

Research Proposal: Improving SHAP Scalability in Credit Scoring

This research proposal explores the challenges of applying Explainable AI (XAI) techniques, specifically SHAP, to credit scoring models. We aim to address the scalability limitations of SHAP in handling large datasets, common in financial institutions.

Introduction

1

AI in Credit Scoring

Improved accuracy but "black box" models.

2

Explainable AI (XAI)

Essential for ensuring understandable and trustworthy credit-scoring decisions.

3

SHAP

Provides detailed, global and instance level explanations

Problem Statement

1. Scalability

2. Time Complexity

3. Lack of Quantitative Data

Research Questions, Hypotheses, and Objectives

Research Questions

1. How does the scalability of SHAP degrade as the volume of prediction instances increases in a tree-based ensemble machine learning model?
2. How can feature selection methods, the optimized version of SHAP (TreeSHAP), or both improve SHAP's scalability?

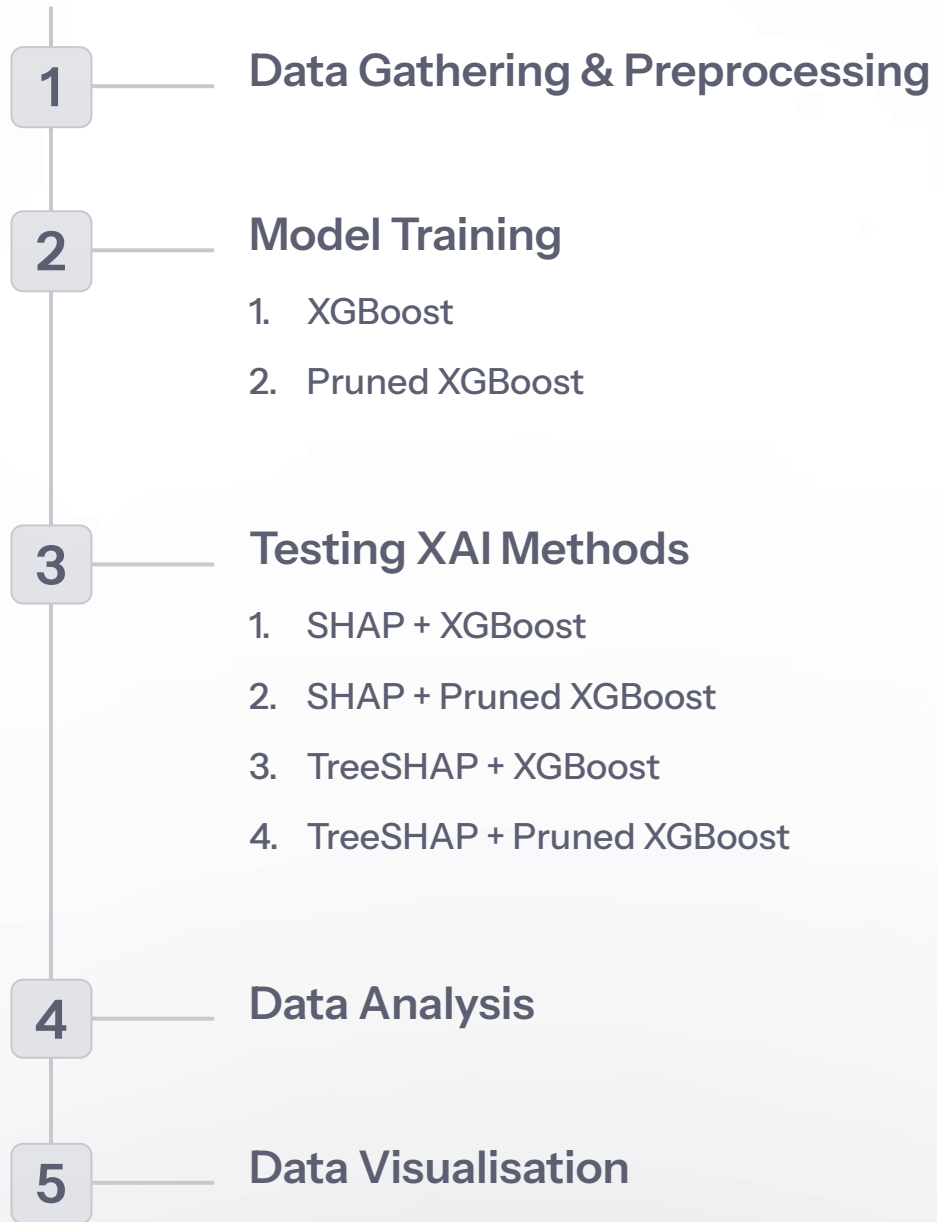
Hypotheses

SHAP's scalability degrades exponentially as the data volume increases, but using feature selection or TreeSHAP will significantly enhance its scalability.

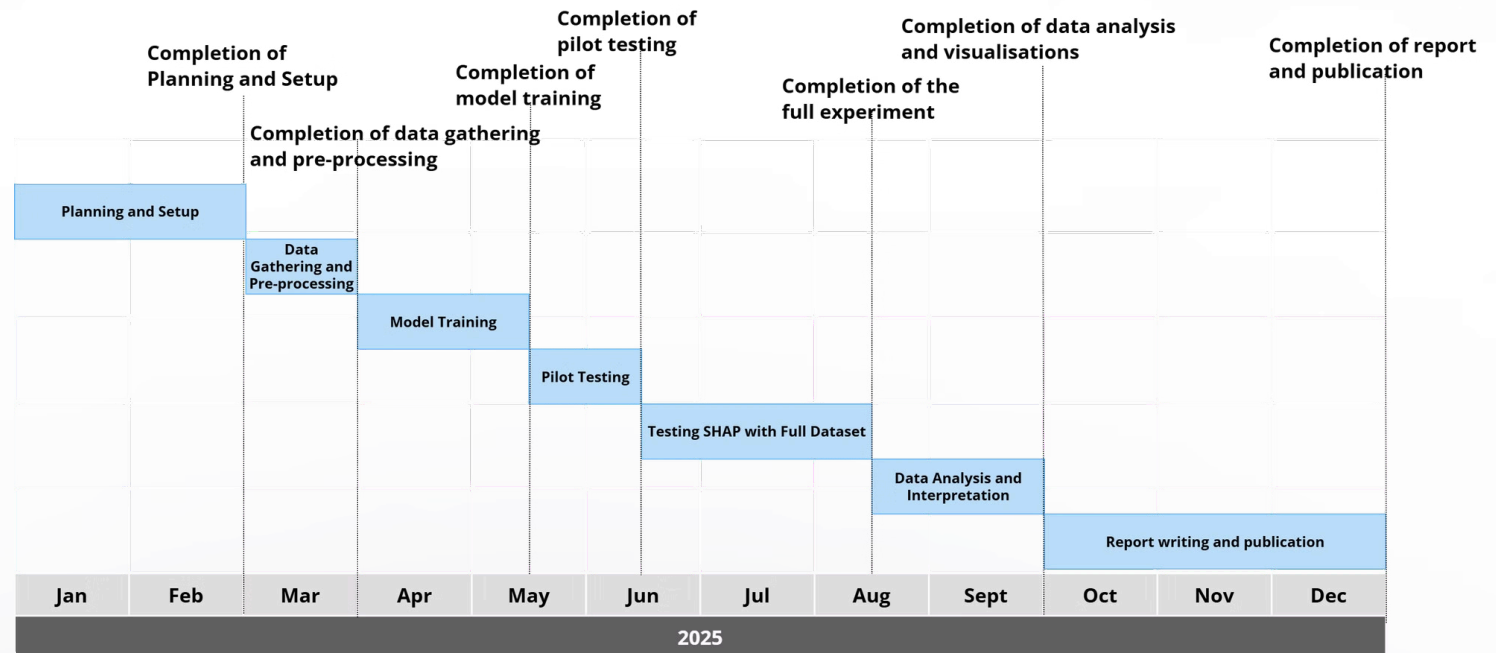
Objectives

1. Quantify SHAP's degradation.
2. Explore how feature selection and TreeSHAP can improve its efficiency.

Research Methodology



Timeline and Milestones



Expected Results and Impact

1

Filling in research gap

2

Identifying effective strategies to improve scalability

3

Offering solid, quantified data for future research

Significance and Contribution

1 Moving research on scalable XAI forward

2 Lowering operational costs of XAI

Conclusion