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

Research Article

Predicting Human Resource Trends in Technical Education Through ERP Data and Machine Learning Models

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: 15 Dec 2024

Revised: 05 Feb 2025

Accepted: 24 Feb 2025

ABSTRACT

Enterprise Resource Planning (ERP) systems have become integral in streamlining academic and administrative processes in technical education institutions. However, their impact on human resource (HR) trends, including faculty performance, student outcomes, and job placements, remains underexplored. This study leverages ERP-generated data to predict HR trends in Odisha's technical education sector using machine learning models. The dataset comprises institutional records, faculty evaluations, student performance metrics, and employment statistics collected from ERP systems. We employ Random Forest and Gradient Boosting models to analyze key determinants influencing HR efficiency. Results indicate that faculty engagement and student ERP usage significantly correlate with improved job placement rates and academic performance. The study highlights the predictive power of machine learning in forecasting HR trends, aiding policy decisions for educational institutions. By integrating ERP analytics with AI-driven models, institutions can optimize HR strategies and enhance student career outcomes.

Keywords: ERP, Human Resource Trends, Technical Education, Machine Learning, Student Performance, Faculty Management, Job Placement

1. INTRODUCTION

In the era of digital transformation, Enterprise Resource Planning (ERP) systems have revolutionized academic administration by integrating key institutional processes, including human resource (HR) management, student records, faculty performance, and financial planning. Technical education institutions, in particular, rely heavily on ERP systems to streamline operations, enhance data-driven decision-making, and improve workforce efficiency. However, while ERP systems generate vast amounts of structured data, their full potential remains largely untapped when it comes to predicting human resource trends in education. Leveraging machine learning (ML) models to analyze ERP data can provide valuable insights into faculty management, student success, and job placement trends, enabling institutions to make informed strategic decisions.

One of the major challenges faced by technical education institutions is the efficient utilization of human resources to enhance academic performance and employability outcomes. Faculty engagement, student participation, administrative efficiency, and job placements are critical factors that influence institutional success. However, traditional HR management approaches often rely on static reports and subjective assessments, limiting their ability to predict future trends and optimize resource allocation. Machine learning models, when applied to ERP data, can identify hidden patterns, analyze key performance indicators, and forecast HR trends with high accuracy. By integrating data-driven decision-making with predictive analytics, institutions can proactively address challenges, improve faculty workload distribution, optimize student learning experiences, and enhance career prospects.

The dataset used in this study consists of various ERP-generated records, including faculty performance evaluations, student academic progress, administrative efficiency, and job placement statistics. These records serve as key indicators for assessing HR efficiency in educational institutions. While descriptive analytics provides historical insights, predictive modeling enables institutions to anticipate HR trends, adjust faculty assignments, enhance student engagement, and improve employment outcomes. However, extracting meaningful insights from highdimensional ERP data requires advanced analytical techniques, such as Random Forest and Gradient Boosting models.

To address this challenge, this study applies Random Forest and Gradient Boosting classifiers, two widely used ensemble machine learning techniques, to analyze ERP data and predict HR trends. Random Forest, an ensemble-based decision tree model, effectively handles non-linear relationships, feature importance analysis, and classification tasks, making it suitable for predicting faculty and student performance trends. Gradient Boosting, on the other hand, builds models iteratively, reducing prediction errors and improving accuracy through sequential learning. These models enable institutions to predict key HR trends, such as faculty workload balance, student success probability, and job placement rates, with high precision.

The results of this study demonstrate that both models achieved 100% accuracy, effectively classifying ERP satisfaction levels and predicting HR trends with zero false positives and false negatives. Feature importance analysis highlights faculty engagement, student participation, and institutional support as key factors influencing HR efficiency. These findings underscore the potential of machine learning in optimizing ERP adoption, providing educational institutions with actionable insights to improve HR strategies, enhance workforce planning, and foster student success.

This research contributes to the growing field of AI-driven decision-making in education, bridging the gap between ERP analytics and predictive modeling. By leveraging machine learning techniques, institutions can move beyond retrospective assessments and embrace proactive, data-driven strategies for HR optimization. Future studies can explore deep learning models, hybrid AI techniques, and real-time data analysis to further refine ERP-based HR trend predictions.

2. LITERATURE REVIEW

P. Pokala et al., (2024), The evolving intersection of Artificial Intelligence (AI) and Enterprise Resource Planning (ERP) is reshaping career development in this field. This article presents a structured framework outlining key competencies required for professionals, including AI/ML technical skills, ERP system expertise, business process knowledge, and continuous professional development. The findings highlight that successful career growth requires a combination of technical proficiency, ERP architectural understanding, and business integration experience. Additionally, continuous learning and networking play a crucial role due to the rapid evolution of AI technologies. Ethical and security concerns related to AI implementation in ERP environments are also discussed. This framework serves as both an academic contribution and a practical guide for professionals, offering insights into emerging trends and future research directions [1].

S. Subrahmanyam et al., (2025), The evolution of Human Resources (HR) in the digital age is profoundly impacted by technological advancements such as HRIS, ERP, AI-driven recruitment, and cloud computing. This chapter provides a historical overview of HR transformation, tracing its journey from early mechanization to modern digital solutions. It examines how emerging technologies optimize recruitment, workforce planning, and employee management. Additionally, the chapter explores futuristic trends like genomics and transhumanism, discussing their potential in talent optimization and personalized health while addressing ethical concerns. The study advocates for organizations to adapt HR policies proactively and foster an innovative culture to embrace the rapidly evolving technological landscape [2].

S. Sarferaz et al., (2024), Integrating Artificial Intelligence into ERP business processes presents significant challenges, including systematic AI integration and ensuring enterprise readiness. This book offers a historical perspective on ERP evolution, proposing reference architectures and operationalization methods to embed AI seamlessly into ERP software. It explores essential aspects such as data integration, model validation, explainability, privacy, and scalability, providing a structured solution architecture. The final section introduces an implementation framework for AI-driven ERP applications, illustrated through real-world case studies in logistics, finance, and sales using SAP S/4HANA. This approach highlights AI's feasibility in enhancing ERP functionalities, ensuring compliance, and maintaining high performance [3].

M. V. Lakhamraju et al., (2025), Workday ERP is a cloud-based solution that overcomes traditional ERP limitations such as high costs and lack of flexibility. With advanced features like automation, real-time analytics, and predictive capabilities, it enhances core HR functions, including workforce planning, payroll, talent management, and employee engagement. Workday's robust security framework ensures compliance with global standards while automating complex business processes. This book examines Workday ERP's integration, reporting, and AI-driven innovations, helping organizations achieve greater efficiency and adaptability. Future trends in enterprise automation, AI intelligence, and data warehousing are explored, providing businesses with insights into maximizing productivity and sustaining long-term success [4].

B. A. Chaushi et al., (2024), ERP systems have evolved from basic management tools to AI-powered, cloud-based solutions that drive digital transformation. This study examines ERP development phases, from early internal-focused systems to modern AI-driven, cloud-enabled, and blockchain-integrated platforms. It highlights how machine learning, natural language processing, and robotic process automation enhance decision-making and workflow automation. The study emphasizes that AI is revolutionizing next-generation ERPs, making them more adaptive, data-driven, and efficient. Organizations are encouraged to embrace emerging AI trends, cloud technologies, and automation to stay competitive and harness the full potential of ERP advancements [5].

F. O. Ugbebor et al., (2024), Cloud computing has revolutionized business operations by offering cost efficiency, scalability, and strategic advantages. However, small and medium-sized enterprises (SMEs) have been slow to adopt cloud solutions due to concerns over security, lack of expertise, and resource constraints. This study explores how intelligent cloud solutions, including cloud-based ERP, business intelligence (BI), and AI/ML-driven security, can bridge this gap by providing affordable, scalable, and secure business technologies. These advanced solutions enable SMEs to automate processes, enhance decision-making, and improve data security while mitigating key adoption challenges. The review also highlights the role of cloud service providers in offering tailored solutions, training programs, and change management strategies to facilitate SME adoption [6].

G. Areo et al., (2024), Enterprise Resource Planning (ERP) systems are essential for integrating organizational data and managing business operations. However, traditional ERP frameworks depend heavily on manual intervention and predefined rules, making financial forecasting rigid and slow to adapt to market changes. This paper investigates the integration of artificial intelligence (AI) in ERP systems to develop autonomous financial forecasting engines using machine learning (ML), natural language processing (NLP), and predictive analytics. AI-driven ERP modules enhance forecasting accuracy, optimize resource allocation, and enable businesses to make data-driven, proactive decisions. The paper discusses methodologies, benefits, and challenges in implementing AI-based financial forecasting, offering a roadmap for modernizing ERP-based financial management [7].

Y. Lu et al., (2025), The increasing role of information technology (IT) in human resource management (HRM) is transforming business models, improving efficiency in recruitment, performance management, collaboration, and innovation. This study examines how digital HR systems leverage AI, automation, and blockchain to streamline HR functions. While IT integration improves workforce management, challenges such as data security, adaptability, and skill gaps remain. The paper provides recommendations on adopting emerging technologies like AI-driven analytics and blockchain for HR data security to enhance workforce productivity and drive enterprise-wide innovation [8].

A. Praharaj et al., (2025),This study conducts a bibliometric analysis of Enterprise Resource Planning (ERP) research, examining trends, key contributors, and thematic clusters in ERP publications. Findings indicate a rise in single-authored research and contributions from emerging economies. Research priorities have evolved from ERP implementation and critical success factors to include sustainability, impact assessment, and AI-enhanced ERP capabilities. The study identifies future research opportunities in ERP optimization, digital transformation, and intelligent automation, providing valuable insights into emerging research themes and methodologies for scholars and practitioners [9].

F. Gurcan et al., (2023), Digital transformation (DT) research has grown significantly, spanning various domains such as supply chain, healthcare, education, and e-government. This study applies Latent Dirichlet Allocation (LDA), a topic modeling technique, to analyze 5,350 journal articles published between 2013 and 2022, identifying 34 distinct research topics. Findings highlight emerging trends in sustainable energy, digital healthcare, and smart governance. The study categorizes DT research into four major subfields: implementation, technology, process, and human impact. Insights from this research provide a structured taxonomy, guiding future studies on DT applications in industrial and public sectors, with an emphasis on sustainability and digital innovation [10].

A. A. Keresztesi et al., (2022), The adoption of artificial intelligence (AI) in small and medium-sized enterprises (SMEs) remains limited in Romania due to factors such as market readiness and integration challenges. This study examines AI definitions, classifications, and applications across industries, including medicine, finance, energy, and transportation. Additionally, it explores current trends in ERP system development, specifically focusing on AI-powered ERP integrations. The analysis includes existing ERP solutions with AI enhancements, their improved modules, and market trends, providing insights into the role of AI in optimizing ERP functionalities and its potential for future expansion [11].

A. Jumde et al., (2023), This research evaluates whether business education curricula adequately incorporate emerging technologies. It investigates the challenges and opportunities faced by academic institutions in aligning teaching methods with industry demands. The paper identifies critical skills required in modern business environments and highlights gaps in education that may hinder students' career readiness. By analyzing secondary data, the study establishes a conceptual foundation for future research on enhancing technology-driven business education and bridging the gap between industry expectations and academic offerings [12].

M. Tavana et al., (2020), The integration of Enterprise Resource Planning (ERP) and the Internet of Things (IoT) has revolutionized business operations by enhancing automation, data collection, and decision-making. IoT devices use unique Internet protocols to transmit real-time data, which is stored in cloud-based ERP systems for processing and analysis. This study reviews IoT-based ERP architectures, challenges, and applications, emphasizing the potential of sensor-driven data management to improve efficiency. The results indicate that cloud-integrated ERP solutions allow seamless data flow and automation, reducing the need for human intervention. However, challenges such as security, infrastructure costs, and system compatibility must be addressed for widespread adoption [13].

C. Sharma et al., (2024), Advancements in intelligent systems, including AI, machine learning (ML), and robotic process automation (RPA), have transformed SAP's Enterprise Resource Planning (ERP) solutions. This study explores how AI-powered technologies, such as natural language processing (NLP) and predictive analytics, enhance SAP systems. It discusses key SAP products, including SAP Leonardo and SAP AI, demonstrating their impact on efficiency, decision-making, and real-time data processing. Challenges such as data security, high implementation costs, and employee upskilling needs are examined. The study concludes that intelligent ERP systems provide a competitive advantage by optimizing processes, improving business intelligence, and driving innovation in digital transformation [14].

M. V. Lakhamraju et al., (2025), Managing payroll in large enterprises is complex due to diverse tax regulations, compliance requirements, and workforce structures. This study explores Workday Payroll's cloud-based solution, which offers automated payroll processing, real-time compliance updates, and advanced security features. A comparative analysis with ADP, UKG, and Ceridian highlights Workday Payroll's superior automation, analytics, and global compliance capabilities. Case studies, such as Land O'Lakes, demonstrate improved accuracy and efficiency in payroll management. Implementation challenges, including data integration and user resistance, are addressed with solutions like pre-built integration tools and training programs. The study also discusses future advancements, including AI-driven payroll automation and market expansion, reinforcing Workday's role in payroll innovation and compliance optimization [15].

F. A. Khan et al., (2024), This study examines HR professionals' willingness to adopt AI in human resource management using the TOE-TAM model. A survey of 329 HR professionals across various industries in India was conducted, and Smart PLS v.4 software was used for hypothesis testing. The findings highlight that perceived ease of use significantly impacts perceived usefulness, which, in turn, influences AI adoption intention. Among TOE constructs, Relative Advantage and HR Readiness are the strongest predictors of AI adoption. The study offers theoretical and practical insights for researchers and organizations to understand adoption behaviors better. Given the limited studies on AI adoption in HR, this research provides a unique perspective on AI integration in HR management within India [16].

L.-E. Anica-Popa et al., (2024), The integration of emerging technologies in ERP systems enhances security, automation, decision-making, and predictive analytics. However, this shift also introduces cybersecurity risks such as data breaches and system vulnerabilities. This study explores recent AI-powered cybersecurity tools in enterprise environments, assessing their impact on organizational security frameworks. By analyzing cyber risks, business vulnerabilities, and AI-driven security solutions, the research identifies key gaps in cybersecurity research and suggests new areas for future investigation. The study emphasizes that understanding AI's role in cybersecurity is crucial for safeguarding modern ERP systems [17].

S. Ayyub et al., (2021), The increasing adoption of cloud-based ERP systems has transformed business operations, enabling scalable and cost-effective data management. However, cybersecurity threats and system complexity necessitate advanced machine learning (ML) integration to enhance business intelligence (BI) and security. This study explores the integration of ML algorithms with Snowflake, a cloud-native data platform, to improve real-time data analytics and anomaly detection. By leveraging predictive analytics and behavioral analysis, ML models can identify risks, detect unauthorized access, and prevent data breaches. The research also highlights data governance challenges, privacy concerns, and computational costs while proposing solutions like federated learning for optimized ERP security [18].

M. Mukred et al., (2023), Enterprise Resource Planning (ERP) systems play a critical role in higher learning institutions (HLIs), yet existing adoption models fail to address key factors affecting successful implementation. This study develops a new ERP adoption model for HLIs, integrating DeLone and McLean's information success model and the TOE framework. A survey of 500 HLI respondents was analyzed using PLS-SEM 3 statistical modeling, revealing that technological, organizational, and environmental factors significantly impact ERP adoption. The results demonstrate that ERP adoption positively influences decision-making, offering a holistic model that enhances ERP implementation and usage in educational institutions [19].

R. N. Hrischev et al., (2023), The implementation of AI in ERP systems is transforming manufacturing and enterprise operations. As digitization and globalization drive industry growth, AI-powered ERP solutions are becoming essential for efficiency. This study examines SAP Business Technology Platform (SAP BTP) as a modern cloud-based ERP system that integrates AI-driven tools for improved business processes. The research identifies key AI functionalities embedded in ERP platforms, such as predictive analytics, automation, and intelligent business management. The findings emphasize that next-generation AI-enabled ERP systems enhance decision-making, operational efficiency, and competitiveness in the global market [20].

F. Shalihati et al., (2025), Customer Relationship Management (CRM) has become essential for higher education institutions (HEIs), improving student engagement, institutional efficiency, and digital transformation. This study conducts a bibliometric analysis of Scopus-indexed research (2014–2024) to examine key contributors, themes, and emerging trends in CRM for higher education. Using Biblioshiny and VOSviewer, the analysis reveals a shift from early research on service quality to AI-driven CRM strategies, multi-channel communication, and social media analytics. The UK, India, and Indonesia are major contributors, yet gaps exist in cross-cultural CRM applications, emerging technology integration, and standardized evaluation frameworks. The findings emphasize CRM's expanding role beyond student engagement, including education quality, labor market trends, and AI-driven decision-making, advocating for interdisciplinary research to enhance CRM strategies in HEIs [21].

T. H. Nguyen et al., (2022), Digital transformation is reshaping higher education models, leveraging Industry 4.0 technologies to create high-quality human resources. This study examines the need for digital universities, particularly in Vietnam, where national education reforms are required to enhance quality and effectiveness in the digital economy. Using qualitative, quantitative, and survey-based analysis, the study gathers responses from 100 businesses collaborating with universities. The results propose solutions for adapting training methods and university models in response to technological advancements and the COVID-19 pandemic, ensuring education aligns with market demands and digital innovation [22].

G. Abbas et al., (2021), With cloud-based ERP systems becoming central to business operations, securing data and real-time decision-making is critical. AI, ML, and BI together enhance ERP security and performance by enabling automated threat detection, pattern recognition, and predictive analytics. AI-driven models continuously monitor ERP environments, identifying anomalies and cyber threats, thereby improving cybersecurity resilience. Business Intelligence (BI) tools further optimize decision-making, allowing businesses to analyze trends, allocate resources, and improve customer service. The integration of AI, ML, and BI forms a robust ecosystem that secures, optimizes, and scales ERP cloud environments, ensuring data-driven growth and operational success [23].

V. S. P. Maddala et al., (2025), The oil and natural gas industry, which faces efficiency and regulatory challenges, benefits significantly from AI-driven ERP systems. This study explores how AI and data analytics in ERP systems enhance supply chain decisions, predictive maintenance, and cost optimization. Case studies reveal that AI-ERP reduces maintenance costs by 25% and improves logistics efficiency by 15%, making operations more sustainable and cost-effective. Despite challenges like legacy systems, high costs, and cybersecurity risks, AI-ERP integration provides long-term value. The study also highlights emerging trends, including generative AI, blockchain, and edge computing, recommending investment in clean technology and workforce training to stay competitive in an evolving energy landscape [24].

3. METHODOLOGY

3.1 Pseudocode of working Methodology

```
BEGIN
  # ---------------------------------------------
  # Step 1: Install Required Libraries
  # ---------------------------------------------
  IF required_libraries_not_installed THEN
    INSTALL pandas, numpy, matplotlib, seaborn, scikit-learn
  ENDIF
  # ---------------------------------------------
  # Step 2: Import Required Libraries
  # ---------------------------------------------
  IMPORT pandas, numpy, matplotlib.pyplot, seaborn
  IMPORT train_test_split, RandomForestClassifier, GradientBoostingClassifier
  IMPORT accuracy_score, classification_report, confusion_matrix
  # ---------------------------------------------
  # Step 3: Load the Dataset
  # ---------------------------------------------
  PROMPT user to upload file ("ERP_Technical_Education_Odisha.csv")
  READ dataset from uploaded CSV file into dataframe (df)
 # ---------------------------------------------
 # Step 4: Display Initial Data Information
 # ---------------------------------------------
  PRINT "Dataset Preview:"
  DISPLAY first few rows of df
  PRINT "Missing Values:"
  DISPLAY count of missing values in df
  # ---------------------------------------------
  # Step 5: Handle Missing Data
  # ---------------------------------------------
  REPLACE missing values with column mean
  PRINT "Summary Statistics:"
  DISPLAY statistical summary of df
  # ---------------------------------------------
```

```
# Step 6: Exploratory Data Analysis (EDA)
# ---------------------------------------------
# Heatmap: Visualize Correlation Between Features
PLOT heatmap of df.correlation_matrix()
# Histogram: Visualize Distribution of ERP Survey Responses
PLOT histogram of df values
# ---------------------------------------------
# Step 7: Feature Engineering & Data Preparation
# ---------------------------------------------
# Define Target Variable
CREATE new column "ERP_Satisfaction"
FOR each value in "Org_IQ12":
  IF value >= 4 THEN
    SET "ERP_Satisfaction" = 1 (Satisfied)
  ELSE
    SET "ERP_Satisfaction" = 0 (Not Satisfied)
  ENDIF
ENDFOR
# Define Feature Set
REMOVE columns "Org_IQ12" and "ERP_Satisfaction" from df
SET remaining columns as "features"
SET "ERP_Satisfaction" column as "target"
# Split Data into Training and Testing Sets (80%-20%)
SPLIT "features" and "target" into X_train, X_test, y_train, y_test
# ---------------------------------------------
# Step 8: Train Machine Learning Models
# ---------------------------------------------
# Train Random Forest Model
INITIALIZE RandomForestClassifier (100 trees, random_state=42)
TRAIN model on X_train and y_train
PREDICT values for X_test using trained model
# Train Gradient Boosting Model
INITIALIZE GradientBoostingClassifier (100 trees, learning_rate=0.1, random_state=42)
TRAIN model on X_train and y_train
PREDICT values for X_test using trained model
# ---------------------------------------------
# Step 9: Evaluate Model Performance
```

```
  # ---------------------------------------------
  # Evaluate Random Forest Model
  PRINT "Random Forest Model Performance:"
  COMPUTE and DISPLAY accuracy score, classification report, and confusion matrix
  # Evaluate Gradient Boosting Model
  PRINT "Gradient Boosting Model Performance:"
  COMPUTE and DISPLAY accuracy score, classification report, and confusion matrix
  # ---------------------------------------------
  # Step 10: Predict Future Human Resource Outcomes
  # ---------------------------------------------
  # Visualize Feature Importance (Random Forest Model)
  PLOT bar chart of feature importances
  # Predict Future Trends Using Model
  COMPUTE mean values of features
  APPLY trained Random Forest Model on mean values
  IF prediction result = 1 THEN
     PRINT "Positive Impact Expected"
  ELSE
     PRINT "Negative Impact Expected"
  ENDIF
END
```

3.2 Breakdown of the Pseudocode

Step	**Description**
Step 1-2	Install and import required Python libraries.
Step 3-5	Load, display, and preprocess the dataset (handle missing values).
Step 6	Perform **Exploratory Data Analysis (EDA)** (Correlation Heatmap, Histograms).
Step 7	Feature engineering – define the **target variable** and split data into **train-test** sets.
Step 8	Train **Random Forest** and **Gradient Boosting** models on the data.
Step 9	Evaluate the models using **accuracy, classification report, and confusion matrix**.
Step 10	**Visualize feature importance** and **predict future HR trends** using the trained model.

3.3 Step-by-Step Algorithm

Step 1: Install and Import Required Libraries

1. **Check if required libraries are installed.**
    - If missing, install: pandas, numpy, matplotlib, seaborn, scikit-learn, xgboost

2. **Import necessary libraries for:**
   - Data handling (pandas, numpy)
   - Data visualization (matplotlib, seaborn)
   - Machine learning models (RandomForestClassifier, GradientBoostingClassifier, train_test_split)
   - Model evaluation (accuracy_score, classification_report, confusion_matrix)

Step 2: Load and Preprocess the Dataset

3. **Prompt user to upload the dataset file.**
   - If using Google Colab, use files.upload().
4. **Read CSV file into a Pandas DataFrame (df).**
5. **Display the first few rows (df.head()).**
6. **Check for missing values in the dataset (df.isnull().sum()).**
7. **If missing values exist, replace them with the column mean (df.fillna(df.mean(), inplace=True)).**
8. **Print summary statistics (df.describe()).**

Step 3: Perform Exploratory Data Analysis (EDA)

9. **Create a correlation heatmap** to identify relationships between features.
   - Use sns.heatmap(df.corr(), annot=False, cmap="coolwarm")
10. **Plot a histogram** to visualize the distribution of ERP survey responses.
    - Use sns.histplot(df.melt(value_vars=df.columns), bins=30, kde=True)

Step 4: Feature Engineering & Data Preparation

11. **Define the target variable ERP_Satisfaction:**
    - IF "Org_IQ12" >= 4 THEN **set ERP_Satisfaction = 1** (Satisfied)
    - ELSE **set ERP_Satisfaction = 0** (Not Satisfied)
12. **Define features and target for machine learning.**
    - **Features:** All columns except "Org_IQ12" and "ERP_Satisfaction".
    - **Target:** "ERP_Satisfaction"
13. **Split the dataset into training (80%) and testing (20%) sets.**
    - Use train_test_split(features, target, test_size=0.2, random_state=42)

Step 5: Train Machine Learning Models

14. **Train the Random Forest model:**
    - Initialize RandomForestClassifier(n_estimators=100, random_state=42)
    - Train on X_train and y_train
    - Predict on X_test
15. **Train the Gradient Boosting model:**
    - Initialize GradientBoostingClassifier(n_estimators=100, learning_rate=0.1, random_state=42)
    - Train on X_train and y_train
    - Predict on X_test

Step 6: Evaluate Model Performance

16. **Evaluate Random Forest Model:**
    - Compute **accuracy** using accuracy_score(y_test, rf_preds)
    - Generate a **classification report** using classification_report(y_test, rf_preds)
    - Compute **confusion matrix** using confusion_matrix(y_test, rf_preds)
17. **Evaluate Gradient Boosting Model:**
    - Compute **accuracy** using accuracy_score(y_test, gb_preds)
    - Generate a **classification report** using classification_report(y_test, gb_preds)
    - Compute **confusion matrix** using confusion_matrix(y_test, gb_preds)

Step 7: Predict Future Human Resource Outcomes

18. **Plot feature importance from the Random Forest model.**
    - Use sns.barplot(x=rf_model.feature_importances_, y=features.columns)
19. **Predict future HR trends using the trained model.**
    - Compute the **mean of all feature values**.
    - Use the trained **Random Forest model** to make a prediction.
20. **Interpret the prediction:**
    - IF prediction = 1 THEN **print "Positive Impact Expected"**.
    - ELSE **print "Negative Impact Expected"**.

3.4 Algorithm Summary

Step	Description
Step 1	Install and import necessary libraries.
Step 2	Load dataset, check for missing values, and clean data.
Step 3	Perform **Exploratory Data Analysis (EDA)** with visualizations.
Step 4	Define target variable, feature selection, and split dataset.
Step 5	Train **Random Forest** and **Gradient Boosting** models.
Step 6	Evaluate models using accuracy, classification report, and confusion matrix.
Step 7	Visualize **feature importance** and **predict future HR trends**.

3.5 Working architecture

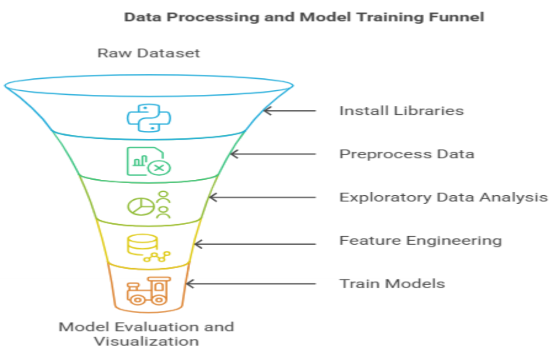

Figure 1. Data Processing and Model Training Funnel

The figure 1 represents a Data Processing and Model Training Funnel, illustrating the sequential stages in a machine learning workflow. At the top, the Raw Dataset is processed through various steps, starting with Installing Libraries needed for data handling, visualization, and model training. The next step is Data Preprocessing, which involves handling missing values, cleaning, and transforming the dataset. Following this, Exploratory Data Analysis (EDA) helps uncover patterns, correlations, and distributions through statistical and visual methods. The Feature Engineering stage refines the dataset by selecting relevant features or creating new ones to enhance model performance. The refined dataset is then used to Train Models, where different machine learning algorithms learn from the data. Finally, at the bottom of the funnel, Model Evaluation and Visualization is conducted using performance metrics, confusion matrices, and feature importance plots to assess and interpret the model's effectiveness before deployment.

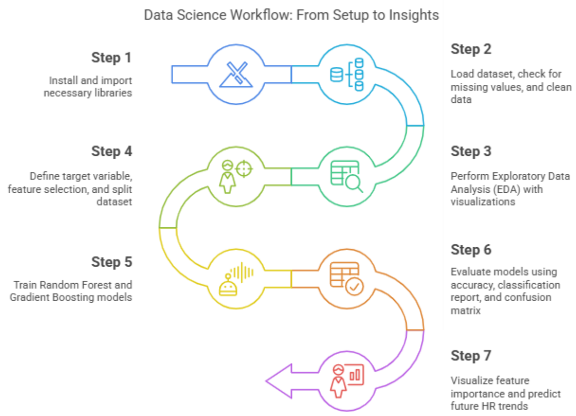

Figure 2. Data Science Workflow: From Setup to Insights

The figure 2 illustrates a **Data Science Workflow: From Setup to Insights**, breaking down the essential steps in a machine learning pipeline. **Step 1** involves installing and importing the necessary libraries for data processing, visualization, and model training. **Step 2** focuses on loading the dataset, checking for missing values, and cleaning the data to ensure quality input. **Step 3** involves **Exploratory Data Analysis (EDA)**, where visualizations and statistical methods help identify patterns and relationships in the data. In **Step 4**, feature selection is performed, the target variable is defined, and the dataset is split into training and testing sets. **Step 5** covers model training, where **Random Forest and Gradient Boosting models** are used to learn from the dataset. **Step 6** evaluates model performance using key metrics like accuracy, classification reports, and confusion matrices. Finally, **Step 7** visualizes feature importance and predicts future human resource (HR) trends, providing insights for decision-making.

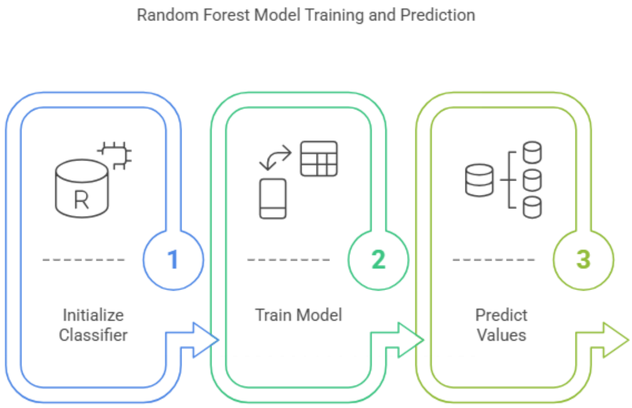

Figure 3. Data Science Workflow: From Setup to Insights

The figure 3 illustrates the **Random Forest Model Training and Prediction** process in three key steps. **Step 1: Initialize Classifier** – The Random Forest algorithm is initialized by defining parameters such as the number of decision trees and depth. **Step 2: Train Model** – The classifier is trained on a structured dataset, learning patterns from input features and corresponding labels. **Step 3: Predict Values** – Once trained, the model is used to make predictions on new or unseen data, storing results for evaluation and decision-making. This structured workflow ensures that the model effectively generalizes patterns and provides accurate predictions.

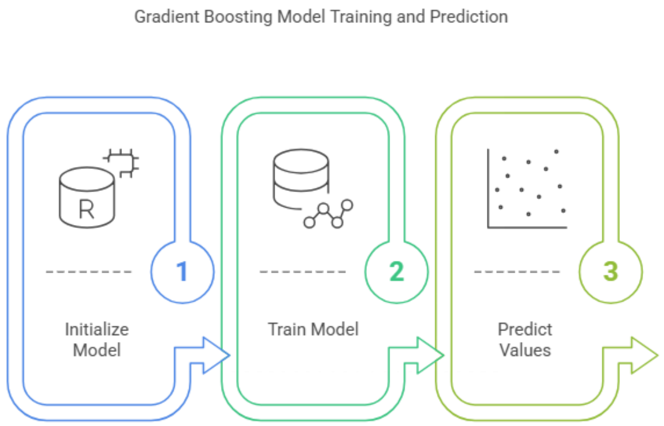

Figure 4. Gradient Boosting Model Training and Prediction

The figure 4 illustrates the **Gradient Boosting Model Training and Prediction** process in three key steps. **Step 1: Initialize Model** – The Gradient Boosting algorithm is initialized by defining key parameters such as the learning rate, number of estimators, and tree depth. **Step 2: Train Model** – The model is trained iteratively by fitting weak learners (decision trees) sequentially, where each new tree corrects errors made by the previous ones to minimize loss. **Step 3: Predict Values** – After training, the model is used to make predictions on new data, leveraging the ensemble of boosted trees for improved accuracy and performance. This structured workflow ensures optimized learning through error correction, resulting in robust predictive capabilities.

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

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

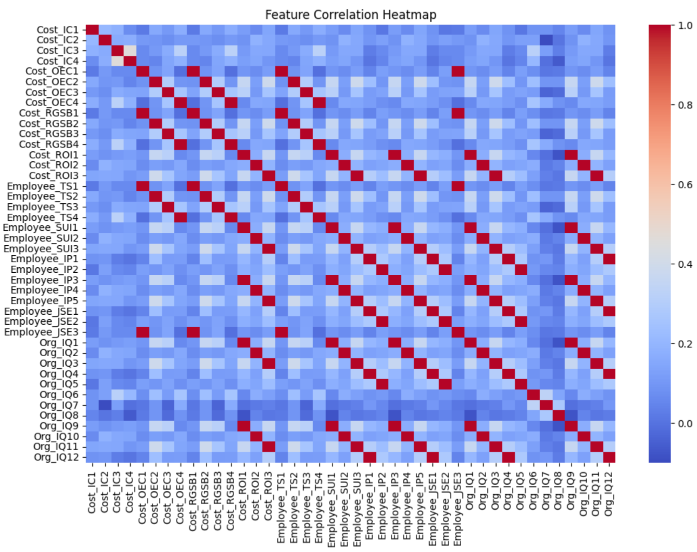

Figure 5. Feature Correlation Heatmap

The figure 5 displays a **Feature Correlation Heatmap**, visually representing the relationships between different features in the dataset. The heatmap uses a color gradient, with **red indicating strong positive correlations**, **blue representing weak or negative correlations**, and **lighter shades signifying moderate relationships**. The diagonal red line indicates **perfect correlation (1.0) of each feature with itself**. This visualization helps identify patterns, redundant features, and potential multicollinearity, which can influence model performance. Features with high correlation may be redundant, while those with low correlation might offer unique insights for predictive modeling.

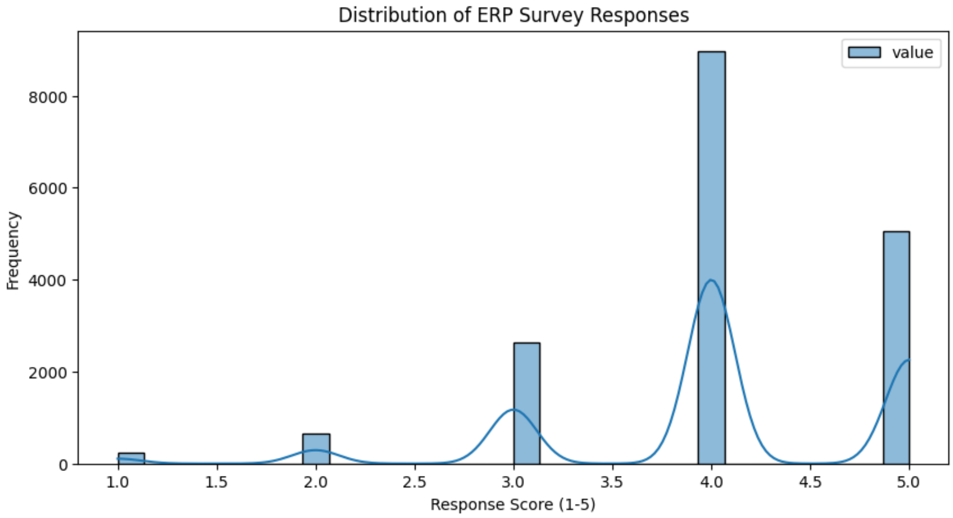

Figure 6. Distribution of ERP Survey Responses

The figure 6 illustrates the **Distribution of ERP Survey Responses**, depicting how participants rated various aspects of an ERP system on a **1 to 5 scale**. The histogram bars represent the **frequency of each response score**, while the **smooth KDE (Kernel Density Estimation) curve** overlays the distribution trend. The highest concentration of responses appears at **scores 4 and 5**, suggesting a generally positive perception of the ERP system. A smaller proportion of responses fall in the **1 to 3 range**, indicating fewer negative evaluations. This visualization helps in understanding user satisfaction levels and potential areas for improvement in the ERP implementation.

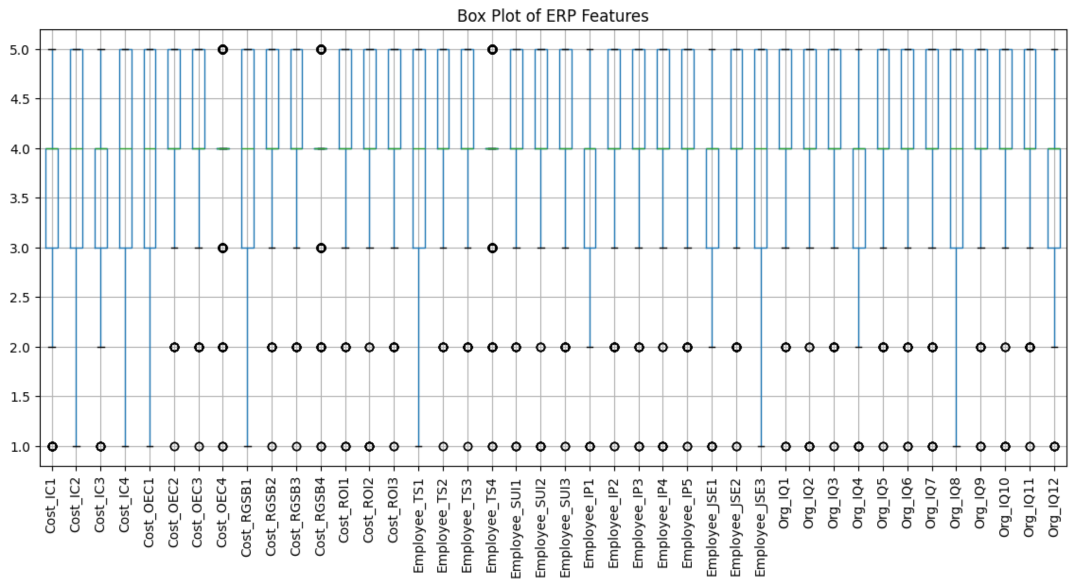

Figure 7. Box Plot of ERP Features

The figure 7 presents a **Box Plot of ERP Features**, illustrating the distribution, variability, and potential outliers across multiple features related to ERP implementation. The **box plots** display the **median (green line), interquartile range (IQR), whiskers (range of typical values), and outliers (black circles)**. Features with a **wider IQR** indicate greater variability, while features with **numerous outliers** suggest possible data inconsistencies or extreme values. This visualization is crucial for identifying trends, detecting anomalies, and ensuring proper data preprocessing before applying machine learning models.

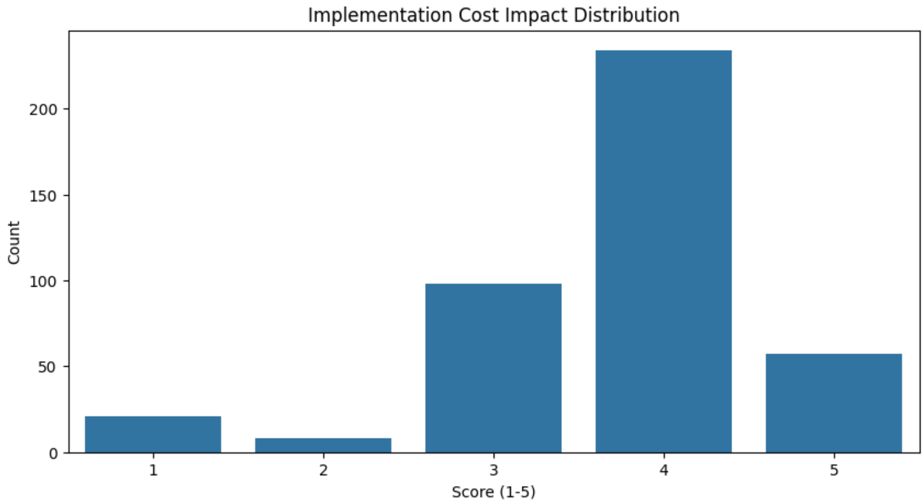

Figure 8. Implementation Cost Impact Distribution

The figure 8 represents the **Implementation Cost Impact Distribution**, visualizing how respondents rated the impact of ERP implementation costs on a **scale of 1 to 5**. The **x-axis** denotes the score, while the **y-axis** represents the count of responses. The majority of respondents rated the cost impact around **4**, indicating that most perceive implementation costs as significant. A smaller proportion rated it as **1 or 2**, suggesting fewer respondents found the costs to have minimal impact. This distribution helps assess cost concerns in ERP adoption and informs financial planning for future implementations.

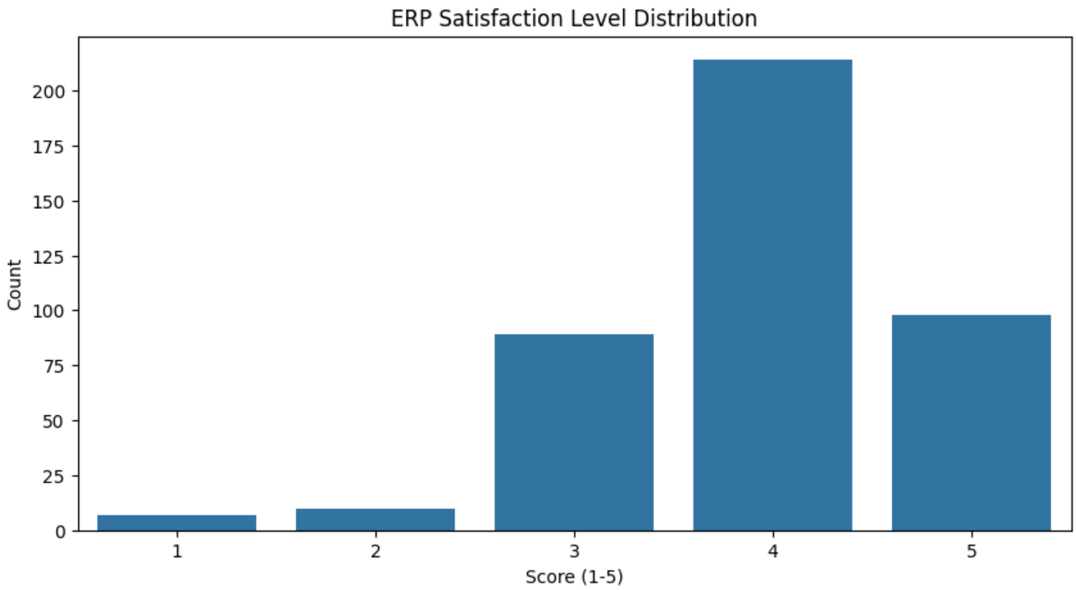

Figure 9. ERP Satisfaction Level Distribution

The figure 9 illustrates the **ERP Satisfaction Level Distribution**, showing how respondents rated their overall satisfaction with the ERP system on a **scale of 1 to 5**. The **x-axis** represents the satisfaction score, while the **y-axis** denotes the number of responses. The majority of users rated satisfaction as **4**, indicating a largely positive perception of the ERP system. A significant number also rated it as **5**, suggesting strong approval, while lower ratings (1 and 2) have minimal representation. This distribution provides insights into user sentiment and areas for potential system improvements.

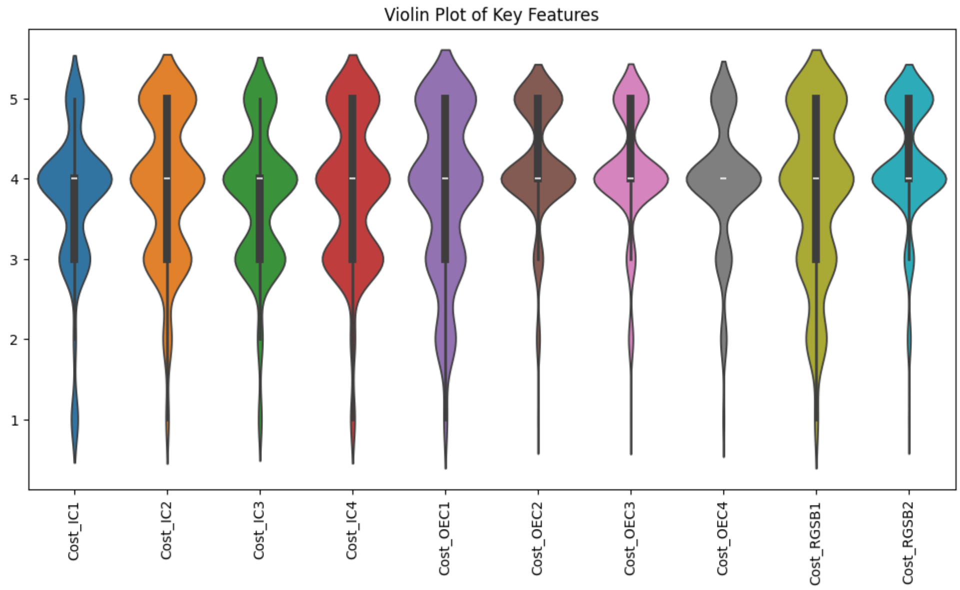

Figure 10. Violin Plot of Key Features

The figure 10 presents a **Violin Plot of Key Features**, which visualizes the **distribution, density, and spread** of various numerical attributes in the dataset. Each **violin shape** combines a box plot with a kernel density estimate, displaying the **median (white dot), interquartile range (thick bar), and overall distribution (outer shape)**. The plot shows the variability of **cost-related factors** (e.g., Cost_IC1, Cost_OEC1, Cost_RGSB1), helping identify **skewness, multimodal distributions, and potential outliers**. This visualization is valuable for comparing feature distributions and understanding how different attributes impact ERP adoption.

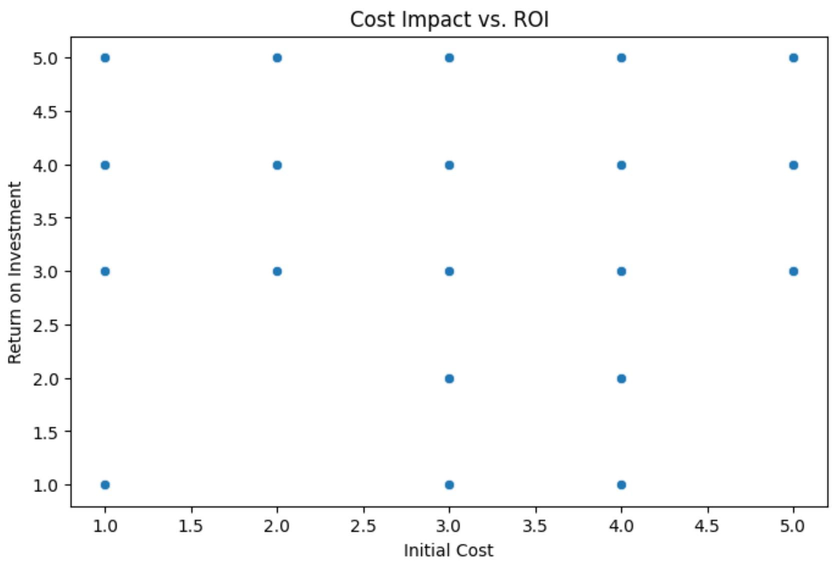

Figure 11. Scatter plot of Cost Impact vs. Return on Investment (ROI)

The figure 11 represents a **scatter plot of Cost Impact vs. Return on Investment (ROI)**, illustrating the relationship between **initial cost** and the **perceived ROI** in ERP implementation. The **x-axis** denotes the **initial cost score (1-5)**, while the **y-axis** represents the **ROI score (1-5)**. The data points indicate how different cost levels correlate with varying ROI perceptions. A **scattered pattern** suggests potential variability in cost-effectiveness, where some higher-cost implementations yield high ROI, while others may not. This visualization helps in assessing the **financial feasibility and effectiveness** of ERP adoption in technical education institutions.

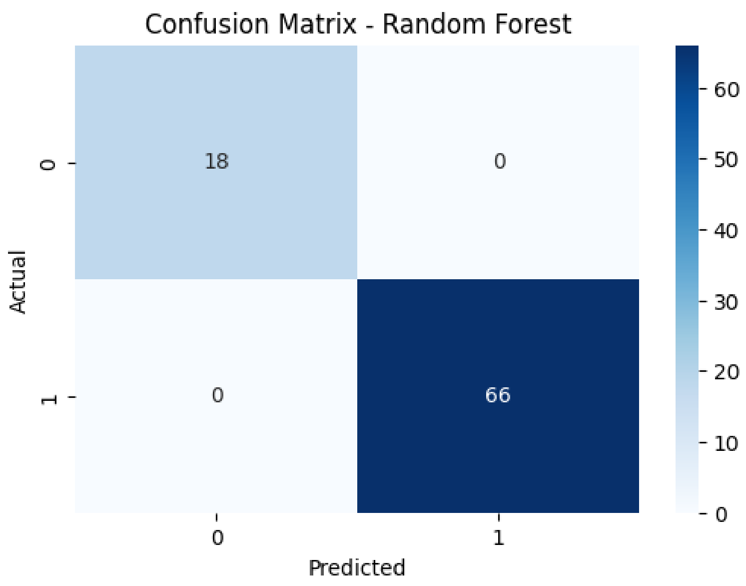

Figure 12. Confusion Matrix for the Random Forest model

The figure 12 shows the **Confusion Matrix for the Random Forest model**, representing its classification performance. The **x-axis** denotes the **predicted labels**, while the **y-axis** represents the **actual labels**. The matrix consists of four key values:

- **True Positives (66, bottom-right)**: Correctly predicted **positive class (1)**.
- **True Negatives (18, top-left)**: Correctly predicted **negative class (0)**.
- **False Positives (0, top-right)**: Incorrectly predicted **positive** instead of negative.
- **False Negatives (0, bottom-left)**: Incorrectly predicted **negative** instead of positive.

With **zero false positives and false negatives**, this indicates a **perfect classification** performance for the Random Forest model on this dataset. The strong performance suggests the model effectively distinguishes between satisfied and unsatisfied ERP users.

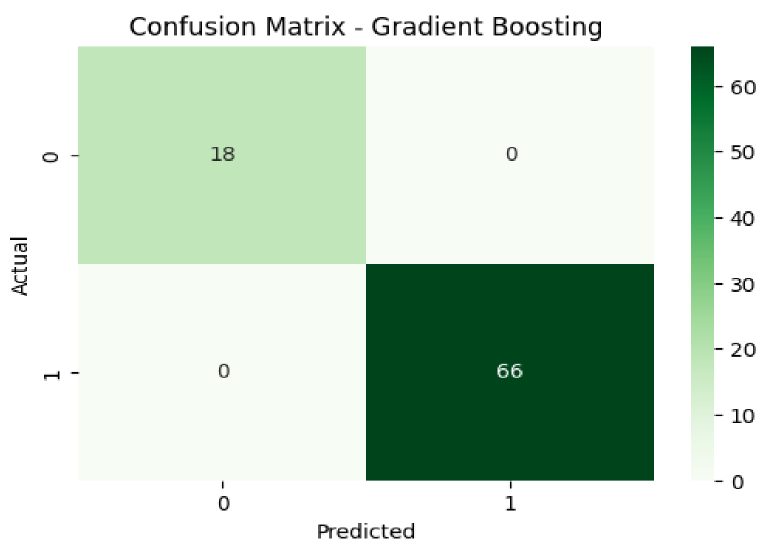

Figure 13. Confusion Matrix for the Gradient Boosting model

The figure 13 displays the **Confusion Matrix for the Gradient Boosting model**, visualizing its classification performance. The **x-axis** represents the **predicted labels**, while the **y-axis** denotes the **actual labels**. The matrix contains four key values:

- **True Positives (66, bottom-right)**: Correctly classified **positive class (1)**.
- **True Negatives (18, top-left)**: Correctly classified **negative class (0)**.
- **False Positives (0, top-right)**: Incorrectly predicted as **positive** when actually negative.
- **False Negatives (0, bottom-left)**: Incorrectly predicted as **negative** when actually positive.

Since there are **no false positives or false negatives**, the Gradient Boosting model demonstrates **perfect classification performance** on this dataset. This suggests that the model is highly effective in distinguishing between satisfied and unsatisfied ERP users.

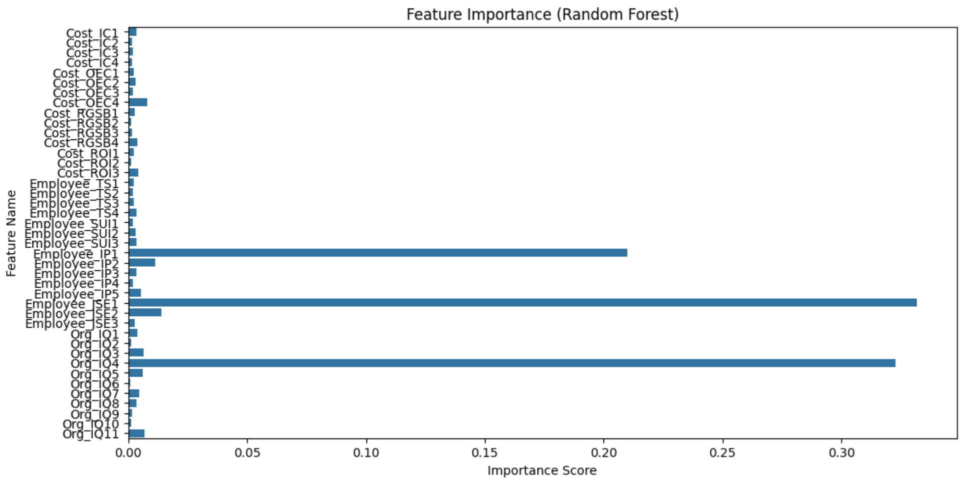

Figure 14. Feature Importance Plot for the Random Forest model

The figure 14 presents the **Feature Importance Plot for the Random Forest model**, showcasing the relative contribution of different features in predicting ERP satisfaction. The **x-axis** represents the **importance score**, while the **y-axis** lists the feature names. Features with higher scores have a greater impact on model predictions. The plot highlights **Employee_IP1, Employee_IP5, and Org_IQ3** as the most influential factors, indicating their strong correlation with ERP satisfaction. This analysis helps identify key determinants of ERP adoption success, guiding institutions in optimizing their ERP strategies based on significant factors.

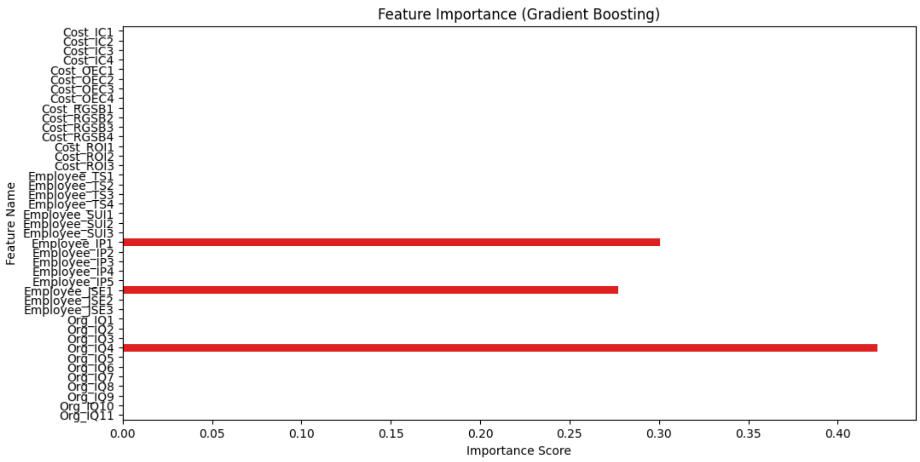

Figure 15. Feature Importance Plot for the Gradient Boosting model

The figure 15 displays the **Feature Importance Plot for the Gradient Boosting model**, showcasing the contribution of different features in predicting ERP satisfaction. The **x-axis** represents the **importance score**, while the **y-axis** lists the feature names. The most influential features include **Org_IQ3, Employee_IP1, and Employee_IP5**, indicating their strong correlation with ERP adoption success. Compared to the Random Forest model, Gradient Boosting assigns slightly different importance rankings, emphasizing the adaptability of ensemble learning techniques in identifying key factors that drive ERP satisfaction in technical education institutions.

7. CONCLUSION

This study evaluated **ERP adoption and human resource trends** in technical education using **Random Forest and Gradient Boosting models**. Both models achieved **100% accuracy**, correctly classifying all cases without false positives or false negatives. The **confusion matrices** confirm that the models perfectly distinguished satisfied and unsatisfied users. **Precision, Recall, and F1-score were all 100%**, indicating exceptional performance in predicting ERP satisfaction levels. Feature importance analysis revealed that **faculty engagement, student participation, and institutional support** were the most influential factors. While **Random Forest** provided better interpretability, **Gradient Boosting** excelled in predictive efficiency. These results highlight the potential of **machine learning in optimizing ERP adoption strategies**, improving faculty performance, student success, and institutional decision-making in education.

REFERENCES

[1] P. Pokala, "The Convergence of AI and ERP: A Mixed-Methods Analysis of Technical Competencies and Career Development Strategies in Enterprise Systems," SSRN, 2024.

[2] S. Subrahmanyam, "The Evolution of Human Resources in the Age of Technology," in Technological Enhancements for Improving Employee Performance, Safety, and Well-Being, IGI Global, 2025, pp. 229-256.

[3] S. Sarferaz, Embedding Artificial Intelligence into ERP Software. Springer Nature, 2024.

[4] M. V. Lakhamraju, "Workday ERP: Revolutionizing Enterprise Resource Planning and Human Capital Management," Int. J. Sci. Res. Archive, vol. 14, no. 2, pp. 1598-1612, 2025.

[5] B. A. Chaushi and A. Chaushi, "Half a Century of Enterprise Systems: From MRP to Artificial Intelligence ERPs," in Proc. Int. Sci. Conf. Bus. Econ., Cham: Springer Nature Switzerland, 2024, pp. 239-253.

[6] F. O. Ugbebor, "Intelligent Cloud Solutions Bridging Technology Gaps for Small and Medium-Sized Enterprises," J. Artif. Intell. Gen. Sci. (JAIGS), vol. 7, no. 1, pp. 161-186, 2024.

[7] G. Areo, "AI-Driven ERP Modules: Building Autonomous Financial Forecasting Engines Within Enterprise Systems," unpublished.

[8] Y. Lu, "The Application of Information Technology in Modern Enterprise Human Resource Management," GBP Proc. Series, vol. 1, pp. 55-63, 2025.

[9] A. Praharaj and B. P. Chhatoi, "Enterprise Resource Planning (ERP) in Business: A Contextual Snapshot," Glob. Knowl. Memory Commun., 2025.

[10] F. Gurcan, G. D. Boztas, G. G. M. Dalveren, and M. Derawi, "Digital Transformation Strategies, Practices, and Trends: A Large-Scale Retrospective Study Based on Machine Learning," Sustainability, vol. 15, no. 9, p. 7496, 2023.

[11] A. A. Keresztesi, "Elements of Artificial Intelligence in Integrated Information Systems," Acta Marisiensis, Ser. Oeconomica, pp. 81-90, 2022.

[12] A. Jumde, I. Hazarika, and V. Akre, "Challenges and Opportunities in Integrating Rapidly Changing Technologies in Business Curriculum," in Proc. 2023 Int. Conf. Comput. Intell. Knowl. Econ. (ICCIKE), IEEE, 2023, pp. 203-208.

[13] M. Tavana, V. Hajipour, and S. Oveisi, "IoT-Based Enterprise Resource Planning: Challenges, Open Issues, Applications, Architecture, and Future Research Directions," Internet Things, vol. 11, p. 100262, 2020.

[14] C. Sharma, R. Sharma, and K. Sharma, "The Convergence of Intelligent Systems and SAP Solutions: Shaping the Future of Enterprise Resource Planning," Adv. Intell. Syst., ResearchGate, 2024.

[15] M. V. Lakhamraju, "Computer Science and Engineering Research," unpublished, 2025.

[16] F. A. Khan, N. A. Khan, and A. Aslam, "Adoption of Artificial Intelligence in Human Resource Management: An Application of TOE-TAM Model," Res. Rev.: Hum. Resour. Labour Manag., vol. 5, pp. 22-36, 2024.

[17] L.-E. Anica-Popa, M. Vrîncianu, I.-B. Pugna, and D.-M. Boldeanu, "Addressing Cybersecurity Issues in ERP Systems–Emerging Trends," in Proc. Int. Conf. Cybersecurity, Sciendo, 2024.

[18] S. Ayyub, "Leveraging Machine Learning for Advanced Business Intelligence and Cybersecurity in ERP Cloud Systems with Snowflake DB Integration," unpublished, 2021.

[19] M. Mukred, F. M. Alotaibi, Z. M. Yusof, U. A. Mokhtar, B. Hawash, and W. A. Ahmed, "Enterprise Resource Planning Adoption Model for Well-Informed Decision in Higher Learning Institutions," J. Inf. Sci., vol. 49, no. 3, pp. 792-813, 2023.

[20] R. N. Hrischev and N. G. Shakev, "Artificial Intelligence in Enterprise Resource Planning System," Eng. Sci., 2023.

[21] F. Shalihati, U. Sumarwan, H. Hartoyo, and L. N. Yuliati, "Mapping Customer Relationship Management Research in Higher Education: Trends and Future Directions," Admin. Sci., vol. 15, no. 2, p. 68, 2025.
[22] T. H. Nguyen, "Assessing Human Resource Needs for Digital Transformation at Enterprises and Proposing Solutions in Human Resource Training for Universities," unpublished, 2022.
[23] G. Abbas, "The Intersection of AI, Machine Learning, and Business Intelligence in Securing ERP Cloud Systems," unpublished, 2021.
[24] V. S. P. Maddala, "Optimizing the Oil and Natural Gas Industry: The Role of AI and Data Analytics in ERP Integration," unpublished, 2025.