Prosperity Prognosticator

Machine Learning for Startup Success Prediction

1. Problem Statement

Startup failure rates are extremely high. This project develops a supervised machine learning model to predict whether a startup will succeed or fail based on financial, operational, and market-related features.

2. Dataset Features

Funding Amount, Funding Rounds, Team Size, Founder Experience, Industry Sector, Revenue Growth, Burn Rate, Market Size, Location, and Success Label (0/1).

3. Machine Learning Pipeline

Data Cleaning → Feature Engineering → Encoding & Scaling → Model Training → Evaluation → Deployment

4. Models Used

Logistic Regression, Random Forest, Gradient Boosting, Support Vector Machine, XGBoost

5. Evaluation Metrics

Accuracy, Precision, Recall, F1-Score, ROC-AUC, Confusion Matrix, Feature Importance

6. Advanced Enhancements

Hyperparameter Tuning (GridSearchCV), Cross-Validation, Feature Importance Visualization, Model Versioning, SHAP Explainability, Class Imbalance Handling (SMOTE)

7. Deployment Options

Flask API, Streamlit Web App, FastAPI Backend, Docker Container, Cloud Deployment (AWS/Azure)

8. Future Improvements

Deep Learning Models, AutoML Integration, Real-time Prediction API, Investor Recommendation System