

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Факультет радиофизики и компьютерных технологий
Кафедра интеллектуальных систем**

Аннотация к магистерской диссертации

**Выявление депрессии по данным социальных сетей с
использованием методов машинного обучения**

специальность 1–98 80 01 «Информационная безопасность»

Нестерович Анастасия Григорьевна

Научный руководитель: Козлова Елена Ивановна, кандидат физико-математических наук, доцент

Минск, 2024

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Магистерская диссертация: 40 страницы, 10 рисунков, 3 таблицы, 40 источников.

Ключевые слова: ВЫЯВЛЕНИЕ ДЕПРЕССИИ, МАШИННОЕ ОБУЧЕНИЕ, ОБУЧЕНИЕ С УЧИТЕЛЕМ, ЗАДАЧА КЛАССИФИКАЦИИ, БИНАРНАЯ КЛАССИФИКАЦИЯ, АНАЛИЗ НАСТРОЕНИЙ, СОЦИАЛЬНЫЕ СЕТИ.

Цель работы – разработать алгоритм, который будет анализировать посты пользователей социальной сети для предсказания депрессивного или нейтрального содержания.

Актуальность магистерской диссертации обусловлена возможностью анализировать большое количество текстовой информации, которая находится в открытом доступе в интернете, что в свою очередь позволяет выявить и предотвратить развитие депрессии - наиболее распространенного нарушения психического здоровья.

Объект исследования – сообщения пользователей социальной сети Twitter.

Предметом исследования является изучение взаимосвязи психоэмоционального состояния человека и его активности в социальных сетях.

Структура магистерской диссертации – диссертация изложена на 40 странице. Данная работа состоит из разделов «Введение» и 3-х трех глав: «Анализ проблемы и исследование предметной области», «Последовательность работ с текстовыми данными» и «Анализ и интерпретация результатов» и основную часть. Список литературы включает в себя 40 источников.

АГУЛЬНАЯ ХАРАКТАРЫСТЫКА ПРАЦЫ

Магістарская дысертацыя: 40 старонак, 10 малюнкаў, 3 табліцы, 40 крыніца.

Ключавыя словы: ВЫЯЎЛЕННЕ ДЭПРЭСІ, МАШЫННАЕ НАВУЧАННЕ, НАВУЧАННЕ З НАСТАЎНІКАМ, ЗАДАЧА КЛАСІФІКАЦЫІ, БІНАРНАЯ КЛАСІФІКАЦЫЯ, АНАЛІЗ НАСТРОЮ, САЦЫЯЛЬНЫЯ СЕТКІ.

Мэта працы – распрацаваць алгоритм, які будзе аналізаваць пасты карыстальнікаў сацыяльнай сеткі для прадказання дэпрэсіўнага або нейтральнага ўтрымання.

Аktуальнасць магістарскай дысертацыі абумоўлівае магчымасць аналізаваць вялікую колькасць тэкставай інфармацыі, якая знаходзіцца ў адкрытым доступе ў інтэрнэце, што ў сваю чаргу дазваляе выявіць і прадухіліць развіццё дэпрэсіі - найбольш распаўсюджанага парушэння псіхічнага здароўя.

Аб'ект даследавання – пасты карыстальнікаў сацыяльнай сеткі Twitter.

Прадметам даследавання з'яўляецца вывучэнне ўзаемасувязі псіхаэмацыйнага стану чалавека і яго актыўнасці ў сацыяльных сетках.

Структура магістарскай дысертацыі – дысертацыя выкладзена на 40 старонак. Дадзеная праца складаецца з раздзелаў «Увядзенне» і 3-х трох галоў: «Аналіз праблем і даследаванне прадметнай вобласці», «Наследаванне работ з тэкставымі дадзенымі» і «Аналіз і інтэрпрэтацыя вынікаў» і асноўнай часткі. Спіс літаратуры ўключае ў сябе 40 крыніц.

SUMMARY

Master's thesis: 40 pages, 10 figures, 3 tables, 40 sources.

Keywords: DEPRESSION DETECTION, MACHINE LEARNING, SUPERVISED LEARNING, CLASSIFICATION, BINARY CLASSIFICATION, SENTIMENT ANALYSIS, SOCIAL NETWORKS.

The goal of the work is to develop an algorithm that will analyze user posts from a social network to predict depressive or neutral content.

The relevance of the master's thesis is due to the ability to analyze a large amount of textual information that is publicly available on the Internet, which in turn makes it possible to identify and prevent the development of depression, the most common mental health disorder.

The object of the study is messages from users of the social network Twitter.

The subject of the study is to study the relationship between a person's psycho-emotional state and his activity on social networks.

The structure of the master's thesis - the dissertation is presented on 40 pages. This work consists of sections "Introduction" and 3 three chapters: "Analysis of the problem and research of the subject area", "Sequence of work with text data" and "Analysis and interpretation of results" and the main part. The bibliography includes 40 sources.