



Bad Smell Detection Using Machine Learning Techniques: A Systematic Literature Review

Ahmed Al-Shaaby¹ · Hamoud Aljamaan¹ · Mohammad Alshayeb¹ 

Received: 20 July 2019 / Accepted: 23 December 2019
© King Fahd University of Petroleum & Minerals 2020

Abstract

Code smells are indicators of potential problems in software. They tend to have a negative impact on software quality. Several studies use machine learning techniques to detect bad smells. The objective of this study is to systematically review and analyze machine learning techniques used to detect code smells to provide interested research community with knowledge about the adopted techniques and practices for code smells detection. We use a systematic literature review approach to review studies that use machine learning techniques to detect code smells. Seventeen primary studies were identified. We found that 27 code smells were used in the identified studies; God Class and Long Method, Feature Envy, and Data Class are the most frequently detected code smells. In addition, we found that 16 machine learning algorithms were employed to detect code smells with acceptable prediction accuracy. Furthermore, we the results also indicate that support vector machine techniques were investigated the most. Moreover, we observed that J48 and Random Forest algorithms outperform the other algorithms. We also noticed that, in some cases, the use of boosting techniques on the models does not always enhance their performance. More studies are needed to consider the use of ensemble learning techniques, multiclassification, and feature selection technique for code smells detection. Thus, the application of machine learning algorithms to detect code smells in systems is still in its infancy and needs more research to facilitate the employment of machine learning algorithms in detecting code smells.

Keywords Code smell · Software quality · Machine learning · Artificial intelligent · Anti-pattern · Bad smell

✉ Mohammad Alshayeb
alshayeb@kfupm.edu.sa

Ahmed Al-Shaaby
g201408620@kfupm.edu.sa

Hamoud Aljamaan
hjamaan@kfupm.edu.sa

¹ Information and Computer Science Department, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia

