РЕЦЕНЗИЯ

на бакалаврскую работу Чистяковой Анны Артуровны на тему
«Создание графовой нейронной сети, предсказывающей пути при символьном исполнении»

Данная работа посвящена разработке графовой нейронной сети для оптимизации выбора путей при символьном исполнении, направленной на максимизацию тестового покрытия программ. Тема работы является актуальной, так как символьное исполнение широко применяется для анализа кода, а использование машинного обучения позволяет существенно улучшить эффективность этого процесса.

В работе реализована графовая нейронная сеть, интегрированная в инфраструктуру PySymGym, и проведены эксперименты по её обучению и оценке эффективности. Особое внимание уделено анализу атрибутов вершин графа, поддержке path condition и сравнению различных подходов к обучению модели. Результаты работы демонстрируют, что предложенное решение позволяет достичь сравнимого с существующими методами процента тестового покрытия при значительном сокращении числа генерируемых тестов.

Работа отличается высокой теоретической и практической проработкой. Автор провела глубокий анализ существующих решений, предложила оригинальные идеи, такие как использование графовых нейронных сетей для анализа path condition, и детально изучила их влияние на качество предсказаний. После прочтения пояснительной записки возникли следующие вопросы и замечания.

Замечания:

1. В экспериментах сравнение проводится только с алгоритмом ExecutionTreeContributedCoverage из V#. Было бы интересно увидеть результаты сравнения с другими современными подходами, например, на основе reinforcement learning.
2. Набор данных, используемый для обучения, содержит ограниченное количество методов, что может повлиять на обобщающую способность модели.

Вопросы:

1. Какой вклад вносит анализ path condition в итоговое качество модели? Можно ли выделить конкретные случаи, где его использование критически важно?
2. Почему в качестве функции потерь была выбрана дивергенция Кульбака-Лейблера? Рассматривались ли альтернативы, например, MSE или кросс-энтропия, и какие результаты они показали?
3. Планируется ли дальнейшее расширение набора данных для обучения модели, включая более сложные примеры из промышленных проектов?

Работа Чистяковой А.А. представляет собой значительный вклад в область применения машинного обучения для символьного исполнения. Она выполнена на высоком научном и техническом уровне, а её результаты имеют практическую ценность. В связи с этим считаю, что бакалаврская работа заслуживает оценки «отлично».

Ведущий разработчик,
ООО «ИИнтелКод» Костицын М.П.