Московский государственный технический университет им. Н.Э. Баумана

Факультет «Информатика и системы управления»

Кафедра «Системы обработки информации и управления»



**Отчет по ЛР №5**

**по курсу «Технологии машинного обучения»**

**«Линейные модели, SVM и деревья решений»**

**ИСПОЛНИТЕЛЬ:**

Кемерова А.М.

Группа ИУ5-61Б

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

"\_\_"\_\_\_\_\_\_\_\_\_\_\_2020 г.

**ПРЕПОДАВАТЕЛЬ:**

Гапанюк Ю.Е.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

"\_\_"\_\_\_\_\_\_\_\_\_\_\_2020 г.

## **Москва 2020**

**Цель лабораторной работы:**

Изучение линейных моделей, SVM и деревьев решений.

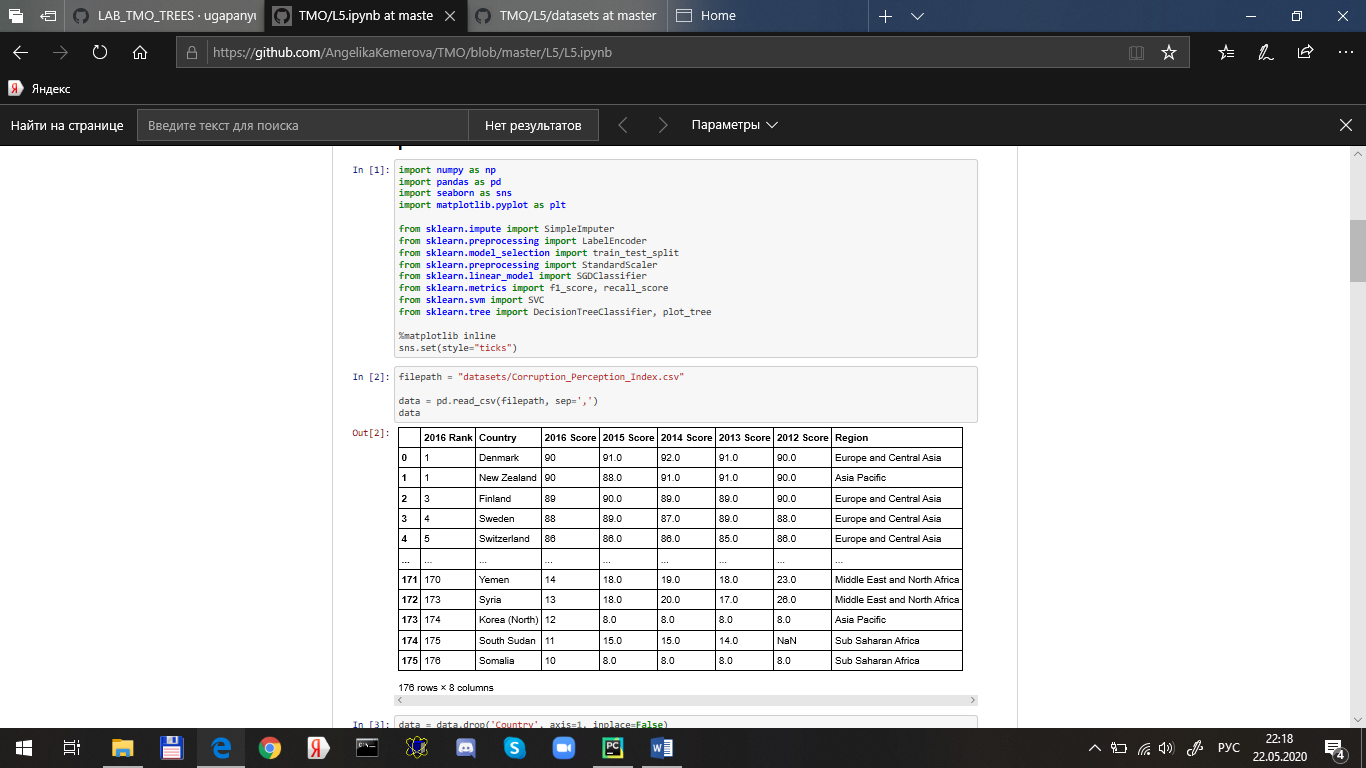
**Задание:**

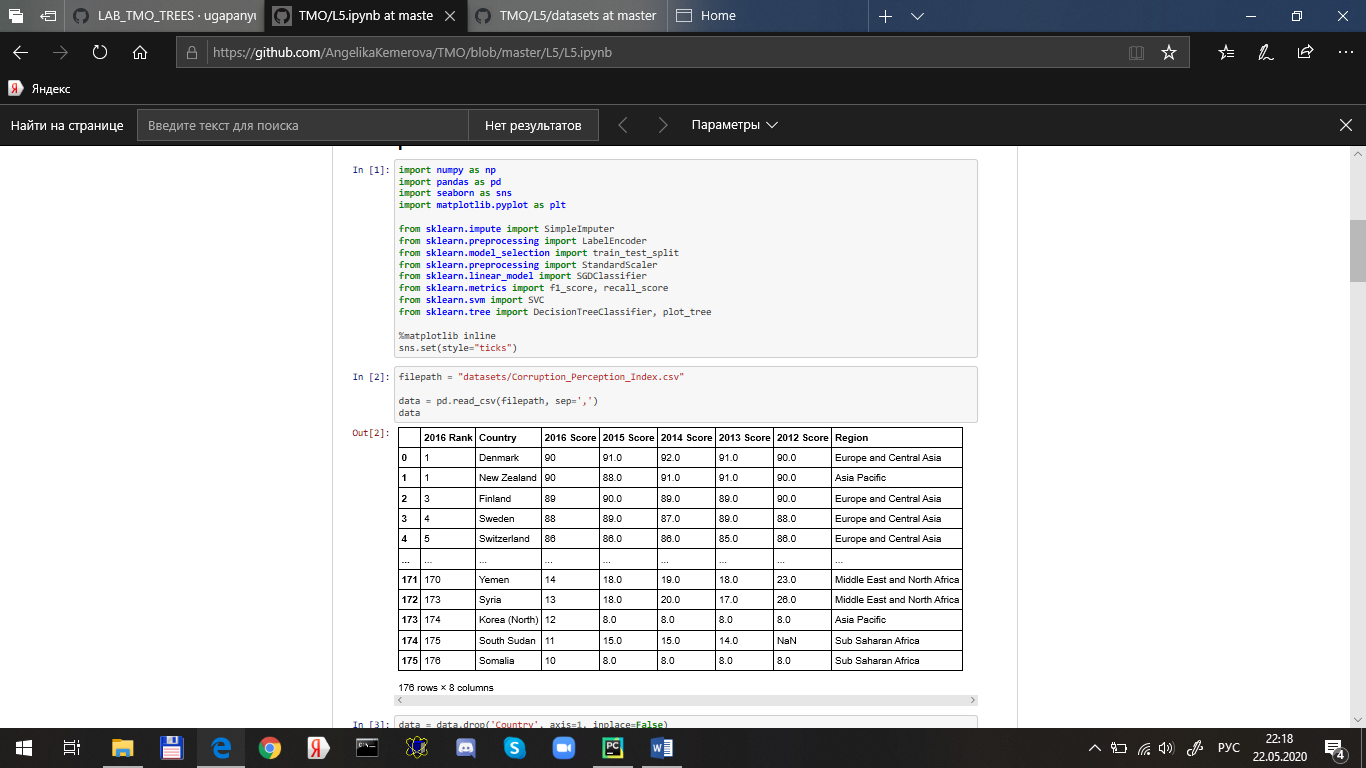
1. Выберите набор данных ([Corruption\_Perception\_Index.csv](https://github.com/AngelikaKemerova/TMO/blob/master/L5/datasets/Corruption_Perception_Index.csv)) для решения задачи классификации или регрессии.
2. В случае необходимости проведите удаление или заполнение пропусков и кодирование категориальных признаков.
3. С использованием метода train\_test\_split разделите выборку на обучающую и тестовую.
4. Обучите следующие модели:
   * одну из линейных моделей;
   * SVM;
   * дерево решений.
5. Оцените качество моделей с помощью двух подходящих для задачи метрик. Сравните качество полученных моделей.

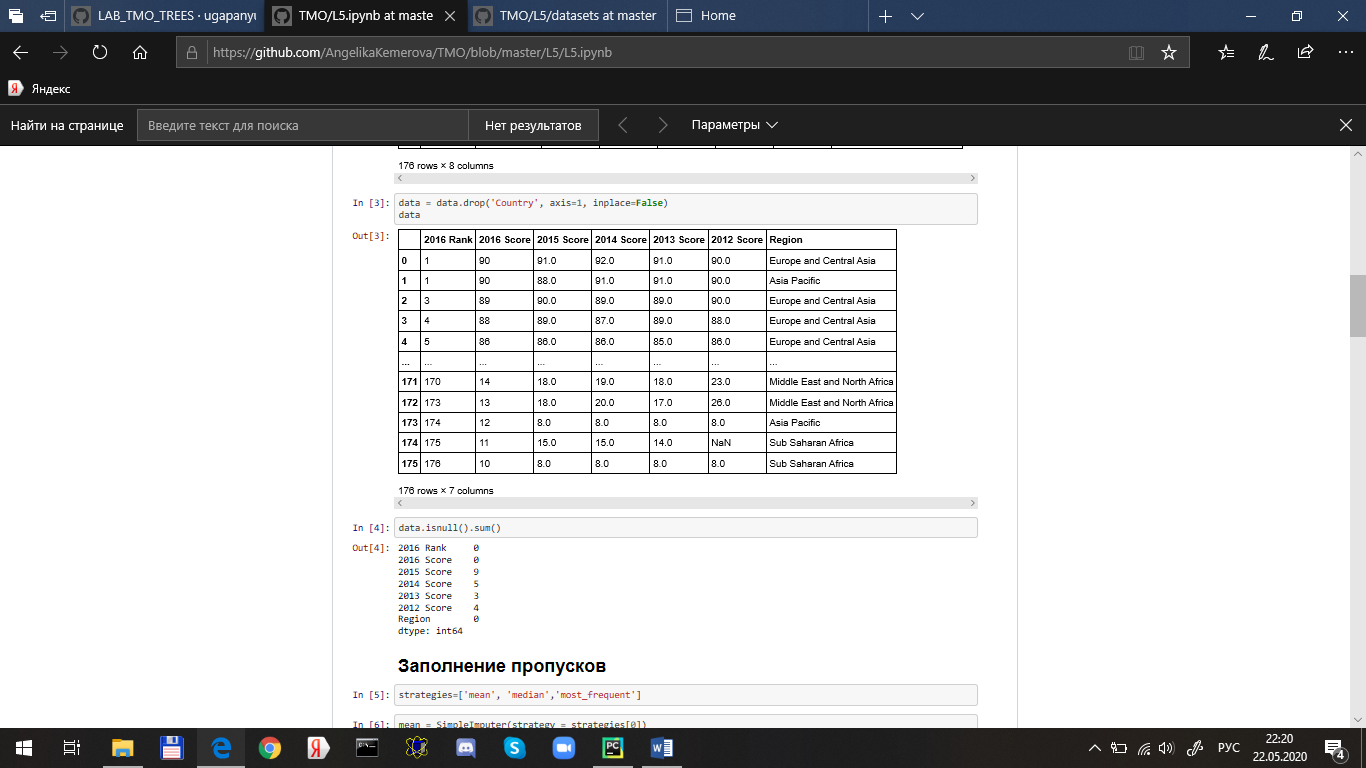
**Дополнительные задания:**

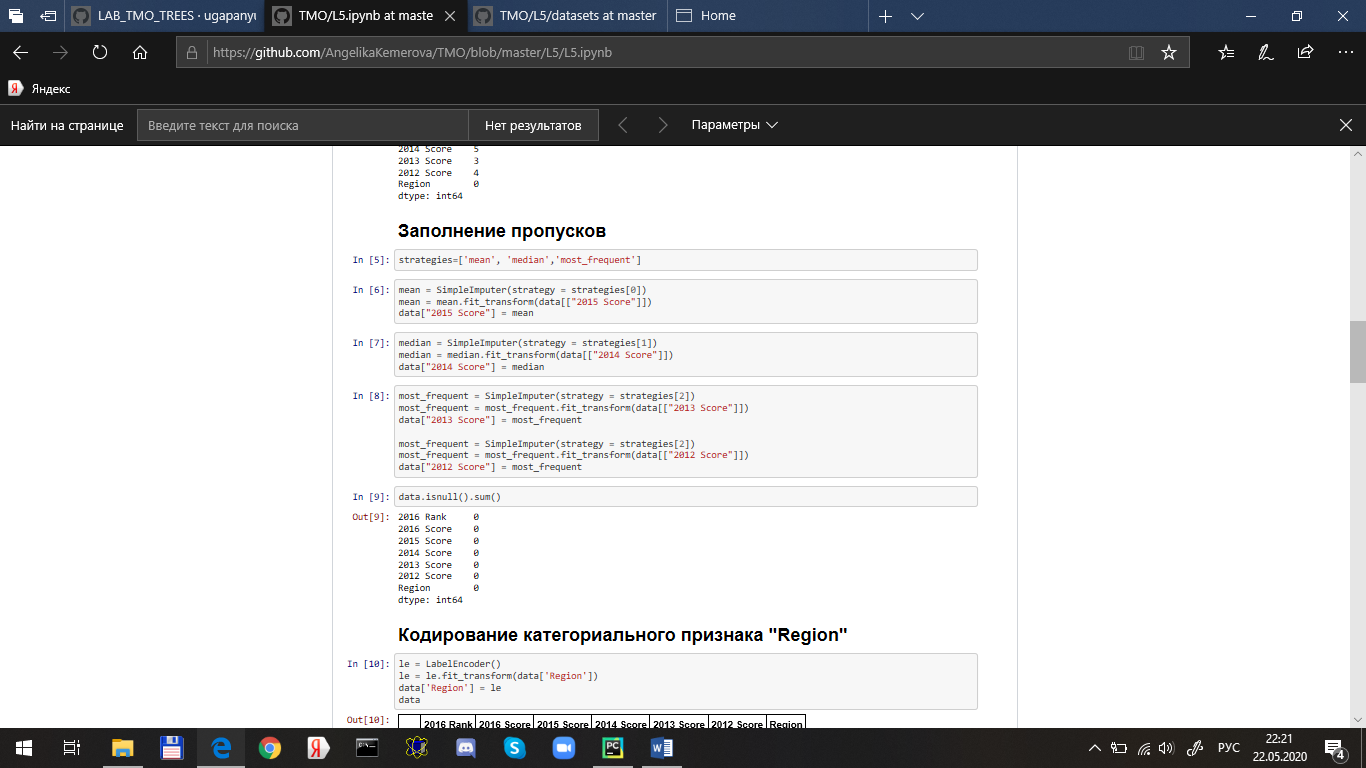
* Проведите эксперименты с важностью признаков в дереве решений.
* Визуализируйте дерево решений.

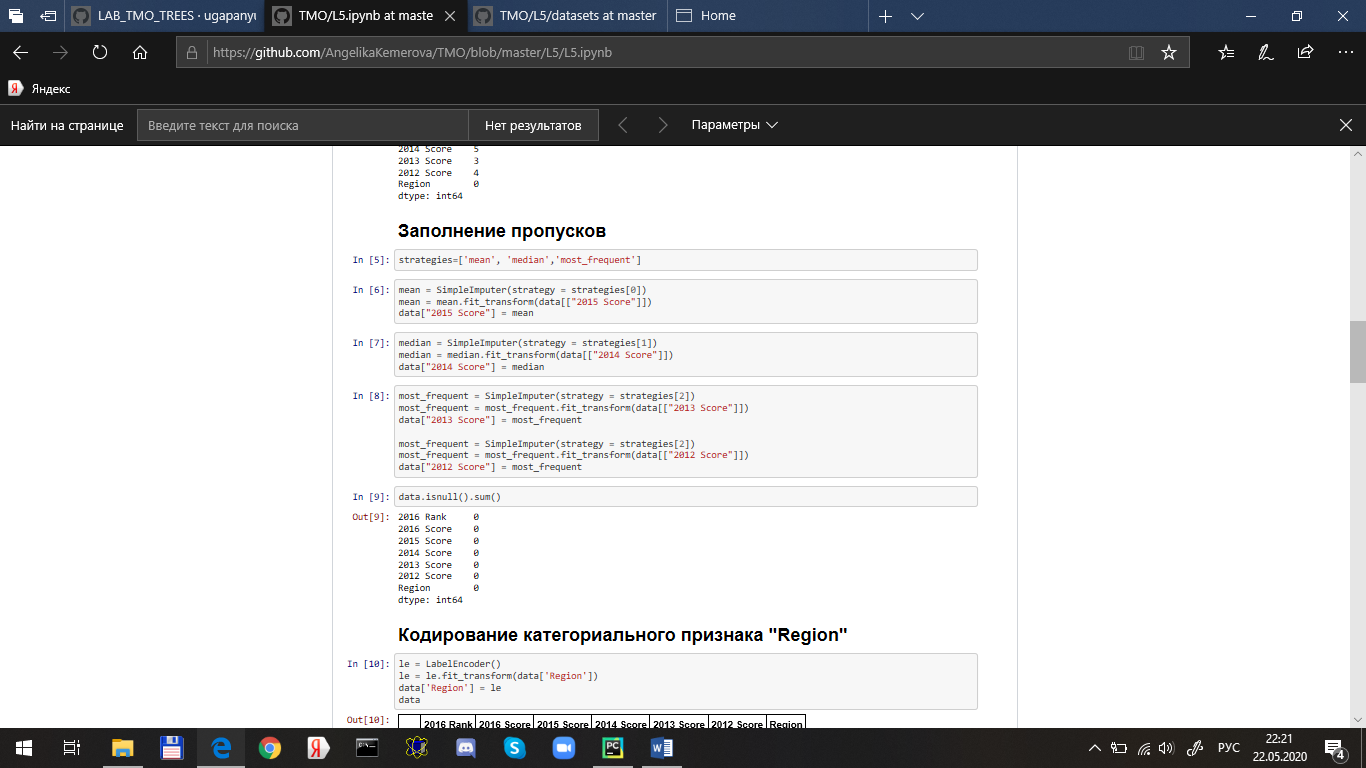
**Выполнение:**

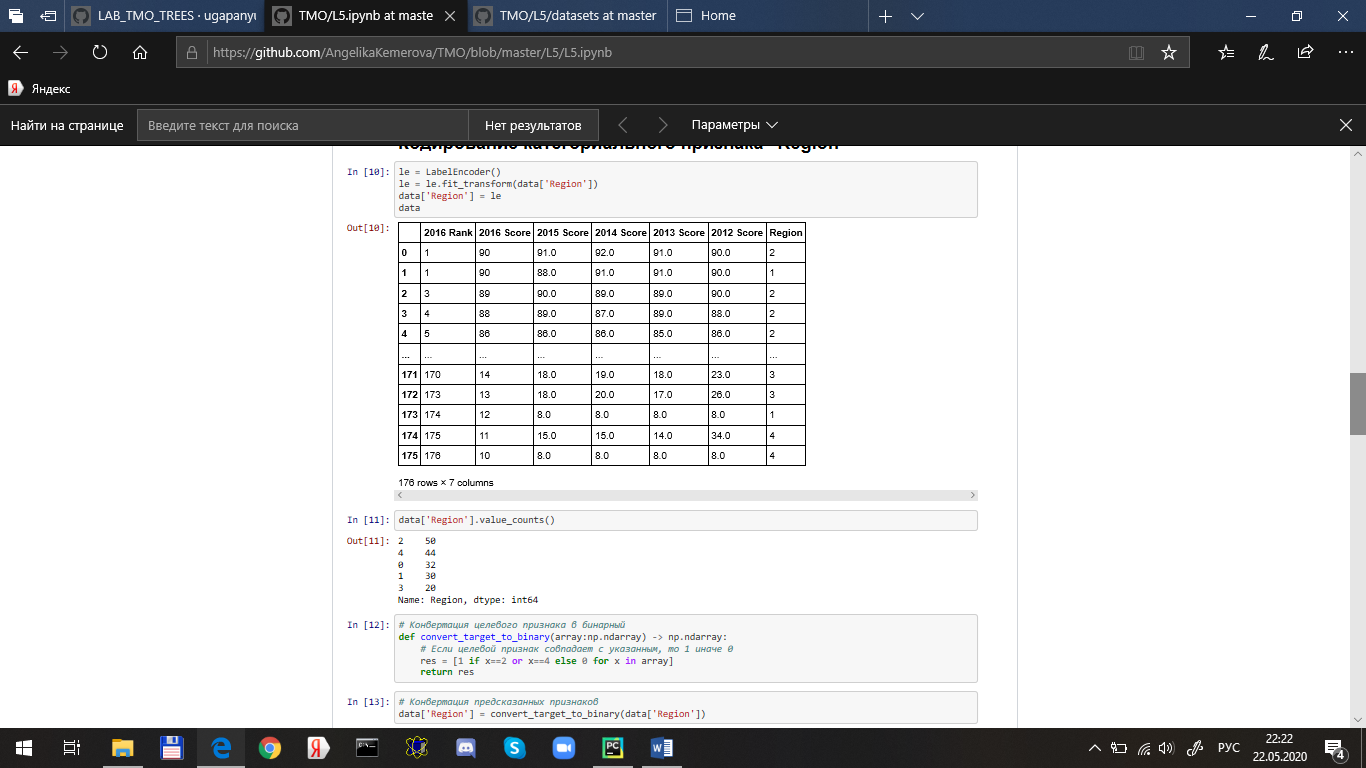


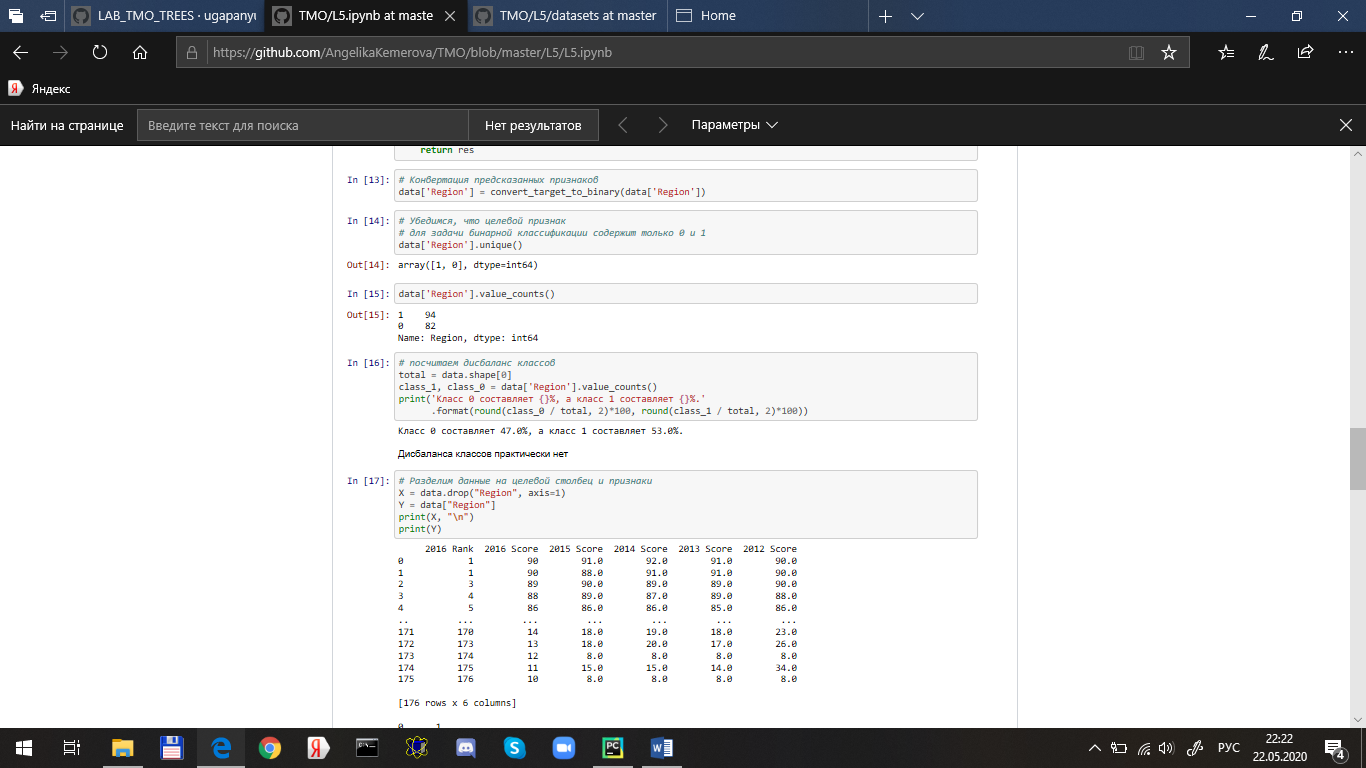


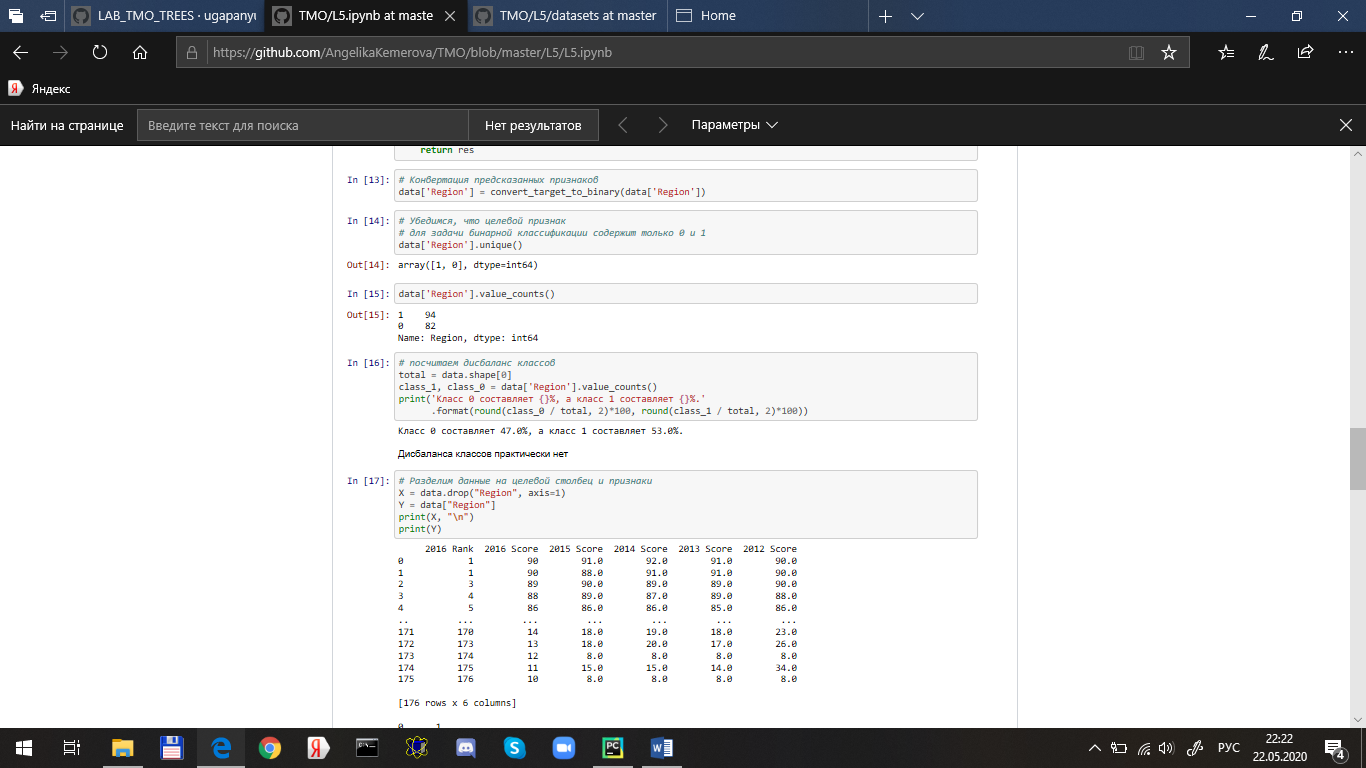


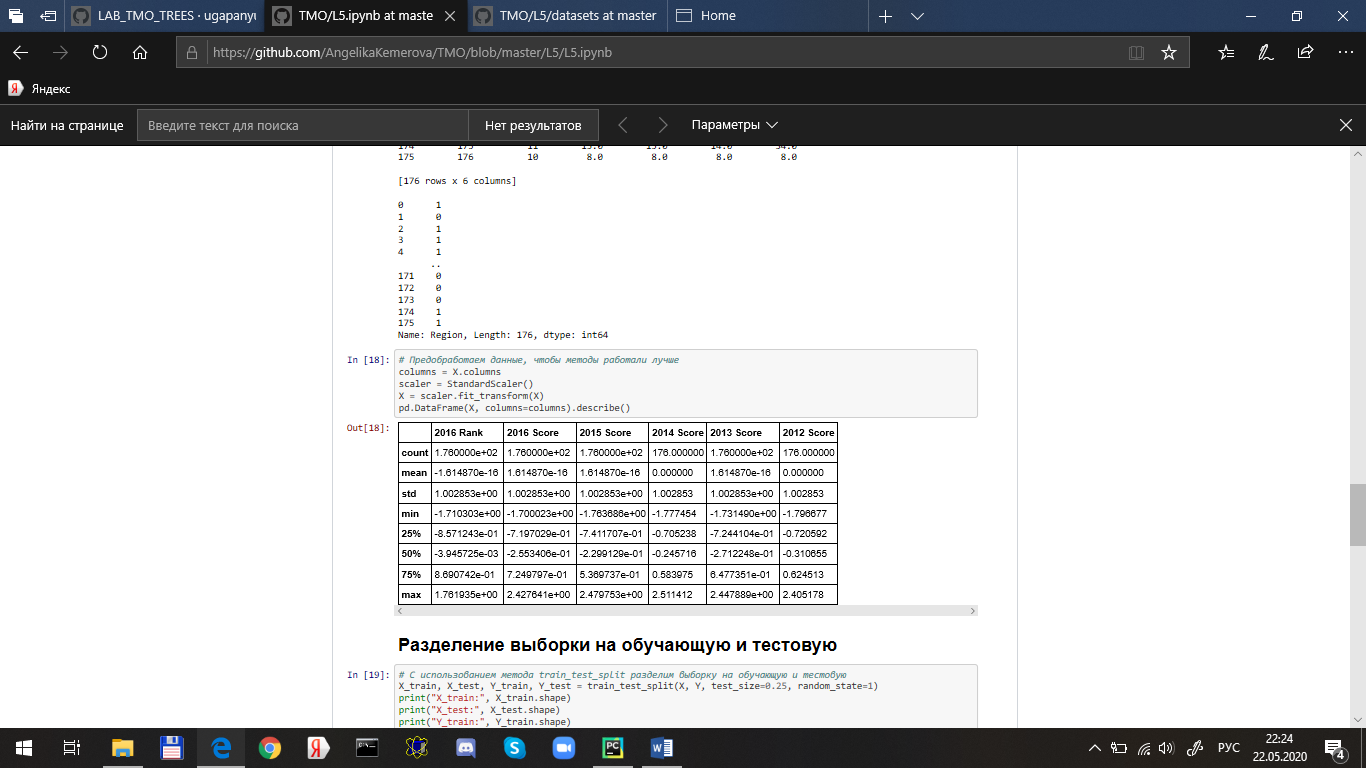


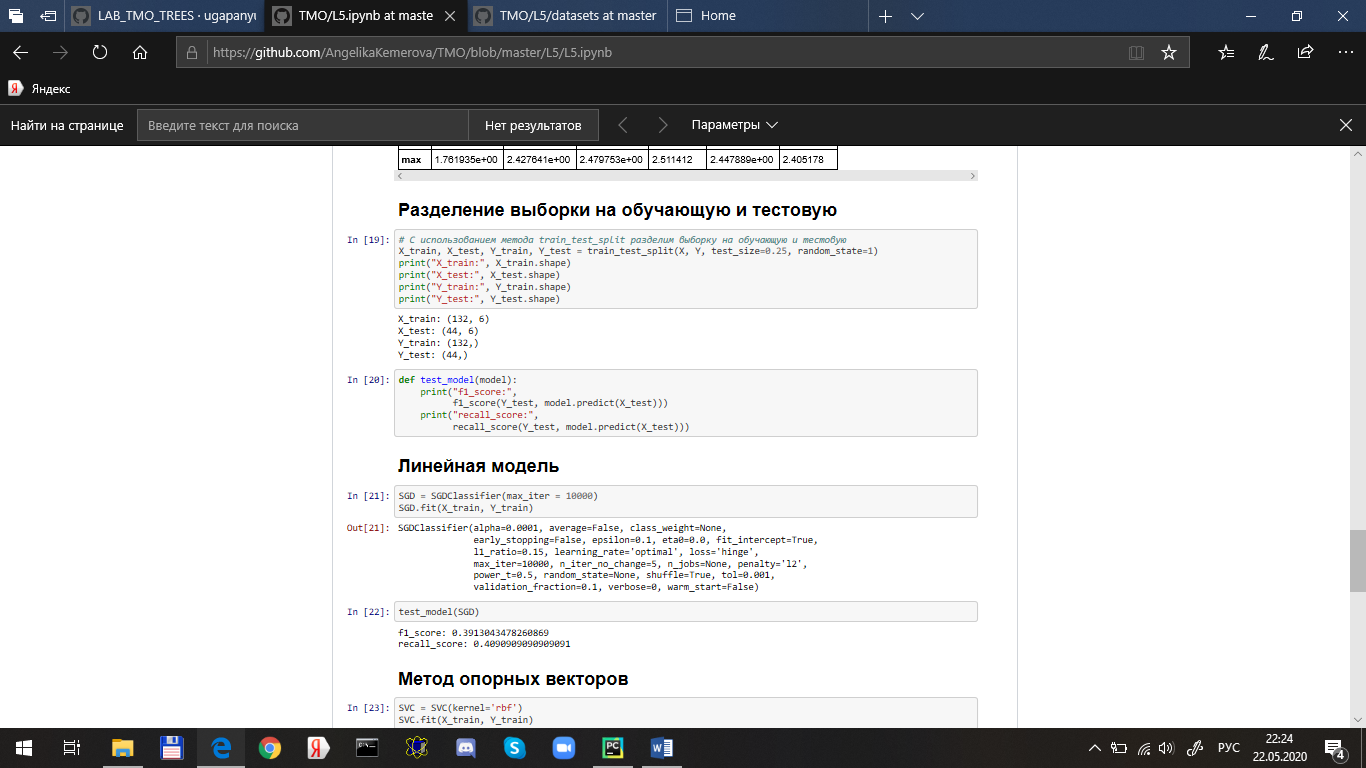


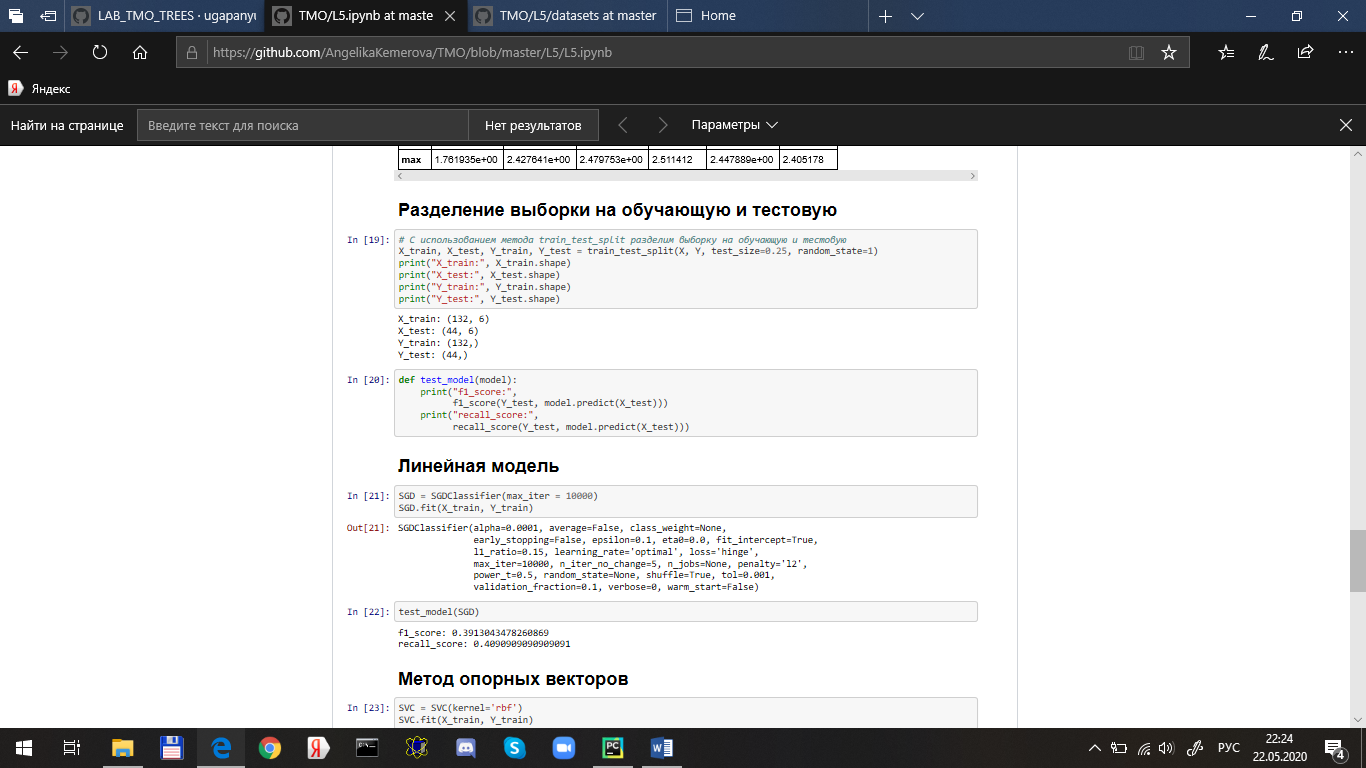


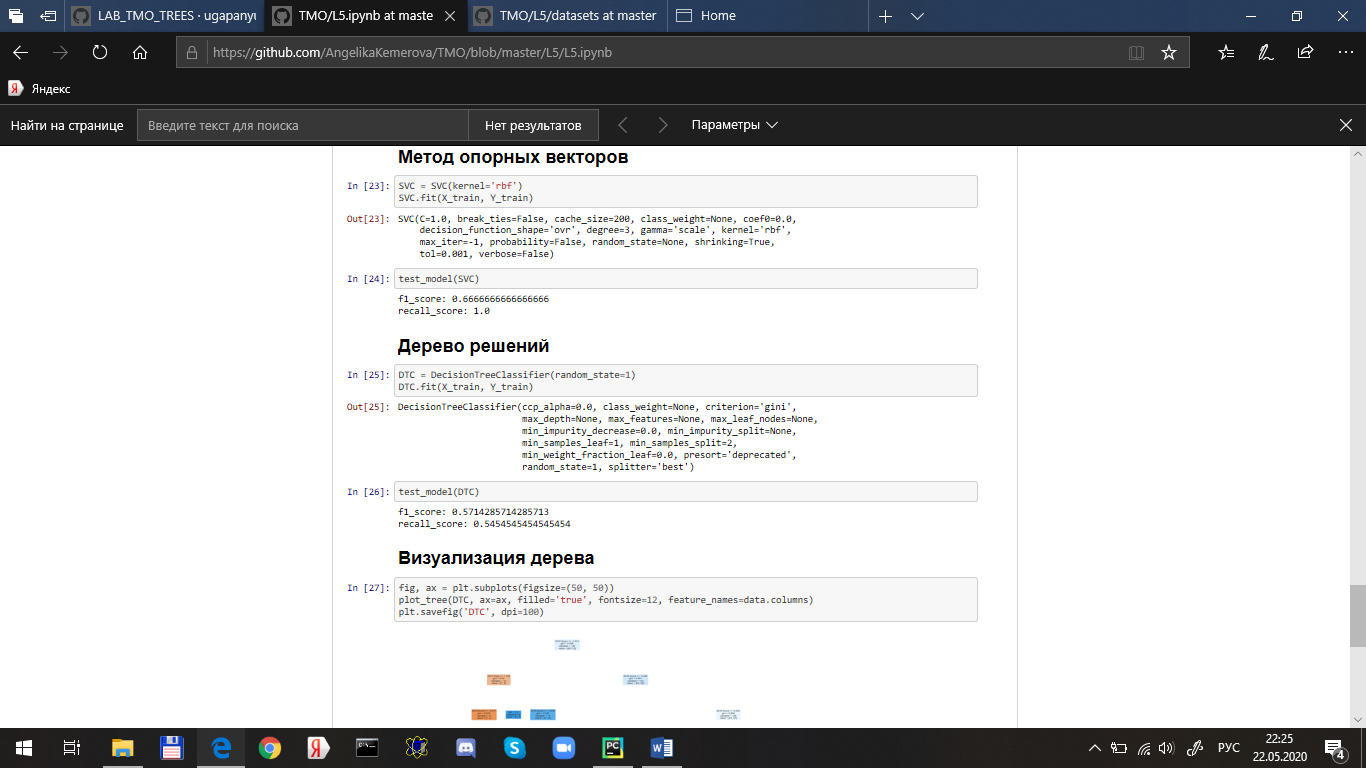


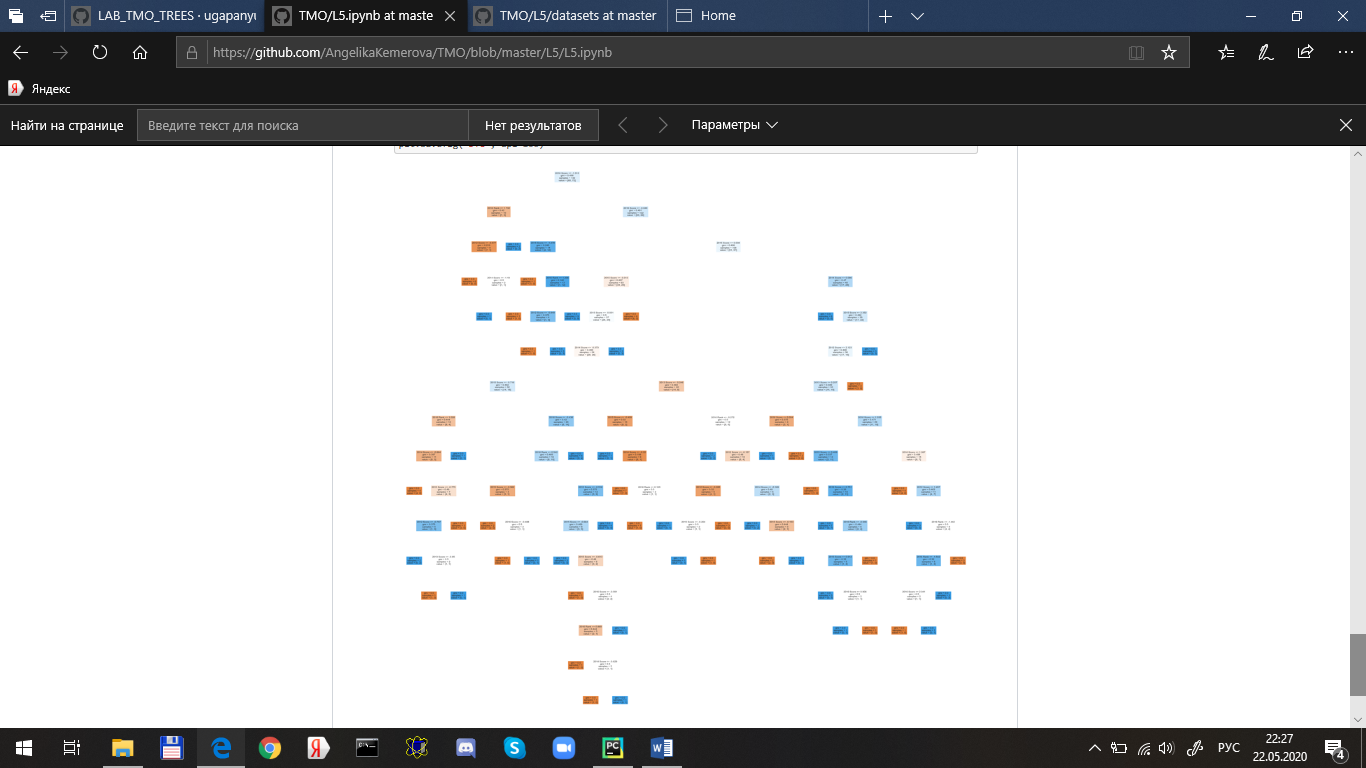












**Вывод:**

Сравнивая качество моделей, можно сделать вывод, что метод опорных векторов (SVM) показал лучшие практические результаты.