

Отзыв научного руководителя
на выпускную квалификационную работу бакалавра по теме
**“Использование сетей глубокого обучения
в задаче локализации ошибок в исходном коде”**

Пословской Элеоноры Дмитриевны

Выпускная квалификационная работа Пословской Элеоноры Дмитриевны посвящена разработке модели, улучшающей существующие подходы к решению задачи локализации ошибок в исходном коде программ. С ростом объемов кодовой базы таких крупных проектов как IntelliJ IDEA, для интеграции в которую создавалась модель, задача все более актуальна и использование подобных моделей может существенно упростить разработку и поддержание качественного программного обеспечения.

Поскольку работа с кодом в силу наличия определенных структур и связей между ними отличается от простой предобработки текста на естественном языке, в работе подробно рассматриваются подходы по векторному представлению кода и процесс сбора, разметки и обработки данных IntelliJ IDEA для получения обучающего набора данных. Студентом был сделан акцент на использовании стеков вызовов как данных, описывающих неудачное завершение любой программы на языке Java и существенно ограничивающих область поиска методов с ошибкой. В качестве модели, решающей задачу автор предлагает использовать архитектуру LSTM и использует дополнительное ранжирование результатов классическими методами. Архитектура и процесс обучения всей модели подробно описан в работе и воспроизводим, приведен способ улучшения качества обучения для отсутствующих данных. Автором диплома были проведены эксперименты и сравнение модели как на данных, предоставленном компанией JetBrains, так и на известном наборе данных, используемом в статьях по данной теме. Методика выбора метрики, определения базового алгоритма методом анализа стеков вызовов и описание открытого датасета для оценки качества также указаны в тексте работы. Результаты разработанной модели показывают существенное улучшение по сравнению с существующими подходами и модель может быть интегрирована в инструмент для обработки ошибок в среде разработки IntelliJ IDEA.

Стоит отметить, что в ходе разработки модели студент продемонстрировал высокий уровень теоретической подготовки и осведомленности в области нейросетевых технологий. Написанный для сбора данных и обучения модели код получился качественным и легко масштабируемым для дальнейшего использования. Принимая вышенаписанное во внимание, считаю что работа заслуживает оценки «отлично».

Научный руководитель:

кандидат физ.-мат. наук, кафедра математической теории игр
и статистических решений, доцент

Панкратова Ярославна Борисовна

Соруководитель:

Руководитель практики в IntelliJ Labs, JetBrains,

Хворов Александр Александрович