models for continuous learning may adjust to changing patient reactions and trial needs, guaranteeing long-term effectiveness.
8. Integration with Telehealth: Using telehealth platforms may increase trial inclusion by reaching out to underprivileged and rural communities with recruiting efforts.
9. Advanced Security Measures: AI-powered security systems can protect patient information, guaranteeing privacy laws are followed and confidentiality is maintained.
10. Cross-Domain Applications: The hybrid model's influence may be increased by adapting it for recruitment in other research fields, including as public health studies and the social sciences.

REFERENCE

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1125-1138 Vol. 21, Issue 2, 2025

[1] M. Sequeira, J. R. Almeida, and J. L. Oliveira, "A comparative analysis of data platforms for rare diseases," in Proc. IEEE 34th Int. Symp. Comput.- Based Med. Syst. (CBMS), Jun. 2021, pp. 366–371.
[2] M. Amiridi, C. Qian, N. D. Sidiropoulos, and L. M. Glass, "Enrollment rate prediction in clinical trials based on CDF sketching and tensor factorization tools," in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), Jun. 2023, pp. 1–5.
[3] P. A. Torres-Saavedra and K. A. Winter, "An overview of phase 2 clinical trial designs," Int. J. Radiat. Oncol. Biol. Phys., vol. 112, no. 1, pp. 22–29, Jan. 2022, doi: 10.1016/J.IJROBP.2021.07.1700.
[4] M. Paul, Y. Dishon-Benattar, Y. Dickstein, and D. Yahav, "Optimizing patient recruitment into clinical trials of antimicrobial-resistant pathogens," JAC-Antimicrobial Resistance, vol. 5, no. 1, pp. 1–6, Dec. 2022, doi: 10.1093/jacamr/dlad005.
[5] M. Brøgger-Mikkelsen, J. R. Zibert, A. D. Andersen, U. Lassen, Z. Ali, and S. F. Thomsen, "Changes in key recruitment performance metrics from 2008–2019 in industry-sponsored phase III clinical trials registered at ClinicalTrials.Gov," PLoS ONE, vol. 17, no. 7, Jul. 2022, Art. no. e0271819, doi: 10.1371/journal.pone.0271819.
[6] J. Liu, C. Yu, C. Li, and J. Han, "Cooperation or conflict in doctor-patient relationship? An analysis from the perspective of evolutionary game," IEEE Access, vol. 8, pp. 42898–42908, 2020.
[7] N. Liu, Z. X. Koh, E. C. Chua, L. M. Tan, Z. Lin, B. Mirza, and M. E. H. Ong, "Risk scoring for prediction of acute cardiac complications from imbalanced clinical data," IEEE J. Biomed. Health Informat., vol. 18, no. 6, pp. 1894–1902, Nov. 2014.
[8] R. Benedek and R. Sarkar, "The Shapley value of classifiers in ensemble games," in Proc. 30th ACM Int. Conf. Inf. Knowl. Manag., 2021, pp. 1558–1567, doi: 10.1145/3459637.3482302.
[9] B. Miyata, B. Tafuto, and N. Jose, "76 methods and perceptions of success for patient recruitment in decentralized clinical trials," J. Clin. Transl. Sci., vol. 7, no. s1, p. 21, Apr. 2023, doi: 10.1017/cts.2023.160.
[10] C. Weng and P. J. Embi, "Informatics approaches to participant recruitment," in Clinical Research Informatics. Cham, Switzerland: Springer, 2023, pp. 219–229.
[11] B. Idnay, Y. Fang, C. Dreisbach, K. Marder, C. Weng, and R. Schnall, "Clinical research staff perceptions on a natural language processingdriven tool for eligibility prescreening: An iterative usability assessment," Int. J. Med. Informat., vol. 171, Mar. 2023, Art. no. 104985, doi: 10.1016/j.ijmedinf.2023.104985.
[12] A. Ahad, M. Tahir, M. A. S. Sheikh, N. Hassan, K. I. Ahmed, and A. Mughees, "A game theory based clustering scheme (GCS) for 5Gbased smart healthcare," in Proc. IEEE 5th Int. Symp. Telecommun. Technol. (ISTT), Nov. 2020, pp. 157–161.
[13] Z. Ning, P. Dong, X. Wang, X. Hu, L. Guo, B. Hu, Y. Guo, T. Qiu, and R. Y. K. Kwok, "Mobile edge computing enabled 5G health monitoring for Internet of Medical Things: A decentralized game theoretic approach," IEEE J. Sel. Areas Commun., vol. 39, no. 2, pp. 463–478, Feb. 2021.
[14] B. S. Raja and S. Asghar, "Disease classification in health care systems with game theory approach," IEEE Access, vol. 8, pp. 83298–83311, 2020, doi: 10.1109/ACCESS.2020.2991016.
[15] V. U. Udeme and U. C. Orumie, "Patients' preferences of healthcare facilities for quality healthcare services in Akwa Ibom state: A game theory approach," Amer. J. Oper. Res., vol. 11, no. 3, pp. 181–198, 2021, doi: 10.4236/AJOR.2021.113011.
[16] A. G. Powell, C. Bowman, C. Brown, R. J. Egan, and W. G. Lewis, "Team strategic philosophy: Requiem for the infinite game," Postgraduate Med. J., vol. 96, no. 1136, pp.

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1125-1138
310–312, Jun. 2020, doi: 10.1136/POSTGRADMEDJ-2019-137441.
[17] T. A. Ahanger, M. Bhatia, and A. Aljumah, "Game-theoryinspired novel mechanism for assessing healthcare quality," IEEE Internet Things J., vol. 9, no. 24, pp. 24856–24865, Dec. 2022, doi: 10.1109/JIOT.2022.3194854.
[18] J. Liu, P. J. Allen, L. Benz, D. Blickstein, E. Okidi, and X. Shi, "A machine learning approach for recruitment prediction in clinical trial design," 2021, arXiv:2111.07407.
[19] D. S. Robertson, K. M. Lee, B. C. López-Kolkovska, and S. S. Villar, "Response-adaptive randomization in clinical trials: From myths to practical considerations," Stat. Sci., vol. 38, no. 2, p. 185, May 2023.
[20] J. C. Brehautet al., "Using behavioral theory and shared decision-making to understand clinical trial recruitment: Interviews with trial recruiters," Trials, vol. 22, p. 298, 2021, doi: 10.1186/S13063-021-05257-X.