МИНИСТЕРСТВО ЦИФРОВОГО РАЗВИТИЯ, СВЯЗИ И МАССОВЫХ КОММУНИКАЦИЙ РОССИЙСКОЙ ФЕДЕРАЦИИ

Ордена Трудового Красного Знамени федеральное государственное бюджетное образовательное учреждение высшего образования

**«Московский технический университет связи и информатики»**

Кафедра «Системное программирование»

Практическая работа №5

**«Spark»**

по дисциплине

Введение в большие данные

Выполнил: студент гр. БЭИ2202

Кулешов А. С.

Проверил: Сергеев А.Е.

Москва, 2025 г

Цель работы. Получить навыки работы с HBase

**ВЫПОЛНЕНИЕ**

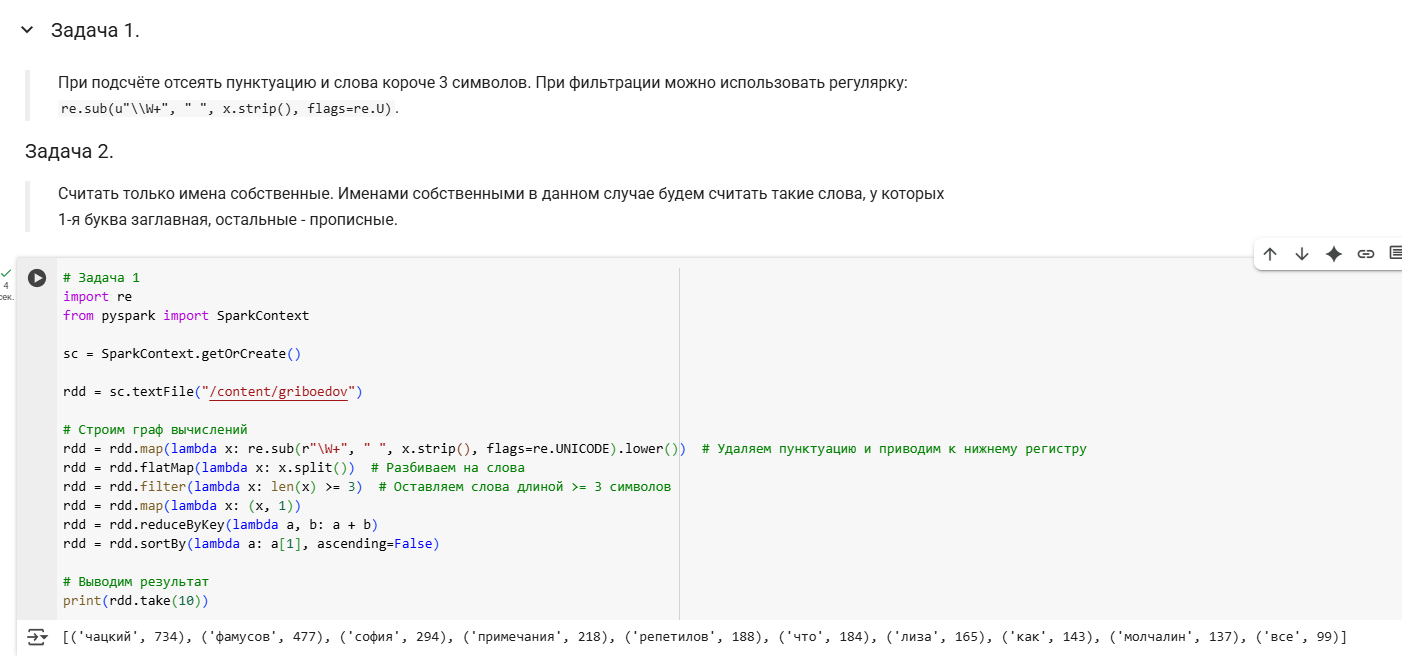


Рисунок 1 – Задача 1

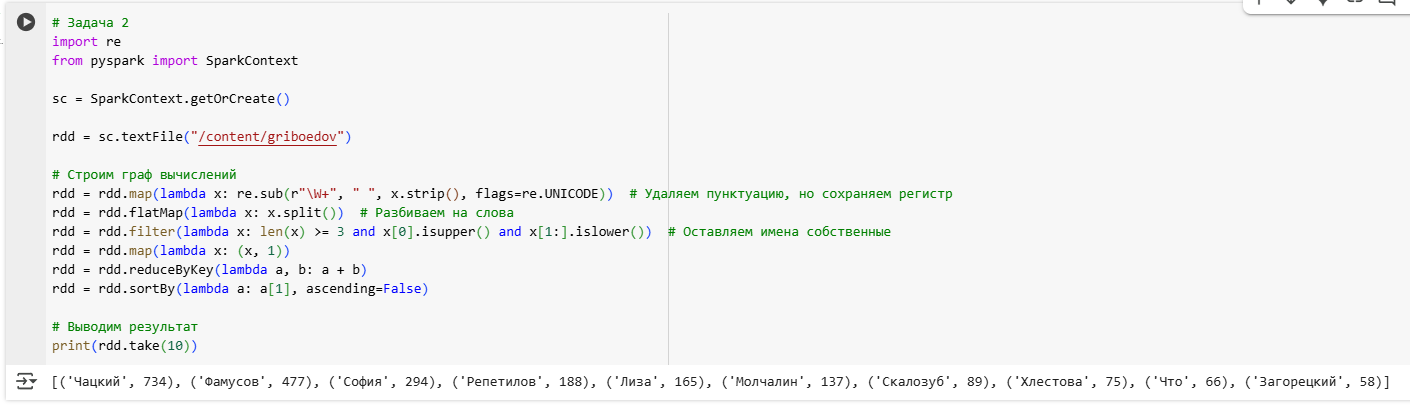


Рисунок 2 – Задача 2



Рисунок 3 – Задача 3

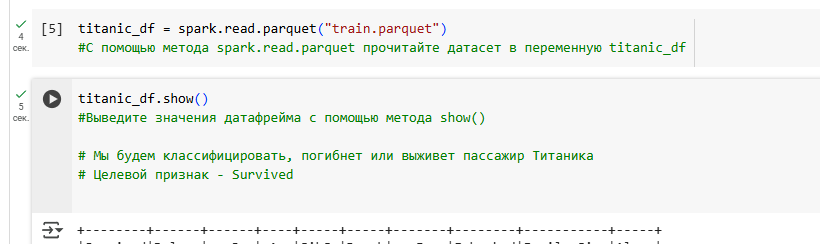


Рисунок 4 – SparkML. Чтение данных

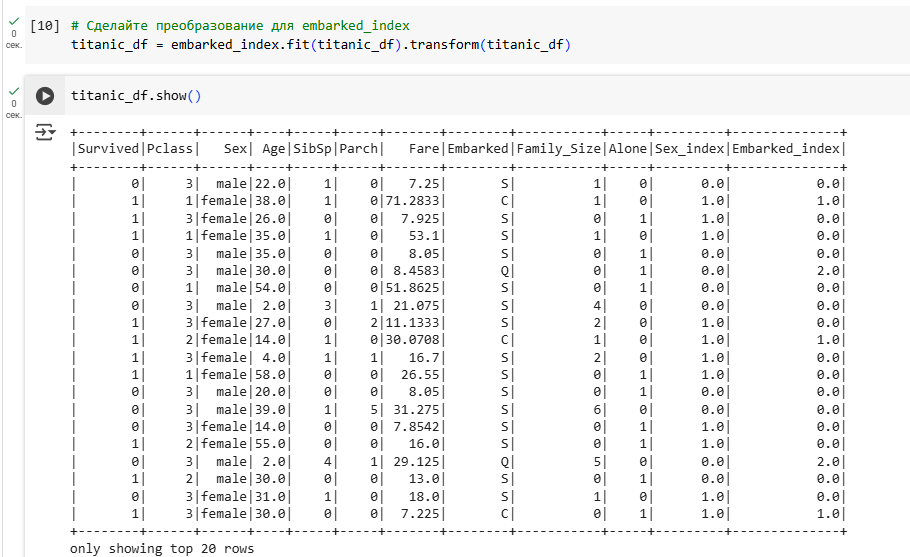


Рисунок 5 – SparkML. Embarked\_index



Рисунок 6 – SparkML. RandomForestClassifier

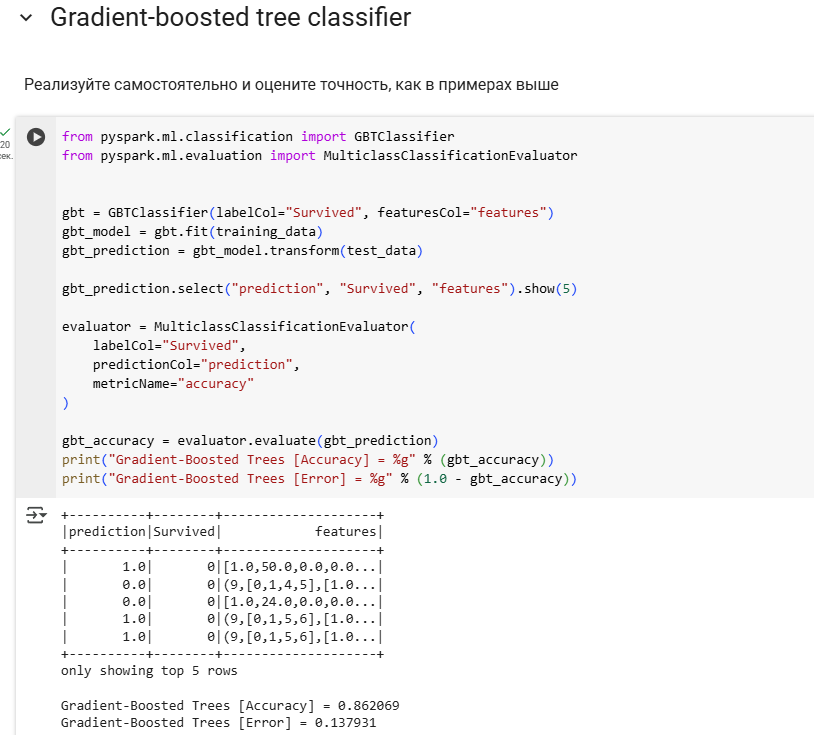


Рисунок 7 – SparkML. GBT

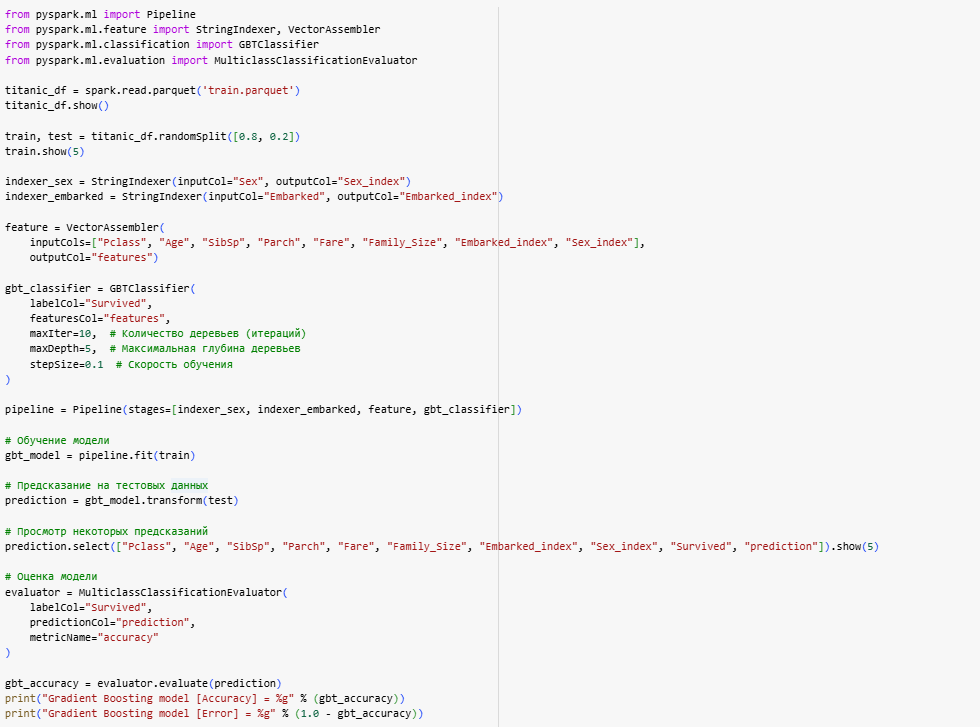


Рисунок 8 – SparkML. Переписал на Градиентный Бустинг

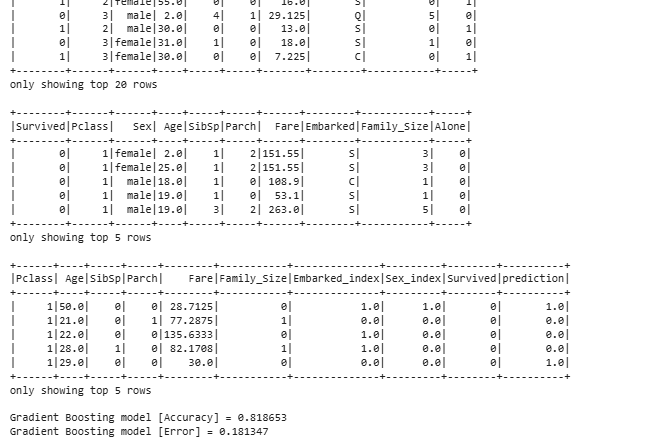


Рисунок 9 – SparkML. Результат