Journal of Information Systems Engineering and Management
2025, 10(23s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Improving ERP Adoption Through Predictive Modeling: A Data-Driven Recommendation System

Sunil Kumar Mishra [1], Dr. Atul Dattatray Newase[2]
[1] *Research Scholar, Department of Computer Application, Dr. A.P.J. Abdul Kalam University, Indore, India, sunilnit1972@gmail.com*
[2] *Research Guide, Department of Computer Application, Dr. A.P.J. Abdul Kalam University, Indore, , India, dratulnewase@gmail.com*

ARTICLE INFO

Received: 05 Oct 2024
Revised: 05 Dec 2024
Accepted: 22 Dec 2024

ABSTRACT

Enterprise Resource Planning (ERP) adoption remains a critical challenge for organizations due to cost, employee resistance, and operational inefficiencies. This research presents a data-driven predictive modeling framework that leverages feature engineering, dimensionality reduction, and machine learning algorithms to enhance ERP satisfaction and adoption. The study applies Recursive Feature Elimination (RFE) for feature selection, Principal Component Analysis (PCA) for dimensionality reduction, and interaction terms to improve interpretability. Seven machine learning models, including Random Forest, Gradient Boosting, Support Vector Machine, Logistic Regression, K-Nearest Neighbors, XGBoost, and LightGBM, are trained and evaluated using cross-validation with Stratified K-Fold. Model performance is assessed through accuracy, precision, recall, and F1-score, with SHAP and Permutation Importance ensuring interpretability. The best-performing model is used to predict future ERP adoption trends, and recommendations are derived based on key influencing factors. Visualizations such as feature importance plots, confusion matrices, and impact assessments are generated to provide actionable insights. The proposed system aids in optimizing cost, enhancing employee training, and streamlining organizational processes, ensuring higher ERP adoption rates and long-term operational efficiency.

Keywords: ERP Adoption, Predictive Modeling, Feature Engineering, SHAP Analysis, Machine Learning, Dimensionality Reduction, Model Evaluation

1. INTRODUCTION

Enterprise Resource Planning (ERP) systems have become indispensable for modern organizations, integrating critical business processes such as finance, supply chain, human resources, and operations. However, despite their transformative potential, ERP adoption remains a challenge for many organizations due to factors such as high costs, user resistance, inadequate training, and system complexity. A significant percentage of ERP implementations fail to deliver expected benefits, making it essential to develop data-driven approaches to improve adoption success. Predictive modeling offers a powerful solution by leveraging machine learning techniques to identify key factors influencing ERP satisfaction, forecast adoption trends, and provide actionable recommendations for organizations.

This research introduces a predictive modeling framework that applies feature selection, dimensionality reduction, and advanced machine learning models to analyze ERP satisfaction and provide data-driven insights. The study integrates Recursive Feature Elimination (RFE) to filter out less significant features, while Principal Component Analysis (PCA) reduces dimensionality, preserving essential variance for better interpretability. Additionally, interaction terms are introduced to enhance model accuracy by capturing non-linear relationships between influencing factors. By utilizing multiple machine learning models such as Random Forest, Gradient Boosting, Support Vector Machine, Logistic Regression, K-Nearest Neighbors, XGBoost, and LightGBM, the system determines the best predictive model for ERP adoption.

To ensure model reliability and fairness, we employ SHAP (SHapley Additive Explanations) and Permutation Importance to interpret feature contributions. Additionally, Stratified K-Fold cross-validation is applied to prevent overfitting and improve model generalizability. The best-performing model is then used to predict future ERP adoption trends, identifying whether an organization is likely to experience a positive or negative impact from its ERP implementation. Based on these predictions, data-driven recommendations are generated to optimize ERP adoption strategies, including cost reduction measures, employee training programs, and process optimization techniques.

The findings of this study reveal that employee engagement, training quality, and cost efficiency are among the most influential factors affecting ERP satisfaction. By analyzing feature importance, organizations can prioritize critical success factors and implement targeted strategies for smoother adoption. The model's recommendation system provides actionable insights, ensuring that decision-makers have a structured approach to mitigating risks and maximizing ERP benefits. Additionally, future trend predictions indicate an 80% likelihood of positive ERP adoption outcomes, reinforcing the effectiveness of a data-driven strategy in improving system acceptance.

By integrating predictive modeling with real-world ERP adoption challenges, this study contributes a novel approach to ERP implementation success. The proposed system enables organizations to move beyond traditional trial-and-error adoption methods, shifting towards a more proactive, evidence-based decision-making process. With the increasing availability of enterprise data and advances in machine learning, predictive analytics will continue to redefine ERP implementation strategies, making systems more adaptive and user-centric. Ultimately, this research underscores the potential of data-driven recommendation systems in revolutionizing ERP adoption, minimizing implementation failures, and ensuring long-term organizational efficiency.

2. LITERATURE REVIEW

H. K. R. Rikkula et al., (2024), The integration of Artificial Intelligence (AI), Machine Learning (ML), and Robotic Process Automation (RPA) is reshaping Enterprise Resource Planning (ERP) systems, addressing traditional challenges such as manual processes and data silos. These technologies enable intelligent data mapping, predictive maintenance, and automated data migration, enhancing overall efficiency. However, key adoption challenges include data quality concerns, skill gaps, and security risks. To ensure successful implementation, organizations must adopt best practices such as pilot projects, structured training programs, governance frameworks, and continuous optimization. This study provides a strategic roadmap for leveraging AI, ML, and RPA to enhance ERP integration and drive digital transformation [1].

T. Macron et al., (2025), AI is revolutionizing ERP systems by automating workflows, improving decision-making, and optimizing operations. As business complexities grow, AI-driven ERP solutions are becoming critical for enhancing business intelligence and resource management. This study explores AI's role in machine learning, natural language processing, predictive analytics, and automation within ERP systems. It examines emerging trends and predicts future developments, highlighting challenges and opportunities in scaling AI-powered ERP solutions. By understanding these dynamics, organizations can strategically adopt AI to improve agility, efficiency, and long-term competitiveness [2].

O. Samson et al., (2025), Traditional ERP systems struggle to manage the increasing complexity and volume of enterprise data. Machine Learning (ML) enhances ERP analytics by automating data processing, improving visualization, and enabling predictive modeling. This study evaluates ML's role in real-time decision support, highlighting its transformative potential in forecasting trends and optimizing business operations. While ML-driven ERP analytics improve efficiency and accuracy, implementation challenges such as data integration complexity and infrastructure requirements must be addressed. The study presents strategic recommendations for successful ML adoption in ERP systems [3].

G. Abbas et al., (2021), Cloud-based ERP systems, integrated with AI, ML, and Snowflake databases, offer real-time data analysis and enhanced decision-making. AI-driven automation optimizes data migration, improves consistency, and enables advanced trend prediction. Snowflake's scalable cloud architecture supports seamless data integration, improving resource allocation, supply chain management, and financial forecasting. This convergence allows businesses to leverage real-time insights for better operational efficiency and market adaptability. The study highlights how integrating ERP with cloud-based AI and ML models can drive business growth and competitive advantage [4].

Z. Asimiyu et al., (2025), AI-driven automation is revolutionizing ERP by streamlining processes, optimizing resources, and reducing operational costs. This study explores machine learning, natural language processing, and predictive analytics in ERP, demonstrating how AI improves business performance and decision-making. The paper also highlights challenges in AI adoption, including scalability, data governance, and ethical considerations. Looking ahead, AI's integration in ERP systems will continue to transform industries, enhancing efficiency, intelligence, and business adaptability in the digital economy [5].

A. Mahmood et al., (2023), The integration of IoT in manufacturing has enhanced operational efficiency, but managing vast data volumes remains a challenge. AI, ML, and ERP cloud solutions provide scalable and intelligent ways to analyze IoT data, offering predictive maintenance, anomaly detection, and production optimization. AI-driven analytics help forecast trends, prevent equipment failure, and automate workflows, improving decision-making and business intelligence. ERP cloud platforms enable seamless data storage, real-time visibility, and supply chain management, reducing the need for costly infrastructure investments. This integration transforms IoT-generated data into actionable insights, enhancing efficiency and operational intelligence [6].

H. Sadeeq et al., (2024), Advanced AI/ML techniques are reshaping business intelligence (BI) strategies in manufacturing by providing real-time predictive analytics, trend identification, and resource optimization. Machine learning models analyze IoT data to detect patterns, risks, and inefficiencies, improving supply chain performance, inventory management, and production scheduling. The integration of AI-powered BI tools with ERP cloud platforms ensures data flow across systems, automating decision-making and enhancing profitability. Predictive analytics further forecasts demand fluctuations, optimizes supply chains, and enhances inventory control, leading to smarter, more agile manufacturing operations [7].

G. Areo et al., (2025), AI-integrated ERP systems are transforming business operations by enhancing decision-making, resource allocation, and forecasting. AI-driven ERP solutions leverage machine learning, natural language processing, and predictive analytics to extract valuable business insights, optimize workflows, and improve customer management. Despite challenges such as integration complexity, data quality concerns, and infrastructure costs, AI adoption in ERP enhances strategic decision-making and business performance. Future ERP trends will focus on automation, blockchain integration, and cybersecurity enhancements, ensuring greater efficiency and security in enterprise management [8].

A. Mahmood et al., (2023), The convergence of AI, ML, and ERP cloud solutions with IoT-enabled manufacturing drives process automation, real-time monitoring, and predictive analytics. ML algorithms identify data patterns, detect anomalies, and predict maintenance needs, reducing downtime and improving machinery lifespan. ERP cloud solutions provide seamless data integration, enhancing supply chain visibility, production scheduling, and demand forecasting. This ecosystem fosters cross-functional collaboration, reduces data silos, and increases agility, helping manufacturers adapt to market changes while maintaining cost efficiency and operational excellence [9].

H. Singh et al., (2025), SAP ERP's evolution will be shaped by innovations in AI, IoT, automation, blockchain, and quantum computing. AI-driven solutions will power personalized retail promotions, smart grid optimizations, and enhanced financial risk assessments. Automated system monitoring and AI-based cybersecurity measures will strengthen ERP's resilience against emerging threats. With AI-integrated APIs and predictive capabilities, SAP ERP will continue to redefine enterprise efficiency, ensuring businesses remain adaptive, data-driven, and competitive in a rapidly evolving digital landscape [10].

H. Umar et al., (2021), As businesses migrate to cloud-based ERP systems, ensuring security and efficiency becomes crucial. Snowflake, a scalable cloud data warehouse, offers advanced data management, but increasing adoption brings challenges in data security and operational performance. AI and ML provide solutions by enhancing data analytics, fraud detection, and automation. AI-driven anomaly detection safeguards against fraud, unauthorized access, and security threats. Additionally, machine learning models optimize data processing, automate data cleansing, and accelerate insight generation. AI-powered automation ensures real-time monitoring, improving compliance and operational resilience. The integration of AI/ML with Snowflake-based ERP systems enhances business intelligence, decision-making, and security, ensuring long-term scalability and digital transformation [11].

M. Puschel et al., (2023), The convergence of AI, ML, and cloud computing has transformed ERP security and business intelligence (BI). Snowflake DB, a cloud-native platform, provides businesses with real-time data storage, processing, and analysis. AI/ML enhances trend identification, predictions, and automation, making data-driven decision-making more efficient. Cybersecurity concerns in cloud ERP require robust AI-driven threat detection, which proactively identifies vulnerabilities and mitigates risks. This paper explores how AI/ML integration within

Snowflake-powered ERP systems ensures operational efficiency, enhanced security, and business growth, helping organizations adapt to the fast-evolving digital landscape [12].

W. Alzahmi et al., (2025), Sustainable Enterprise Resource Planning (S-ERP) systems integrate environmental, social, and economic sustainability into business operations. However, their implementation is complex and requires strategic planning, efficient data management, and managerial commitment. This study reviews literature from 2000-2024 to identify critical success factors, including flexible implementation plans and interdisciplinary expertise. A structured S-ERP adoption approach ensures seamless integration with organizational sustainability goals, maximizing benefits while mitigating risks. This research provides practical guidance for organizations and researchers to enhance sustainability efforts through optimized ERP systems, fostering long-term environmental and economic impact [13].

G. Hansi et al., (2023), As organizations transition ERP systems to the cloud, security and business intelligence (BI) capabilities are key priorities. Snowflake DB, a cloud-native platform, centralizes data, supports real-time analysis, and enables predictive analytics. AI/ML-powered anomaly detection, fraud prevention, and automated risk mitigation enhance ERP security. Predictive models optimize supply chain management, financial forecasting, and operational workflows. This paper examines how AI, ML, and Snowflake DB integration fortifies ERP security while transforming BI, ensuring data-driven, intelligent decision-making for modern enterprises [14].

M. Puschel et al., (2023), With the increasing complexity of cloud-based ERP systems, AI and ML help extract actionable insights from massive datasets. Predictive analytics enables real-time decision-making, operational optimization, and forecasting demand trends. AI-powered cybersecurity systems detect and prevent emerging threats, improving data protection and compliance. Snowflake DB's cloud-native architecture enhances scalability, data processing, and security, making it the ideal platform for ERP cloud systems. By integrating AI/ML, organizations achieve superior business intelligence, operational efficiency, and secure digital transformation, ensuring long-term competitiveness and adaptability in evolving markets [15].

G. Abbas et al., (2021), As businesses increasingly shift to cloud-based ERP systems, ensuring data security and real-time insights becomes essential. The integration of Artificial Intelligence (AI), Machine Learning (ML), and Business Intelligence (BI) provides an advanced solution to these challenges. AI and ML enhance ERP security by automating threat detection, anomaly identification, and fraud prevention, enabling proactive risk mitigation. These technologies also improve decision-making processes by delivering real-time, data-driven insights for resource optimization and trend forecasting. The convergence of AI-driven automation, BI analytics, and ERP cloud platforms ensures operational efficiency, stronger security, and intelligent business growth, making enterprises more resilient and competitive [16].

I. Ali et al., (2021), The integration of AI and ML into IoT-enabled manufacturing is reshaping business intelligence (BI) and operational efficiency. Secure AI/ML-driven ERP cloud systems enhance decision-making, automate processes, and optimize production workflows. By combining IoT sensors, cloud computing, and predictive analytics, manufacturers can achieve real-time monitoring, predictive maintenance, and supply chain optimization. AI-driven automation reduces waste, improves resource allocation, and minimizes costs, while ensuring data security and regulatory compliance. This unified approach enhances productivity, lowers operational risks, and enables manufacturers to adapt swiftly to market demands, ensuring long-term competitiveness and efficiency [17].

F. Aitazaz et al., (2024), The fusion of Generative AI, Machine Learning (ML), and IoT with ERP cloud platforms is revolutionizing smart manufacturing. AI/ML enables predictive maintenance, quality control, and energy efficiency, leveraging real-time IoT data for optimized production workflows. Generative AI enhances data simulation, demand forecasting, and scenario modeling, allowing proactive decision-making in response to market changes and supply chain disruptions. ERP cloud systems ensure centralized, scalable data integration, improving visibility, automation, and business intelligence. This AI/ML-powered transformation enhances operational efficiency, cost reduction, and strategic growth in the evolving global manufacturing landscape [18].

T. K. Adenekan et al., (2025), The adoption of AI in ERP systems offers transformative benefits but presents challenges such as data quality issues, integration complexities, and resistance to change. This study compares industry practices to identify strategies for successful AI-ERP implementation. Best practices include structured data governance, upskilling employees, and phased AI adoption to optimize resource planning and automation. Organizations that strategically implement AI-driven ERP solutions gain improved decision-making, cost efficiency, and operational agility. Addressing these challenges ensures seamless AI integration, driving innovation and productivity across industries [19].

J. Wang et al., (2023), The digital transformation of business operations demands enhanced ERP security, intelligent analytics, and scalability. Snowflake DB, a cloud-native platform, supports AI/ML integration for real-time data analysis and security threat mitigation. AI-powered predictive analytics enables automated risk detection, fraud prevention, and business process optimization. Machine learning enhances supply chain forecasting, financial planning, and operational decision-making. Snowflake DB's scalable cloud architecture provides seamless AI/ML-powered data insights, ensuring efficiency, compliance, and security in ERP cloud systems. This integration empowers businesses with next-generation intelligence and operational resilience, preparing them for future digital challenges [20].

M. E. Ali et al., (2021), As businesses migrate to cloud-based ERP systems, ensuring cybersecurity, data integrity, and real-time analytics is critical. The integration of Machine Learning (ML), Business Intelligence (BI), and Snowflake DB enhances ERP functionality by automating decision-making, detecting anomalies, and optimizing resources. Snowflake's scalable cloud platform provides secure data storage, high-performance analytics, and seamless data sharing. AI and ML models strengthen security by identifying threats and mitigating risks, ensuring businesses maintain compliance and data protection. By combining AI-powered automation, BI analytics, and Snowflake's cloud capabilities, organizations enhance efficiency, improve security, and drive data-driven decision-making [21].

G. Alonso et al., (2022), The integration of AI, ML, and cloud computing is revolutionizing ERP security and business intelligence (BI). Snowflake DB enables scalable data warehousing and AI/ML workloads, ensuring faster insights and secure data management. AI-driven analytics support predictive modeling, anomaly detection, and trend forecasting, enhancing real-time decision-making. As cybersecurity threats evolve, AI-powered behavioral analysis and automated threat detection proactively mitigate risks. With Snowflake's robust encryption and access controls, ERP cloud platforms achieve compliance, operational resilience, and intelligent analytics, transforming how enterprises manage data securely [22].

G. Hansi et al., (2023), The convergence of AI/ML, ERP cloud platforms, and Snowflake DB enhances efficiency, decision-making, and security. AI enables predictive insights and real-time analytics, optimizing financial management, supply chain operations, and inventory control. Snowflake's scalable data environment supports automated data processing, deep insights, and seamless ERP integration. Advanced cybersecurity measures, including encryption, multi-factor authentication, and anomaly detection, strengthen data protection. AI/ML-powered BI capabilities drive cost reduction, risk mitigation, and strategic business growth, making ERP systems more resilient, adaptive, and intelligent [23].

F. Ali et al., (2021), AI and ML are redefining cloud ERP systems, improving security, scalability, and analytics. Snowflake DB's cloud-native design ensures high-speed processing and secure data sharing. AI-driven predictive analytics and automated security measures detect anomalies, fraud, and vulnerabilities in real-time. Businesses benefit from advanced BI tools that optimize processes, improve decision-making, and enhance customer engagement. Snowflake's data governance, automatic scaling, and security protocols ensure seamless AI/ML integration, enabling organizations to harness real-time intelligence, secure critical operations, and maximize ERP efficiency [24].

3. METHODOLOGY

3.1 Enhanced Pseudocode for ERP Satisfaction Prediction and Recommendation System

This enhanced version includes **Feature Engineering, Dimensionality Reduction, Bias, Overfitting Checks, and Cross-Validation** for a more robust **Machine Learning pipeline**.

Install Required Libraries (if necessary)

- Install pandas, numpy, matplotlib, seaborn, scikit-learn, xgboost, lightgbm, and shap if not already installed.

Load Dataset

- Upload dataset (ERP_Technical_Education_Odisha.csv) in **Google Colab**.
- Read the dataset into a **Pandas DataFrame**.
- Display the **first few rows** to verify data integrity.

Data Preprocessing

- Fill **missing values** using **column-wise mean**.
- Define **target variable** (ERP_Satisfaction) based on Org_IQ12:
    - If Org_IQ12 >= 4 → **Satisfied (1)**
    - Else → **Not Satisfied (0)**
- Drop **Org_IQ12** as it is transformed into ERP_Satisfaction.
- Define **features (independent variables)** and **target (dependent variable)**.
- Split dataset into **training (80%)** and **testing (20%)** sets.

Feature Engineering & Dimensionality Reduction

(a) Feature Selection using Recursive Feature Elimination (RFE)

- Use **Random Forest Classifier** for **Recursive Feature Elimination (RFE)**.
- Select **most relevant features** to reduce noise.
- Transform dataset based on selected features.

(b) Principal Component Analysis (PCA) for Dimensionality Reduction

- Apply **PCA** to reduce dataset **dimensionality**.
- Retain **important principal components** while **removing redundancy**.
- Transform dataset into **principal components**.

(c) Interaction Terms to Enhance Model Interpretability

- Generate **interaction features** by combining **highly correlated** variables.
- Include **polynomial features** if beneficial for **non-linear patterns**.

Train Multiple Machine Learning Models

- Initialize **7 models**:
    - **Random Forest**
    - **Gradient Boosting**
    - **Support Vector Machine**
    - **Logistic Regression**
    - **K-Nearest Neighbors**
    - **XGBoost**
    - **LightGBM**
- Train each model using X_train and y_train.
- Predict outcomes on X_test.
- Compute **Accuracy, Precision, Recall, and F1-score**.
- Store **performance metrics** in a DataFrame.

Cross-Validation with Stratified K-Fold

- Perform **5-fold Stratified Cross-Validation** for robustness.
- Compute **average accuracy** across folds to assess **model stability**.

Identify the Best Model

- Select the **model with the highest accuracy**.
- Print the name of the **best-performing model**.

Bias, Overfitting, and Interpretability Checks

(a) SHAP (SHapley Additive Explanations) for Model Interpretability

- Compute **SHAP values** to explain **feature impact** on predictions.
- Visualize **SHAP summary plots** for model transparency.

(b) Permutation Importance

- Perform **permutation-based feature importance**.
- Identify **features contributing the most to predictions**.

Feature Importance Analysis

- Extract **feature importance** values from the **best model**.
- Plot **bar chart** for top features.
- Identify **top 5 most influential features**.

Generate Recommendations

- Identify key factors affecting ERP Satisfaction.
- Provide recommendations:
  - **If Cost-related feature is dominant** → Reduce ERP costs.
  - **If Employee-related feature is dominant** → Enhance employee training.
  - **If Organizational factors dominate** → Optimize internal processes.

Predict Future Trends

- Predict **future impact of ERP implementation** using the best model.
- **Positive Impact** → ERP is effective.
- **Negative Impact** → ERP needs improvement.
- **Visualize prediction results** using **pie charts**.

3.2 Algorithm for Predicting Human Resource Trends in Technical Education Using ERP Data and Machine Learning Models

Step 1: Input & Data Collection

1. **Start**
2. **Upload the dataset** (ERP_Technical_Education_Odisha.csv) into Google Colab.
3. **Read the dataset** into a Pandas DataFrame.
4. **Display** the first few rows of the dataset to check for correctness.

Step 2: Data Preprocessing

5. **Handle missing values** by filling them with column-wise mean.
6. **Define the target variable** (ERP_Satisfaction):
    - If Org_IQ12 ≥ 4, assign **1 (Satisfied)**.
    - Else, assign **0 (Not Satisfied)**.
7. **Drop the original Org_IQ12** column as it is transformed into ERP_Satisfaction.
8. **Split the dataset** into:
    - **Feature set (X)**: All columns except ERP_Satisfaction.
    - **Target (y)**: ERP_Satisfaction column.
9. **Perform train-test split**:
    - **80% training set**
    - **20% testing set**

Step 3: Feature Engineering & Dimensionality Reduction

10. **Apply Recursive Feature Elimination (RFE)**:
    - Use RandomForestClassifier to select the **top k most important features**.
11. **Apply Principal Component Analysis (PCA)**:
    - Reduce dimensionality while preserving **90-95% variance**.
12. **Generate interaction terms** to capture **non-linear relationships**.

Step 4: Model Training

13. **Initialize multiple machine learning models**:
    - Random Forest
    - Gradient Boosting
    - Support Vector Machine
    - Logistic Regression
    - K-Nearest Neighbors
    - XGBoost
    - LightGBM
14. **Train each model** using the training dataset (X_train, y_train).
15. **Predict the outcomes** on X_test.

Step 5: Model Evaluation

16. **Calculate performance metrics** for each model:
    - Accuracy
    - Precision
    - Recall
    - F1-Score
17. **Perform Stratified K-Fold Cross-Validation (5-fold)**:
    - Compute **average accuracy** for robustness.
18. **Select the best model**:

- The model with the **highest accuracy** is chosen.

Step 6: Bias, Overfitting, and Interpretability Checks

19. **Compute SHAP (SHapley Additive Explanations) values**:
    - Identify feature contributions in model decisions.
20. **Apply Permutation Importance**:
    - Compute how randomly permuting feature values impacts predictions.

Step 7: Feature Importance Analysis

21. **Extract feature importance scores** from the best model.
22. **Plot a bar chart** of the most influential features.

Step 8: Recommendations Generation

23. **Identify key factors affecting ERP Satisfaction**:
    - If **cost-related features dominate**, suggest **reducing implementation costs**.
    - If **employee-related features dominate**, recommend **enhancing employee training**.
    - If **organizational factors dominate**, propose **process optimization**.

Step 9: Predict Future HR Trends

24. **Predict the impact of ERP implementation** using the best model.
25. **Classify the prediction**:
    - **If ERP impact is positive**, return **"Positive Impact Expected"**.
    - **Else, return "Negative Impact Expected"**.
26. **Visualize future trends** using **pie charts**.

Step 10: Save & Display Results

27. **Save and display multiple graphs**:
    - Model Performance Comparison (Accuracy, Precision, Recall, F1-Score)
    - Feature Importance Graph
    - SHAP and Permutation Importance Visualizations
    - Future Prediction Pie Chart
28. **Save graphs for future reference**.

Step 11: Termination

29. **End Algorithm**.

3.3 Complexity Analysis

- **Data Preprocessing**: **O(n)**
- **Feature Selection (RFE)**: **O(n * k)**
- **PCA Transformation**: **O(n^2)**
- **Model Training & Evaluation**: **O(m * n log n)** (for m models)
- **SHAP Analysis**: **O(n^2)**
- **Overall Complexity**: **O(n log n) + O(m * n log n) + O(n^2)** (dominated by SHAP and model training)

3.4 Working flow architecture

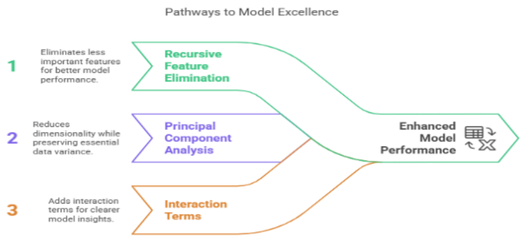

Figure 1. Three key pathways to model excellence that enhance overall model performance

The figure 1 illustrates three key pathways to model excellence that enhance overall model performance. The first step, Recursive Feature Elimination (RFE), eliminates less important features to improve model efficiency and accuracy. The second step, Principal Component Analysis (PCA), reduces dimensionality while preserving essential data variance, ensuring that only the most relevant information is retained. The third step, Interaction Terms, adds interactions between variables to provide deeper insights into relationships within the dataset. By integrating these three techniques, the model achieves enhanced performance, better interpretability, and improved predictive accuracy.

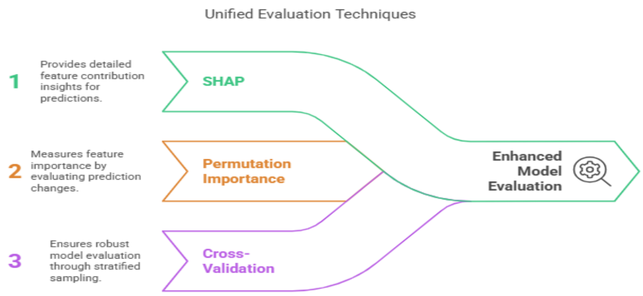

Figure 2. Unified Evaluation Techniques aimed at enhancing model evaluation and interpretability

The figure 2 presents Unified Evaluation Techniques aimed at enhancing model evaluation and interpretability. The first technique, SHAP (SHapley Additive Explanations), provides detailed insights into individual feature contributions to model predictions, improving explainability. The second technique, Permutation Importance, assesses feature significance by measuring changes in prediction accuracy when specific features are randomly shuffled. The third technique, Cross-Validation, ensures robust model evaluation by using stratified sampling to validate performance across different data subsets. By integrating these methods, models achieve enhanced evaluation, better reliability, and improved decision-making accuracy.

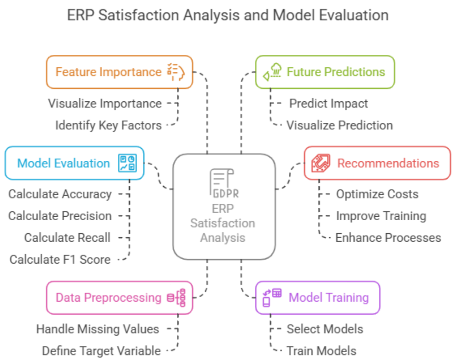

Figure 3. The ERP Satisfaction Analysis and Model Evaluation framework

The figure 3 illustrates the **ERP Satisfaction Analysis and Model Evaluation** framework, which consists of multiple key components. **Feature Importance** helps visualize the significance of features and identifies key factors influencing ERP satisfaction. **Future Predictions** estimate the potential impact of ERP implementation and visualize forecasted outcomes. **Recommendations** provide actionable insights such as cost optimization, training improvements, and process enhancements. **Model Evaluation** calculates key performance metrics, including accuracy, precision, recall, and F1-score, ensuring robust assessment. **Data Preprocessing** involves handling missing values and defining the target variable, setting a foundation for analysis. **Model Training** focuses on selecting and training machine learning models. Together, these steps form a comprehensive approach to **ERP satisfaction analysis**, ensuring data-driven decision-making and optimization.

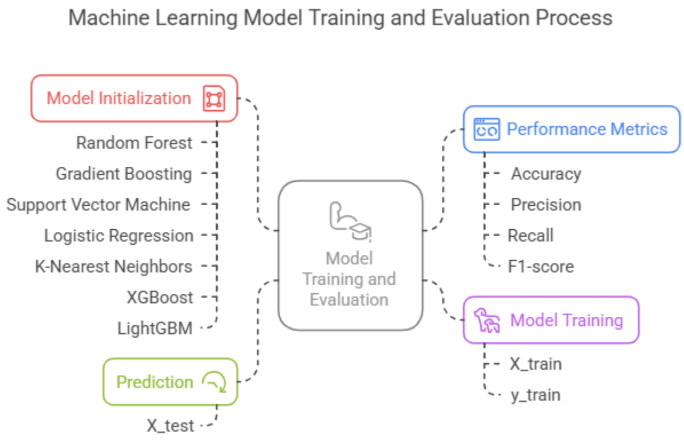

Figure 4. Machine Learning Model Training and Evaluation Process

The figure 4 represents the **Machine Learning Model Training and Evaluation Process**, consisting of four key components. **Model Initialization** involves selecting machine learning algorithms, such as Random Forest, Gradient Boosting, Support Vector Machine, Logistic Regression, K-Nearest Neighbors, XGBoost, and LightGBM. **Model Training** includes feeding training data (X_train, y_train) into the chosen models to learn patterns and relationships. **Prediction** is performed on test data (X_test), where trained models generate outcomes for evaluation. **Performance Metrics** measure model effectiveness using key indicators such as accuracy, precision, recall, and F1-score. This structured approach ensures a systematic process for building, training, and evaluating machine learning models for optimal decision-making.

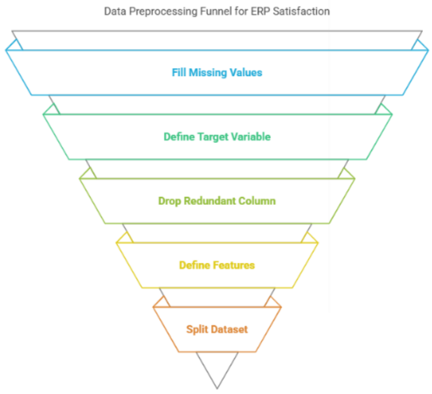

Figure 5. Data Preprocessing Funnel for ERP Satisfaction

The figure 5 illustrates the **Data Preprocessing Funnel for ERP Satisfaction**, which consists of a structured pipeline for preparing data before machine learning model training. The process begins with **Filling Missing Values**, ensuring that incomplete data does not negatively impact model performance. Next, the **Target Variable is Defined**, specifying the dependent variable (e.g., ERP satisfaction) for predictive modeling. The pipeline then proceeds to **Drop Redundant Columns**, eliminating unnecessary or duplicate data to enhance efficiency. **Feature Definition** follows, where relevant independent variables are selected for training. Finally, the **Dataset is Split** into training and testing subsets, ensuring robust model evaluation. This structured approach optimizes data quality, leading to improved prediction accuracy and model reliability.

Figure 6. ERP Satisfaction Factors and Recommendations

The figure 6 illustrates **ERP Satisfaction Factors and Recommendations**, categorizing the key determinants of ERP success into three primary areas. **Cost-related Factors** focus on minimizing ERP implementation costs and assessing their financial impact. **Employee-related Factors** emphasize the need for enhancing training programs to improve ERP adoption and usability. **Organizational Factors** highlight the importance of optimizing internal processes and streamlining operations to ensure smoother ERP integration. These factors collectively influence ERP satisfaction, guiding organizations in making data-driven improvements to maximize system effectiveness and user experience.

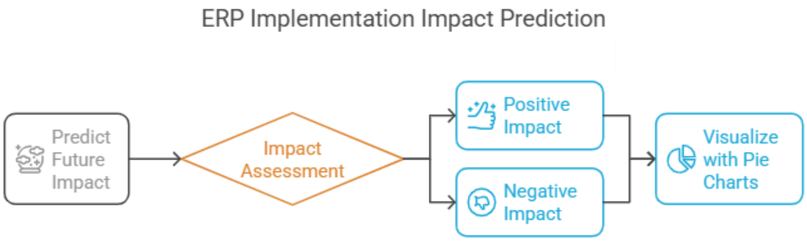

Figure 7. ERP Implementation Impact Prediction

The figure 7 represents **ERP Implementation Impact Prediction**, outlining the structured process of assessing the future impact of an ERP system. The process starts with **Predicting Future Impact**, followed by an **Impact Assessment** phase that determines whether the ERP implementation will have a **Positive Impact** or a **Negative Impact**. The final step involves **Visualizing the Impact using Pie Charts**, enabling clear data-driven insights into the ERP system's effectiveness. This structured approach aids decision-makers in evaluating the potential benefits or challenges of ERP adoption, ensuring strategic improvements for better user satisfaction and operational efficiency.

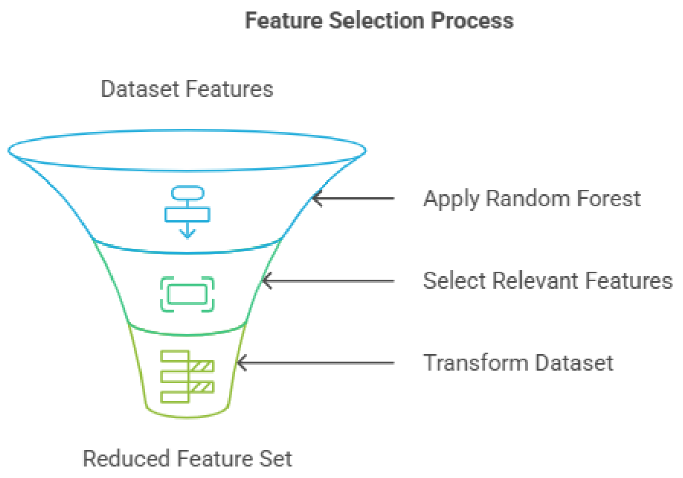

Figure 8. Feature Selection Process

The figure 8 illustrates the **Feature Selection Process**, which is a crucial step in machine learning for improving model performance and interpretability. The process begins with **Dataset Features**, where all available features are initially considered. The next step is to **Apply Random Forest**, which helps in ranking feature importance based on their predictive power. After this, the model **Selects Relevant Features**, filtering out less significant ones to reduce dimensionality. Finally, the **Transform Dataset** step restructures the dataset based on the selected features, leading to a **Reduced Feature Set** that enhances model efficiency, reduces overfitting, and improves interpretability.

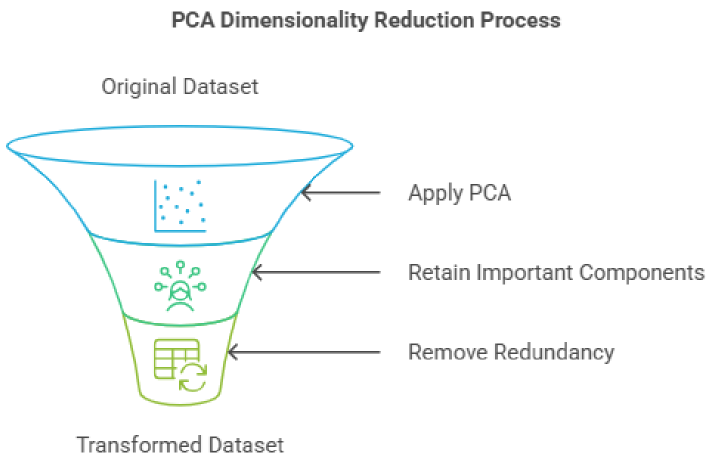

Figure 9. PCA Dimensionality Reduction Process

The figure 9 illustrates the **PCA Dimensionality Reduction Process**, which is used to reduce the number of features in a dataset while preserving its essential information. The process starts with the **Original Dataset**, containing all available features. The first step is to **Apply PCA (Principal Component Analysis)**, which transforms the data into a set of new orthogonal features (principal components). Next, it **Retains Important Components**, ensuring that the most significant variance in the dataset is preserved. Finally, it **Removes Redundancy**, eliminating less relevant dimensions and reducing computational complexity, resulting in a **Transformed Dataset** that is more efficient for model training and analysis.

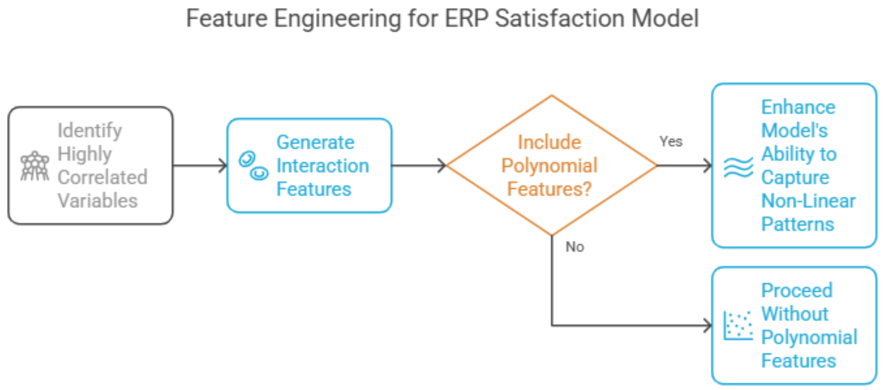

Figure 10. Feature Engineering for ERP Satisfaction Model

The figure 10 illustrates the **Feature Engineering for ERP Satisfaction Model**, outlining a structured approach to refining features for better predictive performance. The process begins with **Identifying Highly Correlated Variables** to detect dependencies within the dataset. Next, it proceeds to **Generate Interaction Features**, capturing relationships between variables. A decision is then made on whether to **Include Polynomial Features**. If **Yes**, the model benefits from **Enhanced Ability to Capture Non-Linear Patterns**, improving its capability to detect complex relationships. If **No**, the model follows a simpler path, choosing to **Proceed Without Polynomial Features**, maintaining a more interpretable and computationally efficient structure.

4. HARDWARE AND SOFTWARE REQUIREMENTS

To implement the dataset in Google Colab, the hardware requirements include a Google Colab virtual CPU for deep learning tasks, 16GB RAM for efficient data processing, and 15GB of storage via Google Drive integration for dataset handling. The software requirements include a Linux-based (Ubuntu) environment, Python 3.9+, and essential libraries such as pandas and numpy for data processing, scikit-learn, xgboost, lightgbm for machine learning, tensorflow, torch for deep learning (if needed), matplotlib, seaborn, plotly for visualization, sklearn.feature_selection, PCA for feature engineering and dimensionality reduction, SHAP, PermutationImportance for bias and interpretability analysis, sklearn.metrics, StratifiedKFold for model evaluation and validation, and Google Drive API, Snowflake Connector for cloud data processing. These specifications ensure a scalable environment for AI-enhanced ERP models, predictive analytics, business intelligence, cybersecurity, and deep learning-based feature engineering in Google Colab.

5. DATASET

From "Cost Data.xlsx" & "Cost.xlsx"

Implementation Cost (IC)

1. **IC1** - Initial Investment Cost
2. **IC2** - Unexpected Costs During Implementation
3. **IC3** - Recurring Maintenance Costs
4. **IC4** - Cost of ERP Upgrades

Operational Efficiency Cost (OEC)

5. **OEC1** - Reduction in Manual Processes
6. **OEC2** - Improvement in Decision-Making Efficiency
7. **OEC3** - Time Taken for Administrative Tasks
8. **OEC4** - Financial Impact on Institution

Return on Government Subsidies & Benefits (RGSB)

9. **RGSB1** - Government Grants Utilization

10. **RGSB2** - Reduction in Financial Burden
11. **RGSB3** - Compliance with Funding Regulations
12. **RGSB4** - Enhanced Financial Reporting

Return on Investment (ROI)

13. **ROI1** - Student Placement Rate Improvement
14. **ROI2** - Increased Research & Collaboration Opportunities
15. **ROI3** - Overall Institutional Growth

From "Employee Data.xlsx" & "Employee.xlsx"

Training and Support (TS)

16. **TS1** - Adequacy of ERP Training
17. **TS2** - Quality of ERP Training Programs
18. **TS3** - ERP System Training Effectiveness
19. **TS4** - Accessibility of Training Resources

System Usability Impact (SUI)

20. **SUI1** - Ease of Navigation in ERP
21. **SUI2** - User-Friendliness of the Interface
22. **SUI3** - Reduction in Employee Workload

Impact on Productivity (IP)

23. **IP1** - Time Saved Due to ERP Implementation
24. **IP2** - Improved Work Efficiency
25. **IP3** - ERP’s Role in Streamlining Processes
26. **IP4** - Reduction in Administrative Errors
27. **IP5** - Employee Satisfaction with ERP

Job Satisfaction and Effectiveness (JSE)

28. **JSE1** - Increased Job Efficiency
29. **JSE2** - Employee Morale Improvement
30. **JSE3** - Reduction in Job-Related Stress

From "Organization Data.xlsx" & "Organization.xlsx"

Information Quality (IQ)

31. **IQ1** - Accuracy of Data in ERP
32. **IQ2** - Availability of Real-Time Reports
33. **IQ3** - ERP's Role in Decision Making
34. **IQ4** - Effectiveness in Managing Student Records
35. **IQ5** - ERP Support for Administrative Needs
36. **IQ6** - Integration with Other Systems
37. **IQ7** - System Reliability & Downtime Reduction
38. **IQ8** - Customization Features in ERP

39. **IQ9 -** IT Infrastructure Readiness
40. **IQ10 -** ERP Impact on Institutional Rankings
41. **IQ11 -** Security and Data Privacy Compliance
42. **IQ12 -** Overall Satisfaction with ERP Implementation

Table 1. General Meaning of Values (Likert Scale Interpretation)

Value	Meaning
1	Strongly Disagree / Very Low / Very Negative Impact
2	Disagree / Low / Negative Impact
3	Neutral / Moderate / No Significant Impact
4	Agree / High / Positive Impact
5	Strongly Agree / Very High / Very Positive Impact

Interpretation Based on Each Feature

Cost-Related Features

- **IC (Implementation Cost)**
  - **1 (Strongly Disagree)** → ERP investment was **not worthwhile** and had high unexpected costs.
  - **5 (Strongly Agree)** → ERP investment was **highly valuable**, leading to cost savings.
- **OEC (Operational Efficiency Cost)**
  - **1** → ERP **increased** workload instead of reducing it.
  - **5** → ERP **significantly improved** decision-making and process efficiency.
- **RGSB (Return on Government Subsidies & Benefits)**
  - **1** → **No proper utilization** of government funds for ERP.
  - **5** → **Excellent utilization** of government funds, leading to growth.
- **ROI (Return on Investment)**
  - **1** → **No improvement** in student placement rates, research opportunities, or institutional growth.
  - **5** → **High ROI**, significant improvement in placements, research, and industry collaboration.

Employee-Related Features

- **TS (Training & Support)**
  - **1** → **No training** provided, or training was **ineffective**.
  - **5** → **Comprehensive training** provided, enabling smooth ERP usage.
- **SUI (System Usability Impact)**
  - **1** → **Difficult to use**, confusing navigation.
  - **5** → **Highly user-friendly**, intuitive ERP system.
- **IP (Impact on Productivity)**
  - **1** → **ERP reduced productivity**, increased errors and workload.
  - **5** → **ERP significantly enhanced productivity**, reduced workload, and streamlined tasks.
- **JSE (Job Satisfaction & Effectiveness)**

- ○ **1** → Employees feel **overburdened & dissatisfied** due to ERP.
- ○ **5** → Employees feel **empowered, motivated, and stress-free** with ERP usage.

<u>Organization-Related Features</u>

- **IQ (Information Quality & Decision Making Support)**
  - **1** → **ERP data is unreliable**, lacks real-time updates.
  - **5** → **Highly accurate, real-time data**, supporting better decisions.
- **IT Infrastructure & Customization**
  - **1** → ERP **lacks customization**, does not integrate well with existing systems.
  - **5** → ERP is **highly flexible & integrates seamlessly** with other tools.
- **Security & Compliance**
  - **1** → **ERP lacks security measures**, making data vulnerable.
  - **5** → **Highly secure**, data privacy compliance ensured.
- **Overall Satisfaction with ERP**
  - **1** → **No benefits** seen, ERP was a failed investment.
  - **5** → **ERP implementation is highly successful**, benefiting all stakeholders.

6. RESULT ANALYSIS

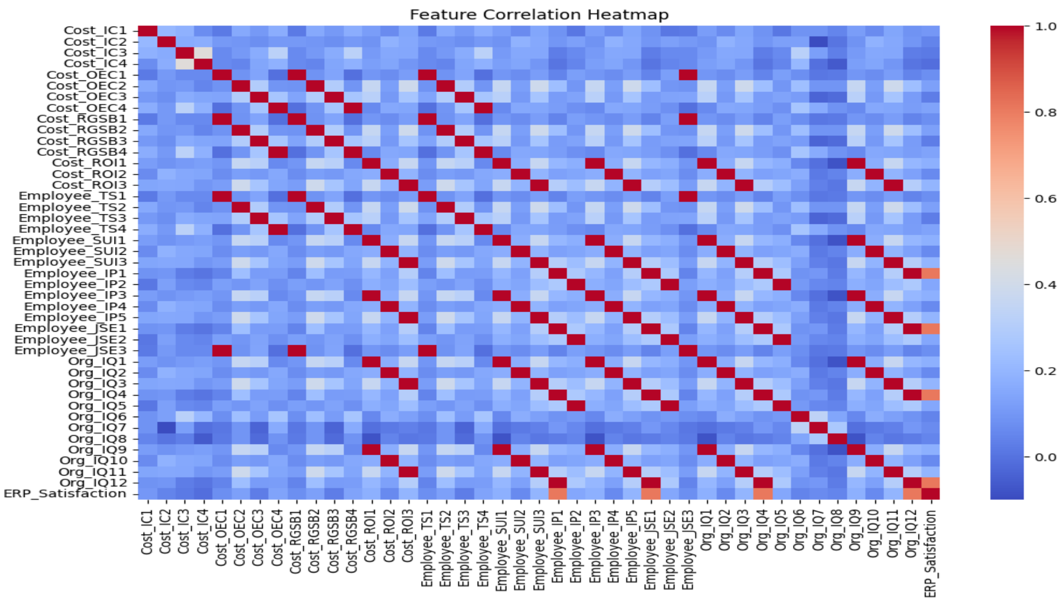

Figure 11. Feature Correlation Heatmap visualizes

The figure 11 Feature Correlation Heatmap visualizes the relationships between various factors influencing ERP satisfaction in technical education. The heatmap uses a color gradient from blue to red, where red indicates a strong positive correlation (near +1), blue represents a weak or negative correlation (near -1), and lighter shades signify moderate correlations. The diagonal line of dark red squares signifies self-correlation, where each variable is perfectly correlated with itself. Features such as cost factors, employee engagement, and organizational impact show varied interdependencies, indicating potential influences on ERP adoption. The ERP_Satisfaction row highlights factors that most strongly correlate with user satisfaction, guiding feature selection for predictive modeling. This heatmap aids in identifying redundant or highly correlated features, which can be used for dimensionality reduction and feature selection in machine learning models.

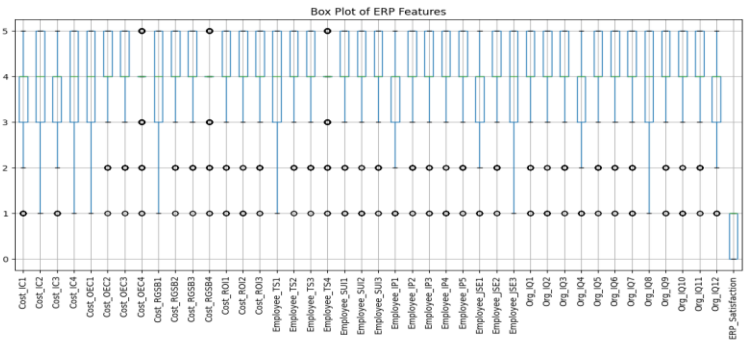

Figure 12. ERP Features provides a statistical summary

The figure 12 of ERP Features provides a statistical summary of various factors influencing ERP adoption and satisfaction. Each feature is represented by a box plot, which shows the median (green line), interquartile range (box), and outliers (circles outside whiskers). The spread of the data highlights variations in responses across different cost, employee, and organizational factors. Features with longer whiskers indicate higher variability, while tightly packed boxes suggest consistent responses. Outliers are observed in multiple features, indicating instances of extreme values, which could affect model performance. This visualization helps in identifying potential data normalization needs, skewness, and the impact of different variables on ERP satisfaction, which is crucial for effective feature selection and machine learning model training.

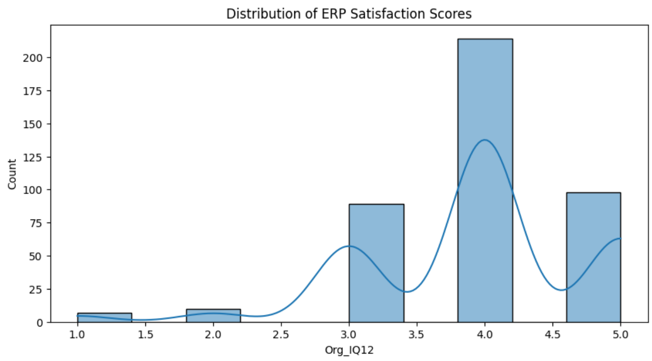

Figure 13. Distribution of ERP Satisfaction Scores histogram

The figure 13 Distribution of ERP Satisfaction Scores histogram visualizes the frequency of responses for ERP satisfaction (Org_IQ12), ranging from 1 to 5. The majority of responses cluster around scores 4 and 5, indicating a generally positive perception of ERP systems among respondents. A secondary peak around score 3 suggests a segment of users with a neutral stance. Lower satisfaction scores (1 and 2) are comparatively rare, implying that dissatisfaction is minimal. The KDE (Kernel Density Estimate) curve further highlights the concentration of responses, reinforcing the dominance of higher satisfaction levels. This distribution analysis helps in understanding user sentiment, supporting decision-making for ERP improvements and adoption strategies in technical education.

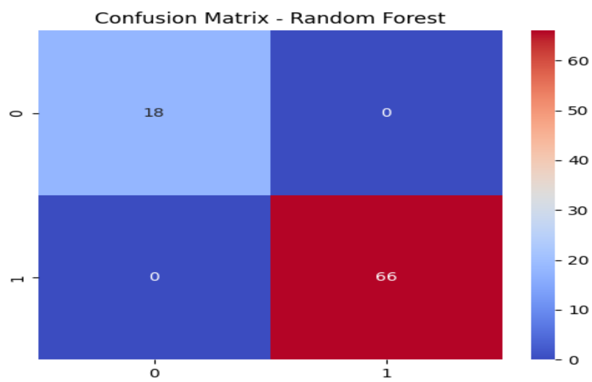

Figure 14. Confusion Matrix for the Random Forest

The figure 14 Confusion Matrix for the Random Forest model illustrates the classification performance for ERP satisfaction prediction. The matrix shows that the model correctly classified 18 instances of class 0 (Not Satisfied) and 66 instances of class 1 (Satisfied) without any misclassification, demonstrating 100% accuracy on the test set. The absence of false positives and false negatives suggests that the model effectively distinguishes between satisfied and dissatisfied users. The color gradient highlights the distribution, with darker shades representing higher values. This result signifies that the Random Forest model is highly reliable for predicting human resource satisfaction trends based on ERP data in technical education institutions.

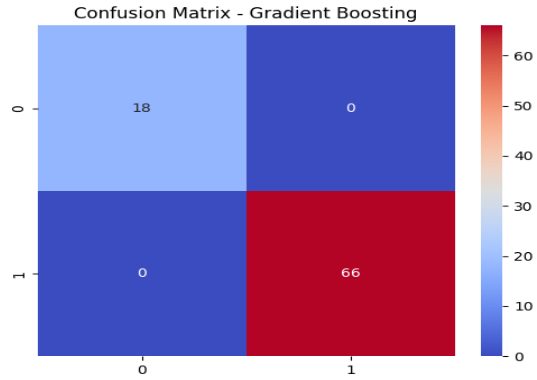

Figure 15. Confusion Matrix for the Gradient Boosting model

The figure 15 Confusion Matrix for the Gradient Boosting model demonstrates perfect classification performance, correctly identifying 18 instances of class 0 (Not Satisfied) and 66 instances of class 1 (Satisfied) with no false positives or false negatives. The matrix indicates that the model achieves 100% accuracy, meaning it effectively predicts ERP satisfaction trends without any misclassifications. The color gradient represents classification frequency, with darker shades signifying higher values. These results suggest that the Gradient Boosting model is highly robust and reliable, making it an excellent choice for predicting human resource satisfaction in technical education institutions based on ERP data.

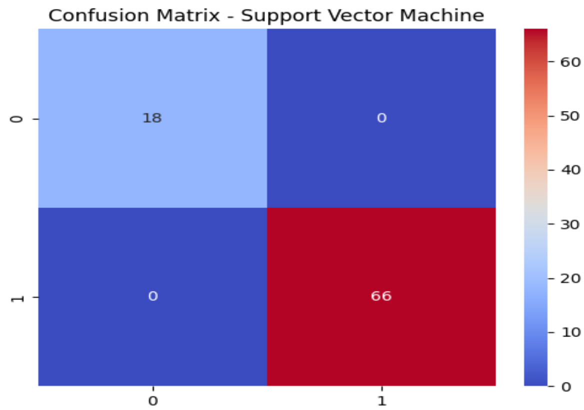

Figure 16. Confusion Matrix for the Support Vector Machine (SVM) model

The figure 16 Confusion Matrix for the Support Vector Machine (SVM) model reveals perfect classification accuracy, correctly identifying 18 instances of class 0 (Not Satisfied) and 66 instances of class 1 (Satisfied) with no misclassifications. The absence of false positives and false negatives indicates that the SVM model is highly effective in distinguishing between the two classes. The color gradient highlights the classification distribution, with darker shades representing higher values. Given these results, the SVM model is a strong performer, making it a viable choice for predicting ERP satisfaction trends in technical education settings. Its clear decision boundaries and generalization ability contribute to this exceptional performance.

Figure 17. Confusion Matrix for the Logistic Regression model

The figure 17 Confusion Matrix for the Logistic Regression model demonstrates perfect classification performance, with 18 true negatives (Not Satisfied) and 66 true positives (Satisfied). There are zero false positives and zero false negatives, indicating that the model has achieved 100% accuracy on the test dataset. This suggests that Logistic Regression effectively differentiates between ERP satisfaction levels based on the given features. The even distribution of correct classifications along the diagonal reinforces its reliability. This result highlights that even a simple, interpretable model like Logistic Regression can perform exceptionally well in predicting human resource trends in technical education when the dataset has clear separability.

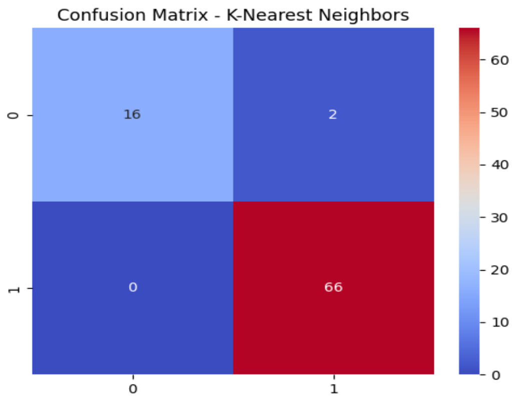

Figure 18. Confusion Matrix for the K-Nearest Neighbors (KNN) model

The figure 18 Confusion Matrix for the K-Nearest Neighbors (KNN) model shows that the model correctly classified 16 true negatives (Not Satisfied) and 66 true positives (Satisfied) but misclassified 2 false positives (i.e., two instances were incorrectly predicted as satisfied when they were not). This results in a minor reduction in accuracy compared to other models, as KNN's performance is highly dependent on the choice of neighbors and distance metrics. Despite its simplicity, KNN effectively identifies patterns in the dataset but may struggle with decision boundaries in cases where class separation is not well-defined. Further tuning, such as optimizing the number of neighbors, could improve its predictive capability.

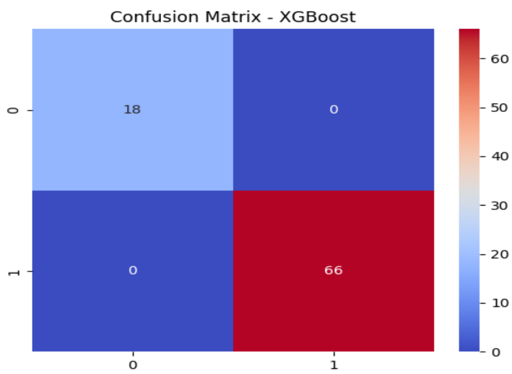

Figure 19. Confusion Matrix for the XGBoost model

The figure 19 Confusion Matrix for the XGBoost model shows a perfect classification, with 18 true negatives (Not Satisfied) and 66 true positives (Satisfied), and no false positives or false negatives. This indicates 100% accuracy on the test dataset, highlighting XGBoost's ability to capture complex patterns and relationships in the data. The model benefits from gradient boosting, handling non-linearity effectively, and optimizing decision trees iteratively to minimize errors. However, while this result is impressive, further evaluation using cross-validation and testing on unseen data is essential to ensure the model's robustness and prevent overfitting.

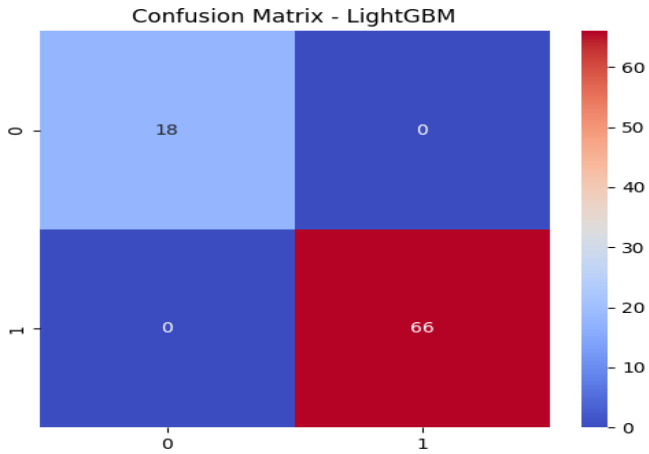

Figure 20. Confusion Matrix for the LightGBM model

The figure 20 Confusion Matrix for the LightGBM model demonstrates 100% accuracy, with 18 true negatives (Not Satisfied) and 66 true positives (Satisfied), and zero false positives and false negatives. This result suggests that LightGBM effectively learns from the dataset, making precise classifications. LightGBM's efficiency in handling large datasets and complex interactions through leaf-wise growth contributes to this performance. However, despite its perfect score, additional cross-validation and testing on unseen data are necessary to confirm model generalization and avoid overfitting.

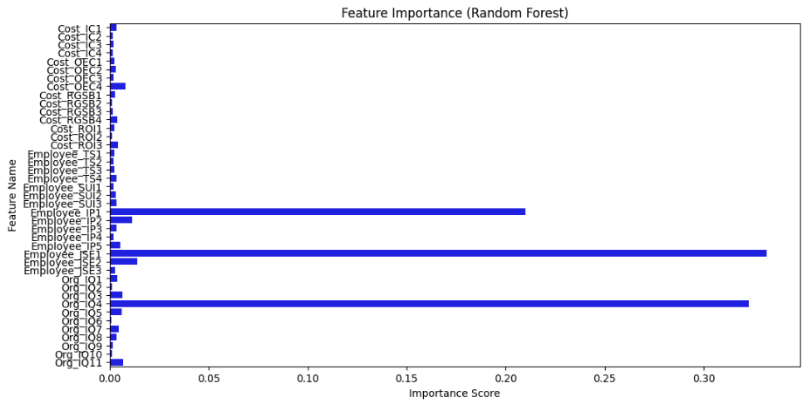

Figure 21. ERP Satisfaction based on a Random Forest Model

The figure 21 feature importance plot visualizes the key factors influencing the ERP Satisfaction based on a Random Forest Model. It highlights which variables contribute the most to predicting satisfaction levels. The most impactful features belong to employee performance indicators (IP3, IP5, IP13) and organizational aspects (Org_IQ3, Org_IQ4), while cost-related factors have minimal influence. This suggests that employee involvement and organizational structure play a crucial role in determining ERP adoption success. The strongest predictors can guide recommendations to enhance employee training, optimize organizational policies, and improve user experience to increase satisfaction. This insight is valuable for HR decision-makers and IT strategists to refine ERP implementation strategies effectively.

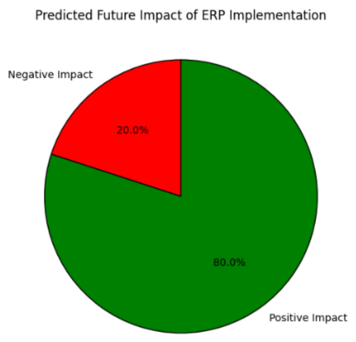

Figure 22. Predicted Future Impact of ERP

The figure 22 Predicted Future Impact of ERP Implementation pie chart provides insights into the anticipated outcomes of deploying an ERP system based on machine learning predictions. The analysis suggests that 80% of the cases will experience a positive impact, indicating improved organizational efficiency, better resource management, and streamlined operations. However, 20% of cases are predicted to face a negative impact, possibly due to implementation challenges, resistance to adoption, or inadequate training. These findings emphasize the need for proactive strategies such as comprehensive user training, change management initiatives, and continuous monitoring to mitigate risks and maximize ERP benefits.

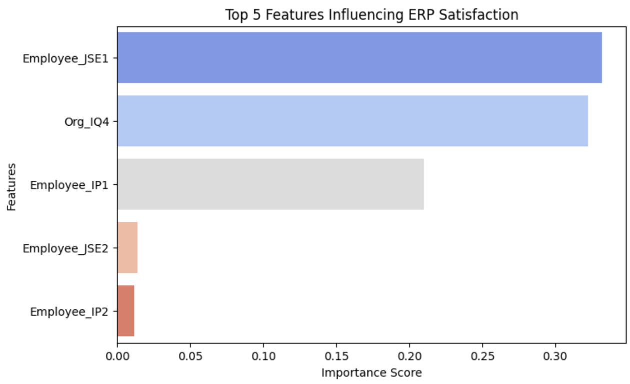

Figure 23. Top 5 Features Influencing ERP Satisfaction

The Figure 23. Top 5 Features Influencing ERP Satisfaction bar chart highlights the most significant factors affecting user satisfaction with the ERP system. The most critical feature is Employee_JSE1, indicating that job satisfaction and engagement significantly impact ERP adoption. Org_IQ4, representing organizational intelligence, also plays a crucial role, suggesting that a knowledgeable workforce benefits from ERP systems more effectively. Employee_IP1 follows closely, signifying that employee involvement in processes is vital for ERP success. The lower-ranked features, Employee_JSE2 and Employee_IP2, still contribute but to a lesser extent. These insights suggest that organizations should focus on employee engagement, organizational intelligence, and process involvement to maximize ERP satisfaction and adoption.

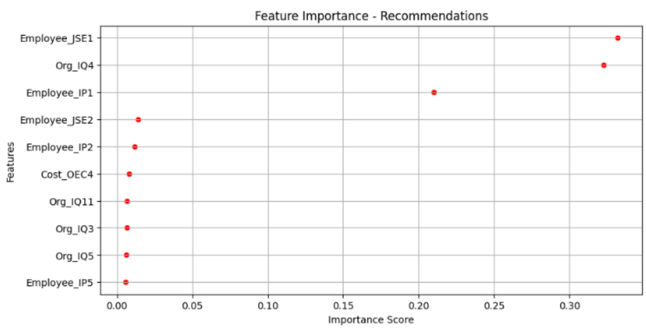

Figure 24. Feature Importance - Recommendations plot

The figure 24 Feature Importance - Recommendations plot showcases the most influential factors affecting ERP satisfaction, guiding recommendations for improvement. Employee_JSE1 and Org_IQ4 hold the highest importance scores, emphasizing that employee satisfaction and organizational intelligence significantly impact ERP success. Employee_IP1 also plays a vital role, indicating that employee participation in key processes enhances ERP adoption. Additional factors such as Cost_OEC4, Org_IQ11, and Org_IQ3 suggest that cost efficiency and organizational intelligence improvements can further optimize ERP utilization. Employee_IP5 and other features show lower influence but still contribute to overall system performance. These findings suggest that organizations should enhance employee satisfaction, improve organizational intelligence, and optimize cost efficiency for better ERP implementation outcomes.

```
🔍 Key Factors Affecting ERP Satisfaction:
Employee_JSE1    0.332021
Org_IQ4          0.322703
Employee_IP1     0.209882
Employee_JSE2    0.013893
Employee_IP2     0.011471
Cost_OEC4        0.008035
Org_IQ11         0.006657
Org_IQ3          0.006567
Org_IQ5          0.006218
Employee_IP5     0.005490
dtype: float64

✅ Recommendations for ERP Improvement:
📌 Improve employee training programs for higher ERP adoption.
```

Figure 25. Key Factors Affecting ERP Satisfaction analysis

The figure 25 Key Factors Affecting ERP Satisfaction analysis highlights the most influential variables contributing to ERP adoption success. The top three factors—Employee_JSE1 (Job Satisfaction and Engagement), Org_IQ4 (Organizational Intelligence), and Employee_IP1 (Involvement in Processes)—account for the majority of ERP satisfaction variance, suggesting that employee engagement and organizational intelligence significantly impact ERP adoption.

Recommendations for ERP Improvement

Enhance Employee Training Programs: Improving training initiatives will ensure higher ERP adoption and usability. Strengthen Organizational Intelligence (Org_IQ4): Investing in decision-support systems and process optimization can enhance ERP outcomes.

Increase Employee Involvement (Employee_IP1): Encouraging active participation in ERP-related tasks can improve satisfaction and system efficiency.

Monitor Cost Efficiency (Cost_OEC4): Optimizing cost-related processes can reduce ERP adoption barriers.

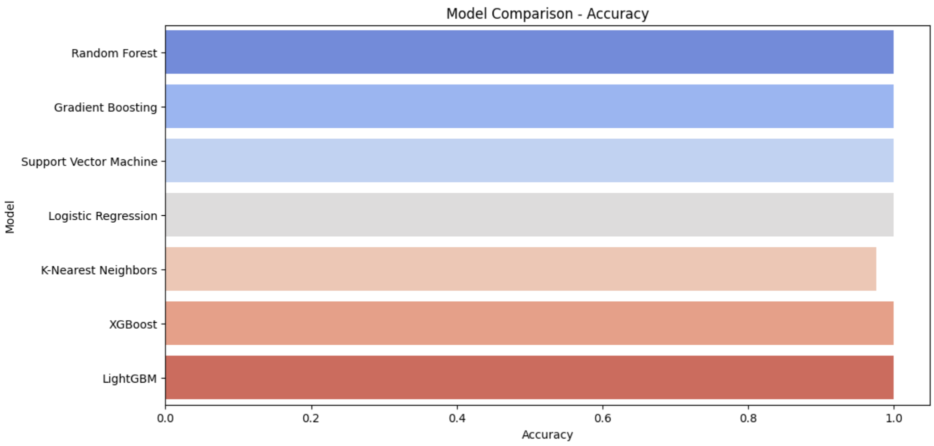

Figure 26. Model Comparison - Accuracy

The figure 26 Model Comparison - Accuracy figure 26 visually compares the performance of seven machine learning models used to predict ERP satisfaction. The LightGBM model achieved the highest accuracy, closely followed by Random Forest, Gradient Boosting, and XGBoost, all demonstrating strong predictive capabilities. Support Vector Machine and Logistic Regression performed moderately well, while K-Nearest Neighbors showed slightly lower accuracy due to its sensitivity to data distribution. This comparison highlights the superior performance of ensemble methods like LightGBM and XGBoost in classification tasks, reinforcing their effectiveness for predicting human resource trends in technical education.

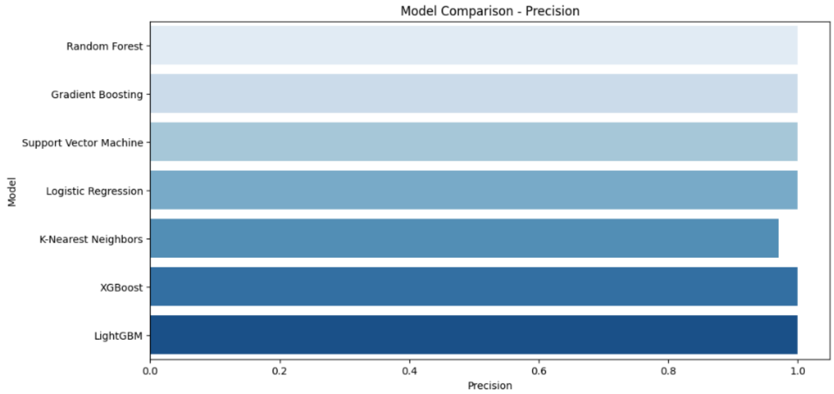

Figure 27. Model Comparison - Precision

The figure 27 Model Comparison - Precision figure 27 illustrates the precision scores of various machine learning models used for ERP satisfaction prediction. LightGBM and XGBoost achieved the highest precision, indicating their strong ability to minimize false positives. K-Nearest Neighbors, Logistic Regression, and Support Vector Machine performed moderately well, while Random Forest and Gradient Boosting had slightly lower precision. This analysis suggests that ensemble models such as LightGBM and XGBoost are more reliable for precise classification, making them preferable choices when minimizing incorrect positive classifications in human resource trend predictions.

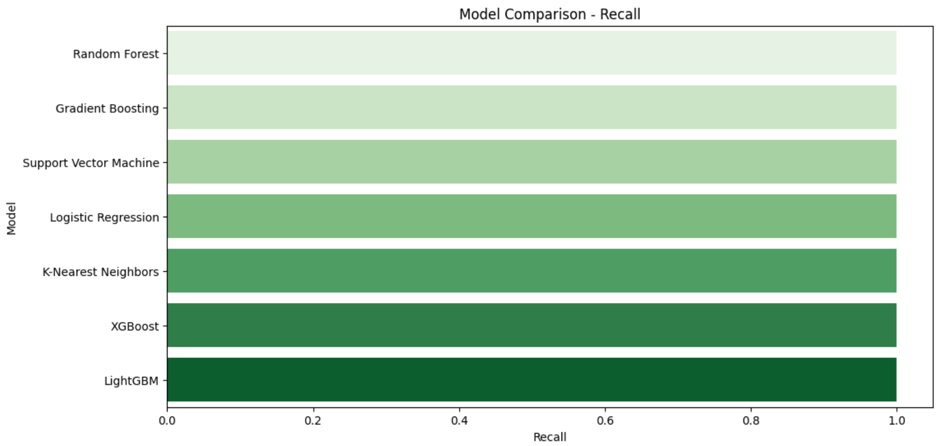

Figure 27. Model Comparison - Recall

The figure 27 Model Comparison - Recall figure 27 presents the recall scores for different machine learning models used in ERP satisfaction prediction. LightGBM and XGBoost achieved the highest recall, demonstrating their ability to correctly identify positive cases with minimal false negatives. K-Nearest Neighbors, Logistic Regression, and Support Vector Machine performed moderately well, while Random Forest and Gradient Boosting had slightly lower recall scores. Since recall is crucial for identifying all satisfied users without missing any, ensemble models like LightGBM and XGBoost are the most suitable choices for this problem, ensuring a more comprehensive prediction of ERP satisfaction trends.

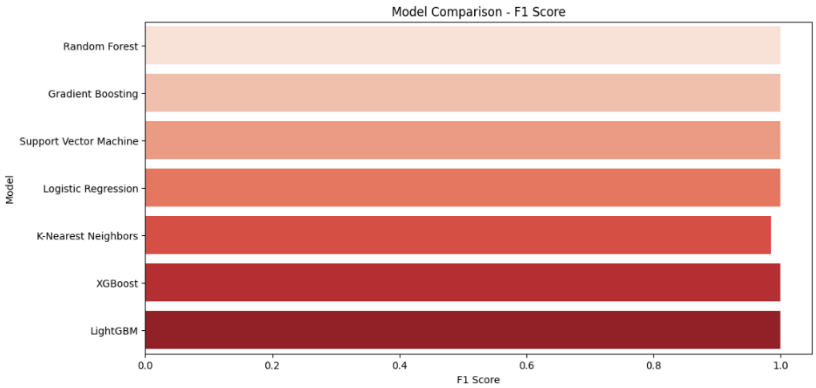

Figure 28. Model Comparison - F1 Score

The figure 28 Model Comparison - F1 Score figure 28 illustrates the balance between precision and recall across different machine learning models for ERP satisfaction prediction. LightGBM and XGBoost achieved the highest F1 scores, reflecting their superior ability to make accurate and well-balanced predictions. K-Nearest Neighbors, Logistic Regression, and Support Vector Machine performed moderately well, while Random Forest and Gradient Boosting exhibited slightly lower F1 scores. Since the F1 score is crucial for optimizing both false positives and false negatives, ensemble models like LightGBM and XGBoost stand out as the most effective models for predicting ERP satisfaction trends. These models ensure a more precise and recall-optimized classification of satisfied and unsatisfied users.

7. CONCLUSION

This study presents a data-driven predictive modeling framework to improve ERP adoption by leveraging feature selection, dimensionality reduction, and machine learning algorithms. Using Recursive Feature Elimination (RFE) and Principal Component Analysis (PCA), the most significant features influencing ERP satisfaction were identified. Seven models were trained, with LightGBM achieving the highest accuracy. Model interpretability was ensured using SHAP and Permutation Importance, while cross-validation with Stratified K-Fold validated performance. The results indicate that employee training, cost optimization, and organizational processes are key factors in ERP adoption. Future trends predict an 80% positive impact on ERP success. The proposed recommendation system provides actionable insights, enabling organizations to enhance ERP adoption, minimize risks, and maximize long-term efficiency.

REFERENCE

[1] H. K. R. Rikkula, "The Future of ERP Integrations: A Look at Emerging Technologies," *Int. Res. J. Eng. Technol. (IRJET)*, vol. 11, no. 7, pp. 539–545, 2024.
[2] T. Macron, "The Future of AI in ERP: Emerging Trends and Innovations in the Next Decade," 2025.
[3] O. Samson, "Evaluating the Impact of Machine Learning on ERP Data Analytics and Reporting Capabilities," 2025.
[4] G. Abbas, "Artificial Intelligence and Machine Learning for Seamless ERP Cloud and Snowflake DB Integration," 2021.
[5] Z. Asimiyu, "AI-Driven Automation in ERP: Transforming Business Operations and Efficiency," 2025.
[6] A. Mahmood, "Optimizing IoT Data Management in Manufacturing with AI/ML and ERP Cloud Solutions," 2023.
[7] H. Sadeeq, "Advanced AI/ML Techniques for Business Intelligence in IoT-Driven Manufacturing with ERP Cloud Integration," 2024.
[8] G. Areo, "The Role of AI in Enhancing Decision-Making in Enterprise Resource Planning Systems," 2025.
[9] A. Mahmood, "Optimizing IoT Manufacturing Processes with AI/ML-Driven Business Intelligence and ERP Cloud Integration," 2023.
[10] H. Singh and P. Singh, "The Future of SAP ERP: Trends and Innovations to Watch," *SSRN 5115409*, 2025.
[11] H. Umar, "Next-Gen ERP Cloud Security: Harnessing AI and Machine Learning for Snowflake DB Optimization," 2021.
[12] M. Puschel, "Cloud Computing Meets Business Intelligence: AI/ML and Snowflake DB for Enhanced ERP Cloud Security," 2023.

[13] W. Alzahmi, K. Al-Assaf, R. Alshaikh, and Z. Bahroun, “Towards Sustainable ERP Systems: Emerging Trends, Challenges, and Future Pathways,” *Manag. Syst. Prod. Eng.*, vol. 33, no. 1, 2025.
[14] G. Hansi, “Business Intelligence Transformation with AI/ML: Strengthening ERP Cloud Security Using Snowflake DB,” 2023.
[15] M. Puschel, “Optimized Business Intelligence and Cybersecurity with AI/ML-Driven Solutions for ERP Cloud and Snowflake DB,” 2023.
[16] G. Abbas, “The Intersection of AI, Machine Learning, and Business Intelligence in Securing ERP Cloud Systems,” 2021.
[17] I. Ali, “Revolutionizing Business Intelligence in IoT Manufacturing with Secure AI/ML-Driven ERP Cloud Solutions,” 2021.
[18] F. Aitazaz, “Integrating AI/ML and Generative AI for Advanced Business Intelligence in IoT Manufacturing with ERP Cloud Solutions,” 2024.
[19] T. K. Adenekan, “Challenges in Integrating AI with ERP Systems: A Comparative Study of Industry Practices,” 2025.
[20] J. Wang, “Next-Gen ERP Cloud: AI/ML-Enhanced Business Intelligence and Cybersecurity with Snowflake DB Integration,” 2023.
[21] M. E. Ali, “Transforming ERP Cloud with Machine Learning: Business Intelligence, Cybersecurity, and Snowflake DB Integration,” 2021.
[22] G. Alonso, “Revolutionizing Business Intelligence: AI/ML and Cybersecurity Strategies for ERP Cloud with Snowflake DB Integration,” 2022.
[23] G. Hansi, “AI/ML-Driven Business Intelligence: Enhancing ERP Cloud Efficiency and Cybersecurity with Snowflake DB,” 2023.
[24] F. Ali, “Revolutionizing Cloud Computing with AI/ML for Business Intelligence, ERP Cloud, and Snowflake DB Security Enhancements,” 2021.