

Code smells detector for GitHub



Группа: SE 504

Студент: Дергун Карина

Руководитель: Брыксин Тимофей

Введение

При написании кода люди часто допускают ошибки, выкладывают такой код на просторах интернета, например на GitHub-репозиториях, в связи с этим появилась идея: создать такой инструмент, который бы мог находить такие общепринятые «плохие» места в коде, и оставлять пометки в соответствующем Issue.

Идея:

Написать анализатор C-кода, который бы скачивал заданные репозитории с GitHub, анализировал, оставлял запись в Issue.

Выбор языка:

Python3

Выбор библиотеки:

Clang API vs PyCParser

Задачи и технологии

Используемые технологии:

- Python3 (PyCParser)
- Bash(shell)

Задачи:

- Выкачка заданных репозиториев с GitHub
- Нахождение и сборка Makefile'ов с необходимыми параметрами из полученных данных
- Обработка препроцессинга при помощи Анализатора
- Вывод результатов работы в Issue для соответствующих репозиториев
- + Объединение всех подзадач в один программный продукт
- + Подключение машинного обучения

Процесс работы

- 1) Написать Анализатор С-кода, на Python3, используя библиотеку PyCParser.
- 2) Нахождение и выкачка необходимых репозиторий с GitHub через API
- 3) Препроцессирование полученных данных
- 4) Передача данных в Анализатор. Запись результатов в Issue репозитория





Результаты

- Анализатор С-кода который проверяет на наличие в коде мест, где в функции передаются в качестве аргументов по значению переменные типов, отличных от базовых (то есть структур).
- При нахождении передачи аргументов в функции по значению, оставлять запись в Issue соответствующего репозитория.
- Дальнейшая работа будет направлена на увеличение возможностей парсинга, подключения дополнительных языков, и дополнена машинным обучением.



Спасибо за внимание