# NLP Selection of Tests - ...

Marius Bosler

Seminar: Software Quality
Advisor: Roland Würsching
Technical University of Munich
marius.bosler@tum.de

**Abstract.** Even though there have been many advancements in recent years, a significant amount of testing still needs to be conducted manually. This means that with each software release requiring numerous manual tests, a considerable amount of time is spent executing these tests. Meanwhile, many of these tests tend to overlap in bug detection, which could be avoided if a method were found to execute only the necessary number of tests, as multiple tests may detect the same bug. One way to achieve this is by clustering them through an analysis of the language used in their description. Then, by carefully selecting only a subset of all tests based on said clustering, a lot of time can be saved, especially if the selected tests cover a significant portion of the detected bugs.

**Keywords:** Test Selection, Natural Language Processing, Clustering

## 1 Introduction

## 2 Related Work

Li et al. [1] has proposed a method, [2] another and [3] yet another

## 3 Methodology

## 4 Evaluation

### 4.1 Research Questions

To test our approach, we want to use our dataset of XWiki to answer the following Research Questions (RQs):

- RQ1: Does clustering test cases by test steps cluster similar tests? Each test consists of a number of test steps to execute when executing the tests. It is not given that clustering these test steps actually gives us a clustering that clusters similar test cases together.
- RQ2: How much time/many tests can be saved by selecting tests based on clustering test cases by test steps? The goal of this paper is to demonstrate wether clustering tests in said method can produce helpful results. But is there actually a substantial amount of time/tests that can be saved when only selecting a portion of all tests?
- RQ3: Does the proposed method pose a threat to fault coverage if only a subset of tests are executed?
- RQ4: What is an appropriate amount of tests to be executed without threatening the fault coverage too much?

alternatively (shorter):

- RQ1: Does clustering test cases by steps effectively group similar tests and reduce test count or time?
- RQ2:Does clustering test cases by steps effectively reduce test count or time?
- RQ3: What balance ensures optimal fault coverage with minimum tests after clustering?

## 5 Discussion

## 6 Threats to Validity

## 7 Conclusion

You can also reference other parts of the document, e.g., sections or subsections. In Section **??** we briefly introduced something, whereas in Subsection **??**, we motivated something else.

Make sure to capitalize chapters, sections or subsections when referencing them.

## References

1. Li, L., Li, Z., Zhang, W., Zhou, J., Wang, P., Wu, J., He, G., Zeng, X., Deng, Y., Xie, T.: Clustering test steps in natural language toward automating test automation. In: Proceedings of the 28th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering. p. 1285–1295. ESEC/FSE 2020, Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3368089.3417067, `https://doi.org/10.1145/3368089.3417067`
2. Sutar, S., Kumar, R., Pai, S., BR, S.: Regression test cases selection using natural language processing. In: 2020 International Conference on Intelligent Engineering and Management (ICIEM). pp. 301–305 (2020). https://doi.org/10.1109/ICIEM48762.2020.9160225
3. Viggiato, M., Paas, D., Buzon, C., Bezemer, C.P.: Identifying similar test cases that are specified in natural language. IEEE Transactions on Software Engineering **49**(3), 1027–1043 (2023). https://doi.org/10.1109/TSE.2022.3170272