

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

Аннотация к дипломной работе

**ПРОГНОЗИРОВАНИЕ СТОИМОСТИ ПОЕЗДКИ В ТАКСИ С
ПОМОЩЬЮ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ**

Марус Анна Сергеевна

Научный руководитель:
кандидат физ.-мат. наук,
доцент Л. Л. Голубева

2024

В дипломной работе 64 страницы, 23 иллюстрации, 2 таблицы, 11 источников, 1 приложение.

МАШИННОЕ ОБУЧЕНИЕ, ПРЕДСКАЗАНИЕ СТОИМОСТИ ПОЕЗДОК ТАКСИ, ПРЕДОБРАБОТКА ДАННЫХ, ОБУЧЕНИЕ С УЧИТЕЛЕМ, ЛИНЕЙНАЯ РЕГРЕССИЯ, ДЕРЕВО РЕШЕНИЙ, СЛУЧАЙНЫЙ ЛЕС, ГРАДИЕНТНЫЙ БУСТИНГ, МЕТОД К-БЛИЖАЙШИХ СОСЕДЕЙ, МЕТОД ОПОРНЫХ ВЕКТОРОВ

Объектом исследования является набор размеченных данных, содержащих сведения о поездках в такси в Нью-Йорке за 2023 год.

Целью дипломной работы является проведение интеллектуального анализа статистических данных поездок такси в Нью-Йорке и предсказание стоимости поездок в такси методами машинного обучения.

Для достижения поставленной цели использовались методы машинного обучения.

В дипломной работе получены следующие результаты:

- 1) описаны методы и задачи машинного обучения;
- 2) выполнена предобработка статистических данных, найдены новые признаки, проанализировано влияние различных факторов на стоимость проезда в такси;
- 3) построены прогнозы стоимости поездок в такси различными методами машинного обучения, выполнен сравнительный анализ их работы.

Дипломная работа носит практический характер. Ее результаты могут быть использованы для прогнозирования стоимости поездок такси по заданным параметрам.

Дипломная работа является завершенной, поставленные задачи решены в полной мере, присутствует возможность дальнейшего развития исследований.

Дипломная работа выполнена автором самостоятельно.

Thesis project is presented in the form of an explanatory note of 64 pages, 23 figures, 2 tables, 11 references, 1 application.

MACHINE LEARNING, TAXI FARE PREDICTION, DATA PREPROCESSING, TEACHER TRAINING, LINEAR REGRESSION, DECISION TREE, RANDOM FOREST, GRADIENT BOOSTING, K-NEAREST NEIGHBOR METHOD, SUPPORT VECTOR MACHINE LEARNING

The research object of this thesis project is to study a set of marked-up data containing information about taxi rides in New York in 2023.

The purpose of this work is to conduct an intelligent analysis of statistical data on taxi trips in New York and predict the cost of taxi trips using machine learning methods.

Machine learning methods were used to achieve the goal.

The main results of the thesis project are as follows:

- 1) methods and tasks of machine learning are described.
- 2) pre-processing of statistical data was performed, new signs were found, the influence of various factors on the cost of taxi travel was analyzed;
- 3) forecasts of the cost of taxi trips using various machine learning methods were constructed, and a comparative analysis of their work was performed.

This thesis project is a practical one. Its results can be used to predict the cost of taxi rides according to the specified parameters.

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

The thesis project was done solely by the author.