To achieve more effective assignment of bugs to teams in large scale industrial contexts use ensemble-based machine learning to automate bug assignment

Related work quantifies the scale of the problem in real projects

Problem

Labour-intensive and error-prone bug assignment in two companies from the telecom and the automation domain respectively

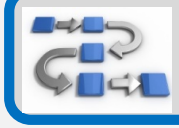
Application of solution to bug data.

Solution

Stacked generalization (SG), combining several classifiers, automates bug assignment

Applied state-of-the art ensemble learner

Problem observed in real projects: Eclipse Platform (Anvik and Murphy 2011), the Mozilla foundation (Bhattacharya et al. 2012), and at Ericsson (Jonsson et al. 2012). Evaluated on data from Telecom and Automation domains.



Evaluated in 5 real projects across 2 companies/domains, on 50 k bug reports, using K-fold cross-validation and sliding window validation.

Precision in automated bug assignment on par with manual (50-89%), which makes it useful in practice, saving cost and time. SG consistently outperforms individual classifiers. When training SG, aim for at least 2,000 bug reports in the training set. Relying only on K-fold cross-validation is not enough to evaluate automated bug assignment.