



ICSE 2022

Pittsburgh, California, USA

수영이 형하고 같이 작업한 논문 “SymTuner: “가 2번의 실패 후 ICSE 2022에 붙어서 2022년 5월 20일 ~ 2022년 5월 26일 ICSE 학회 참여를 위해 미국의 피츠버그에 다녀왔다. 두 번이나 해당 논문이 떨어졌었고 2저자로서도 많이 아쉬웠었지만, 그 두 번의 실패에서 얻은 리뷰어들의 코멘트들을 반영하여 결국 ICSE 2022에서 “Distinguished Paper”에 뽑히게 되었다. 근 2년동안 코로나 바이러스로 인해 해외 학회들은 항상 비대면인 화상 영상으로 진행을 하였지만, 최근에 코로나 바이러스가 많이 안정화됨에 따라 직접 in-person으로 참석할 수 있는 기회가 생겨 이번에 참석하게 되었다. 2년만의 in-person 학회라 그 전의 학회 참석하고는 다른 느낌이 있었고, 이러한 느낌들을 글로써 남기고자 한다.

ICSE 2022



Figure 1. ICSE 2022 학회 내부 모습

개인적으로 해외 학회를 오랜만에 virtual이 아닌 in-person으로 참석해서 그랬을 수도 있지만, 코로나 바이러스가 안정화되었다고는 하지만 그래도 걱정이었는데 약 600여명 정도로 예전만큼은 아니지만 생각했던 것보다는 많은 수의 사람들이 참석하였었다. 이번 ICSE 학회에서 느꼈던 점은 이론적인 논문들도 있지만, 대체로 실용성이 높거나 혹은 Software Engineering 학회답게 학술적인 것들만 있지는 않았다는 점이다. 좋은 의미로 “이런 주제로도 연구/논문이 나올 수 있구나” 싶었던 논문들도 많았었고, 평소에 관심이 없었기에 잘 몰랐던 부분들에 대해서 알 수 있었던 점이 좋았었다.

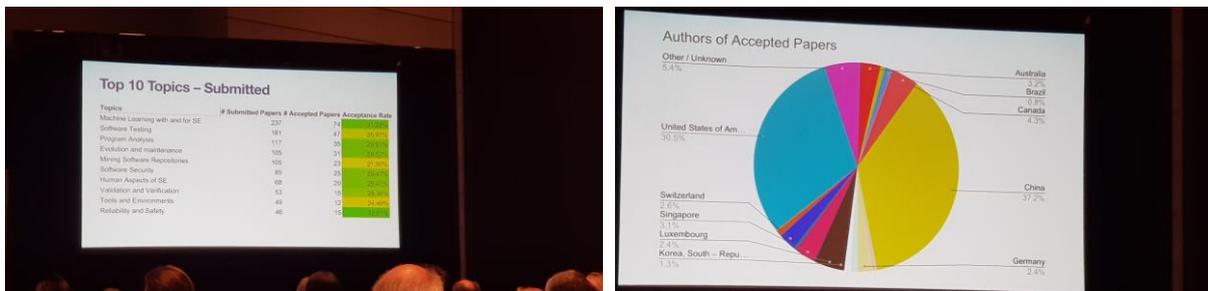


Figure 2. 학회에 제출된 논문들에 관한 여러 지표들

학회가 본격적으로 시작하기 전에, 간단한 지표들을 보여주었는데 이 지표들이 생각보다 재미있었다. 총 몇 개국에서 얼마만큼의 논문을 냈는지, 그 중 accept된 비율이 얼마나 되는지도 보여주었고, 깨알같이 우리나라가 있어서 기분이 괜찮았었다. 또한, 어떠한 주제들이 있었는지도 보여줌으로써 현재 SE 분야에서 트렌드를 이끌고 있다고 생각되는 주제들이 어떠한 것들이 있는지 등을 확인할 수가 있었다.

이번에 내가 직접 발표를 했던 것은 아니지만, 이번 ICSE에서 다른 사람들의 발표를 보면서 느꼈던 점이 몇 가지 있었다. 코로나의 영향인지, 아니면 사람 자체가 많아서 그런 것이었는지는 정확히 모르겠지만, 매 발표들이 5분 정도로 매우 짧았다는 것이 그 첫 번째다. 각 사람에게 5분밖에 시간이 할당되지 않다 보니, 많은 발표자들이 발표 자료에 엑시스들만을 넣으려고 하였던 것이 잘 느껴졌으나, 그만큼 세세한 디테일들이 누락이 될 수밖에 없었던 것이 아쉬웠다.

흥미로웠던 논문들

여기서는 ICSE 2022에 직접 참여해서 들었던 재밌었던 논문들을 설명하고자 한다. 앞서 말한대로 5분이라는 짧은 발표들이라 사전에 논문을 제대로 파악하지 못했던 발표들에 대해서는 발표 중에는 세세한 디테일들에 대해서 넘길 수밖에 없어서 들으면서 개인적으로 아쉬운 마음이 들었다.

1. Free Lunch for Testing: Fuzzing Deep-Learning Libraries from Open Source



Figure 3. Free Lunch for Testing: Fuzzing Deep-Learning Libraries from Open Source

아무래도 Fuzzing 관련 연구를 진행하고 있다 보니 Fuzzing 논문들에 더 많은 관심이 갈 수밖에 없었고, 이번 ICSE에서는 많은 Fuzzing 관련 논문(16개 논문)들이 나왔다. 여기서는 그 중 하나인 FreeFuzz에 관해 얘기하고자 한다.

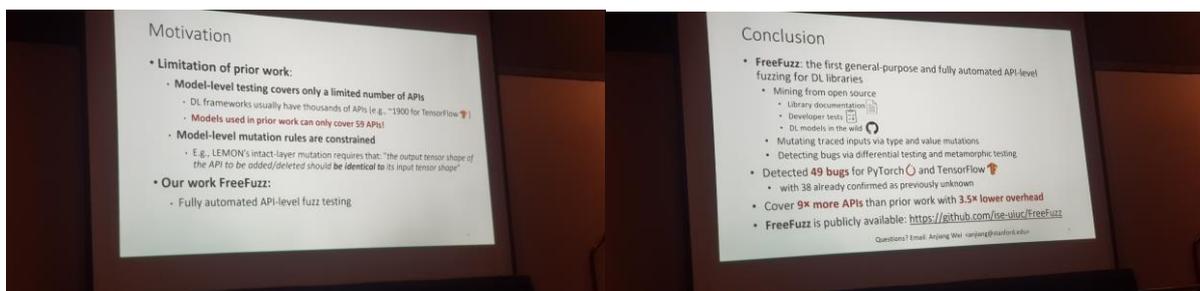


Figure 4. FreeFuzz 발표 슬라이드 (Motivation 및 Conclusion)

해당 논문은 학술과 산업을 막론하고 많은 분야에서 활용되고 있는 머신 러닝 관련 라이브러리 함수 (ML API)들을 대상으로 하는 Fuzzing 기법을 제시하고 있다. 요즘 많은 부분들에 대해서 머신 러닝 기법이 활용되는 만큼, 머신 러닝을 위한 API들이 끼치는 영향은 무시하지 못하는데, 현존하는 기술들은 그 중 일부만을 대상으로 하고 있으며, 이는 실제로 많이 쓰이는 머신 러닝 API들을 제대로 테스트하기에는 부족함을 해당 논문의 동기 중 하나로써 설명을 하였다. 실제로 많이 쓰인다는 PyTorch나 TensorFlow에서 많은 버그를 찾음으로써 해당 연구의 중요성을 잘 보여주었다고 생각한다.

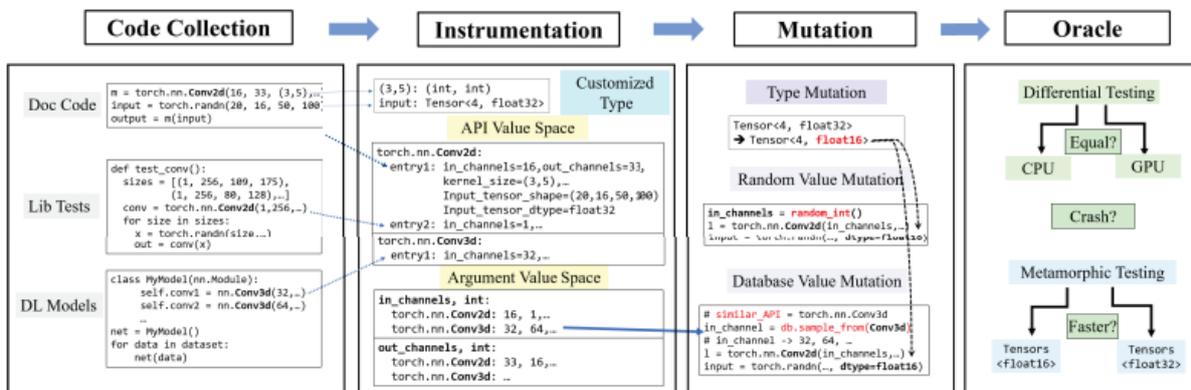


Figure 5. FreeFuzz의 overview

이 논문에서는 머신 러닝 API Fuzzing을 위해, mutation 기반의 fuzzing 기법을 활용하였고, 파이썬 API를 대상으로 하기 때문에 이에 특화된 기술 및 엔지니어링을 기본적으로 진행하였다. 일반적인 mutation-based fuzzing과 비슷한 방식을 취했는데, 먼저 API의 소스코드, Library test 소스 코드 및 DL 모델의 소스코드를 구하고, 해당 코드들에 대해 instrumentation 과정을 거친다. 성공적으로 Fuzzing 진행 중 performance를 확인하고 fuzzing 방향성을 가이드하기 위해서, 머신 러닝 API fuzzing을 위해 특화된 instrumentation을 소개하는데, 간단히 말해서 일반적인 coverage-based fuzzer에서의 branch/code coverage와 같이 어떠한 타입이나 부분들을 확인할 것인지를 제시한다. 그 후, 2가지 mutation rule(type mutation, value mutation)을 활용하여 mutation-based fuzzing을 진행하게 된다.

Fuzzing이라는 기술 자체는 단순히 일반적인 프로그램에서 쓰이는 것이 아니라,

테스팅이 필요한 다른 분야에서도 충분히 활용될 수 있고 좋은 성능을 보일 수가 있는데, 해당 연구는 거의 전 분야에서 활용될 수 있는 머신 러닝을 대상으로 했다는 점에 있어서 충분히 큰 영향력을 끼칠 수 있었겠다라는 생각이 들었다. Fuzzing 분야에 대해 연구를 진행하는 입장에서, 지금은 어떻게 보면 단순히 일반적인 프로그램을 대상으로 기존에 존재하는 기술을 더 좋게 보완하는 식으로 진행하고 있지만, 추후에는 나만의 Fuzzer를 만들고 싶다는 생각이 들었다.

2. Collaboration Challenges in Building ML-enabled Systems.



Figure 6. Collaboration Challenges in Building ML-enabled System

이 논문은 학술/논리적인 논문이라기 보다는 기업이나 실제 소프트웨어를 개발하는데 있어서 생기는 문제점 등을 파악한, Software Engineering 관련 논문이었다. 평소에는 접해보지 못하는 분야 및 주제에 관한 논문이고 꽤 재밌었기에 이번 기회에 설명하고자 한다.

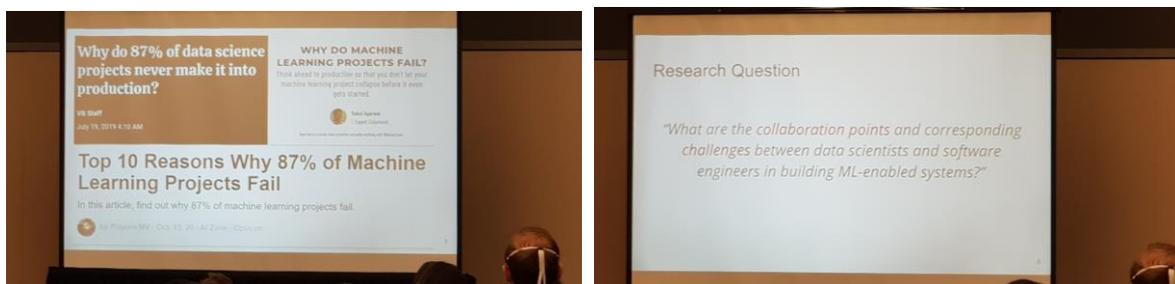


Figure 7. 87%의 머신 러닝 프로젝트가 실패하고 왜 실패하는가?

이 논문에서는 하나 이상의 머신러닝 시스템을 활용하는 규모가 큰 소프트웨어 (ML-enabled systems) 개발 프로젝트를 진행함에 있어서 존재하는 애로사항 등이 무엇인지 조사한 논문이라고 간단히 요약할 수가 있다. 이 발표의 초반에 87%

에 해당하는 데이터 과학 프로젝트가 실패한다는 사실을 말하는데, 생각보다 매우 높은 수치에 놀랐던 기억이 있다. 해당 연구는 실제 프로젝트를 진행했던 그룹/기업들을 대상으로 한 인터뷰를 기반으로 진행되었기에, 단순히 학술적인 측면에서가 아닌 실제 산업에서의 측면을 더 자세히 알 수 있는 기회가 되었다.

이렇게 많은 프로젝트가 불발이 되는 이유로, 해당 논문은 해당 프로젝트를 진행하는 실제 개발자(software engineer)와 해당 머신 러닝 기법을 활용하게 되는 데이터 과학자(data scientist) 간의 협업 및 의사소통, 문서화 방식과 그 외 다양한 부분들에 대해서 많은 뼈격거림이 있음을 보인다. 실제 개발자와 데이터 과학자들은 서로 다른 그룹으로서 서로 협업을 하면서 하나의 프로젝트를 진행하게 되는데, 이러한 것들로 인해서 성공적인 프로젝트를 하기가 힘들었음을 보인다.

하나의 예로, 머신 러닝에서 쓰이는 데이터와 관련하여 데이터를 제공하는 팀과 해당 데이터를 활용하여 모델을 만드는 팀이 따로 존재하는데, 모델을 제공하고 활용하는 팀이 다르다 보니 서로 원하는 지향점이 다르고 그러한 모델을 구성하는 과정에서 둘 혹은 그 이상의 그룹간의 협상 및 협업이 제대로 진행되지 않아서 많은 실패를 하였음을 보였다.

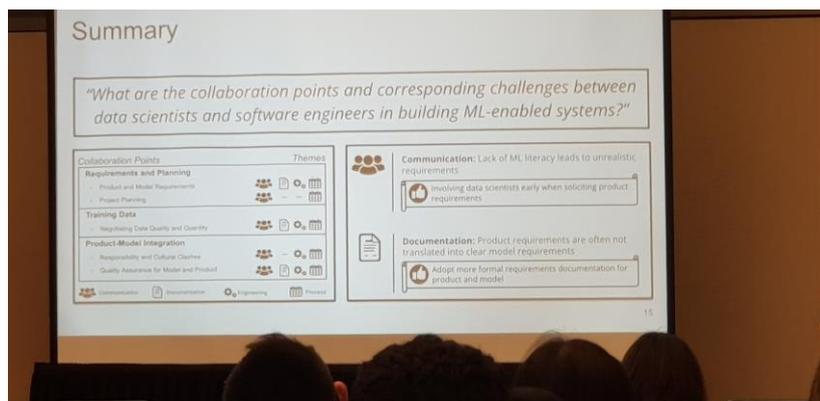


Figure 8. Summary

학술/기술적인 논문들이 매우 중요한 것은 분명하나, 해당 기술들이 실제 산업에서 활용되기 위해서는 이렇게 학술과 산업 간의 불일치함을 해결하는 방안의 연구들 또한 그 영향력이 대단할 것이라는 생각이 들었다. 나중에 기회가 된다면 이러한 방식의 연구 또한 진행해봐도 재밌을 것 같다는 인상을 받았었다.

Life in Pittsburgh

미국에서 생활하는 친구들과 다녀온 연구실 동료들에게 많이 들었던 점으로 미국에서는 마스크를 이제 안 쓴다는 것이 있었는데, 실제로 많은 사람들이 길거리에서 마스크를 벗고 다니는 모습이 인상적이었다. 건물 실내처럼 폐쇄적인 공간이 아닌 이상 거의 대다수의 사람들이 마스크를 벗고 다니고 있었고 그러한 모습들이 실제로 많이 편해보였다.

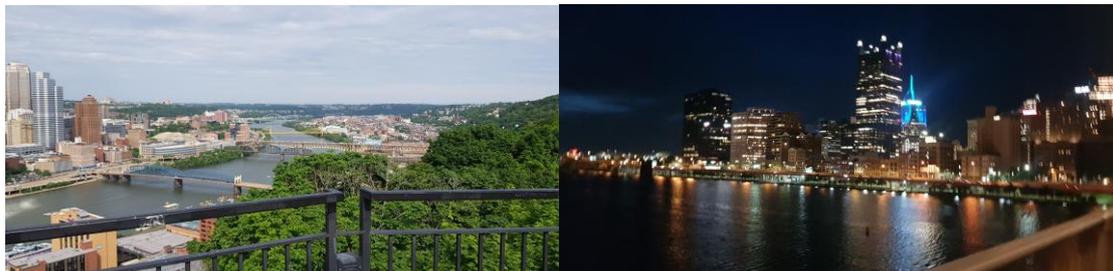


Figure 9. 피츠버그 전경

시간이 있을 때, 일행들과 같이 피츠버그 스카이라인에 올라갈 일이 있었다. 날씨가 좋아서 그런지 실제 도시 전경은 매우 인상 깊었고, 덕분에 여러 사진들을 많이 찍을 수 있었다. 본래 해외를 어떤 이유에서든 가게 되면 그 나라의 통속 음식과 같은 신기한 것들을 주로 먹어보려고 하는데, 아쉽게도 미국/피츠버그에서는 통속 음식이라고 할 만한 것이 없어 아쉬웠으나 피츠버그 내 양조장들은 종류가 많아서 다양한 맛의 맥주를 맛볼 수 있었던 점은 좋았다.



Figure 11. 소방차와 1층 로비에서 대기하는 투숙객들

이번에 피츠버그에 있으면서 재밌었던 사건이 몇 가지가 있었는데 그 중 하나는,

우리가 머물던 호텔에서 화재 경보가 울렸다는 것이다. 수영이 형의 발표를 하루 앞두고, 저녁(12시경)에 수영이 형의 발표 리허설을 도와주고 방으로 돌아와서 자려고 자리를 잡는데, 영화에서만 보던 것처럼 갑자기 “위이잉” 하더니 화재 경보가 울렸다. 자고 있던 원석이도 일어나고 급하게 핸드폰 정도만 챙겨서 비상구 계단(엘리베이터 쓰지 말라고 방송이 계속 나왔다)을 통해 1층 로비로 내려가니 이미 많은 사람들이 나와 있었다.

상황을 주위 사람들과 호텔 관리인에게 물어서 확인해보니 11층 위로 화재 경보가 울렸고, False Alarm으로 판단된다는 소식을 들었다. 그렇게 소방관들이 와서 확인을 하고 실제 위험은 없다는 결론이 나서 다시금 자리로 들어갈 수 있었는데, 학교에서 종종 하던 소방훈련이 아닌 실제 상황은 처음 겪어보는 것이라 많이 놀랐었다. 그래도 다행히 아무 일 없이 웃고 넘길만한 헤프닝으로 좋게 기억이 되었다.

마치며

해외 학회에 in-person으로 참석하는 것은 오랜만이라 생각보다 더 기대되고 설렘이 많았다. 아쉽게도 이번 ICSE에서는 제2저자로서 참석을 하였고, 발표 자체도 메인 학회 전 영상 발표를 한 번만 할 수 있는 기회밖에 없었지만, 다음에는 1저자로서 직접 발표 자료를 만들고 실제 발표 또한 진행할 수 있도록 더 노력을 할 것이다.