Project: EasyVisa

Analyzed the data of Visa applicants, build a predictive model to facilitate the process of visa approvals, and based on important factors that significantly influence the Visa status recommend a suitable profile for the applicants for whom the visa should be certified or denied.

Covered Skills and Tools

- Conducted comprehensive Exploratory Data Analysis (EDA) on visa application data, including problem definition, univariate and bivariate analysis. Utilized a variety of visualizations to uncover patterns, deriving key insights into individual variables and their interrelationships.
- Implemented rigorous data preprocessing steps to ensure data quality. This involved removing duplicate entries, treating missing values, and addressing outliers. Engaged in feature engineering to enhance the dataset's predictive power, preparing it effectively for modeling.
- Developed predictive models using ensemble techniques, starting with Decision Trees. Expanded to Bagging and Random Forest classifiers to leverage collective decision-making, enhancing model robustness and accuracy. Provided detailed performance analysis for each model, underlining their effectiveness in visa application processing.
- Further improved model performance through meticulous hyperparameter tuning of Bagging techniques, including Decision Trees, Bagging, and Random Forest classifiers. Utilized GridSearchCV for optimization, significantly enhancing model precision and reliability.
- Constructed advanced ensemble models using Boosting techniques, specifically AdaBoost and Gradient Boosting classifiers. Evaluated baseline performance and engaged in hyperparameter tuning to refine model efficacy. Explored the potential of a Stacking Classifier to integrate multiple model predictions, achieving superior accuracy.
- Undertook a second round of performance enhancement for Boosting models, focusing on AdaBoost and Gradient Boosting classifiers. Applied strategic hyperparameter adjustments to optimize performance metrics, further solidifying the models' predictive capabilities.
- Maintained a high standard of presentation throughout the project, ensuring the notebook was well-structured with a logical flow. Emphasized clarity and visual appeal through clean, well-commented code and engaging visualizations. Concluded with actionable business recommendations based on model findings, demonstrating the practical value of the analysis and predictive modeling.
- Other Covered Skills: EDA, Data Data Preprocessing, Customer Profiling, Bagging Classifiers (Bagging and Random Forest), Boosting Classifier (AdaBoost, Gradient Boosting, XGBoost), Stacking Classifier, Hyperparameter Tuning using GridSearchCV, Business insights.