

Use of reservoir computing for classification of issues in Github

Anton Kovin*

September 24, 2022

Modern software is developed and maintained using "bug-tracking systems". Often, users take part in the development and testing of such software, especially when we speak about "open" software. Users create bug reports (type "bug"), descriptions of desired changes (type "enhancement"), or ask questions (type "question"), often forgetting to put a label of correct issue type. However, such labels are necessary for successful processing of such reports by developers. The number of unmarked reports can reach several thousand per project. To mark them, developers have to spend their "precious" time. Automatic classification of issue reports would significantly simplify the process of developing and maintaining such software. In this study, we evaluate work of LogNNet neural network, through training and testing it on the database we collected from repository of javascript framework "Cypress" on Github. This network is working by the paradigm of "reservoir computing", which is a new approach in artificial recurrent neural network (RNN) training, where an RNN (the reservoir) is generated randomly and only a readout is trained. This is greatly facilitated the practical application of RNNs and outperformed classical fully trained RNNs in many tasks. We received working application which can classify Github Issues reports by there type: bug, enhancement, question. The results of classification were not high, about 54%. But implementation of this technology may reduce the consumption of RAM using on computer or server, performing this task.

*Institute of Applied Mathematical Research, KarRC RAS, akovin@list.ru