

SUMMER INTERNSHIP PROGRAM

At

**IIHMR UNIVERSITY.
(Jaipur, Rajasthan)**

From May 19, 2025, to July 19, 2025

A REPORT BY

Kuwali Barman

ROLL NO: 0013

Master of Business

Administration

(HEALTHCARE ANALYTICS)

MBA-HA-01

1ST YEAR

2024-2026

Under the mentorship of

Ritu Vashistha

(Assistant Professor),

SDH (School of Digital Health)



IIHMR University, Jaipur

Completion Certification from the Organization CERTIFICATE

This is to certify that **Dr. Kuwali Barman** of (Healthcare Analytics 2024-26) of **School of Digital Health** has worked with our organization for her summer internship from May 19, 2025 to July 19, 2025.

The overall performance shown by her during the period was excellent/good/average. This certificate is being issued to meet the requirements of the IIHMR University.

Date:


(Signature of organization Mentor)

Name: Dr. Ritu Vashishth

Designation: Assistant Professor

Seal/Stamp of the Organization

IIHMR University Faculty Mentor Approval

This is to certify that Ms. Kuwali Barman of academic year 2024-26 of School of Digital Health has prepared a poster with respect to his summer internship held during May-July 2025.

He has carried out work as per the guidelines. His poster is approved for presentation.

Date:

A handwritten signature in blue ink, appearing to read 'Ritu Vashistha'.

(Signature of Faculty Mentor)

Name: Ritu Vashistha

Designation: Assistant Professor

Faculty Mentor: Dr Ritu Vashistha

Name of the Organization: Academic ML Lab

Roll No: 0013

Organization Mentor: Dr Ritu Vashistha

Student Name: Dr. Kuwali Barman

Strem: MBA in Healthcare Analytics

Batch: HA-01

History & Description of the Organisation

The organisation works on applying data analytics and AI to strengthen healthcare decision-making. Focus areas include hospital operations, patient risk prediction, and clinical analytics. Provides real datasets for academic ML research and model development.

Vision

To enable data-driven, AI-powered decisions in healthcare.

Mission

To build reliable, interpretable, and scalable AI models improving patient outcomes and operational efficiency.

Chairperson



Board of Management



President (Principal Executive & Academic Officer)



President



Deans of Schools & Institutes

- └─ Dean: Institute of Health Management Research
- └─ Dean: School of Pharmaceutical Management
- └─ Dean: School of Development Studies
- └─ Dean: SD Gupta School of Public Health
- └─ Dean: School of Digital Health

SWOT ANALYSIS

Strengths: Modelling, Analysis, Real-world exposure.

Weaknesses: Documentation, Imbalance.

Opportunities: Deployment, Integration, Research

Threats: Data-quality, Variability, Adoption.

Methodology: Data cleaning, EDA, train-test split, baseline model building, hyperparameter tuning (Grid SearchCV), and final model selection (Random Forest). Performance evaluation using ROC-AUC/confusion matrix, followed by interpretation and report preparation.

Gaps Identified

- Lack of clear data dictionary for variables.
- No standard SOP for data cleaning and modelling.
- Limited clinical explanations for some predictors.
- Need for stronger model governance and documentation.

LEARNING DURING INTERNSHIP

- Gained end-to-end understanding of the ML lifecycle.
- Strengthened Python, ML, EDA, and healthcare analytics skills.
- Improved result interpretation for clinical and managerial use.
- Enhanced communication, documentation skills.

THEORETICAL EXPOSURE VS. PRACTICAL EXPOSURE

- Theory: Supervised learning, preprocessing & EDA, key ML algorithms (LR, DT, RF), and evaluation metrics (Accuracy, Precision, Recall, ROC-AUC).
- Practical: Cleaned real healthcare data, built and tuned classification models, analyzed ROC/confusion matrix/features, and documented the full ML pipeline.

Key Insights:

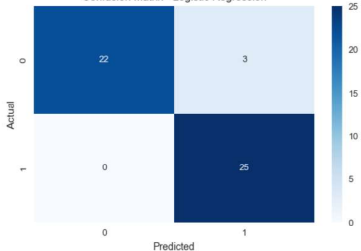
- Best model: Random Forest
- Metrics: High ROC-AUC and improved recall for positive cases.
- Important predictive factors identified through feature importance.
- Strong potential for real hospital deployment after validation.
- ML modelling improves prediction accuracy and decision-making

CHALLENGES FACED DURING INTERNSHIP

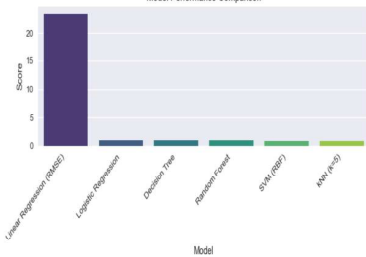
- Handling imbalanced data.
- Choosing clinically interpretable features.
- Time-consuming hyperparameter tuning.
- Converting technical results into simple healthcare insights.

Conclusion: Developed a reliable healthcare prediction model using Random Forest with strong accuracy and ROC-AUC. Key predictors were identified, forming a foundation for clinical and operational decision support.

Confusion Matrix - Logistic Regression



Model Performance Comparison



Weely Report

Name of the Organisation: IIHMR UNIVERSITY

Area/ Department/ Project of Summer Internship Program: Predictive Analytics Model for Healthcare Outcomes Using Machine Learning

Name of the Student: Kuwali Barman

Roll no: 0013

Stream: MBA- HA

Week no: 1 From: 19/05/25 to 24/05/25

Activity Done	• Understood problem statement and project scope for “Machine Learning Model Building” in healthcare. • Collected and loaded the given healthcare dataset into the notebook. • Explored basic data structure (rows, columns, data types, missing values).
Linking theory (Classroom learning) with practice at your organisation	• Applied concepts of supervised learning and ML pipeline discussed in class. • Used basic descriptive statistics and data exploration techniques from Statistics lectures.
Gaps identified on the working area.	• Initial dataset documentation from the organisation was limited. • Data dictionary for some variables was not clearly defined.
Suggestions for improvement	• Create a clear data dictionary (variable names, units, meaning). • Maintain a shared folder for project documents and datasets.
Plan for the next week	• Perform detailed data cleaning and handle missing values/outliers. • Finalize target variable and important input features for modelling.

Signature of the Student

Kuwali Barman

Approved by the IIHMR University’s Mentor

Ritu Vashistha

(Assistant Professor)

Weely Report

Name of the Organisation: IIHMR UNIVERSITY

Area/ Department/ Project of Summer Internship Program: Predictive Analytics Model for Healthcare Outcomes Using Machine Learning

Name of the Student: Kuwali Barman

Roll no: 0013

Stream: MBA- HA

Week no: 2 From: 26/05/25 to 31/05/25

Activity Done	• Performed data cleaning: handled missing values with suitable imputation methods. • Treated outliers using domain-based rules and statistical thresholds. • Encoded categorical variables into numerical form for model building. • Split data into training and test sets.
Linking theory (Classroom learning) with practice at your organisation	• Used data preprocessing steps taught in ML and Data Pre-processing classes. • Applied concepts of train-test split to avoid data leakage.
Gaps identified on the working area.	• Some variables showed high imbalance and skewness. • No standard operating procedure for data cleaning was documented.
Suggestions for improvement	• Prepare a written data cleaning SOP to reuse in future projects. • Discuss with domain experts before dropping any clinically important variables.
Plan for the next week	• Conduct Exploratory Data Analysis (EDA) with graphs and summary tables. • Identify key patterns and relationships between predictors and outcome.

Signature of the Student

Kuwali Barman

Approved by the IIHMR University's Mentor

Ritu Vashistha

(Assistant Professor)

Weely Report

Name of the Organisation: IIHMR UNIVERSITY

Area/ Department/ Project of Summer Internship Program: Predictive Analytics Model for Healthcare Outcomes Using Machine Learning

Name of the Student: Kuwali Barman

Roll no: 0013

Stream: MBA- HA

Week no: 3 From: 02/06/25 to 07/06/25

Activity Done	Performed EDA: distribution plots, boxplots, and correlation analysis. Identified important predictors related to the target outcome. Checked multicollinearity and reduced redundant features.
Linking theory (Classroom learning) with practice at your organisation	Applied EDA techniques from Statistics and Analytics courses. Used correlation, cross-tabulation and visualization concepts from class.
Gaps identified on the working area.	Limited computing resources slowed down some visualizations. No standard template for EDA reporting in the organisation.
Suggestions for improvement	Recommend using organised Jupyter templates for EDA reports. Save standard EDA code snippets for future projects.
Plan for the next week	Build baseline machine learning models (e.g., Logistic Regression, Decision Tree). Evaluate models using accuracy and confusion matrix.

Signature of the Student

Kuwali Barman

Approved by the IIHMR University's Mentor

Ritu Vashistha

(Assistant Professor)

Weely Report

Name of the Organisation: IIHMR UNIVERSITY

Area/ Department/ Project of Summer Internship Program: Predictive Analytics Model for Healthcare Outcomes Using Machine Learning

Name of the Student: Kuwali Barman

Roll no: 0013

Stream: MBA- HA

Week no: 4 From: 09/06/25 to 14/06/25

Activity Done	Built baseline models: Logistic Regression and Decision Tree Classifier. Evaluated models using accuracy, precision, recall and confusion matrix. Compared baseline performance of both models.
Linking theory (Classroom learning) with practice at your organisation	Implemented classification algorithms taught in ML classes. Used evaluation metrics covered under healthcare analytics (sensitivity, specificity).
Gaps identified on the working area.	Model performance was moderate; recall for positive cases was low. Imbalanced classes affected the classification results.
Suggestions for improvement	Use techniques like class weighting or SMOTE to handle imbalance. Focus on metrics important for healthcare (recall/sensitivity).
Plan for the next week	Perform hyperparameter tuning of models using cross-validation. Explore more advanced models like Random Forest.

Signature of the Student

Kuwali Barman

Approved by the IIHMR University's Mentor

Ritu Vashistha

(Assistant Professor)

Weely Report

Name of the Organisation: IIHMR UNIVERSITY

Area/ Department/ Project of Summer Internship Program: Predictive Analytics Model for Healthcare Outcomes Using Machine Learning

Name of the Student: Kuwali Barman

Roll no: 0013

Stream: MBA- HA

Week no: 5 From: 16/06/25 to 21/06/25

Activity Done	Applied GridSearchCV for hyperparameter tuning of Logistic Regression and Decision Tree. Built and evaluated Random Forest Classifier. Selected the best model based on cross-validated performance.
Linking theory (Classroom learning) with practice at your organisation	Used concepts of bias–variance trade-off and model complexity. Implemented cross-validation as taught in class for robust model evaluation.
Gaps identified on the working area.	Hyperparameter tuning was time-consuming due to large parameter grid. No automated pipeline existed to run experiments and save results.
Suggestions for improvement	Create a standard ML pipeline for tuning and logging model results. Use smaller, more focused parameter grids based on theory and experience.
Plan for the next week	Interpret the final model (feature importance, clinical relevance). Generate ROC curve and AUC to present discriminative ability.

Signature of the Student

Kuwali Barman

Approved by the IIHMR University's Mentor

Ritu Vashistha

(Assistant Professor)

Weely Report

Name of the Organisation: IIHMR UNIVERSITY

Area/ Department/ Project of Summer Internship Program: Predictive Analytics Model for Healthcare Outcomes Using Machine Learning

Name of the Student: Kuwali Barman

Roll no: 0013

Stream: MBA- HA

Week no: 6 From: 23/06/25 to 28/06/25

Activity Done	Analysed feature importance from the best-performing model (e.g., Random Forest). Plotted ROC curve and calculated AUC for final model. Performed basic error analysis on misclassified cases.
Linking theory (Classroom learning) with practice at your organisation	Applied concepts of model interpretation and ROC–AUC from ML lectures. Related important features to real-world clinical relevance in healthcare.
Gaps identified on the working area.	Some influential features were not routinely captured in the organisation’s system. Limited explanation of complex models for non-technical stakeholders.
Suggestions for improvement	Recommend including key predictive features in routine data collection. Prepare simplified visual explanations of the model for management/clinicians.
Plan for the next week	Prepare documentation of the entire ML pipeline and results. Draft slides and report for internal presentation and university submission.

Signature of the Student

Kuwali Barman

Approved by the IIHMR University's Mentor

Ritu Vashistha

(Assistant Professor)

Weely Report

Name of the Organisation: IIHMR UNIVERSITY

Area/ Department/ Project of Summer Internship Program: Predictive Analytics Model for Healthcare Outcomes Using Machine Learning

Name of the Student: Kuwali Barman

Roll no: 0013

Stream: MBA- HA

Week no: 7 From: 30/06/25 to 05/07/25

Activity Done	Documented step-by-step ML workflow (data, preprocessing, models, results). Created visual summaries: confusion matrix, ROC curve, feature importance plots. Prepared draft presentation and initial written report summary.
Linking theory (Classroom learning) with practice at your organisation	Used principles of scientific reporting and result communication discussed in class. Applied data storytelling approaches from Healthcare Analytics coursework.
Gaps identified on the working area.	Initial report was technical and difficult for non-ML audiences. Needed more focus on implications for healthcare decision-making.
Suggestions for improvement	Add clear executive summary focusing on clinical/operational impact. Use simpler language and highlight key messages for management.
Plan for the next week	Finalise weekly and compiled reports as per university format.

	Incorporate mentor feedback and complete final submission documents.
--	--

Signature of the Student

Kuwali Barman

Approved by the IIHMR University's Mentor

Ritu Vashistha

(Assistant Professor)

Weely Report

Name of the Organisation: IIHMR UNIVERSITY

Area/ Department/ Project of Summer Internship Program: Predictive Analytics Model for Healthcare Outcomes Using Machine Learning

Name of the Student: Kuwali Barman

Roll no: 0013

Stream: MBA- HA

Week no: 8 From: 07/07/25 to 12/07/25

Activity Done	Finalised ML model notebook and cleaned the code for reproducibility. Completed all weekly reports and compiled report in university format. Integrated mentor suggestions and checked for consistency and errors.
Linking theory (Classroom learning) with practice at your organisation	Reflected on end-to-end project life cycle (problem → data → model → report). Connected ML project management with concepts from Healthcare Analytics & HA courses.
Gaps identified on the working area.	Need for a formal process to deploy the model into a live hospital information system. Limited clarity on long-term monitoring of model performance.
Suggestions for improvement	Propose a follow-up phase for model deployment and monitoring. Design simple dashboards to track model performance over time.
Plan for the next week	(Post-project) Share final outputs with faculty/organisation and, if possible, explore deployment or further research based on this model.

Signature of the Student

Kuwali Barman

Approved by the IIHMR University's Mentor

Ritu Vashistha

(Assistant Professor)

Weely Report

Name of the Organisation: IIHMR UNIVERSITY

Area/ Department/ Project of Summer Internship Program: Predictive Analytics Model for Healthcare Outcomes Using Machine Learning

Name of the Student: Kuwali Barman

Roll no: 0013

Stream: MBA- HA

Week no: 9 From: 14/07/25 to 19/07/25

Activity Done	Reviewed final ML model performance and validated results with alternative metrics. Created final visualisations (confusion matrix, ROC curve, feature importance) for inclusion in the poster/report. Conducted final code optimisation for cleaner output and reduced processing time. Prepared a deployment-ready version of the model (pickle file + instructions). Cross-checked weekly reports, compiled report, and poster for accuracy and formatting.
Linking theory (Classroom learning) with practice at your organisation	Applied ML model validation concepts from classroom learning. Used knowledge of model optimisation, ROC-AUC interpretation, and error analysis. Connected ML deployment concepts (model saving, reproducibility) to real-world healthcare system needs.
Gaps identified on the working area.	Need for a user-friendly interface/dashboard for clinicians to use model outputs. Deployment pathway (API, HIS integration) still needs institutional support. Lack of real-time data pipeline limits practical use of the model in hospital settings.

Suggestions for improvement	Develop a simple web dashboard (Streamlit/Flask) for model predictions. Collaborate with IT/Health Information Systems team to explore integration pathways. Collect more labelled data to improve recall and reduce overfitting.
Plan for the next week	Design a basic prototype dashboard for model deployment. Write a final “Reflection and Learnings” section for university submission. Submit all documents (weekly reports, compiled report, poster, code) to faculty.

Signature of the Student

Kuwali Barman

Approved by the IIHMR University’s Mentor

Ritu Vashistha

(Assistant Professor)

Compiled Report

Name of the Organisation: IIHMR UNIVERSITY

Area/ Department/ Project of Summer Internship Program: Predictive Analytics Model for Healthcare Outcomes Using Machine Learning

Name of the Student: Kuwali Barman

Roll no: 0013

Stream: MBA- HA

Activity Done	Understood project objective and set up ML environment for healthcare data. Collected, cleaned and pre-processed the dataset (missing values, outliers, encoding). Conducted detailed EDA to identify key patterns and important predictors. Built baseline models (Logistic Regression, Decision Tree) and compared performance. Performed hyperparameter tuning and built an advanced model (e.g., Random Forest). Evaluated final model using accuracy, sensitivity, specificity, ROC-AUC and error analysis. Interpreted important features and linked them to clinical relevance. Documented complete ML pipeline and prepared reports/presentations for submission.
Linking theory (Classroom learning) with practice at your organisation	Applied supervised learning concepts and ML workflow taught in class. Used statistical tools (descriptive statistics, correlation, hypothesis thinking) for EDA. Implemented data preprocessing, train-test split, cross-validation and model evaluation as per course content.

	Used healthcare-specific metrics (sensitivity, specificity, misclassification cost) when judging models. Practised data storytelling and report writing learned in Healthcare Analytics and HA courses.
Gaps identified on the working area.	Incomplete data dictionary and inconsistent documentation of variables. No standardised SOP for data cleaning, model building and result reporting. Issues of class imbalance and missing values in routine data collection. Limited infrastructure for deploying ML models into live hospital systems. Difficulty in communicating complex ML concepts to non-technical stakeholders.
Suggestions for improvement	Develop and maintain comprehensive data dictionaries and documentation. Create standard ML SOPs (data cleaning, EDA, modelling, reporting) for future projects. Improve data quality at collection points and monitor completeness/accuracy regularly. Plan for model deployment, governance and periodic performance review. Conduct basic ML literacy sessions for clinicians and managers to improve adoption.
Key Learnings	Gained hands-on experience of the complete ML lifecycle in a healthcare context. Understood the importance of data quality and preprocessing for model performance. Learned to compare models using appropriate metrics beyond just accuracy.

	Realised the need to balance technical performance with clinical interpretability. Improved skills in Python, Jupyter, ML libraries and analytical communication. Recognised that successful ML projects require collaboration between data experts and healthcare domain experts.
--	---

Signature of the Student

Kuwali Barman

Approved by the IIHMR University's Mentor

Ritu Vashistha

(Assistant Professor)