CS454 Coursework #2

20130143 Yihan Kim

# 개요

주어진 메소드들의 오류 지표를 바탕으로 메소드의 오류 랭킹 함수를 만드는 것을 목표로 함. 유전 프로그래밍(GP)을 이용한 Fault localization으로 오류가 되는 함수를 찾아내어 보자.

모든 지표를 합한 sum 함수는 50% 가량의 ACC@1 정확도를 가진다. 이보다 정확한 fitness를 찾는 찾기 위해 Symbolic regression을 사용한다. 41개의 데이터를 바탕으로 오류가 발생하는 코드가 ranking에서 앞에 있도록 하는 expression을 찾는다. 각 테스트 케이스에 대해 발생하는 오류율을 최소화하는 값을 찾아낸다.

# 실행 방법

**Training :** $ python3 main.py training Math\_1.csv Lang\_25.csv …

**Validate :** $ python3 main.py validate Lang\_17.csv Math\_31.csv …

# 의존성

Python 3.6.2

외부 라이브러리 : NumPy 1.13.1, DEAP 1.1.0