

Отчёт о выполнении работы

Выполнил: Солдатенков Антон

Вариант: 5 (MongoDB)

Цель работы:

Получить навыки выгрузки исходных данных и отправки результатов модели с использованием различных источников данных согласно варианту задания.

Задачи работы:

1. Обеспечить выгрузку данных при каждом запуске модели.
2. Обеспечить загрузку данных сразу по завершении работы модели.
3. Необходимо разработать протокол взаимодействия между моделью и источником данных.
4. Необходимо разработать формат хранения данных исходя из особенностей источника данных.
5. Рекомендуется использование docker контейнеров.

Ход работы

1. Из дампа OpenFoodFacts на локальном хосте развёрнута база данных MongoDB
2. Установлены необходимые .jar пакеты для взаимодействия Spark и MongoDB
3. Настроены чтение и запись в развёрнутую базу данных
4. Средствами SparkML разработан пайплайн обучения модели К-средних
5. Выполнена запись предсказаний модели в базу данных для дальнейшего использования в лабораторной работе №4.