Title : Constructing Software Defect Prediction Models From Static Code Metrics

Author(s) : Napat Bootmeechai **Student ID :** 650510609

Major : Computer Science

Advisor(s) :
1. Assistant Professor Dr. Matinee Kiewkanya
2. *Assistant Professor Benjamas Panyangam*

Type of presentation* *(choose 1)* **:**

- ☐ Oral Presentation
- ☑ Poster
- ☐ Cooperative Education

ABSTRACT

Software defects affect system quality and increase development cost, especially in large systems where manual code inspection is inefficient. This study aims to construct and compare software defect prediction models using static code metrics from the KC1 dataset (NASA MDP), which contains approximately 2,100 records.

Feature selection was performed using Pearson Correlation and ANOVA F-test, and SMOTE was applied to address class imbalance. Five machine learning algorithms were implemented: K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Random Forest, Gradient Boosting, and XGBClassifier. Model performance was evaluated using Accuracy, Precision, Recall, and F1-Score, with emphasis on F1-Score due to class imbalance. Experimental results show that ensemble methods, particularly XGBClassifier and Random Forest, achieved the best performance, with F1-scores of 0.842508 and 0.838659 , respectively.

The selected model was deployed as a web application for predicting defect-prone modules based on static code metrics. The results demonstrate that combining static code metrics with ensemble learning techniques can effectively support early software defect prediction.