

УДК 004.67

ПОДХОД К ЗАДАЧЕ АНАЛИЗА ВРЕМЕННЫХ РЯДОВ МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ С ИСПОЛЬЗОВАНИЕМ ФРЕЙМВОРКА APACHE SPARK В ПЛАТФОРМЕ SDLITICA

Дашкова М.С.(Университет ИТМО, Санкт-Петербург)
Научный руководитель – к.т.н., доцент Перл И.А.
(Университет ИТМО, Санкт-Петербург)

В работе рассмотрены основные методы анализа временных рядов с использованием машинного обучения, проведен обзор фреймворка Apache Spark, как инструмента работы с временными рядами на платформе sdLitica .

Введение. Нас окружает множество явлений, которые можно описать временными рядами. Это способ сказать, что многие вещи или события можно описать как наборы наблюдений, которые происходят в течение определенного периода. Использование такого представления данных на сегодня играет важную роль во многих областях, таких как экономика, бизнес, энергетика, транспорт, торговля, спорт, медицина и многое другое. Для получения важной информации данные подвергаются анализу. Проект sdLitica занимается разработкой платформы для анализа временных рядов, которая будет содержать спектр инструментов, основанных на привычных статистических методах и на современных подходах с использованием машинного обучения.

Для разработки такой платформы есть задача обработки временных рядов методами машинного обучения с использованием фреймворка Apache Spark. В ходе предыдущего исследования были отобраны некоторые методы машинного обучения, подходящие для платформы, один из них градиентный бустинг.

Целью работы является изучение способа реализации градиентного бустинга с помощью Apache Spark и его интеграции в платформу sdLitica.

Основная часть. В ходе исследования проводится обзор методов анализа временных рядов, фреймворка Apache Spark и реализация модели градиентного бустинга для работы с данными временных рядов.

Выводы. Данная работа позволяет ознакомиться с основными подходами к задаче анализа временных рядов методами машинного обучения, а также демонстрирует основные идеи работы с Apache Spark.

Дашкова М.С. (автор)

Подпись

Перл И.А. (научный руководитель)

Подпись

Контакты автора: dashkova_m@inbox.ru, 8(981)801-96-13