

## Loan Default Prediction Application

This is a web-based decision support system for predicting the likelihood of loan default. Built with Streamlit, the application allows users to input borrower and loan information, applies preprocessing and dimensionality reduction, and predicts whether a loan is likely to be paid back or default.

### Features

- Accepts key borrower and loan parameters as input:
  - Interest Rate Spread
  - Upfront Charges
  - Rate of Interest
  - Property Value
  - Loan-to-Value Ratio (LTV)
  - Credit Type
  - Submission of Application
  - Debt-to-Income Ratio (DTI)
  - Annual Income
  - Loan Amount
- Preprocesses inputs using a StandardScaler and PCA transformer.
- Predicts loan outcome using a trained Random Forest classifier.
- Provides real-time probability estimates for loan being paid back or defaulting.
- Interactive visualization of input data and prediction confidence.

### Installation

1. Clone this repository:  
git clone <https://github.com/Ekow24/Advancing-Credit-Risk-Modeling-A-Loan-default-Prediction-Approach-using-Machine-Learning.git>  
cd Advancing-Credit-Risk-Modeling-A-Loan-default-Prediction-Approach-using-Machine-Learning
2. Install required packages:  
pip install -r requirements.txt

### Usage

1. Run the Streamlit app:  
streamlit run app.py in your terminal
2. Use the sidebar to input borrower and loan information.
3. Click Predict Loan Outcome to view predictions and confidence scores.
4. Optionally, check Show input data to view the processed input.

## Files

- app.py — Main Streamlit application.
- scaler.pkl — Pretrained StandardScaler for input normalization.
- pca.pkl — PCA transformer for dimensionality reduction.
- trained\_rf\_model.pkl — Trained Random Forest classifier.

## Notes

- The app is currently for local use only and has not been deployed publicly.
- Multiple models were evaluated during development; the app uses one selected high-performing model. Other saved models can be swapped as needed for experimentation.

## References

- Sayed, E. H., Alabrah, A., Rahouma, K. H., Zohaib, M. and Badry, R. M. (2024), 'Machine learning and deep learning for loan prediction in banking: Exploring ensemble methods and data balancing', IEEE Access.
- Naik, K. (2021), 'Predicting credit risk for unsecured lending: A machine learning approach', arXiv preprint arXiv:2110.02206 .