# **Analyses of software project characteristics on pull request acceptance in distributed software development**

With the continued growth of the web and the advent of distributed version control systems, distributed software development has become a mainstream development approach, especially as social coding tools such as GitHub [1] have changed the way software is developed collaboratively and publicly on the World Wide Web [2]. Instead of pushing changes to a central repository, development developers are pulling them from other repositories and merging them locally [3]. This work builds on the dataset of Xunhui Zhang, Ayushi Rastogi, Yue Yu et al.'s study [4] to investigate the impact of item features in pull requests on whether pull requests can be successfully merged. There are six main project features, namely Programming languages, Popularity of project, Age of project, Workload of a project, Activeness of project, and Openness of a project. Using the project characteristics provided in the dataset [4], data cleaning as well as data analysis and data visualisation were carried out to find out the relationship between project characteristics and the success rate of pull requests.

## **Reference**

[1] GitHub: Social Coding. [Online] Available:http://github.com/

[2] Khadke, N., Teh, M. H., & Shen, M. (2012). Predicting acceptance of github pull requests. *Stanford–CS 229, Tech. Rep*.

[3] Gousios, G., Pinzger, M., & Deursen, A. V. (2014, May). An exploratory study of the pull-based software development model. In Proceedings of the 36th International Conference on Software Engineering (pp. 345-355).

[4] Zhang, X., Rastogi, A., & Yu, Y. (2020, June). On the Shoulders of Giants: A New Dataset for Pull-based Development Research. In *Proceedings of the 17th International Conference on Mining Software Repositories* (pp. 543-547).