

딥 러닝 소프트웨어 개발자들의 코드 냄새 도입 행위에 대한 연구

김정일, 강순주*
경북대학교, *경북대학교

2009307043@knu.ac.kr, *sjkang@ee.knu.ac.kr

A Study on Code Smell Introduction Behavior of Deep Learning Software Developers

Jungil Kim, Soonju Kang*
Kyungpook National Univ, * Kyungpook National Univ.

*Corresponding author

요 약

딥 러닝 (Deep Learning : DL) 기술이 많은 주목을 받으면서 DL 소프트웨어 개발에 참여하는 개발자들이 늘어나고 있다. 이 현상은 DL 소프트웨어 개발 프로젝트들의 생산성을 향상시키는 긍정적인 영향을 끼칠 수 있음과 동시에 소프트웨어 개발 모범 사례를 잘 준수하지 않는 DL 개발자들이 생산되는 DL 소프트웨어의 코드를 이해하기 어렵게 하는 부정적인 영향을 끼칠 수 있다. 따라서 DL 소프트웨어 프로젝트 개발에 참여하는 DL 개발자들이 소프트웨어 개발 모범 사례를 얼마나 잘 준수하는지를 이해할 필요가 있다. 본 논문에서는 DL 개발자들의 코드 냄새 도입 행위에 대한 경험 연구를 소개한다. 이 연구를 위해서 공개 DL 소프트웨어 저장소를 임의로 100 개를 수집했다. 임의로 수집한 그 DL 소프트웨어 저장소에서 활동하는 DL 개발자들이 작성한 코드를 검사해서 코드 냄새를 도입한 집단과 도입하지 않은 집단으로 그 개발자들을 나누었다. 그런 다음 두 집단의 규모를 비교했다. 이 관찰 결과로 전체 개발자 집합 가운데 코드 냄새를 도입하는 개발자들이 그렇지 않은 개발자들보다 더 많이 있다는 것을 확인했다.

I. 서 론

복잡하게 작성되어서 이해하기 어려워 고치기 힘든 코드를 '코드 냄새'라고 한다 [1]. 코드 냄새는 소프트웨어 공학에서 강조하는 모범 코딩 사례를 위반한 대표적인 요소이다. 소프트웨어 시스템에 코드 냄새가 많을 수록 그 소프트웨어 시스템은 유지 보수하기가 힘들어진다.

딥 러닝은 기계 학습 기법의 한 분야로 산업계와 학계 양쪽에서 많은 주목을 받고 있다 [2,3,4]. 이 기법이 많은 주목을 받으면서 공개 DL 소프트웨어 시스템 개발에 참여하는 개발자들 (DL 개발자들)이 많이 늘어나고 있다 [5]. 기본적으로 공개 소프트웨어 시스템 개발에 참여하는 DL 개발자들이 늘어나면 그 DL 소프트웨어 시스템의 코드 생산성 증가를 기대할 수 있다. 하지만 소프트웨어 공학 지식이 모자란 DL 개발자들이 무분별하게 들어오는 것으로 생산된 코드의 품질이 크게 떨어질 수 있는 우려도 있다. 예를 들어, 개발자들이 모범이 되는 소프트웨어 코딩 사례를 잘 준수하지 않아 코드 냄새를 도입한다면 작성된 그 코드는 이해하기 어렵게 될 것이다. 공개 DL 소프트웨어 시스템 개발에서 DL 개발자들이 코드 냄새 도입에 얼마나 주의하는지를 알아보는 것은 공개 DL 소프트웨어 시스템의 품질을 향상시킬 수 있는 개발 전략을 세우는 데에 중요하다. 따라서 DL 개발자들의 코드 냄새 도입 행위를 파악해보는 연구가 필요하다.

이 논문에서는 DL 개발자들의 코드 냄새 도입 행위에 대한 연구를 소개한다. 이 연구를 하기 위해서 깃허브

(GitHub¹)에 있는 오픈 소스 DL 소프트웨어 프로젝트 저장소 100 개를 연구 데이터 집합으로 임의로 선택하고 그 연구 데이터 집합에 포함된 DL 소프트웨어 프로젝트 저장소들의 커밋 이력 (Commit history)를 모두 조사했다. 각 커밋 이력에서 DL 개발자들이 도입한 코드 냄새를 검출하기 위해서 코드 냄새 검출 도구인 Pylint²를 썼다. Pylint 로 얻은 코드 냄새 검출 결과를 바탕으로 그 연구 데이터 집합에 포함된 모든 개발자들 1398 명을 코드 냄새 도입 집단과 도입하지 않은 집단으로 나누었다. 그 결과로 1398 명 가운데 952 명 (68%)이 코드 냄새 도입 집단으로 446 명이 코드 냄새 도입하지 않은 집단으로 분류된 것을 확인했다. 이 결과는 DL 개발자들은 코드 냄새를 도입하는 경향이 대체로 더 많다는 것을 나타낸다. 여기서 더 나아가 DL 개발자들이 어떤 코드 냄새들을 도입하는지, 도입한 코드 냄새가 어떻게 진화하는지, 코드 냄새를 제거하는 경향이 있는지에 대해서 앞으로 추가 연구를 해서 알아볼 필요가 있다.

II. 본론

이 연구에 필요한 공개 DL 소프트웨어 저장소들은 모두 깃허브에서 수집했다. 깃허브는 소프트웨어 개발을 지원하는 가장 유명한 온라인 호스트 사이트이다. 많은 공개 소프트웨어들이 깃허브에서 개발되고 배포되고 있다. 깃허브에서 공개 DL 소프트웨어 저장소들을 찾기

¹ <https://github.com/>

² <https://pypi.org/project/pylint/>

위해서 깃허브 검색 API³를 썼다. 깃허브 검색 API 는 중요 단어들(Keywords)이 주어지면 그 단어들과 가장 관련 있는 소프트웨어 저장소들을 검색한다. 이 연구에서는 DL 과 관련된 다음 중요 단어들을 입력으로 썼다.

Keras, Tensorflow, Torch, Deep-learning, deep-neural network, CNN, RNN

위 중요 단어들을 입력으로 한 깃허브 검색 API 결과로 1000 개 공개 소프트웨어 저장소들을 얻었다. 이 저장소들에서 진짜 DL 소프트웨어 저장소를 찾기 위해서 그 저장소들을 DL 관련된 소프트웨어 저장소들과 DL 비관련된 소프트웨어 저장소들로 직접 분류했다. 각 저장소의 소개 글 (Readme text)을 눈으로 직접 읽어서 진짜 DL 소프트웨어 저장소인지 아닌지를 검사했다. 이 분류 작업으로 DL 튜토리얼 코드 프로젝트나 DL 관련 책 예제 코드 저장소 같이 DL 소프트웨어가 아닌 저장소들을 추려내고 최종적으로 DL 소프트웨어 저장소 100 개를 연구 대상 데이터로 결정했다. 그런 다음 연구 대상 데이터 집합에 있는 각 DL 소프트웨어 저장소들의 커밋 이력을 모두 추적해서 코드 냄새가 도입된 커밋 이력을 찾았다. 이 연구에서는 코드 냄새를 검출하기 위해서 pylint 를 썼다. Pylint 는 파이썬 소스 코드에 코드 냄새가 있는지 검사하는 정적 코드 분석 도구이다. 이 도구를 써서 각 커밋 이력에 있는 파이썬 소스 파일들에 있는 코드 냄새를 검출했다. 이 검출 작업으로 연구 대상 데이터에 있는 저장소들의 각 커밋 이력에서 코드 냄새가 있는 파이썬 소스 파일들과 그 파일들을 작성한 개발자 1398 명의 정보를 찾았다.

연구 대상 개발자 집단에서 코드 냄새를 도입하는 비율이 어느 정도 되는지 알아보기 위해서 그 개발자 집단을 코드 냄새 도입 집단과 코드 냄새 비도입 집단으로 분류했다. 연구 대상 개발자가 작성한 파이썬 소스 파일들 안에서 코드 냄새가 검출되었다면 그 개발자를 코드 냄새 도입 집단으로 분류했다. 반대로 연구 대상 개발자가 작성한 모든 파이썬 소스 파일에 코드 냄새가 검출되지 않았다면 그 개발자를 코드 냄새 비도입 집단으로 분류했다.

그림 1 에서 이 분류 결과를 보여준다. 전체 연구 대상 개발자 집단 1398 명 가운데 952 명 (68%)이 코드 냄새 도입 집단으로 분류되었고 446 명이 코드 냄새 비도입 집단으로 분류되었다. 연구 대상 개발자 집단 70% 가가이 코드 냄새를 도입한 적이 있다는 것을 나타낸다. 이 결과는 DL 개발자들이 코드 냄새를 도입하는 데에 주의하지 않는다는 것을 나타낸다.

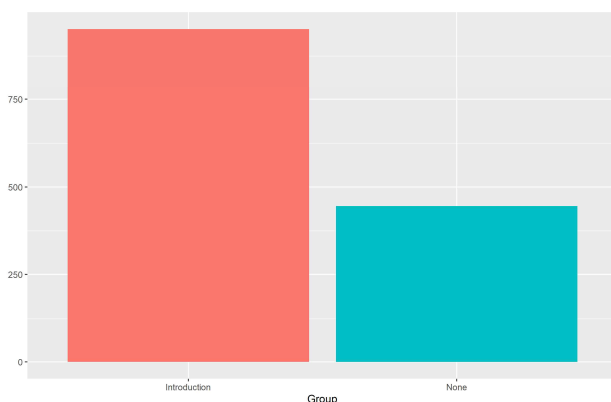


그림 1. 연구 대상 개발자 집단 분류 결과.

III. 결론

이 연구에서는 DL 개발자들의 코드 냄새 도입 행위에 대해 조사했다. 임의로 선정한 연구 대상 DL 개발자들 가운데 70%에 가까운 개발자들이 자신이 작성한 파이썬 소스 코드에 코드 냄새를 도입한다는 것을 밝혔다. DL 개발자들이 코드 냄새를 도입하는 경향을 줄일 수 있는 방법을 찾기 위해서 향후 추가 연구를 할 계획이다. 추가 연구로 DL 개발자들이 자주 도입하는 코드 냄새 종류, 코드 냄새가 자주 도입되는 파이썬 파일 유형, 코드 냄새 제거 경향 등을 알아볼 것이다.

ACKNOWLEDGMENT

이 논문은 2018 년도 정부 (교육부)의 재원으로

한국연구재단의 지원을 받아 수행된 기초연구사업임 (NRF-2018R1A6A1A03025109)

참 고 문 헌

- [1] H. Jebnoun, et. al., "The scent of deep learning code: An empirical study" Proc. Min. Soft. Repo., pp. 420-430, Jun. 2020.
- [2] Q. Guo, et. al., "An empirical study towards characterizing deep learning development and deployment across different frameworks and platforms" Proc. Auto. Soft. Eng., pp. 810-22, Nov. 2019.
- [3] M. J. Islam, et. al., "A comprehensive study on deep learning bug characteristics" Proc. Found. Soft. Eng., pp.510-520, Aug. 2019.
- [4] J. Han, et. al., "An empirical study of the dependency networks of deep learning libraries" Proc. Soft. Main. and Evol., pp. 868-878, Sep. 2020.
- [5] T. Zhang, et. al., "An empirical study of common challenges in developing deep learning applications", Proc. Soft. Reli. Eng., pp. 104-115, Oct. 2019.

³ <https://docs.github.com/en/rest/search>