Contribution ID: **429**   Type: **Oral**

Assessing Software Defect Prediction on WLCG Software: a Study with Unlabelled Datasets and Machine Learning Techniques

Tuesday 5 November 2019 17:00 (15 minutes)

Software defect prediction aims at detecting part of software that can likely contain faulty modules - e.g. in terms of complexity, maintainability, and other software characteristics - and therefore that require actual attention. Machine Learning (ML) has proven to be of great value in a variety of Software Engineering tasks, such as software defects prediction, also in the presence of unlabelled datasets that contain a set of features (i.e. software metrics) for the various software modules (such as files, classes and functions) but lack of modules classification like their defectiveness. To accomplish these tasks, datasets have to be collected for the various modules and properly preprocessed before the application of ML techniques: these activities are essential to manage missing values and/or removal inconsistencies amongst data and to make labelled datasets.

Unlabelled datasets represent the vast majority of software datasets. The extraction of the complete set of features (defectiveness included) and the labeling of the various modules imply effort and time. In literature there exist various approaches to build a prediction model on unlabelled datasets that entail a high number of permutations. Cloud computing infrastructure, GPU-equipped resources and adequate ML framework can give the chance to build software defect prediction model within a reasonable computation time.

This new study describes the analysis of new unlabelled datasets from WLCG software, coming from HEP-related experiments and middleware, with ML techniques by implementing models in different available frameworks, such as Weka, R and python-based frameworks. We have evaluated these frameworks by considering four aspects: learning curve, extensibility, hardware utilization and speed. This study also includes new approaches to label the various modules due to the heterogeneity of software metrics distribution. Our results suggest that predictive accuracy is generally above 96%; furthermore, our procedure keeps trace of the predict defective modules.

A major objective of this work is to reduce the distance between theory and practice in software quality, by providing strengths and limitations of the considered frameworks and methods. This will enable developers in WLCG and other scientific communities to assess the applicability of this study to other software, with the ultimate goal to better understand and reduce software defects in complex projects.

Consider for promotion

Yes

Authors: Dr RONCHIERI, Elisabetta (INFN CNAF); Dr CANAPARO, Marco (CNAF); Dr SALOMONI, Davide (INFN CNAF); MARTELLI, Barbara (INFN CNAF)

Presenter: MARTELLI, Barbara (INFN CNAF)

Session Classification: Track 5 –Software Development

Track Classification: Track 5 –Software Development