Explainable AI-Driven Code Smell Detection and Refactoring: A Hybrid Framework Using LIME and SHAP for Automated Software Code Quality Maintenance

Shikha Sharma
Shikha Sharma Jawaharlal Nehru University, New Delhi, India

V B Singh
Shikha Sharma Jawaharlal Nehru University, New Delhi, India

V. B. Singh
Jawaharlal Nehru University, New Delhi, India

Abstract—Generally, code smells are symptoms in source code that indicate potential design and implementation problems that could hinder maintainability, readability, and software quality. Nevertheless, advances in software maintenance, code smell identification, and refactoring indicated that automated solutions for explaining smelly-code detection approaches remain challenging. Current methods do not have a methodical open approach in decision-making and thus lead to low developer trust and adoption rate which urges the need to incorporate the XAI framework in current approaches in software development. This research introduces a new technique Smart Machine-Enabled Learning and Logic for eXplainable AI Code Smell Detection (SMELLXAI) for detecting and refactoring code smells through explainable artificial intelligence (XAI) so as to improve software code quality. The methodology employs a three-phase approach: Step [1] involves extraction of features from Abstract Syntax Tree (AST) and Graph Neural Networks (GNN) for structure and code dependency analysis in combination with BERT-based embeddings for semantic analysis. Step [2] for smell detection through an ensemble of Gradient Boosting Decision Trees (GBDT) and Convolutional Neural Networks (CNN) to identify complex code patterns in a precise manner. Step [3] provides explainable refactoring recommendations using Local Interpretable Model-agnostic Explanations (LIME) and SHapley Additive exPlanations (SHAP) to generate developer-friendly justifications. The aforementioned system incorporates automated refactoring suggestions based on Genetic Algorithms (GA) which is optimized for maintainability metrics. Experimental evaluation of the Code Smell Dataset in Kaggle shows a 98.45% accuracy in smell detection with 97.41% precision. It outperforms other SOTA approaches while ensuring the interpretability of an attention visualization and feature importance analysis.

Keywords—Code Smell, Deep Learning, XAI, Genetic Algorithm, Refactoring, SHAP, LIME