

ОТЗЫВ

рецензента на выпускную квалификационную работу обучающегося 18.Б11-мм группы
математико-механического факультета СПбГУ

Цирикова Семена Алексеевича

«Улучшение качества эпитопного картирования методами глубокого обучения»

Выпускная квалификационная работа (ВКР) С.А. Цирикова направлена на создание и реализацию модели глубокого обучения для решения задачи эпитопного картирования, занимающую важное место в биоинформатике при проведении исследований по компьютерному моделированию молекул. В настоящее время существует ряд подходов к решению исследуемой проблемы, однако, повышение точности остается открытым вопросом.

В работе поставлены следующие задачи: составить обзор существующих и используемых компьютерных моделей эпитопного картирования на основе алгоритмов машинного обучения, выбрать адекватный набор данных для обучения, спроектировать и реализовать модель глубокого обучения, провести эксперименты, сравнить результаты с другими моделями, обеспечить решение задачи эпитопного картирования с метрикой качества Precision, превосходящей аналоги, разработать прототип сервиса для использования предобученной модели. По мнению рецензента, все поставленные задачи работы выполнены. По тексту работы имеются следующие замечания:

- 1) Формально введены метрики Recall и Precision, но не дано понимание их сущности. Например, Precision можно интерпретировать как долю объектов, названных классификатором положительными и при этом действительно являющимися положительными, а recall показывает, какую долю объектов положительного класса из всех объектов положительного класса нашел алгоритм. В целом, эта информация может быть найдена в других источниках, но желательно ее наличие и в тексте работы.
- 2) Как правило, в работах по машинному обучению представлено графическое описание предлагаемых моделей, которое дает детальное понимание их структурных составляющих и взаимодействия между блоками. В рассматриваемой работе приведено только текстовое описание, которого, в общем случае, достаточно, но графическая информация улучшила бы понимание новизны предложенной модели.
- 3) Прототип системы не включает в себя описания структурных блоков программы (например, классов). Однако, приведенный недостаток нивелируется наличием ссылки на исходный код системы с описанием ее назначения и инструкций для запуска.
- 4) Недостаточно хорошо раскрыты результаты проведенных экспериментов. Предложенная модель повышает метрику Precision, но ценой ухудшения метрики Recall. В итоге получается, что предложенная модель имеет наилучшие показатели по одной метрике, а вторая модель по другой. Вследствие этого возникает вопрос: какое соотношение этих метрик лучше для практики применительно к рассматриваемой задаче и почему.

Стоит отметить, что представленные замечания не снижают значимости проведенной работы. Считаю, что выпускная квалификационная работа Цирикова С.А. «Улучшение качества эпителиального картирования методами глубокого обучения» удовлетворяет всем требованиям, предъявляемым к выпускным квалификационным работам бакалавров, и может быть оценена на «отлично».

Кандидат физико-математических наук

Ерофеева В. А.

27.05.2022

