?

Aliaksei Bialiauski

aliaksei.bialiauski@hey.com

?

Minsk, Belarus

Abstract

This paper is about something new.

Keywords: Machine Learning, Text Classification, Random-Forest, Transformers

1 Introduction

TBD.. Author [1]

2 Related Work

3 Research Method

The goal of this study is to understand whether GitHub repositories can be classified as sample or real. This leads to the following research questions:

RQ1 Do text transformers can predict classes based on text?

RQ2 Which technique performs better in task to classify GitHub repositories on real and sample?

First, we prepared a dataset of 1,000 public GitHub repositories. It is important to have both: real projects and repositories with examples. We distributed number of repositories between real and samples 750 and 250 respectively.

Second, we label our dataset,...

4 Results

TBD..

5 Limitations

6 Discussion

7 Conclusion

8 Acknowledgements

References

[1] Test Author. 2024. Test Article. *Test Journal* (2024).