

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Факультет прикладной математики и информатики

КАФЕДРА ИНФОРМАЦИОННЫХ СИСТЕМ УПРАВЛЕНИЯ

АННОТАЦИЯ К МАГИСТЕРСКОЙ ДИССЕРТАЦИИ

«Конструктивный анализ подходов к обработке больших и малых данных»

Олифирук МАКСИМ ВАЛЕРЬЕВИЧ

Научный руководитель – доктор физико-математических наук, профессор,
заведующий кафедры технологий программирования ФПМИ Курбацкий А.Н.

МИНСК, 2024

РЕФЕРАТ

Магистерская диссертация, 52 страниц, 19 рисунков, 2 таблицы, 13 формул, 12 источников.

Ключевые слова: обработка больших данных, обработка малых данных, методы анализа данных, машинное обучение, глубокое обучение, конструктивный анализ, нейронные сети, классификация текста, обработка естественного языка.

Цель работы - является проведение конструктивного анализа подходов к обработке как больших, так и малых данных в контексте методов анализа данных, включая машинное обучение и глубокое обучение.

Объект исследования – нейронные сети в применении классификации текста.

Предмет исследования – методы обработки данных, применяемые для анализа как больших, так и малых объемов информации.

В работе проводится анализ различных методов и подходов к классификации текста в контексте обработки естественного языка. Происходит изучение и сравнение различных моделей машинного обучения, включая классические алгоритмы и глубокие нейронные сети, применяемых для классификации текста.

Кроме того, проводится обзор различных архитектур нейронных сетей, таких как рекуррентные нейронные сети (RNN), сверточные нейронные сети (CNN), а также трансформеры, и изучаются их применимость и эффективность в задачах классификации текста.

Основной фокус работы направлен на анализ эффективности различных подходов к классификации текста на различных наборах данных, а также выявление их преимуществ и недостатков в различных сценариях применения. В результате проведенного исследования делаются выводы о наиболее эффективных методах и подходах к классификации текста в различных условиях и задачах.

ABSTRACT

Master's thesis, 52 pages, 19 figures, 2 tables, 13 formulas, 12 sources.

Keywords: big data processing, small data processing, data analysis methods, machine learning, deep learning, constructive analysis, neural networks, text classification, natural language processing.

The goal of the work is to conduct a constructive analysis of approaches to processing both large and small data in the context of data analysis methods, including machine learning and deep learning.

The object of study is neural networks in the application of text classification.

The subject of the research is data processing methods used to analyze both large and small amounts of information.

The paper analyzes various methods and approaches to text classification in the context of natural language processing. Various machine learning models are studied and compared, including classical algorithms and deep neural networks, used for text classification.

In addition, various neural network architectures such as recurrent neural networks (RNNs), convolutional neural networks (CNNs), and transformers are reviewed and their applicability and effectiveness in text classification tasks is studied.

The main focus of the work is aimed at analyzing the effectiveness of various approaches to text classification on various data sets, as well as identifying their advantages and disadvantages in various application scenarios. As a result of the study, conclusions are drawn about the most effective methods and approaches to text classification in various conditions and tasks.