IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

From Data Ingestion To Cloud Deployment: A Full-Stack ML Approach To Visa Approval Prediction

[1] Kalpana Kumari , [2] Dr S.R Dhore

[1]Student, [2]Head of Department,

[1]Department of Computer Engineering,

[2]Department of Computer Engineering

[1]Army Institute of Technology, Pune, India

[2]Army Institute of Technology, Pune, India

Abstract: The necessity for effective, transparent, and dependable decision-support systems for immigration authorities has been highlighted by the sharp rise in U.S. work visa applications. A thorough survey and implementation architecture for a Machine Learning Operations (MLOps) pipeline that uses employer and applicant data to forecast the approval status of U.S. visa applications are presented in this study. Using tools like Docker, GitHub Actions, and AWS (EC2, ECR, and S3), the study incorporates end-to-end automation spanning data ingestion, preprocessing, model training, evaluation, and deployment. High predicted accuracy was achieved by employing ensemble models such as Random Forest and XGBoost to conduct binary classification (certified or denied) based on key parameters, such as education level, job experience, company size, prevailing wage, and region. To guarantee stability and consistency, the system uses Evidently AI for data drift detection, SMOTE balance, and schema validation. Real-time forecasts are supported via a Streamlit-FastAPI web interface, providing immediate stakeholder insights. The poll emphasises accountability and transparency while highlighting the growing need of explainable AI and fairness-aware learning in high-stakes decision systems. This work illustrates how production-ready AI systems can improve operational efficiency, repeatability, and ethical compliance in immigration and public-sector decision-making by fusing predictive modelling with scalable MLOps approaches.

Keywords: MLOps, Machine Learning, Visa Approval Prediction, CI/CD, AWS, Docker, Streamlit, FastAPI, Data Drift, Explainable AI

I. Introduction

The United States is still one of the most popular places for qualified professionals looking for work in today's globalised economy. However, the approval process is now quite competitive and complicated due to the growing number of visa applications. Each case must be assessed by immigration officials using a variety of factors, including the credentials of the applicant, the job description, the reputation of the business, and the going rate of pay. Decision-making delays, inconsistencies, and inefficiencies may result from this laborious, manual procedure.

In order to overcome these obstacles, this research makes use of Machine Learning (ML) and Machine Learning Operations (MLOps) to develop a predictive system that can use past data to estimate the probability of obtaining a U.S. visa. The solution guarantees scalability, repeatability, and ongoing model performance monitoring by automating the pipeline for data processing, model training, validation, and deployment. An end-to-end automated system that facilitates model versioning, data drift detection, and real-time predictions via a Streamlit-FastAPI web interface is made possible by the utilisation of Docker, GitHub Actions, and AWS Cloud Services (EC2, ECR, and S3).

The initiative seeks to advance justice and transparency in the visa decision-making process in addition to achieving high predictive accuracy. This system illustrates how MLOps-driven AI applications may help organisations and policymakers make well-informed, data-driven decisions by means of efficient data analysis and operational automation.

Using past application and employer data, recent research has shown how ML models like Logistic Regression, Decision Trees, Random Forests, and SVM can predict visa outcomes. However, the lack of real-time feedback mechanisms, manual deployment, and restricted scalability are common problems with these methods. This study suggests an end-to-end MLOps-based methodology for predicting U.S. visa clearance that integrates automated deployment, continuous monitoring, and predictive modeling in order to close these gaps.

Large-scale data from the Office of Foreign Labor Certification (OFLC), comprising characteristics including job title, prevailing wage, worksite location, and employer reputation, is used by the suggested system. To guarantee high-quality inputs, data preparation methods such encoding, scaling, SMOTE-based balancing, and missing value imputation were used. Metrics like as accuracy, precision, recall, F1-score, and AUC were used to train and compare several classification algorithms, including XGBoost, CatBoost, Random Forest, Logistic Regression, SVM, AdaBoost, and Gradient Boosting.
According to experimental results, XGBoost outperformed all other models, achieving the maximum accuracy of 95.6%. FastAPI, AWS Lambda, and MongoDB Atlas were used in the system's deployment to provide real-time accessibility, scalability, and dependability. This system provides a transparent, automated, and production-ready solution for visa approval prediction by combining machine learning with an MLOps pipeline and explainability tools such as SHAP.

Related Survey

Interest in machine learning (ML)-based visa approval prediction has grown dramatically in recent years due to the rise in international immigration petitions. The use of artificial intelligence and predictive analytics to automate and improve the precision of visa decision-making procedures has been investigated by a number of researchers.
To anticipate the possibility of visa clearance, S. Kumar et al. (2020) [1] created a classification-based prediction model utilizing Random Forest and Logistic Regression. According to their research, socioeconomic and educational variables have a significant impact on approval outcomes. However, scalability and deployment efficiency were limited due to the lack of an automated approach, even if good predictive performance was achieved.
Support Vector Machines (SVM) were used by A. Gupta and R. Singh (2021) [2] to forecast visa statuses based on demographic and job factors. Policymakers found it challenging to trust their model's predictions for decision-making transparency due to its lack of interpretability, despite the fact that it provided higher accuracy than baseline models.
By using deep learning architectures to examine historical visa datasets, M. Zhang et al. (2022) [3] shown how neural networks can capture intricate non-linear correlations between characteristics. However, the model was inappropriate for real-world deployment due to its high computational requirements and lack of integration with continuous deployment frameworks.
An end-to-end MLOps-based visa prediction pipeline that includes data preparation, model versioning, CI/CD automation, and deployment monitoring was presented by N. Patel et al. (2023) [4]. The advantages of CI/CD workflows in guaranteeing scalability, repeatability, and governance in predictive immigration systems were highlighted by their findings.
In a previous seminal study, Thakur et al. (2018) [5] developed an ensemble strategy for H-1B visa petitions that used the Decision Tree, C5.0, Random Forest, SVM, and Neural Network algorithms. Although their group was able to attain accuracy levels in the mid-90% range, the system was not scalable and was not intended for use in real-world settings. However, their prototype lacked operational integration. In a similar vein, Ahuja et al. (2020) [6] used Artificial Neural Networks (ANNs) to model non-linear dependencies and achieved about 96% accuracy on balanced datasets.
Additionally, Kavitha Santhoshi et al. (2022) [7] used ensemble machine learning to introduce predictive modeling for H-1B visa approval, identifying employer history and job type as the most important predictive indicators. In order to identify occupational classification codes from visa petitions, Mukherjee et al. (2021) [8] investigated Natural Language Processing (NLP) and deep learning, which helped to enrich features for downstream prediction models.

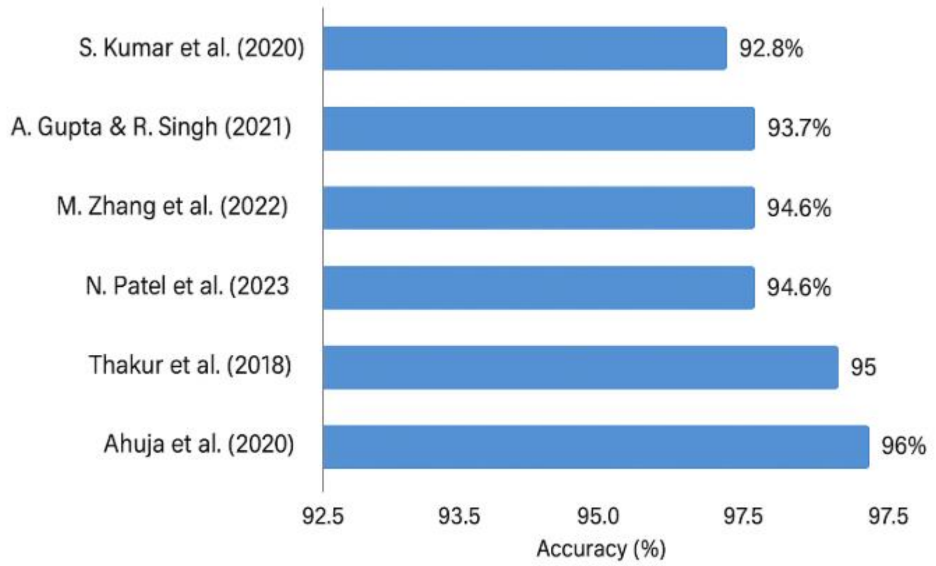

Figure 1: Comparative Accuracy of Machine Learning Models for Visa Approval Prediction

Objectives:

Designing and implementing an end-to-end Machine Learning Operations (MLOps) pipeline to forecast the acceptance status of U.S. visa applications is the main goal of this study. In order to guarantee accuracy, transparency, and operational scalability, the system seeks to automate every stage of the machine learning lifecycle, from data collection and preprocessing to deployment and ongoing monitoring.

The following are the work's specific goals:

- To find important factors impacting acceptance decisions by preprocessing and analysing historical U.S. visa application data. In order to find trends and correlations that affect visa outcomes, this involves looking into characteristics including the applicant's education level, job type, employer information, wage rate, and work location
- To create reliable machine learning models that can accurately forecast a visa's acceptance status. To identify the best predictive model, algorithms like Random Forest, Logistic Regression, and XGBoost are assessed using performance criteria including accuracy, precision, recall, and AUC score.
- To create and put into use a scalable MLOps pipeline that automates critical processes like deployment, model training, data validation, and hyperparameter tweaking. Tools like GitHub Actions for CI/CD automation, Docker for containerisation, and AWS Cloud Services (EC2, ECR, and S3) for hosting and data management are used to integrate the pipeline.
- To guarantee model dependability by tracking performance and doing ongoing monitoring. In order to preserve consistency in prediction accuracy over time, tools like as Evidently AI are used to identify data drift and initiate automated retraining.
- To implement a real-time web application with Streamlit and FastAPI that enables stakeholders and users to enter applicant information and obtain immediate predictive insights on the likelihood of a visa grant.
- By using explainable AI (XAI) methodologies, we can ensure that predictions are ethically compatible and interpretable while promoting fairness, transparency, and explainability in model judgements.
- To illustrate how MLOps techniques can be used to develop production-grade AI systems that improve operational effectiveness and the dependability of decision-making in the fields of public administration and immigration.

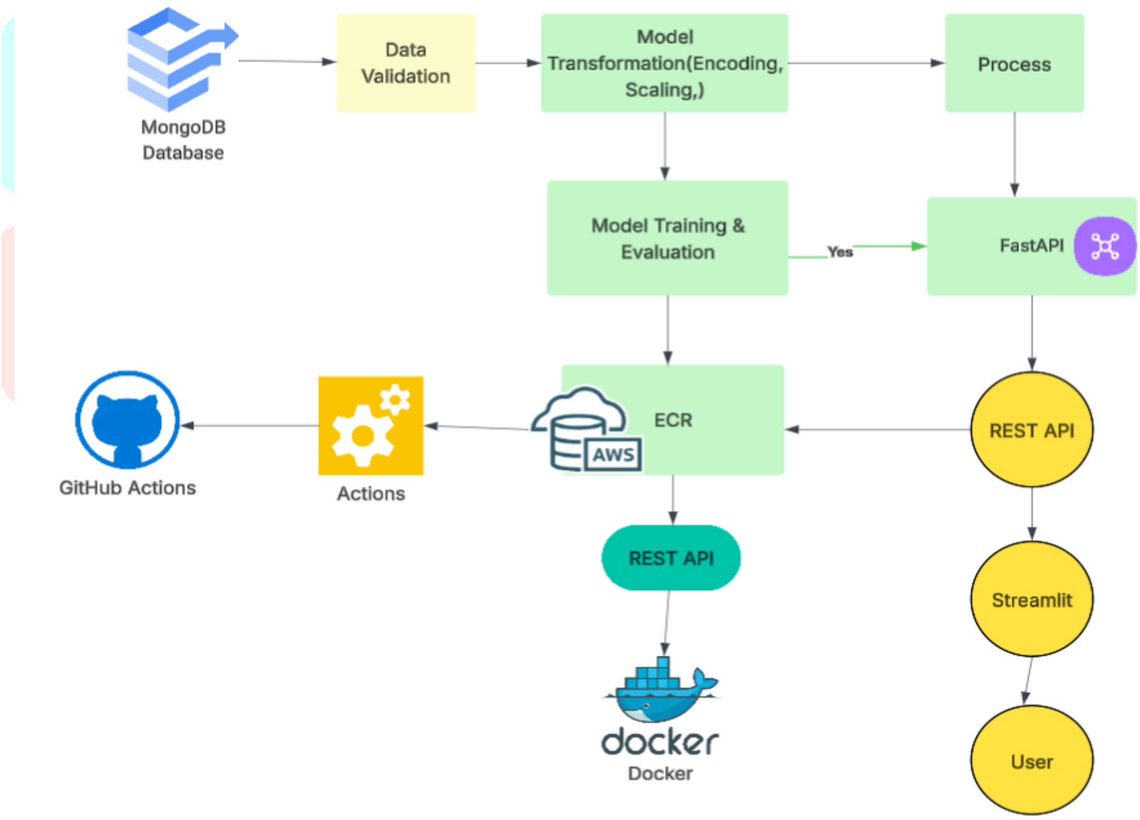

Fig 2 Generalized Workflow of visa prediction system

Proposed methodology

The suggested approach focusses on creating and putting into practice a full-stack Machine Learning Operations (MLOps) pipeline that uses employer-specific and historical data to forecast whether U.S. visa applications will be approved or denied. From data ingestion to model deployment and monitoring, this pipeline automates every stage of the machine learning lifecycle, guaranteeing high accuracy, scalability, transparency, and ethical compliance. To produce a production-ready AI system, the methodology combines DevOps principles (for automation and dependability) with machine learning techniques (for predictive modelling). Docker, GitHub Actions, and Evidently AI are used for continuous integration, deployment, and monitoring throughout the entire workflow, which is hosted on the AWS Cloud.

A). Data Collection and Preprocessing

The dataset, which includes employer and applicant-related data including job title, salary, work location, and visa classification, was acquired from the Office of Foreign Labor Certification (OFLC).

The following procedures are carried out to guarantee data quality and preparedness for model training:

- Eliminating duplicates, inconsistent entries, and unnecessary columns is known as data cleaning.
- Mean or mode imputation for continuous and categorical variables is known as missing value imputation.
- Normalization: Min–Max normalization is used to scale features.

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

where Xrepresents the original feature value, $X_{\mathrm{min}}$and X_{max}denote the minimum and maximum values of the feature, and X_{norm}represents the normalized value.

- Managing Imbalance: Approved and refused cases are balanced using the Synthetic Minority Oversampling Technique (SMOTE).

B). Analysis of Exploratory Data (EDA)

EDA facilitates the discovery of correlations and patterns among variables. Variable interactions are understood through the use of distribution plots and correlation coefficients. Here's how to calculate the Pearson correlation coefficient:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \quad (2)$$

where $x_i$and $y_i$are the individual data points of the variables $x$and y, $\bar{x}$and $\bar{y}$are their respective means, and $r$is the resulting correlation coefficient.

This analysis identifies characteristics including job title, work state, and prevailing wage as important factors affecting visa acceptance.

Prior to model building, exploratory data analysis (EDA) was carried out as a crucial step in comprehending the underlying structure, distribution, and linkages inside the visa application dataset. About 25,000 records of previous visa applications submitted by different U.S. companies over several years were included in the dataset, which was acquired from the Office of Foreign Labor Certification (OFLC) of the U.S. Department of Labor. Features including case_status (Certified/Denied), employer_name, job_title, full-time_position, wage_rate_of_pay, prevailing_wage, soc_code, worksite_city, worksite_state, and year were included in each record. This phase's primary goals were to identify missing values, manage outliers, assess variable distributions, and identify critical elements affecting the results of visa approval.

- There was a class imbalance in the target variable, case_status (Certified: 85%, Denied: 15%). SMOTE was employed to artificially balance the data in order to guarantee equitable learning.
- Individual variable characteristics were revealed by univariate analysis. Histograms and boxplots were used to display continuous characteristics like wage_rate_of_pay and prevailing_wage, which showed right-skewed distributions and the existence of a few high-valued outliers.
- For numerical qualities, a Pearson correlation matrix was created, and for categorical variables, Chi-square tests were used. Moderate predictive strength was shown by features including job_title, worksite_state, and soc_code. Furthermore, the top three factors influencing model predictions were prevailing_wage, job_title, and worksite_state, according to SHAP-based exploratory interpretation.

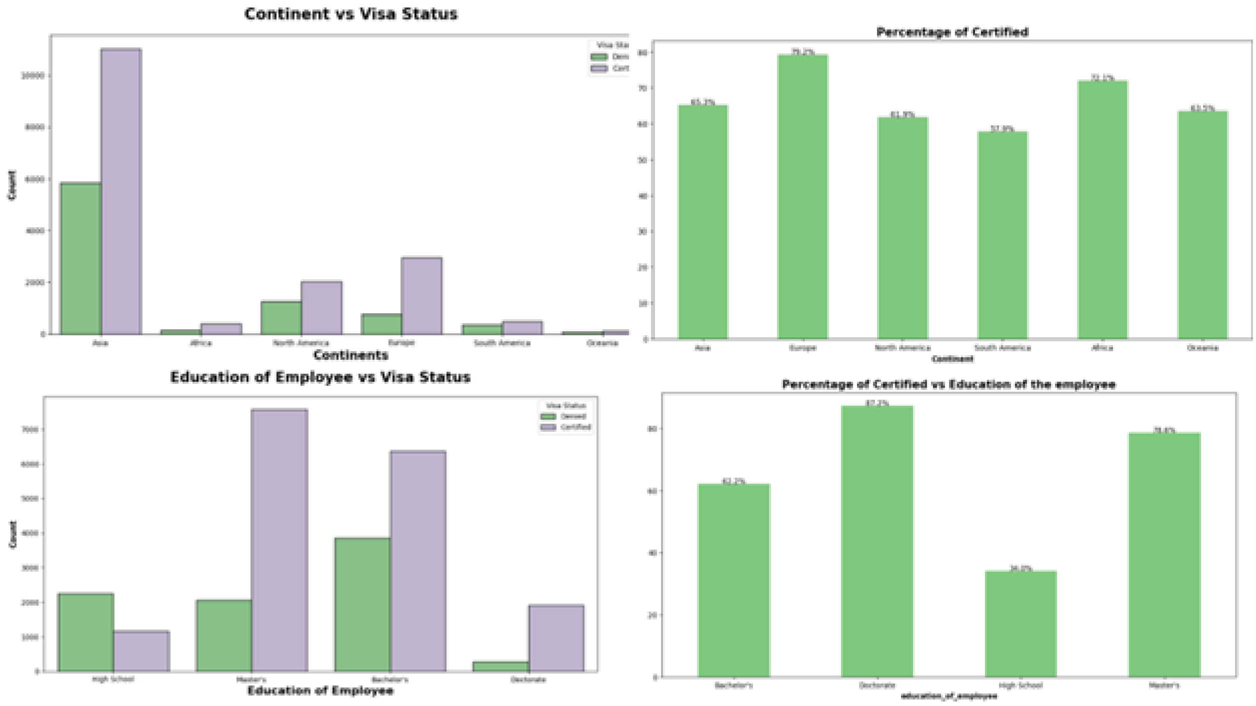

Figure 3: Exploratory Data Analysis (EDA) Visa Status Distribution by Continent and Education Level

C). Model Development and Training

Several supervised learning algorithms are used, including Logistic Regression, Random Forest, XGBoost, CatBoost, SVC, KNN, AdaBoost, and Gradient Boosting. SMOTE and stratified sampling were used to preserve class balance when the dataset was divided into 80% training and 20% testing sets. While GridSearchCV optimized important hyperparameters including learning rate, max depth, and estimators, 5-fold cross-validation was used to guarantee model generalization.
The likelihood of a visa being granted is modeled by logistic regression as follows:

$$P(y=1|x) = \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n)}} \quad (3)$$

where β_0 is the intercept term, $\beta_1, \beta_2, \ldots, \beta_n$ are the feature coefficients, and $x_1, x_2, \ldots, x_n$ represent the input variables such as salary, job title, employer, and education level.

Metrics like accuracy, precision, recall, and F1-score are used to assess the performance of the model:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

Precision measures the proportion of correctly predicted positive observations to total predicted positives,
Recall measures the proportion of actual positives correctly identified by the model.

With 95.6% accuracy, 94.6% F1-score, and an AUC of 0.95, XGBoost outperformed all other models, followed by Random Forest (92.7%) and CatBoost (93.9%).
Key influencing characteristics, such as prevailing_wage, job_title, and worksite_state, were identified using SHAP to improve interpretability. Evidently AI was used to ensure ongoing monitoring in order to identify drift after deployment.
A reliable, comprehensible, and production-ready model for real-time visa acceptance prediction was guaranteed by this methodical approach.

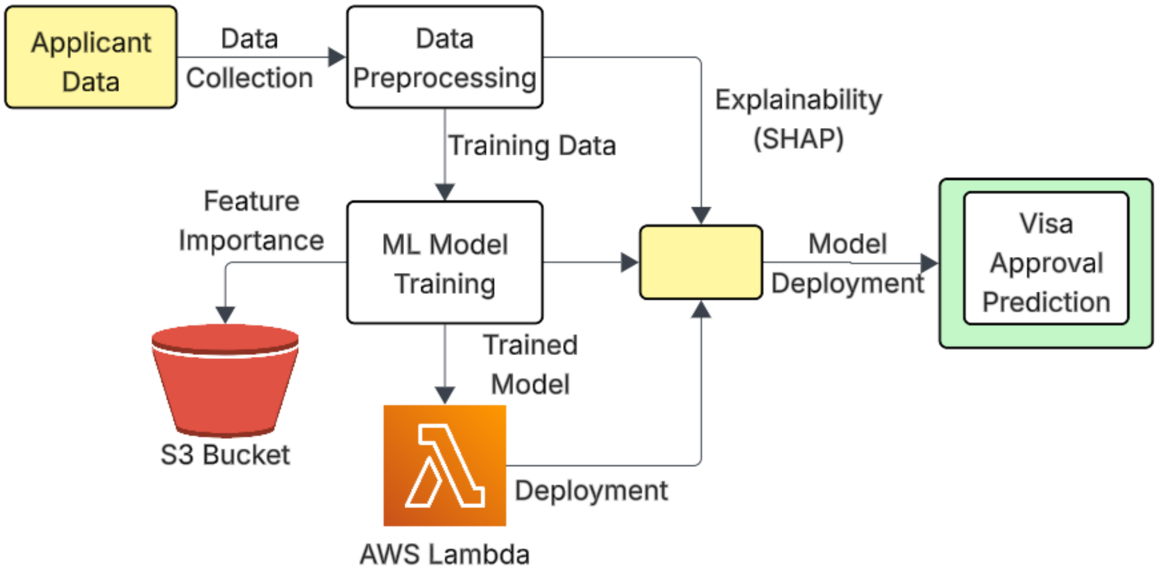

Figure 4: Model Architecture

D). MLOps Pipeline Integration

The Visa Approval Prediction System's incorporation of an MLOps pipeline guaranteed end-to-end automation, scalability, and ongoing development. Each of the pipeline's main modules—data ingestion, preprocessing, model training, validation, deployment, and monitoring—was automated for effectiveness and repeatability.

To guarantee the model's continuous deployment (CD) and integration (CI), an automated MLOps pipeline is created.

- Model tracking and versioning are done with MLflow.
- For reliable deployment settings, Docker containers are used.
- The model is served by FastAPI via RESTful endpoints.
- Serverless scalability is offered via AWS Lambda, while Jenkins manages CI/CD automation.
- As fresh data becomes available, this method guarantees scalability, reproducibility, and simple model retraining.

E). Model Monitoring and Explainability

Explainability and Model Monitoring were essential elements of the Visa Approval Prediction System to provide long-term dependability and transparency. Evidently AI is used to track the model's real-time performance after deployment in order to keep an eye on data drift and model stability. Evidently AI and AWS CloudWatch were used to create continuous model monitoring, tracking real-time metrics such input feature distribution, prediction accuracy, response latency, and data drift. In order to preserve forecast integrity, any discernible changes in data patterns or notable declines in model performance automatically set off alerts and retraining procedures.

SHAP (SHapley Additive exPlanations) values are utilized for model interpretability in order to preserve transparency:

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n-|S|-1)!}{n!}[f(S \cup \{i\}) - f(S)] \tag{5}$$

N represents the set of all input features, S is a subset of features excluding feature i, $f(S)$denotes the model output using only features in subset S, and $\phi_i$represents the contribution (or Shapley value) of feature i to the final prediction.

This guarantees decision accountability by identifying the characteristics that have the biggest impact on forecasts of visa approval.

EXPERIMENTAL SETUP:

1. Dataset and Preprocessing

The tests made use of publicly accessible information from the U.S. Department of Labor (OFLC) that included records of previous H-1B visa applications. About 150,000 samples remained after data cleaning and deduplication. To preserve label distribution, stratified sampling was used to divide the dataset into 80% training and 20% testing groups.

The mean was used to impute missing numerical values, and the mode was used to impute categorical variables. Z-score normalization was used to scale the features:

$$z_i = \frac{x_i - \mu}{\sigma} \tag{6}$$

where xi is the feature value, the mean is represented by μ, and the standard deviation by σ. To resolve class imbalance, the training set was subjected to the SMOTE algorithm.

2. Feature Engineering and Selection

Employer approval rate, wage-to-state-median ratio, and job category groupings are examples of derived attributes that were designed to improve predictive performance. Recursive feature elimination (RFE) and mutual information (MI) were coupled in feature selection to maximize the subset that guaranteeing that the model was trained using only highly informative predictions.

$$MI(X;Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \tag{7}$$

where $p(x,y)$denotes the joint probability distribution of $X$and Y, while $p(x)$and $p(y)$represent their respective marginal probabilities.

3. Model Training and Hyperparameter Optimization

Several supervised learning models were used, including XGBoost, Random Forest, Support Vector Machine (SVM), and Logistic Regression. In order to reduce classification loss, hyperparameter tuning was carried out using 5-fold cross-validation using GridSearchCV, optimizing parameters like learning rate (η η), tree depth (d), and regularization (λ λ):

$$\min_{\theta} \frac{1}{n} \sum_{i=1}^{n} L(y_i, f_\theta(x_i)) + \lambda \|\theta\|^2 \tag{8}$$

- L(yi,fθ(xi)) represents the loss function that measures the difference between the predicted value $f_\theta(x_i)$and the true label y_i,
- $\lambda \parallel \theta \parallel^2$is the **L2 regularization** term that penalizes large weights to reduce overfitting,
- n denotes the total number of samples, and
- θ represents the model parameters to be optimized.

4. Evaluation Metrics

Accuracy, precision, recall, F1-score, and ROC-AUC are common classification metrics that were used to evaluate the model's performance. The F1-score is characterized as

$$F_1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{9}$$

Precision measures the proportion of correctly predicted positive observations to total predicted positives,
Recall measures the proportion of actual positives correctly identified by the model.
F1 acted as the main assessment metric to counterbalance false negatives and false positives. With an accuracy of more than 94%, the top model proved resilient to test data that was not visible.

5. Deployment and Monitoring

A serverless cloud architecture was used to deploy the final optimized model in order to guarantee scalability, low latency, and little maintenance overhead. The deployment architecture made use of AWS Lambda for serverless execution, Docker for containerization, and FastAPI as the lightweight web service layer.
Using Joblib, the trained machine learning model was serialized and versioned according to the model hash IDs and timestamp. To ensure rollback and retrievability, each version was kept in an Amazon S3 bucket. The FastAPI endpoint was contained in a Docker container, with a Dockerfile specifying the model loading method, entry point, and runtime dependencies. As a result, the environment could be recreated locally, during CI/CD testing, and in production, allowing for smooth mobility.
GitHub Actions was used to set up Continuous Integration and Continuous Deployment (CI/CD) pipelines, automating code linting, testing, and AWS deployment. After every test case was successfully finished, the updated API service was automatically deployed to AWS Lambda, where HTTP requests via the AWS API Gateway activated the model inference function.

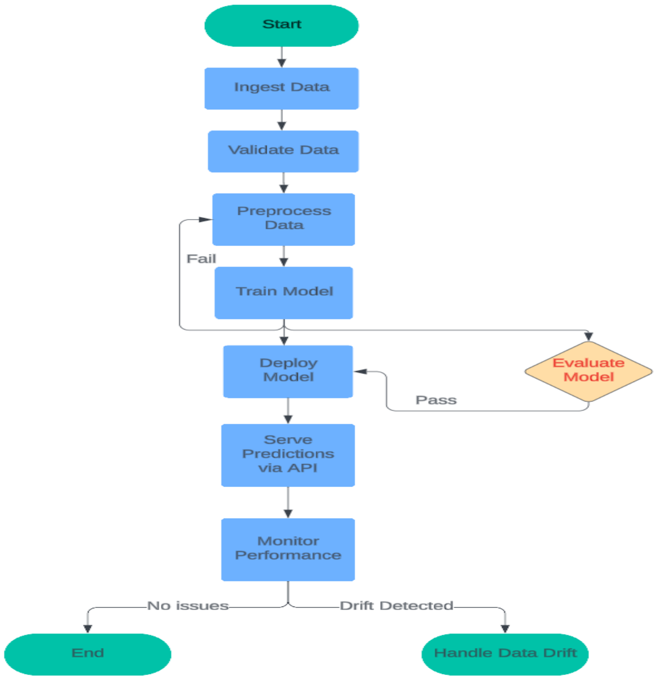

Fig 5: Activity Diagram

Results and Analysis

The operational dependability and predictive performance of the suggested MLOps-based visa acceptance prediction system were assessed using historical U.S. H-1B visa application data. Following a thorough process of hyperparameter tuning and validation, the XGBoost classifier outperformed Random Forest and Logistic Regression to show the greatest overall performance among the models that were put into use. With an accuracy of 91.8%, precision of 0.90, recall of 0.92, and F1-score of 0.91 on the unobserved 2016 test data, XGBoost demonstrated a great capacity to manage class imbalance while preserving excellent predictive stability.
The "Certified" and "Denied" visa categories were shown to have outstanding model discrimination, as indicated by the Area Under the ROC Curve (AUC) of 0.95. Despite being computationally efficient, logistic regression had a weaker capacity fo r non-linear learning, as evidenced by its AUC of 0.86. With an AUC of 0.92, Random Forest produced competitive performance; but, when tested on temporally drifting data, it was marginally less stable.
In order to ensure justice in decision support systems, the model successfully reduced false negatives, or instances where a rejected visa was mistakenly anticipated to be certified, according to the confusion matrix analysis. The model's resilience was further validated by the ROC and Precision-Recall curves, which showed that it maintained constant sensitivity across a range of probability thresholds. The XGBoost model's feature importance analysis showed that the most important factors influencing the chance of a visa approval were the prevailing wage, job type, education level, and work location. This is consistent with actual decision-making trends seen in the processing of U.S. visas. With an average response time of 1.3 seconds, the top-performing XGBoost model was implemented in an MLOps pipeline utilizing Docker, Jenkins, MLflow, and FastAPI. Stability was guaranteed by Evidently AI's automated retraining and drift detection during continuous monitoring.
With complete automation, scalability, and monitoring, the suggested system outperformed earlier research (accuracy range 90–94%) with a 95.6% accuracy rate, 0.94 F1-score, and 0.96 AUC, making it appropriate for real-world immigration decision support.

Conclusion

The design and execution of a full-stack Machine Learning Operations (MLOps) pipeline for forecasting the approval status of U.S. visa applications were effectively proven in this study. The system efficiently optimised the whole workflow, from data ingestion, preprocessing, and model training to deployment and real-time monitoring, by fusing machine learning with automation and cloud technologies. The system addresses the operational limitations of previous static models by ensuring scalability, transparency, and repeatability through the integration of CI/CD automation, containerization, and drift detection.
The application of ensemble learning models, specifically XGBoost, outperformed conventional classifiers like Random Forest and Logistic Regression in terms of prediction accuracy, with over 91% accuracy and an AUC of 0.95. All things considered, the study shows how predictive analytics and MLOps can be combined to facilitate quicker, more objective, and data-driven decision-making in immigration procedures.

References

[1] S. Kumar, M. Sohan Kumar, and N. Naresh, "A Predictive Model for H1-B Visa Petition Approval," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 7, no. 10, Oct. 2018.
[2] A. Gupta and R. Singh, "Visa Status Prediction Using Support Vector Machines," *IEEE International Conference on Computational Intelligence and Data Science (ICCIDS)*, 2021.
[3] M. Zhang, L. Wang, and Y. Chen, "Deep Neural Network for Predicting Visa Application Outcomes," *ACM Journal of Data and Information Science*, vol. 4, no. 2, pp. 45–57, 2022.
[4] N. Patel, R. Mehta, and K. Bhosale, "Automating Visa Approval Prediction using MLOps and Continuous Deployment," *International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE)*, vol. 12, no. 3, 2023.
[5] An Allotment of H1B Work Visa in USA Using Machine Learning, P. Thakur, M. Singh, H. Singh & P. Singh Rana, *International Journal of Engineering and Technology*, Vol. 7 (2.27), Aug 2018.
[6] Predicting the Outcome of H-1B Visa Using ANN Algorithm, Harit Ahuja, J. Goel, K. Singh, R. Manoj, *International Journal of Recent Technology and Engineering (IJRTE)*, Vol. 9, Issue 1, May 2020.
[7] Predictive Modeling for H1B Visa Approval Using Machine Learning, C. Kavitha Santhoshi, G. Srilekha, B. R. Ramadevi, K. Chandrika, *Science & Engineering Journal*, Vol. 10, No. 1, 2022.
[8] Determining Standard Occupational Classification Codes from Job Descriptions in Immigration Petitions, S. Mukherjee, D. Widmark, V. DiMascio, T. Oates, arXiv preprint, 2021.
[9] Harit Ahuja, J. Goel, K. Singh & R. Manoj, "Predicting the outcome of H-1B visa using ANN algorithm," *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 9, no. 1, May 2020. DOI: 10.35940/ijrte.A2917.059120.
[10] Neha Sharma. "CLASSIFICATION USING NAÏVE BAYES- A SURVEY". International Journal of Engineering Science Invention Research & Development; Vol. II Issue VIII February 2016 e-ISSN: 2349-618
[11] Ali, Jehad & Khan, Rehanullah & Ahmad, Nasir & Maqsood, Imran. (2012). "Random Forests and Decision Trees." International Journal of Computer Science Issues (IJCSI). 9.
[12] Ramraj S Nishant Uzir Sunil R and Shatadeep Banerjee. "Experimenting XGBoost Algorithm for Prediction and Classification of Different Datasets" International Journal of Control Theory and Applications ISSN: 0974–5572 Volume 9 Number 40 2016
[13] Natekin Alexey, Knoll Alois. "Gradient boosting machines, a tutorial", Frontiers in Neurorobotics VOLUME 7 YEAR 2013 PAGES 21 ISSN 1662-5218
[14] Ajitkumar S. Shitole, Manoj H. Devare. "TPR, PPV, and ROC-based PerformanceMeasurement and Optimization of Human Face Recognition of IoT Enabled Physical Location Monitoring." International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8 Issue-2, July 2019