

РЕЦЕНЗИЯ

на выпускную квалификационную работу студента IV курса бакалавриата
факультета Прикладной математики – процессов управления
ФГБОУ ВПО «Санкт-Петербургский государственный университет»

Стригулина Кирилла Анатольевича

«Анализ тональности высказываний в Twitter»

Выпускная квалификационная работа Стригулина Кирилла Анатольевича посвящена проблеме определения тональности русскоязычных высказываний пользователей в социальной сети Twitter. Тематика работы является актуальной.

Рецензируемая работа состоит из введения, постановки задачи, трех глав и заключения. Во введении говорится о цели работы – анализ тональности отзывов в социальной сети Twitter для определения репутации компаний, при этом отсутствует обзор литературы по данной предметной области. Далее идет постановка задачи, которая, согласно написанному, состоит в построении классификатора $f \in F$, минимизирующего математическое ожидание штрафа по распределению π . Однако, приведенная формула в работе больше не используется, о каких-либо штрафах и распределениях речь не идет. В итоге поставленная задача не решена.

Первая глава посвящена обзору задач и методов машинного обучения, видам предварительной обработки данных, способам анализа тональности текстов. При этом обзор проведен поверхностно, без ссылок на источники, содержит ошибки. Например, на стр. 16 сказано: «Признаки, которые попадают в промежуток между плоскостями являются опорными», что является неверным. А также на стр. 17 добавлено «... сложность затрудняет четкое понимание, как именно он (алгоритм) работает ...».

Вторая глава называется «Исследования» и состоит всего из 2 страниц! Однако, как таковые исследования отсутствуют, приведены данные с сайта конференции www.dialog-21.ru. Далее предлагается использовать некоторую F-меру без достаточного основания и объяснения, не говоря уже о строгом математическом доказательстве.

Третья глава названа автором «Программная реализация», однако не приводится ни алгоритма, ни кода программы, т.е. программная реализация отсутствует. Глава очень мала по объему – 3 страницы, содержит только общие описания выполненных действий без детального объяснения. Остается неясным следующее: каким образом была осуществлена выгрузка твитов из базы данных и их предварительная обработка, каким образом было выполнено представление высказываний в виде n-грамм, с какой целью было осуществлено уменьшение размерности матрицы и почему для оценки выбрана именно F₁-мера и т.д. Абсолютно непонятно, что было в действительности сделано автором. Более того, в заключении говорится о сравнении «результатов теоретических исследований и экспериментальных результатов», что на самом деле отсутствует.

Также в тексте допущены орфографические, стилистические и пунктуационные ошибки, список литературы не оформлен должным образом. Объем работы крайне мал. На самом деле задача, заявленная в параграфе «Постановка задачи» не решена.

На основании вышесказанного, считаю, что выпускная квалификационная работа бакалавра Стригулина Кирилла Анатольевича «Анализ тональности высказываний в Twitter» заслуживает оценки «неудовлетворительно».

Рецензент:

к.ф.-м.н.,

доцент кафедры теории управления



Захаров А.О.