

Technology used

Pandas

- Data analysis library for structured data.
- Used for data cleaning, preprocessing, and feature engineering.

NumPy

- Numerical computing library for arrays.
- Helps in mathematical computations and data transformations.

Seaborn

- Statistical visualization library built on Matplotlib.
- Used for creating histograms, box plots, and correlation matrices.

Matplotlib

- Plotting library for data visualization.
- Helps in visualizing trends, distributions, and model performance.

Scikit-Learn

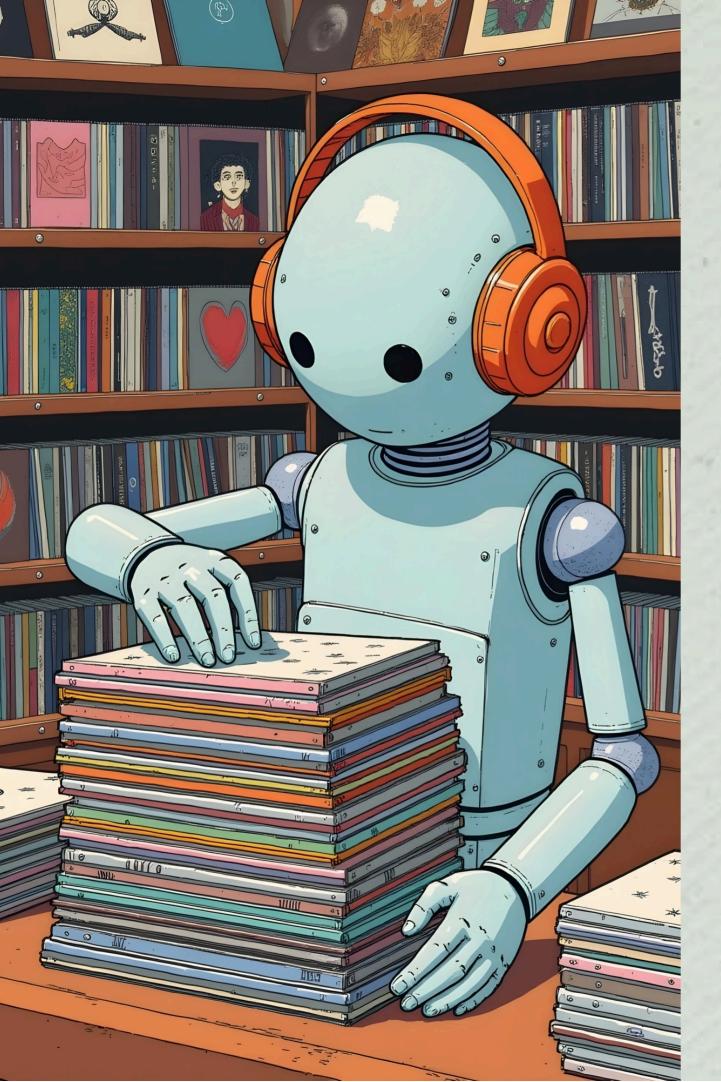
- Machine learning library for classification and regression.
- Used for model training, evaluation, and feature selection.

Imbalanced-Learn

- Library for handling imbalanced datasets.
- Uses RandomOverSampler to balance data and improve predictions.

XGBoost

- High-performance gradient boosting library.
- Enhances prediction accuracy and efficiently handles large datasets.



FEATURES

Data Preprocessing:

- Handles missing values, categorical data encoding, and feature scaling.
- Uses Pandas and Scikit-Learn for data transformation.
- Ensures data is clean and suitable for machine learning models.



Feature Selection

- Identifies key attributes that influence credit approval decisions.
- Uses RandomForest feature importance to select the most relevant features.
- Helps improve model accuracy by reducing unnecessary variables.



Multi-Model Comparison

- Compares different machine learning models like Logistic Regression, Decision Tree, Random Forest, and XGBoost.
- Evaluates models based on accuracy, performance, and efficiency.
- Helps in selecting the best algorithm for credit approval prediction.

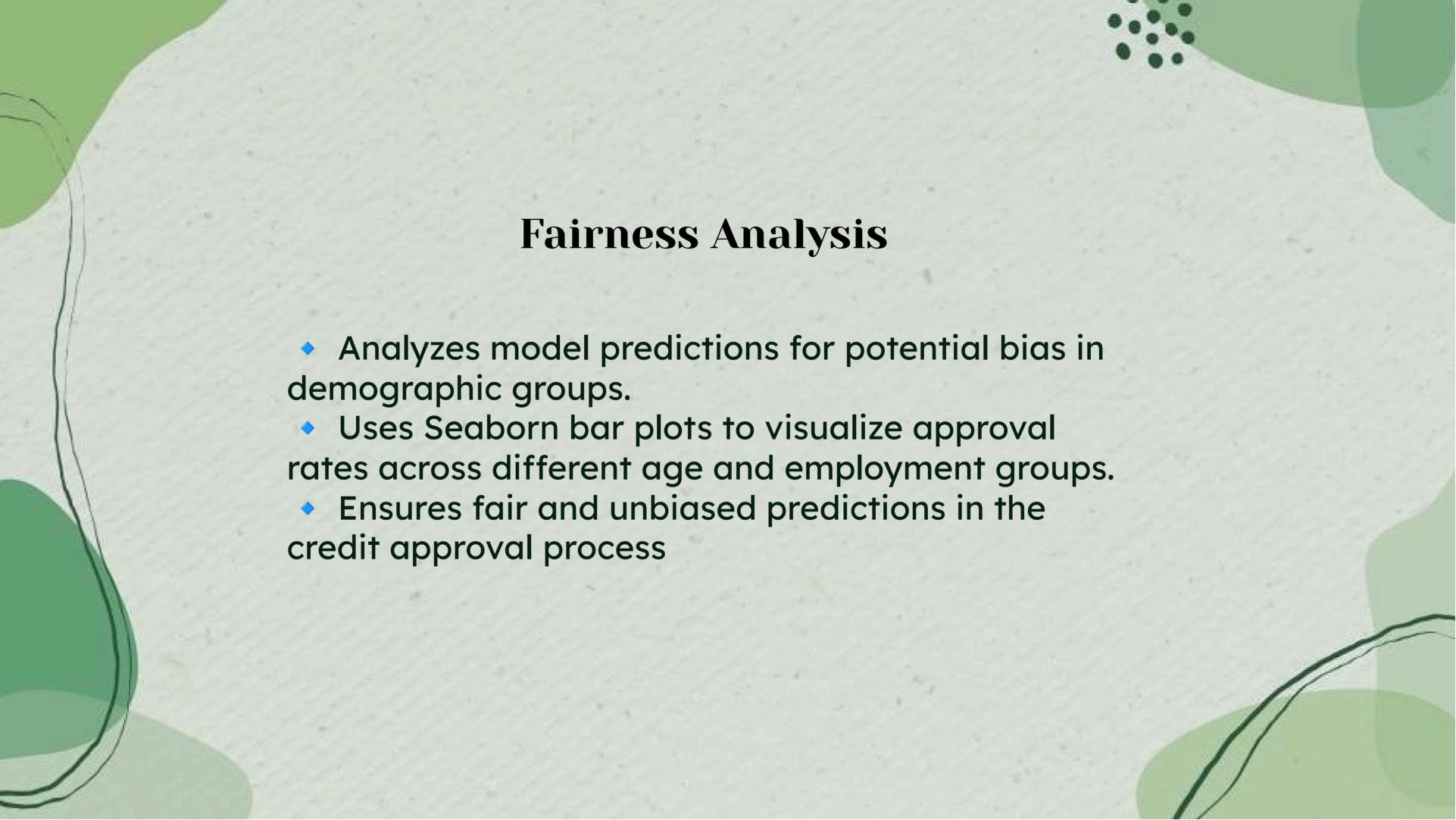




Interactive Predictions

- Allows users to input applicant details such as income and credit history.
- Predicts approval status using the trained model in real-time.





Thanks!

Do you have any questions?

THANK YOU by SRI THARSHINI K

@CYBERNAUT









