Combined Report: Credit Risk Assessment & Student Performance Prediction using LIME

Introduction

This project explores the application of **Machine Learning (ML)** models combined with **LIME (Local Interpretable Model-agnostic Explanations)** to build interpretable systems in two domains:

1. **Credit Risk Assessment (Finance domain)**
   Dataset: *German Credit dataset*
   Objective: Predict loan risk (Good/Bad credit) and use LIME to explain the classification decisions.

2. **Student Performance Prediction (Education domain)**
   Dataset: *UCI Student Performance dataset*
   Objective: Predict whether a student will **pass/fail** based on academic and socio-economic features, with LIME explaining individual predictions.

Both problems demonstrate how black-box models (like Random Forests) can be explained using LIME for better decision-making.

Problem 1: Credit Risk Assessment

Objective

- Predict whether a loan applicant is a **good credit risk** or **bad credit risk**.
- Provide interpretable explanations for individual predictions using LIME.

Steps Performed

1. **Data Loading**: The *German Credit dataset* was used.
2. **Preprocessing**: Encoded categorical variables and split into training/testing sets.
3. **Model Training**: A Random Forest Classifier was trained to predict credit risk.
4. **Evaluation**: Model performance was evaluated using confusion matrix and classification metrics.
5. **Explainability with LIME**: LIME was applied to understand which features (e.g., credit history, income, loan purpose) influenced individual predictions.

Results

- The model achieved good classification performance.
- LIME explanations revealed factors such as **credit history length, age, existing loans, and savings** were crucial in decision-making.
- Finance-focused interpretation: A rejected loan could be explained transparently to the applicant by highlighting contributing factors (e.g., insufficient credit history).

Problem 2: Student Performance Prediction

Objective

- Predict whether a student will **pass or fail** final exams.
- Provide interpretable insights for each student's prediction.

Steps Performed

1. **Data Loading**: Used the *UCI Student Performance dataset* (`student-mat.csv`).
2. **Preprocessing**: Converted categorical variables to numerical using one-hot encoding. Defined the target variable as **Pass (>=10 marks)**.
3. **Model Training**: Random Forest Classifier trained with socio-economic and academic features.
4. **Evaluation**: Performance measured using confusion matrix and classification report.
5. **Explainability with LIME**: LIME was applied to explain why certain students were predicted as pass/fail.

Results

- The Random Forest model showed strong predictive ability.
- LIME explanations highlighted key influencing factors such as:
- **Study time, absences, parental education, previous grades (G1, G2)**
- Lifestyle factors like alcohol consumption and family support
- Educational insight: Teachers and administrators can use these explanations to identify at-risk students and provide targeted interventions.

Comparative Insights

- **Credit Risk Model (Finance)** helps banks make fair and explainable lending decisions.
- **Student Performance Model (Education)** helps educators identify at-risk students and act proactively.
- LIME proved to be a versatile tool in both domains, ensuring **trust and transparency** in machine learning predictions.

Deliverables

- **Code**: Python scripts implementing preprocessing, model training, evaluation, and LIME visualizations.
- **Outputs**: Confusion matrices, classification reports, and LIME explanations (plots).
- **Reports**: Domain-specific insights for Finance (Credit Risk) and Education (Student Prediction).

Conclusion

This project successfully demonstrated: - The use of machine learning models (Random Forest) in high-stakes decision-making scenarios. - The power of LIME in providing **local interpretability**, making black-box predictions transparent. - Practical applications: - In Finance: Lending institutions can justify loan approval/rejection. - In Education: Educators can design interventions for students at risk of failure.

Thus, this work bridges **predictive accuracy with interpretability**, a critical requirement in modern AI applications.