

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
МЕХАНИКО-МАТЕМАТИЧЕСКИЙ ФАКУЛЬТЕТ
Кафедра дифференциальных уравнений и системного анализа

Аннотация к дипломной работе

**ОБНАРУЖЕНИЕ СПОЙЛЕРОВ В ОТЗЫВАХ НА КНИГИ И
ФИЛЬМЫ С ИСПОЛЬЗОВАНИЕМ КОНТЕКСТНОЙ
ИНФОРМАЦИИ**

Мацкевич Александр Владимирович

Научный руководитель:
ассистент, А. П. Тишуров

Минск, 2022

В дипломной работе 30 страниц, 5 рисунков, 20 источников.

НЕЙРОННЫЕ СЕТИ, ОБРАБОТКА ЕСТЕСТВЕННОГО ЯЗЫКА, АНАЛИЗ ТЕКСТА, ОБНАРУЖЕНИЕ СПОЙЛЕРОВ, РЕКУРРЕНТНЫЕ НЕЙРОННЫЕ СЕТИ

Целью дипломной работы является разработка модели для обнаружения спойлеров, а также анализ того, насколько сильно контекстная информация влияет на качество классификации.

Объектом исследования являются тексты отзывов на произведения, в данной работе рассматривались отзывы с платформ Goodreads и IMDb.

В дипломной работе получены следующие результаты:

- Разработана модель, которая позволяет обнаруживать спойлеры на уровне предложений с качеством, превышающим все модели, натренированные на аналогичных данных
- В ходе экспериментов выявлено значительное влияние контекстной информации на качество итоговой модели.

Дипломная работа носит исследовательский характер. Для использования результатов работы на агрегаторах отзывов требуется доработка и модификация моделей.

Дипломная работа завершена, все задачи успешно выполнены, есть возможность для дальнейших исследований и разработок.

Дипломная работа выполнена автором самостоятельно.

The thesis consists of 30 pages, 5 figures, 20 references, 1 application.

NEURAL NETWORKS, NATURAL LANGUAGE PROCESSING, TEXT ANALYSIS, SPOILER DETECTION, RECURRENT NEURAL NETWORKS

The purpose of this work is to develop a model for detecting spoilers, as well as to analyze how much contextual information affects the quality of classification.

The research object of this thesis project is the texts of reviews for works, in this work reviews from the Goodreads and IMDB platforms were considered.

The following results were obtained in the thesis work:

- A model has been developed that allows detecting spoilers at the sentence level with a quality that exceeds all models trained on similar data.
- During the experiments, a significant influence of contextual information on the quality of the final model was revealed.

The diploma work is of a research nature. To use the results of work on review aggregators, the models need to be finalized and modified.

The thesis project is complete, all tasks have been successfully done, there is a possibility for further research and development.

The thesis work was done by the author independently.