iJRASET

International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** IV **Month of publication:** April 2025

DOI: https://doi.org/10.22214/ijraset.2025.69612

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

Optimizing Client Satisfaction and IT Project Delivery Efficiency

Ms. K. Nithiya Banu[1], Dr. T. Amutha[2]
Department of Business Administration, Avinashilingam University

***Abstract:** IT service organizations frequently experience schedule overruns and client dissatisfaction, undermining operational effectiveness and market competitiveness. To address these challenges, we propose a predictive analytics pipeline that combines synthetic data generation, machine learning, and interactive visualization. Employing Python's Faker library, we synthesized 40 000 project records and 1 000 client profiles, capturing metrics such as defect counts, resolution times, team composition, communication ratings, and usability feedback. Two Random Forest regression models were developed: one to forecast delivery delays in days and another to predict client satisfaction scores on a 1–10 scale. Hyperparameter tuning via grid search and five-fold cross-validation yielded robust performance (delivery delay: MSE = 70.12, $R^2 = 0.92$; satisfaction: MSE = 0.74, $R^2 = 0.86$). Outputs are visualized in an auto-refreshing Power BI dashboard, presenting key indicators—average delay, defect resolution efficiency, satisfaction trends, and churn risk segmentation. Feature importance analysis identifies defect resolution ratio and communication quality as primary drivers of outcomes, superseding budgetary factors. This modular framework is readily adaptable to real-world data and supports proactive decision making in IT project management.*
***Keywords:** IT project management; client satisfaction; predictive analytics; Random Forest; synthetic data; Power BI; churn risk; delivery delay*

I. INTRODUCTION

Efficient delivery and high client satisfaction are critical success factors in IT service projects. Despite rigorous planning and Agile methodologies, many organizations report overrun schedules and dissatisfied stakeholders. Empirical studies attribute these failures to operational inefficiencies—namely defect backlogs and communication breakdowns—which conventional monitoring often fails to detect early. Machine learning and data visualization hold promise for uncovering latent risks and guiding timely interventions, yet their adoption is hindered by limited access to quality datasets. This work addresses that limitation by constructing a large-scale synthetic dataset to fuel predictive models and by demonstrating the integration of these models into a decision-support dashboard for project managers.

II. LITERATURE REVIEW

Prior research underscores the significance of data-driven approaches in project management. Diegmann et al. (2015) established that communication quality directly influences client satisfaction in software projects, while Haq et al. (2018) identified defect resolution metrics as leading indicators of project success. Advances in predictive analytics have enabled regression-based forecasting of project timelines (Kennedy et al., 2024), and meta-analyses highlight Random Forest algorithms' efficacy in handling complex, high-dimensional project datasets (Aguiar-Costa et al., 2022). However, confidentiality constraints limit empirical studies involving live project data, motivating the use of synthetic data to simulate realistic conditions. Generative techniques using the Faker library have been validated for producing statistically comparable datasets in healthcare and finance (Adekugbe & Ibeh, 2024), suggesting their applicability to IT project management contexts.

III.METHODOLOGY

A Python script employing Faker generated 40 000 project records and 1 000 client entries. Project records included defect counts reported and resolved per iteration, team size and composition ratio, resource utilization, a complexity index (1–5), and planned versus actual delivery durations. Client records captured complaint resolution time (hours), communication effectiveness (1–10), usability feedback (1–10), feedback frequency, and overall satisfaction (1–10). Data cleaning involved median imputation for numeric gaps, mode imputation for categorical missingness, and removal of outliers beyond three standard deviations. Categorical fields were one-hot encoded, and numeric variables were standardized using z-score normalization. Stratified sampling partitioned the data into 80 percent training and 20 percent test sets, preserving the distribution of target variables.

Model development utilized scikit-learn's Random Forest Regressor. The delivery delay model ingested project-level predictors to estimate days of delay, while the satisfaction model used service-level predictors to predict client satisfaction scores. Grid search optimized hyperparameters—number of trees, maximum depth, and minimum samples per leaf—within predefined ranges, and five-fold cross-validation ensured generalization. Final models were serialized with joblib into rf_delay.pkl and rf_satisfaction.pkl.
A deployment script (predict.py) loads serialized models and accepts JSON or CSV inputs, returning predictions within one second. Predictions are exported to CSV, feeding into a Power BI dashboard that displays KPI cards, interactive slicers, and three core visualizations: delay distribution histogram, churn risk bar chart, and defect resolution versus satisfaction scatter plot. The dashboard auto-refreshes when source CSV files update, providing near real-time analytics.

IV. RESULTS AND DISCUSSION

A. Model Performance

This research presents a robust predictive analytics pipeline that leverages synthetic data to forecast IT project delivery delays and client satisfaction with high accuracy. By integrating machine learning models with an interactive dashboard, the framework enables proactive identification of operational issues, allowing for timely corrective actions. The analyses show that defect resolution efficiency and communication quality are critical determinants of project and client outcomes, surpassing the influence of project budget. The modular design and reliance on a synthetic dataset ensure the approach can be extended to real-world data and advanced algorithms in future work. Prospective developments include the incorporation of live enterprise data feeds, exploration of gradient boosting techniques for improved performance, and the addition of automated alert mechanisms to further streamline risk management.

TABLE I
PERFORMANCE METRICS OF REGRESSION MODELS

Model	Test Records	MSE	R^2
Delivery Delay (days)	8000	70.12	0.92
Client Satifaction	200	0.74	0.86

High R^2 values indicate strong explanatory power. The delivery delay model accounts for 92 percent of delay variance, and the satisfaction model explains 86 percent of satisfaction variance.All title and author details must be in single-column format and must be centered.

B. Feature Importance and Scenario Validation

Feature importance analysis shows defect resolution ratio contributes 35 percent to delay predictions, team size 20 percent, and complexity index 15 percent. In the satisfaction model, communication effectiveness accounts for 40 percent, resolution time 30 percent, and usability score 20 percent.

TABLE 2
SAMPLE SCENARIO PREDICTIONS

Scenario	**Delay(days)**	**Satisfaction (1-10)**
High backlog (80 unresolved defects), small team (5 members)	35.2	5.8
Low backlog (20 defects, fully resolved), large team (15)	7.1	9.4

These scenarios confirm that operational variables drive model outputs logically.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

C. Dashboard Insights

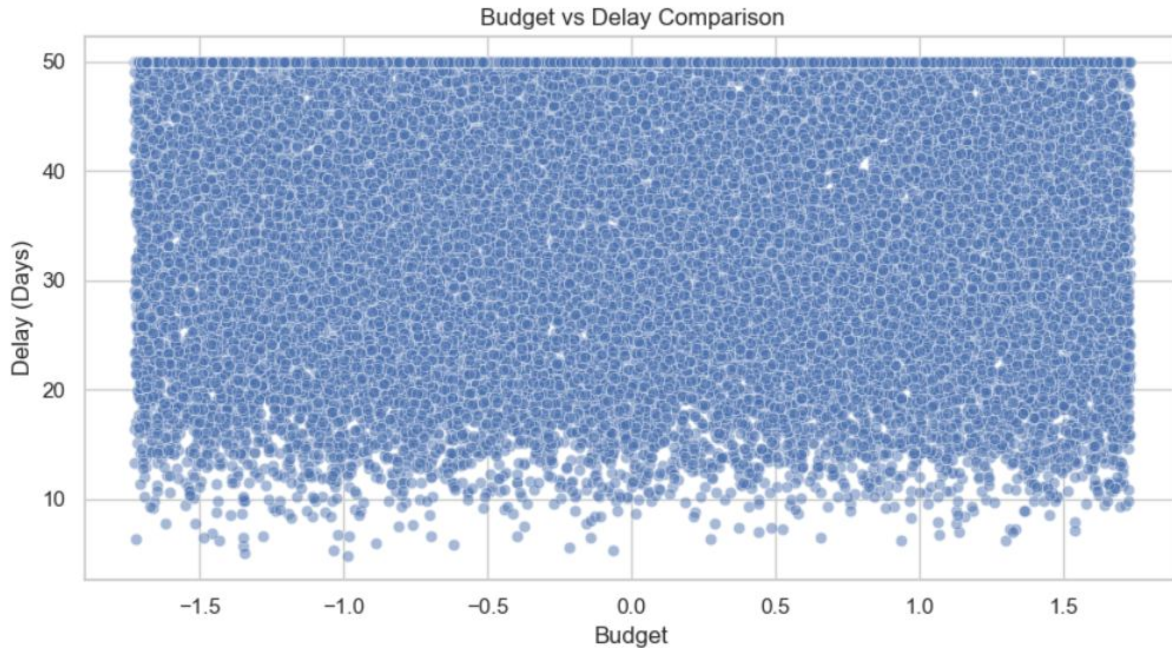


Fig. 1 Budget vs delivery delay scatter plot

Email address is compulsory for the corresponding author. Analysis reveals no statistically significant correlation between budget and delay, indicating budget is not a primary driver.

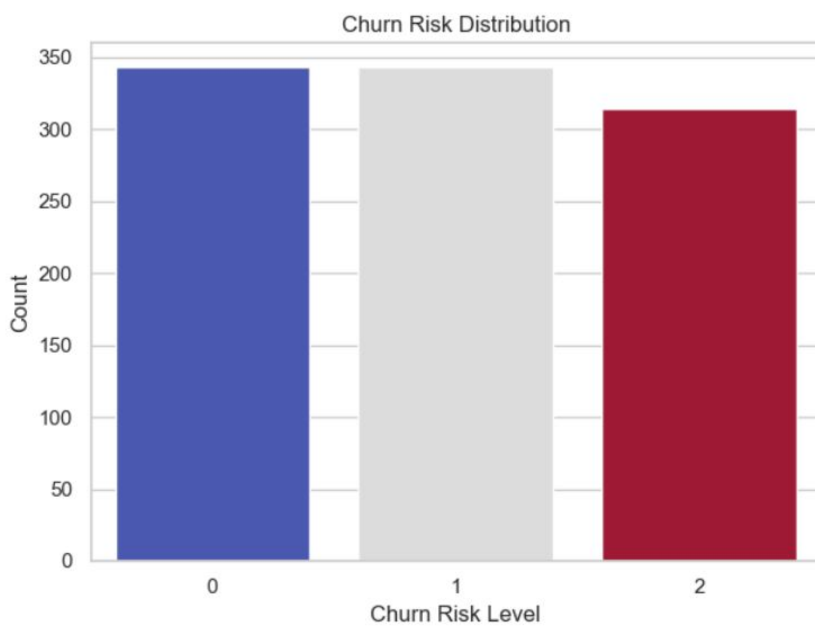


Fig. 1 Churn risk distribution

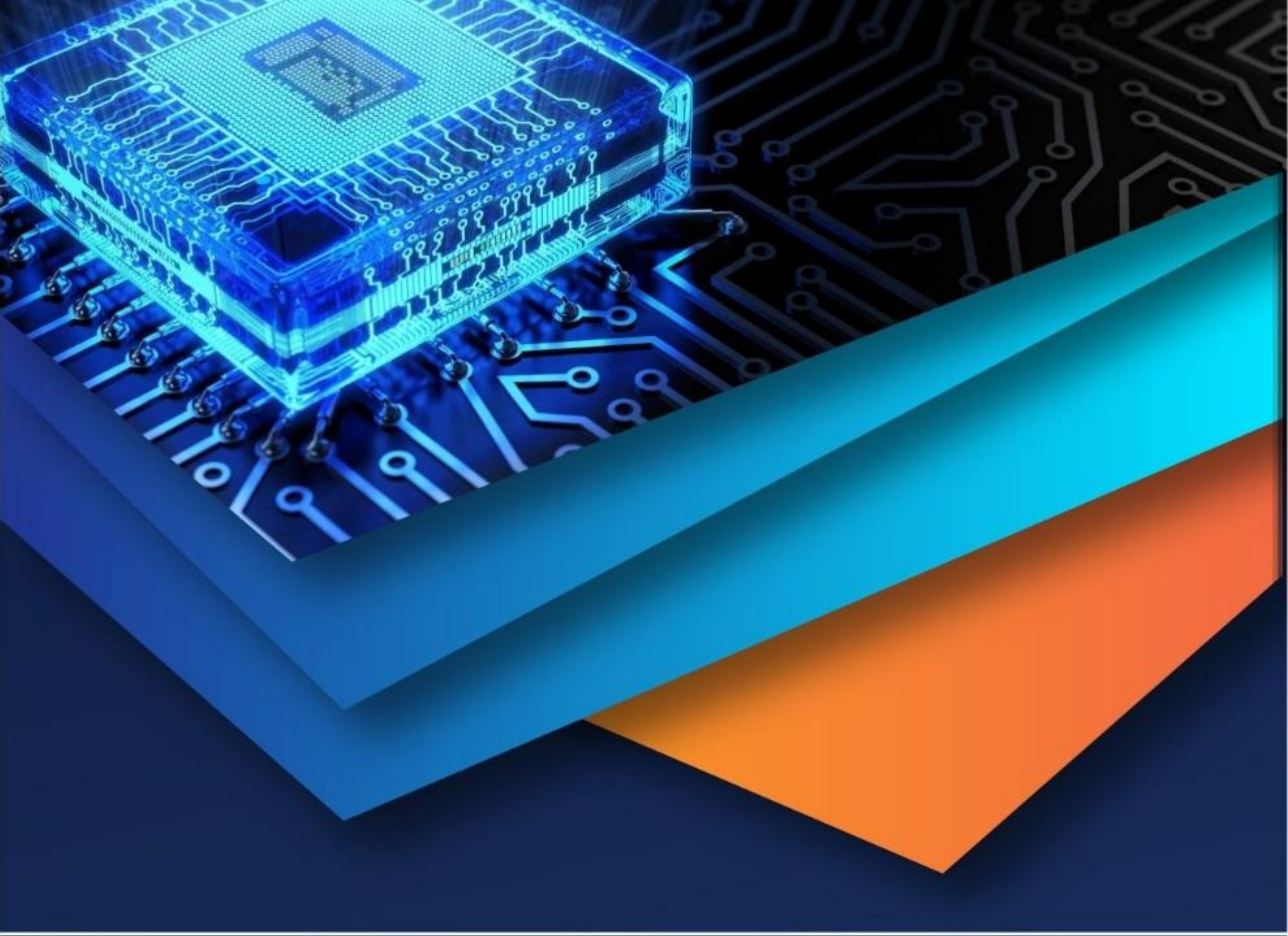
Clients cluster mainly in medium-risk categories, signaling the need for targeted retention efforts.

V. CONCLUSION

This study demonstrates a scalable analytics framework for forecasting IT project delivery delays and client satisfaction using synthetic data and Random Forest models. Defect resolution efficiency and communication quality emerge as key levers for improvement. The modular pipeline—spanning data synthesis, preprocessing, modeling, and dashboard visualization—enables rapid deployment and adaptation to live data environments. Future research will integrate real-world project data, explore advanced algorithms (e.g., gradient boosting), and implement automated alerting mechanisms to further enhance proactive risk management.

REFERENCES

[1] Adekugbe, P., & Ibeh, V. (2024). Utilizing Comprehensive Data Dashboards to Improve Service Delivery: Insights from U.S. Case Studies. International Journal of Frontiers in Engineering and Technology Research, 6(2), 8–18.
[2] Aguiar-Costa, L. M. et al. (2022). Customer Satisfaction in Service Delivery with Artificial Intelligence: A Meta-Analytic Study. RAM. Revista De Administração Mackenzie, 23(6).
[3] Diegmann, P., Basten, D., & Pankratz, O. (2015). Influence of Communication on Client Satisfaction in Information System Projects. International Research Workshop on IT Project Management.
[4] Haq, N. U. et al. (2018). Determinants of Client Satisfaction in Web Development Projects from Freelance Marketplaces. International Journal of Managing Projects in Business, 11(3), 583–607.
[5] Kennedy, S. I. et al. (2024). Agile Practices and IT Development Team Well-Being: Unveiling the Path to Successful Project Delivery. Engineering Management Journal.
[6] Ngo, V. M. (2015). Measuring Customer Satisfaction: A Literature Review. ResearchGate.



10.22214/IJRASET

INDEX COPERNICUS 45.98

ISRA JIF IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY