Problem Statement – 3rd

Predicting completion of clinical studies with explainability

Team Name - pushadapumohansai

Name - Pushadapu Mohana Venkata Siva Naga Sai
Contact Number – 7416698427
Email ID - pushadapumohansai@gmail.com

Approach & methodology

Overview

- Clinical trials often face delays or incompletion due to various influencing factors.
- Predicting trial completion helps optimize design, investments, and resource allocation.
- Historical data reveals patterns impacting trial success or failure.
- Both structured and unstructured data require analysis for effective predictions.
- Uncompleted trials can be classified as Suspended, Withdrawn, or Terminated.
- Validate model results to ensure alignment with clinical domain insights.
- Enable better trial design and decision-making using explainable AI solutions.
- Evaluate models using precision, recall, F1, confusion matrix, and AUC-ROC.

Methodology

- ClinicalTrials.gov data (~450,000 trials) includes both structured and unstructured trial-related features for analysis.
- Missing data is handled, text fields cleaned, numerical variables normalized, and class imbalance addressed.
- Features like trial phase, conditions, criteria, and amendments frequency are crucial for predictive insights.
- Precision, Recall, F1 Score, and AUC-ROC metrics manage imbalanced data and evaluate model performance.
- Predict trial status while improving trial design efficiency and reducing risks in R&D processes.
- Models compared include baseline algorithms, advanced methods, and explainable AI frameworks like SHAP and Causal Inference.

Framework / tools used

- TensorFlow is used for building deep learning models due to its flexibility and scalability.
- PyTorch is leveraged for its dynamic computation graph, ideal for experimentation and NLP tasks.
- scikit-learn provides robust tools for preprocessing, feature selection, and baseline model comparisons.
- Transformers (Hugging Face) are employed for processing unstructured text fields like criteria and descriptions.
- SHAP explains model predictions by calculating feature contributions, improving interpretability and trust.
- Matplotlib and Seaborn are utilized for EDA and visualizing insights from data and model results.

Model choice & setup

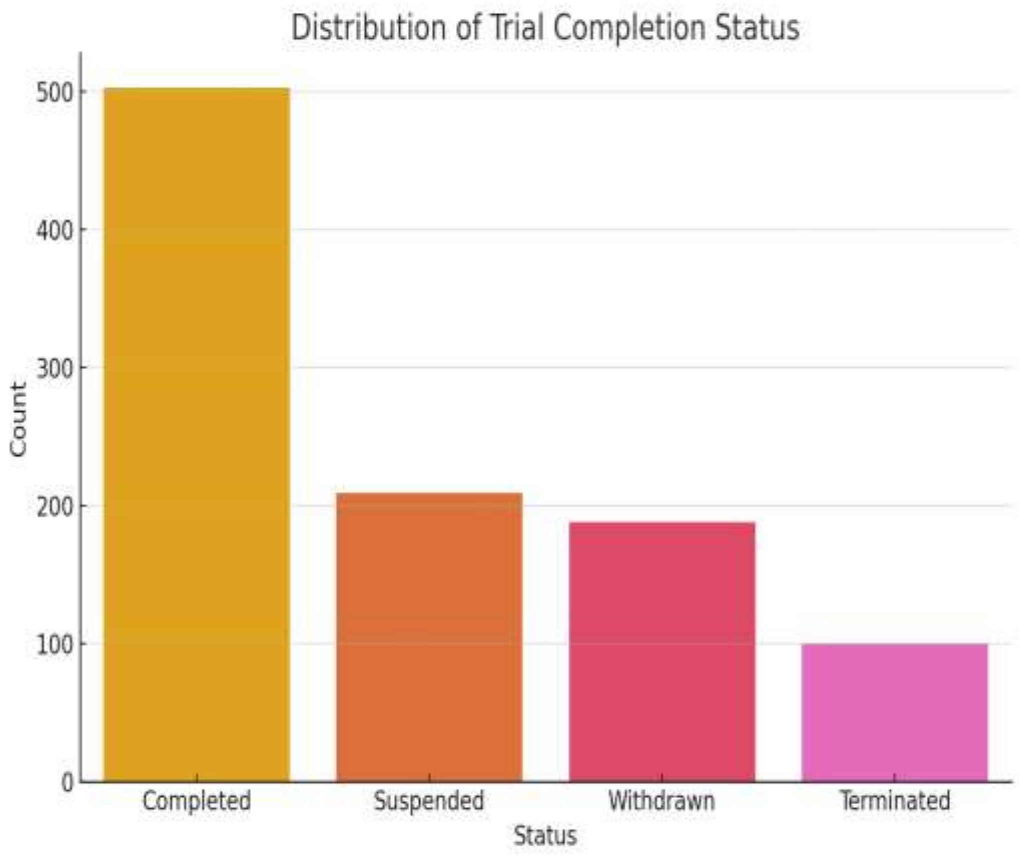
Model Selection	Model Architecture
• Logistic Regression is chosen as a baseline model for its simplicity and interpretability in classification. • Random Forest handles structured data well and provides feature importance for explainability. • Deep Learning Models (e.g., Feedforward Networks) are used for their ability to capture complex relationships. • Transformer-based Models (e.g., BERT) process unstructured text like criteria and descriptions effectively. • Ensemble Methods combine multiple model predictions to improve accuracy and robustness. • Explainable AI Tools (e.g., SHAP) ensure model outputs align with clinical trial domain insights.	• Data Ingestion: Load ClinicalTrials.gov data into the pipeline • Preprocessing Layer: Handle missing values and normalize data. • Feature Engineering: Create features like complexity scores, duration, and embeddings for text. • Data Splitting: Split data into training, validation, and test sets • Model Layer: Apply models like XGBoost and Neural Networks. • Explainability Module: Use SHAP and causal inference • Evaluation: Assess performance using Precision, Recall, F1, and AUC-ROC. • Deployment: Package the pipeline for integration with clinical workflows.

Model Training & Evaluation

Evaluation Metrics

- Model Training Process: Split data into training and validation sets, use cross-validation for model selection.
- Preprocessing: Apply feature scaling, handle missing data, and encode categorical variables before training.
- Model Fitting: Train models like XGBoost, Random Forest, and Neural Networks using the training set.
- Hyperparameter Tuning: Use GridSearchCV or RandomizedSearchCV for optimizing model hyperparameters.
- Evaluation Criteria: Evaluate performance on the validation set, considering overfitting/underfitting.
- Key Metrics: Assess model performance using Precision, Recall, F1, AUC-ROC, and accuracy.
- Root Mean Square Error (RMSE): Measures the model's prediction error.
- Mean Absolute Error (MAE): Provides average prediction error, easy to interpret and outliers.
- R-squared (R^2) Score: Represents the proportion of variance explained by the model, measuring fit quality.
- Final Evaluation: Test the model on the test set, assess generalization using the selected metrics.

Reports and Visualizations



Model Outcoes

- **Model Interpretation: Highlight key features influencing predictions, supported by SHAP or feature importance charts.**
- **Key Findings: Trials with complex criteria or higher amendments correlate with "Not Completed" status.**
- **Model Performance: Present Precision, Recall, F1 Score, and AUC-ROC values to demonstrate model reliability.**
- **Comparison Insights: Compare models (e.g., XGBoost vs. Transformers) to identify the best-performing approach.**
- **Implications: Findings aid in improving trial design, reducing failures, and optimizing R&D investments.**
- **Visual Aids: Use ROC Curve, confusion matrix, and feature importance plots to explain outcomes clearly.**
- **Error Analysis: Present insights from misclassified cases to refine model understanding and domain alignment.**
- **Summary Graphs: Use bar and pie charts to summarize prediction distributions and key metrics visually.**