Usama rafique

2k18/csme/33

Paper 1

Title : The Scent of Deep Learning Code: An Empirical Study

Authors:

Hadhemi Jebnoun

Houssem Ben Braiek

Mohammad Masudur Rahman

Foutse Khomh

Conference: Msr,20 october seoul

Introduction:

Sculley et al. [22] examined the long-term maintenance costs of machine learning (ML)-based software systems at Google and reported that ML-based systems encounter all the maintenance issues of traditional software systems. However, they noted that ML-based software systems suffer from an additional set of issues that arise from their statistical and data-driven nature. There have been a few earlier works that examined software bugs in deep learning frameworks [13] and analysed the software engineering practices followed by DL practitioners [1]. They suggest that poor coding practices and quick solutions often result in low-quality code containing various code smells. The presence of code smells within the software systems might incidentally degrade their quality and performance, and thus hinder their maintenance and evolution. While there have been a number of studies on the code quality of traditional software systems and a few on ML-based systems, to date, no investigation has been performed on the code quality of DL-based software systems.

Research Methodology:

It was three major steps. First, deep learning-based and traditional software systems are carefully selected from GitHub for the study (Fig. 1-(a)). Each of the software systems (a.k.a., repositories) is pre-processed and prepared for code smell detection. Second, we detect code smells using PySmell tool from each of the releases of DL-based and traditional systems (Fig. 1-(b)). Third, we collect bug-fixing commits and their changed source files to determine the co-existence between code smells and software bugs (Fig. 1-(c)). The following subsections discuss these steps in details.

Subject System Collection & Filtration

System Collection: We attempt to contrast between DL-based and traditional software systems in terms of their code quality (e.g., presence of code smells). Thus, we need to collect both types of systems for our study. In order to collect deep learning systems, we perform keyword search with GitHub Search API [9]. In particular, we choose a set of popular keywords related to various deep learning technology and frameworks as follows

Result

In this paper, we perform a comparison of smells occurrences between traditional and deep learning applications.We analyze a total of 118 repositories (59 deep learning + 59 traditional). We make the following observations:

1) No significant difference: There is no statistically significant difference in the code smell occurrences between deep learning and traditional software systems.

2) Prevalence of code smells in deep learning projects: The most frequent smell types found are Long Ternary Conditional Expression, Complex Container Comprehension, and Long Lambda Function.

3) Violations of the best practices: DL practitioners might not be aware of the code smells in their code, which possibly explains the increasing trend of smell occurrences across the software releases.

4) Code smells lead to bugs: Our findings confirm that the presence of code smells may increase the chances of bugs occurrence. Ours is the first work that extensively investigates the code quality.