

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Факультет прикладной математики и информатики

Кафедра дискретной математики и алгоритмики

Аннотация к магистерской диссертации

«Интеллектуальная визуальная аналитика сложных данных»

Кобзарь Ульяна Алексеевна

Научный руководитель – кандидат физико-математических наук, доцент
кафедры дискретной математики и алгоритмики ФПМИ Мушко В. В.

Минск, 2022

Реферат

Магистерская диссертация, 56 страниц, 35 рисунков, 10 таблиц, 25 источников.

Ключевые слова: ВИЗУАЛИЗАЦИЯ ДАННЫХ, СИСТЕМА РЕКОМЕНДАЦИИ ВИЗУАЛИЗАЦИЙ, ДЕРЕВО ПРИНЯТИЯ РЕШЕНИЙ, ОВЕРСЭМПЛИНГ, ДИАГРАММА, МАШИННОЕ.

Объект исследования — системы рекомендации визуализаций, взаимосвязи между набором данных и типом визуализации.

Предмет исследования — алгоритмы машинного обучения для решения задачи визуальной аналитики.

Цель работы — разработка программного обеспечения для рекомендации визуализаций предложенного набора данных.

В ходе работы были рассмотрены существующие системы рекомендации визуализаций и изучены алгоритмы, на которых они основаны. Выбраны признаки, извлекаемые из данных, по которым с помощью алгоритмов машинного обучения определяется подходящий тип визуализации. Алгоритмы обучены и протестированы на наборах данных и соответствующих им визуализациях, полученных из Plotly Community Feed. Продемонстрирована возможность их использования для рекомендации типа визуализаций.

Полученный результат можно использовать для проведения разведочного анализа данных и создания базовых визуализации во многих областях.

Abstract

Master thesis, 56 pages, 35 figures, 10 tables, 25 resources.

Keywords: DATA VISUALIZATION, VISUALIZATION RECOMMENDATION SYSTEM, DECISION TREE, OVERSAMPLING, CHART, MACHINE LEARNING.

The object of research is visualization recommender systems, their interactions between data set and types of visualization.

The subject of study is machine learning algorithms for solving the problem of visual analytics.

The aim of this work is the software development of visualization of the suggested data set.

During the study, existing visualization recommender systems and their algorithms, on which they are based, have been studied. There have been chosen the features extracted from the data set and according to which the exact type of visualization is chosen by machine learning algorithms. The algorithms have been learned and tested on the concrete datasets and corresponding visualizations obtained from Plotly Community Feed. The possibility of the usage of these algorithms is clearly demonstrated.

The result can be applied to preliminary exploratory data analysis and the implementing basic visualization in many spheres.