

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Факультет прикладной математики и информатики

Кафедра математического моделирования и анализ данных

Аннотация к магистерской диссертации

**«Применение компьютерного анализа данных для прогнозирования
поведения пользователей в социальных сетях»**

Полянский Александр Анатольевич

Научный руководитель – физико-математических наук
Мальцев М.В.

Минск, 2022

Реферат

Магистерская диссертация, 51 страниц, 31 рисунок, 3 таблицы, 19 источников.

Ключевые слова: СОЦИАЛЬНЫЕ СЕТИ; РЕКОМЕНДАТЕЛЬНАЯ СИСТЕМА; ИНФОРМАЦИОННЫЙ ПОИСК; ОБРАБОТКА ЕСТЕСТВЕННОГО ЯЗЫКА; МАШИННОЕ ОБУЧЕНИЕ; АНАЛИЗ ДАННЫХ;

Объектом исследования методы и алгоритмы анализа данных социальных сетей.

Предметом исследования являются методы и алгоритмы предсказания поведения пользователя на основании данных социальной сети Telegram.

Целью работы было поставлено построение системы, позволяющей получать и накапливать данные из социальных сетей. Построение методики предсказания поведения пользователей путем отнесения их к определенному классу.

В ходе работы:

- 1) Приведен общий обзор методов и библиотек для компьютерного анализа текстов.
- 2) Написано программное обеспечение для получения, обработки и хранения данных из социальной сети Telegram.
- 3) Предложена методика классификации постов по определенным классам для определения модели поведения пользователей.
- 4) Проверены предположения о возможности улучшения метрик при исследовании схожести текстов в векторном пространстве на основе учета положения в векторе.

Полученные результаты можно применить для построения рекомендательных систем, исследований содержания информации социальной сети Telegramm.

Abstract

Master thesis, 51 pages, 31 figures, 3 tables, 19 resources.

Keywords: SOCIAL NETWORKS; RECOMMENDATION SYSTEM; INFORMATION SEARCH; NATURAL LANGUAGE PROCESSING; MACHINE LEARNING; DATA ANALYSIS;

The object of research is methods and algorithms of data analysis of social networks.

The subject of study is methods and algorithms for predicting user behavior based on data from the Telegram social network.

The aim of this work to build a system that allows you to receive and accumulate data from social networks. Construction of a methodology for predicting user behavior by assigning them to a certain class.

The methodology:

- 1) A general overview of methods and libraries for computer text analysis is given.
- 2) Software has been written to receive, process and store data from the Telegram social network.
- 3) A method of classifying posts by certain classes is proposed to determine the user behavior model.
- 4) Assumptions about the possibility of improving metrics in the study of the similarity of texts in vector space based on taking into account the position in the vector are tested.

The results are can be used to build recommendation systems, research the content of the information of the Telegram social network.