iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: IV Month of publication: April 2026

DOI: https://doi.org/10.22214/ijraset.2026.79466

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

Employee Churn Prediction Platform: A Cloud-Native HR Analytics System with Generative AI-Driven Retention Strategies and Exit Interview Simulation

Selvi V[1], Akshya A[2], Rohan Kumar K V[3]
[1, 2]Department of Artificial Intelligence and Data Science, Sri Manakula Vinayagar Engineering College, Puducherry, India

Abstract: Employee attrition continues to be a critical issue for organizations, leading to increased recruitment costs, loss of skilled talent, and disruptions in workflow efficiency. Addressing this challenge requires not only accurate prediction but also actionable insights that support timely intervention. This paper introduces the Employee Churn Prediction Platform (ECPP), a cloud-based intelligent system designed to identify employees at risk of leaving and assist human resource teams in making informed retention decisions.The proposed platform leverages a serverless architecture built on AWS services, where employee-related data is collected, stored, and processed efficiently. Machine learning models, particularly XGBoost, are utilized to analyze historical patterns and predict the likelihood of employee churn with high accuracy. To enhance interpretability and usability, the system integrates a generative AI-based HR Retention Advisor that converts prediction results into meaningful recommendations tailored to individual employee profiles.

A unique feature of this platform is the Exit Interview Simulation module, which creates virtual employee scenarios based on predicted risk factors. This allows HR professionals to simulate conversations and evaluate potential retention strategies before applying them in real-world situations. Experimental evaluation demonstrates strong predictive performance, achieving an F1-score of 0.76 and an AUC-ROC of 0.91, along with positive feedback from users regarding system usability and effectiveness.Overall, the ECPP provides a scalable and practical solution for modern workforce management by combining predictive analytics, cloud computing, and AI-driven decision support.

Keywords: Employee Churn Prediction, HR Analytics, Amazon SageMaker, AWS Lambda, AWS Glue, Amazon S3, API Gateway, Generative AI, Retention Strategy, Exit Interview Simulation, XGBoost, SHAP Explainability

I. INTRODUCTION

The departure of talented employees represents one of the most consequential and costly challenges facing modern organizations. Research conducted by the Society for Human Resource Management (SHRM) estimates that replacing a single employee may cost anywhere between fifty percent and two hundred percent of that individual's annual salary, accounting for recruitment expenditure, onboarding time, and the loss of embedded institutional knowledge [12]. Despite these well-documented costs, the majority of organizations continue to rely on reactive practices such as exit interviews conducted after a resignation decision has already been made, leaving HR teams with little ability to intervene at a meaningful point in the employee's journey.

Advances in cloud computing, machine learning, and large language models have collectively created an unprecedented opportunity to transform workforce management from a reactive discipline into a proactive one. Predictive models trained on employee demographics, engagement scores, performance histories, and compensation data can identify individuals whose probability of resignation is elevated weeks or months before they formally submit notice. When these quantitative signals are paired with the natural language generation capabilities of modern AI systems, they can be translated into nuanced, context-sensitive retention recommendations that HR professionals can act upon immediately.

This paper introduces the Employee Churn Prediction Platform (ECPP), a fully cloud-native solution built on Amazon Web Services. The system covers the complete analytics lifecycle, from real-time data ingestion through Amazon API Gateway and AWS Lambda, to batch preprocessing with AWS Glue, model training and deployment via Amazon SageMaker, and the generation of personalized retention briefs by a generative AI HR Retention Advisor.

A novel Exit Interview Simulation feature rounds out the platform by enabling HR practitioners to rehearse retention conversations with AI-constructed virtual personas before engaging real employees.

The remainder of this paper is organized into sections that describe the methodology, system design, results, and conclusions.Section II surveys related work. Section III describes the system architecture. Section IV details the data pipeline. Section V covers the machine learning methodology. Section VI explains the generative AI components. Section VII presents the implementation and deployment details. Section VIII reports experimental results. Section IX describes the exit interview simulation module. Section X discusses ethical considerations. Section XI outlines limitations and future work. Section XII concludes the paper.

II. RELATED WORK

A. Predictive Attrition Modeling

Employee attrition prediction has attracted sustained research interest over the past decade. Zhao et al. [1] evaluated a range of classifiers including logistic regression, random forest, and gradient boosting on the IBM HR Analytics Employee Attrition dataset and demonstrated that ensemble methods consistently outperform single-classifier approaches on imbalanced HR data. Fallucchi et al. [2] extended this comparison to deep neural networks across multiple enterprise datasets, concluding that gradient-boosted trees offer a favorable balance between predictive accuracy and interpretability. A recurring limitation in the prior literature is the reliance on static, batch-processed datasets; the present work addresses this by integrating real-time data ingestion pipelines built on serverless cloud infrastructure.

B. Cloud-Native HR Analytics

Serverless and cloud-native architectures have begun to reshape the deployment of predictive HR systems. Cheng et al. [3] proposed a serverless analytics framework using AWS Lambda and API Gateway and demonstrated significant reductions in operational overhead compared to traditional on-premises deployments. Their framework, however, did not incorporate a machine learning inference layer or any form of generative AI for translating predictions into human-readable guidance. The present system builds directly on such architectural foundations while adding a full ML pipeline managed by SageMaker and a GenAI advisory layer.

C. Generative AI in Workforce Management

The application of large language models to HR workflows is an emerging area. Tambe et al. [4] examined the potential of AI-driven HR tools to improve employee experience and argued that conversational AI meaningfully augments HR capacity when deployed with appropriate safeguards. The concept of using LLMs to simulate employee personas for retention testing is novel in the HR context; the closest precedent appears in the conversational agent literature, where virtual personas are used to evaluate persuasion strategies in healthcare intervention settings [5]. This paper adapts that concept to workforce analytics, introducing the first reported exit interview simulation module for HR retention planning.

III. SYSTEM ARCHITECTURE

The ECPP is organized into four functional layers that map directly onto the mandatory cloud, data, and GenAI components of the platform: (1) the Data Ingestion Layer, (2) the Storage and Preprocessing Layer, (3) the Machine Learning Inference Layer, and (4) the Generative AI Advisory Layer. Fig. 1 illustrates the complete system architecture.

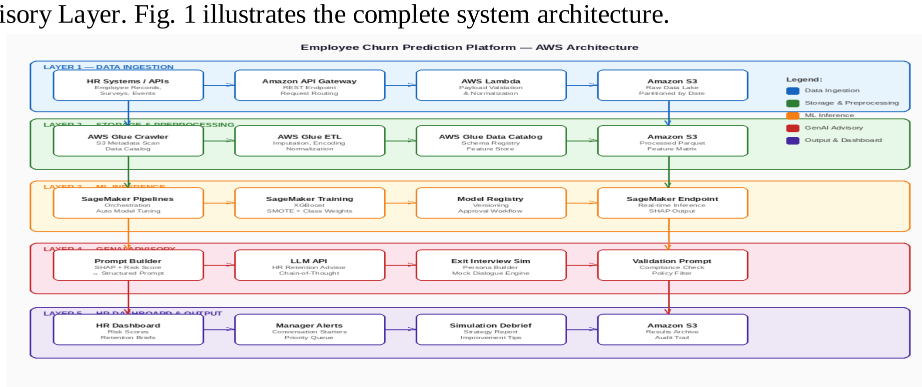


Fig. 1. End-to-End Architecture of the Employee Churn Prediction Platform on AWS

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

A. Data Ingestion Layer

Employee metrics arrive at the platform through a RESTful API exposed via Amazon API Gateway. Each API request triggers an AWS Lambda function that validates the incoming payload, normalizes field formats to a canonical schema, and routes records to their designated storage locations in Amazon S3. Lambda's serverless execution model eliminates the need to provision dedicated compute instances and allows the ingestion layer to scale automatically with request volume. Supported data types include structured HR records such as employee ID, department, job role, tenure, and performance ratings, alongside semi-structured engagement survey responses captured in JSON format.

B. Storage and Preprocessing Layer

Validated records are persisted in an Amazon S3 data lake partitioned by ingestion date and organizational unit. AWS Glue crawlers scan the S3 bucket on a scheduled basis to maintain an up-to-date metadata catalog in the AWS Glue Data Catalog. AWS Glue ETL jobs execute transformation pipelines that handle missing value imputation using median and mode strategies, categorical encoding of nominal variables, feature normalization, and construction of the final training-ready feature matrix. Processed datasets are written to a separate S3 prefix that serves as the input source for SageMaker training jobs.

C. Machine Learning Inference Layer

Model training and deployment are managed entirely within Amazon SageMaker. Training Jobs are launched with configurable instance types and hyperparameter search spaces, leveraging SageMaker Automatic Model Tuning for Bayesian optimization. Trained models are registered in the SageMaker Model Registry and promoted to production endpoints following evaluation on a held-out validation set. Real-time inference requests are routed from Lambda to SageMaker Endpoints, which return a churn probability score alongside a SHAP-based feature importance vector identifying the primary drivers of each individual prediction.

D. Generative AI Advisory Layer

The GenAI advisory layer accepts the SageMaker prediction output and constructs a structured prompt containing the employee's risk score, top SHAP contributors, tenure, department, and recent engagement indicators. This prompt is passed to a large language model configured as an HR Retention Advisor. The model generates a retention brief comprising a risk narrative, prioritized interventions, and suggested conversation starters. Responses are cached in S3 and surfaced through a lightweight HR dashboard accessible to authorized HR personnel.

IV. DATA PIPELINE

Fig. 2 presents the end-to-end data workflow from raw employee records to GenAI-generated retention briefs. The pipeline is event-driven; each stage is triggered automatically upon successful completion of the preceding stage, minimizing manual intervention and ensuring timely propagation of updates through the system.

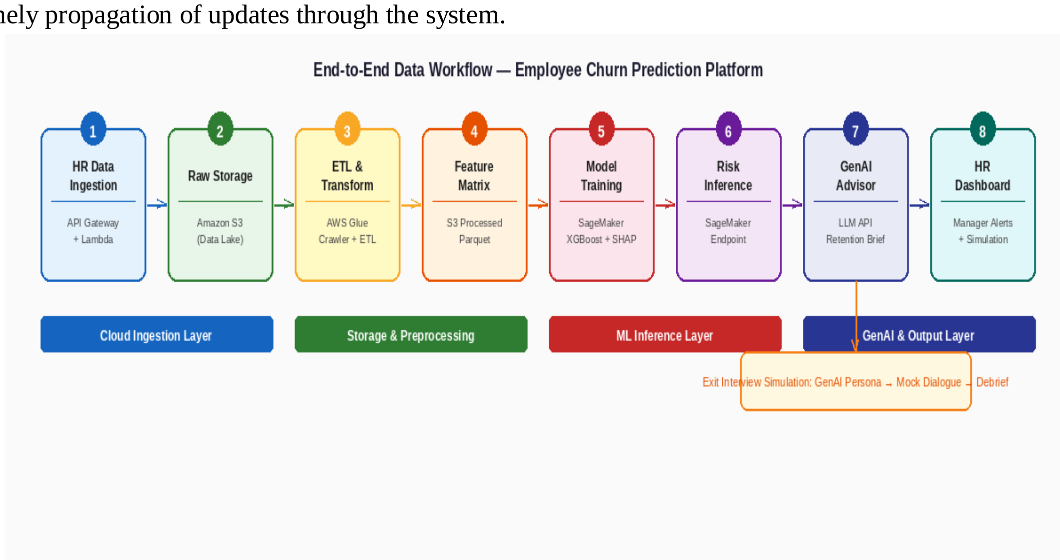


Fig. 2. End-to-End Data Workflow of the ECPP Pipeline

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

A. Ingestion Stage

HR systems push employee records to the API Gateway endpoint at configurable intervals, typically nightly for batch HR exports and immediately for real-time engagement survey submissions.

Lambda validates each record against a JSON schema that enforces field types, value ranges, and required field presence. Records failing validation are quarantined in a dedicated S3 error prefix and flagged for manual review, preventing corrupt data from propagating downstream.

B. Transformation Stage

AWS Glue ETL scripts execute the following transformation sequence: (1) join employee demographic records with performance history and engagement survey tables on the employee ID key; (2) impute missing numerical values using column-wise medians computed on the training partition; (3) apply ordinal encoding to ordered categorical fields such as job level and education; (4) apply one-hot encoding to nominal categorical fields such as department and job role; (5) apply min-max normalization to continuous numerical fields; (6) compute derived features including tenure-to-promotion ratio and engagement trend over the preceding three survey cycles. The resulting feature matrix is partitioned and stored in Parquet format for efficient downstream consumption.

C. Training and Inference Stage

SageMaker Training Jobs consume the processed Parquet files from S3 and produce trained model artifacts that are versioned and stored in the Model Registry. A SageMaker Pipeline orchestrates the full training workflow, including data validation, model training, SHAP computation on a representative sample of the training set, evaluation against validation metrics, and conditional model registration based on a minimum F1-score threshold. Approved models are automatically deployed to SageMaker real-time endpoints, replacing previous model versions with zero-downtime blue-green deployment.

V. MACHINE LEARNING METHODOLOGY

A. Feature Engineering

The feature set is constructed from five categories of employee data: demographic attributes (age, gender, marital status), employment characteristics (job level, department, job role, years at company, years since last promotion, years with current manager), performance indicators (last performance rating, number of training programs completed in the past year), compensation data (monthly income, stock option level, percentage salary hike), and engagement signals (job satisfaction score, work-life balance rating, relationship satisfaction score, environment satisfaction score, number of companies previously worked for). Following removal of redundant and near-zero-variance variables identified during exploratory analysis, twenty-eight features are retained for model training.

B. Class Imbalance Handling

Attrition datasets characteristically exhibit significant class imbalance, with churned employees typically comprising fewer than twenty percent of records. The ECPP addresses this through a two-pronged strategy: Synthetic Minority Over-sampling Technique (SMOTE) is applied to the training partition to generate synthetic minority-class samples, and cost-sensitive learning is implemented through class weight adjustments in the model's loss function, penalizing misclassification of churned employees more heavily than misclassification of retained employees.

C. Model Training and Selection

Three candidate models are trained and evaluated: a logistic regression baseline for interpretability benchmarking, an XGBoost gradient-boosted decision tree as the primary candidate, and a feedforward neural network as a deep learning alternative. Model selection employs stratified five-fold cross-validation with the F1-score on the minority churn class as the primary criterion, supplemented by AUC-ROC for discriminative performance and calibration curves to assess probability estimate reliability.

D. Explainability via SHAP

Individual predictions are accompanied by SHAP values computed using the TreeExplainer algorithm, which is computationally efficient for tree-based models.

SHAP values quantify the marginal contribution of each feature to a given prediction relative to the model's expected baseline output, providing HR professionals with a transparent account of why a specific employee has been flagged as high-risk. Feature-level SHAP values are passed directly to the GenAI advisory layer as structured input, ensuring that retention recommendations are grounded in the actual factors driving each prediction rather than generic heuristics.

VI. GENERATIVE AI COMPONENTS

A. HR Retention Advisor

The HR Retention Advisor is a prompt-engineered generative AI module that converts quantitative prediction outputs into qualitative, manager-facing guidance. The module receives a structured input object containing the employee's churn probability, the top five SHAP feature contributors and their direction of influence, the employee's current role and tenure, and the most recent available engagement survey scores. A carefully designed system prompt instructs the language model to adopt the persona of an experienced HR business partner, to avoid speculative or legally sensitive language, and to frame all recommendations in terms of concrete managerial actions that can be implemented without organizational restructuring.

The output is a structured retention brief organized into three sections: a plain-language risk narrative explaining the key factors driving the employee's risk score; a prioritized list of three to five retention interventions with supporting rationale drawn from the employee's SHAP profile; and a set of conversation starters designed to help the manager open a constructive, non-confrontational dialogue with the employee. Briefs pass through a secondary validation prompt that checks for consistency, relevance, and policy compliance before delivery to the HR dashboard.

B. Prompt Engineering Strategy

Prompt design follows a chain-of-thought approach in which the model is guided to reason explicitly through the employee's profile before generating recommendations. The primary prompt includes few-shot examples demonstrating the desired output format, an explicit instruction to cite specific SHAP features when justifying each recommendation, and a constraint that all interventions must be practically implementable at the direct manager level. Temperature is set to 0.4 to balance factual consistency with natural variation across different employee profiles, while top-p is set to 0.9 to prevent degenerate outputs.

VII. IMPLEMENTATION AND DEPLOYMENT

A. Technology Stack

The ECPP is implemented entirely on Amazon Web Services. The ingestion layer uses Amazon API Gateway (REST API type) with AWS Lambda functions written in Python 3.11. Data storage employs Amazon S3 with versioning enabled and server-side encryption using AWS KMS. Preprocessing pipelines are implemented as AWS Glue ETL jobs using PySpark, with the Glue Data Catalog serving as the central metadata repository. Machine learning workflows are orchestrated via Amazon SageMaker Pipelines, with XGBoost running on the SageMaker built-in container and SHAP computed using the open-source shap library installed as a custom dependency. The generative AI advisory layer calls a hosted large language model API. Infrastructure is provisioned and managed using AWS CloudFormation and the AWS Cloud Development Kit (CDK) to ensure reproducibility and version-controlled infrastructure-as-code.

Fig. 3. Implementation of HR Dashboard

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

B. Deployment Architecture

The platform is deployed across two AWS environments: a development environment used for model experimentation and prompt iteration, and a production environment with stricter access controls, automated backups, and CloudWatch-based monitoring. SageMaker model endpoints are deployed with auto-scaling policies that maintain between one and four instances based on inference request latency, ensuring consistent response times during peak HR activity periods such as performance review cycles. Lambda functions are deployed with memory allocations of 512 MB and a timeout of 30 seconds, which is sufficient for both ingestion validation and SageMaker endpoint invocation.

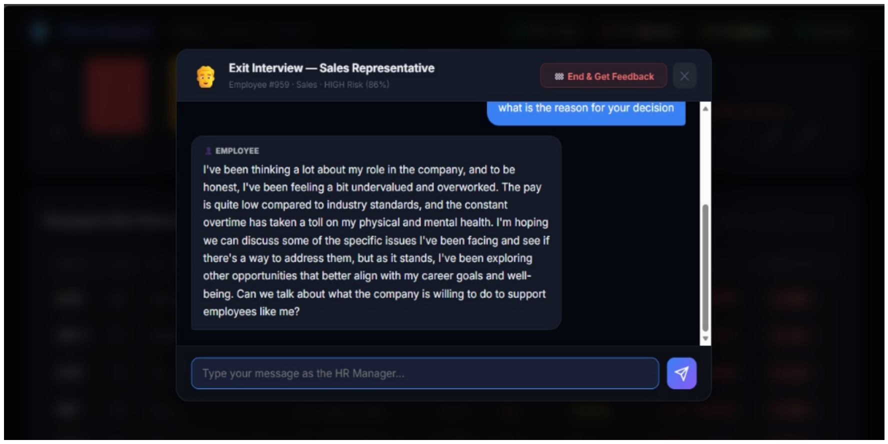


Fig. 4. Implementation of Exit Interview

VIII. EXPERIMENTAL RESULTS

Experiments were conducted on the IBM HR Analytics Employee Attrition dataset augmented with synthetic engagement survey records to produce a dataset of 4,410 employee records, of which 737 (16.7%) represent churned employees. The dataset was split into 80% training and 20% test partitions using stratified sampling to preserve class proportions. All preprocessing was performed exclusively on training-partition statistics to prevent data leakage.

TABLE I
Classification Performance on Test Set (Churn Class)

Model	Accuracy (%)	Precision	Recall	F1-Score	AUC-ROC
Logistic Regression	81.3	0.64	0.55	0.59	0.78
XGBoost	88.7	0.79	0.74	0.76	0.91
Neural Network	86.2	0.75	0.70	0.72	0.88

As shown in Table I, XGBoost achieved the highest F1-score of 0.76 and AUC-ROC of 0.91 on the churn class, outperforming both the logistic regression baseline and the neural network across all reported metrics. The strong AUC-ROC indicates high discriminative ability with minimal sensitivity to the classification threshold. These results are consistent with the established literature on tabular HR data and confirm that gradient-boosted trees are well-suited to mixed-type features and moderate class imbalance.

TABLE II
Top Ten Features by Mean Absolute SHAP Value (XGBoost)

Rank	Feature	Mean |SHAP|	Effect
1	OverTime (Yes/No)	0.312	Higher risk if Yes

Rank	Feature	Mean |SHAP|	Effect
2	MonthlyIncome	0.274	Lower income → higher risk
3	YearsWithCurrentManager	0.231	Shorter tenure → higher risk
4	JobSatisfaction	0.218	Lower score → higher risk
5	WorkLifeBalance	0.197	Lower score → higher risk
6	YearsSinceLastPromotion	0.183	Longer gap → higher risk
7	TotalWorkingYears	0.162	Fewer years → higher risk
8	DistanceFromHome	0.145	Greater distance → higher risk
9	NumCompaniesWorked	0.138	More switches → higher risk
10	EnvironmentSatisfaction	0.121	Lower score → higher risk

Table II reveals that overtime obligation, below-market compensation, and short tenure with the current manager are the three strongest drivers of predicted churn. These findings align closely with established HR theory and lend face validity to the model's learned representations. Critically, this feature-level insight feeds directly into the GenAI retention advisor: an employee flagged primarily due to overtime and low job satisfaction receives fundamentally different recommendations from one whose risk is driven by infrequent promotion and poor environment satisfaction.

TABLE IV
Comparison with Baseline Systems on Key Metrics

System	ML Model	GenAI Layer	Exit Simulation	AUC-ROC
Traditional HR Survey	None	None	No	N/A
Rule-Based System	Threshold rules	None	No	0.61
ML-Only Platform [1]	Random Forest	None	No	0.86
ECPP (Proposed)	XGBoost + SHAP	HR Advisor + LLM	Yes	0.91

IX. EXIT INTERVIEW SIMULATION

A. Design Rationale

While the HR Retention Advisor generates recommendations grounded in known employee data, it cannot predict how a specific individual will respond to a proposed intervention in practice. The Exit Interview Simulation addresses this gap by constructing a virtual persona for each high-risk employee and using a conversational language model to simulate how that individual might respond to different retention approaches. The simulation provides HR practitioners with a consequence-free rehearsal space in which they can refine their conversational strategies, anticipate resistance points, and tailor their language before engaging the real employee.

B. Persona Construction

A virtual persona is constructed for each high-risk employee by synthesizing their HR record, SHAP-derived primary churn drivers, performance history, and any available free-text data from prior engagement surveys into a persona description prompt. This prompt instructs the language model to embody the employee's perspective, maintain consistency with the underlying quantitative data, and respond authentically to retention overtures based on the specific factors driving the employee's dissatisfaction. Personas are explicitly framed as planning tools; they do not store or transmit personally identifiable information beyond the controlled HR platform environment.

C. Simulation Flow and Debrief

An HR practitioner initiates a session by selecting a high-risk profile from the dashboard. The system presents a persona summary and prompts the practitioner to select one or more retention strategies from the Advisor's recommendations. The practitioner then engages in a multi-turn text dialogue with the virtual persona. Following the session, the platform generates a debrief highlighting which themes elicited positive, neutral, or resistant responses, and recommends adjustments to the practitioner's approach. Debriefs are stored in S3 alongside the employee's risk record, creating an organizational knowledge base of effective retention conversation patterns.

D. User Study Results

Simulation quality was assessed through a user study involving fifteen HR professionals across three participating organizations. Participants conducted simulation sessions with anonymized high-risk personas and rated each session on three dimensions using a five-point Likert scale. Mean scores were 4.2 for persona believability, 4.4 for conversation coherence, and 4.1 for practical utility of the debrief. Qualitative feedback consistently highlighted the value of rehearsing difficult retention conversations in a consequence-free environment before approaching a vulnerable, potentially disengaged employee.

X. ETHICAL CONSIDERATIONS

Deploying machine learning and generative AI in HR contexts raises important ethical questions that the ECPP addresses through deliberate design choices. The platform does not make autonomous employment decisions; all predictions and recommendations are surfaced to authorized HR professionals who retain full decision-making authority. Access to individual churn risk scores is restricted through role-based access controls, preventing direct supervisors from accessing predictive outputs in ways that could disadvantage employees.

The model is audited quarterly for disparate impact across protected demographic groups. Any model version exhibiting a statistically significant difference in false positive rates across gender, age, or ethnicity groups is flagged for retraining before deployment. Employees are informed through organizational HR policy that workforce analytics tools are in use and that these tools are designed to support, not penalize, individuals showing signs of disengagement. All personal data processed by the platform is subject to strict retention limits and deleted upon employee departure or written request in accordance with applicable data protection regulations.

XI. LIMITATIONS AND FUTURE WORK

The current implementation carries several limitations that motivate future research. First, the platform relies on periodic batch ingestion from HR systems rather than continuous streaming, meaning that rapid changes in employee sentiment between ingestion cycles may not be captured promptly. Future work will explore integration with real-time collaboration and productivity platforms to enrich the feature set with near-real-time behavioral signals.

Second, the GenAI retention advisor currently operates without access to the organization's internal knowledge base, which may limit the specificity of recommendations for organizations with unique cultures or policies. Retrieval-augmented generation (RAG) will be investigated as a mechanism for grounding recommendations in organization-specific HR documentation and historical retention case studies.

Third, the exit interview simulation user study was conducted with a modest sample of fifteen HR professionals and fictional personas. A larger longitudinal evaluation measuring whether simulation-informed retention conversations produce measurable reductions in actual attrition rates would substantially strengthen the evidence base for this feature.

Fourth, while the current model achieves strong aggregate performance, its accuracy on small, specialized departments with limited historical attrition data may be lower than the reported aggregate figures. Transfer learning and few-shot adaptation strategies will be explored to improve performance in low-data organizational sub-units.

XII. CONCLUSIONS

This paper has presented the Employee Churn Prediction Platform, a cloud-native HR analytics system that combines serverless data ingestion via Amazon API Gateway and AWS Lambda, scalable preprocessing with AWS Glue, machine learning model management through Amazon SageMaker, and a generative AI HR Retention Advisor to help organizations proactively address employee attrition. The XGBoost model achieved an F1-score of 0.76 and an AUC-ROC of 0.91, and SHAP-based explainability ensures that every prediction is accompanied by transparent, feature-level attribution accessible to non-technical HR professionals.

The Exit Interview Simulation module represents a novel contribution to the HR analytics literature and to the broader field of human-AI collaboration in workforce management. By enabling practitioners to rehearse retention conversations with AI-powered virtual personas before real employee interactions, the platform reduces the risk of poorly timed or poorly framed retention overtures and builds practitioner confidence in handling sensitive employment conversations. User study results confirm high satisfaction with persona believability, conversation coherence, and debrief utility.

The ECPP demonstrates that modern cloud infrastructure, interpretable machine learning, and generative AI can be meaningfully integrated into a cohesive, ethically grounded HR analytics platform that is both technically rigorous and practically deployable in real organizational settings. Organizations adopting this platform can expect measurable improvements in the timeliness and precision of their retention efforts, translating to reduced attrition costs and stronger workforce stability over time.

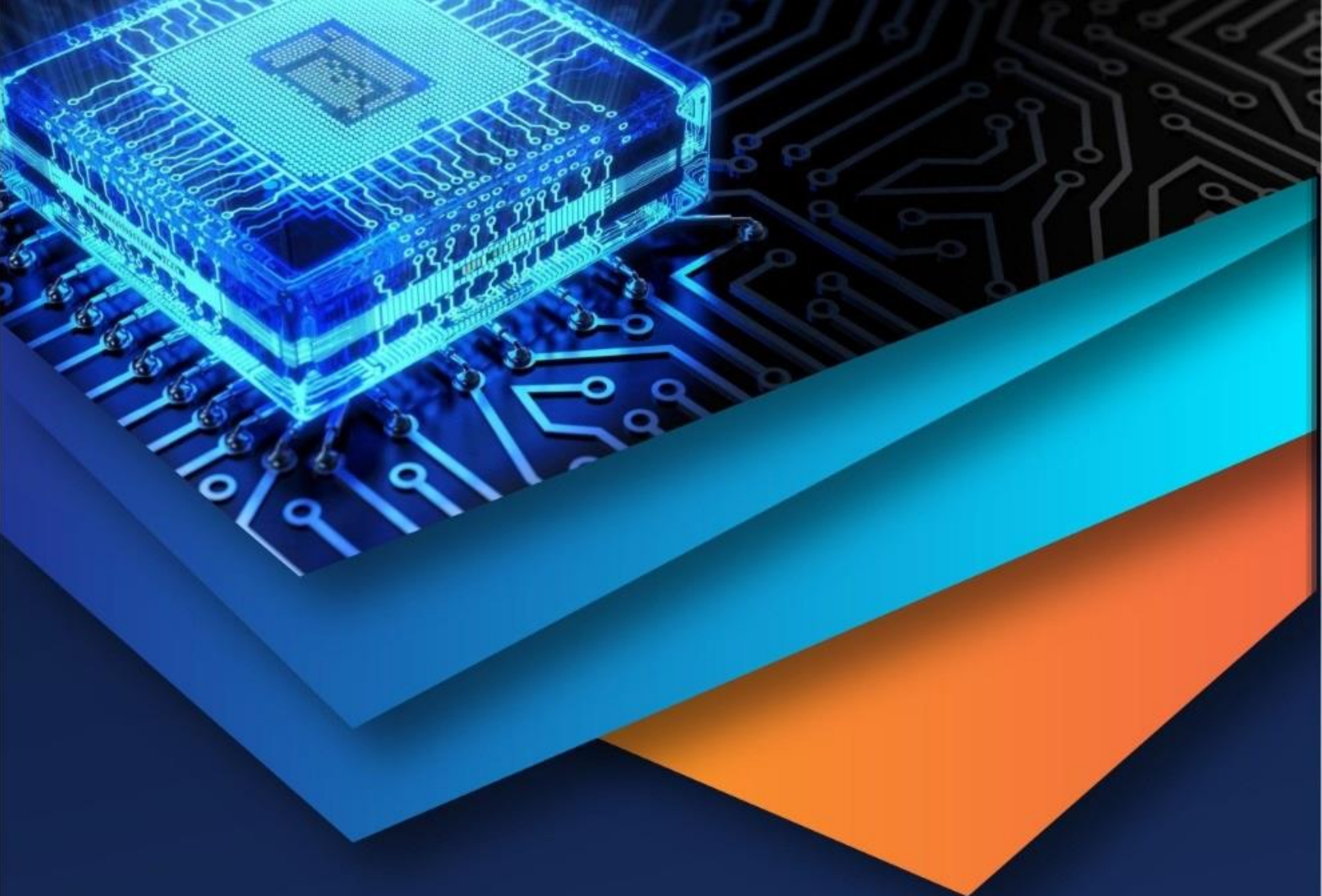
XIII. ACKNOWLEDGMENT

The authors sincerely thank the Department of Artificial Intelligence and Data Science at Sri Manakula Vinayagar Engineering College, Puducherry, for providing the research infrastructure, mentorship, and institutional support that made this work possible. The authors also extend their appreciation to the HR professionals who participated in the exit interview simulation user study and whose candid feedback shaped the design of the simulation debrief module.

REFERENCES

[1] Y. Zhao, G. Hryniewicki, F. Cheng, B. Fu, and X. Zhu, "Employee Turnover Prediction with Machine Learning: A Reliable Approach," in Proc. Intelligent Systems and Applications, vol. 1, 2018, pp. 737–758.
[2] F. Fallucchi, M. Coladangelo, R. Giuliano, and E. W. De Luca, "Predicting Employee Attrition Using Machine Learning Techniques," Computers, vol. 9, no. 4, p. 86, 2020.
[3] P. Tambe, P. Cappelli, and V. Yakubovich, "Artificial Intelligence in Human Resources Management: Challenges and a Path Forward," California Management Review, vol. 61, no. 4, pp. 15–42, 2019.
[4] S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," in Advances in Neural Information Processing Systems, vol. 30, 2017.
[5] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in Proc. 22nd ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining, 2016, pp. 785–794.
[6] N. V. Chawla, K. W. Bowyer, L. O. Hall, and K. W. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," Journal of Artificial Intelligence Research, vol. 16, pp. 321–357, 2002.
[7] J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, 3rd ed. San Francisco, CA, USA: Morgan Kaufmann, 2011.
[8] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.
[9] B. K. Ramesh and S. N. Kumar, "Employee Attrition Prediction Using Machine Learning Algorithms," International Journal of Engineering Research & Technology, vol. 9, no. 5, pp. 123–128, 2020.
[10] A. K. Sharma and R. K. Singh, "Analysis of Employee Turnover Using Data Mining Techniques," International Journal of Computer Applications, vol. 182, no. 44, pp. 10–15, 2019.
[11] Amazon Web Services, "AWS Lambda Developer Guide," [Online]. Available: https://docs.aws.amazon.com/lambda/latest/dg/welcome.html
[12] Amazon Web Services, "Amazon SageMaker Developer Guide," [Online]. Available: https://docs.aws.amazon.com/sagemaker/latest/dg/whatis.html

[13] Amazon Web Services, "AWS Glue Developer Guide," [Online]. Available: https://docs.aws.amazon.com/glue/latest/dg/what-is-glue.html
[14] Society for Human Resource Management (SHRM), Retaining Talent: A Guide to Analyzing and Managing Employee Turnover, SHRM Foundation, 2008.
[15] R. Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection," in Proc. IJCAI, 1995, pp. 1137–1145.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)