Московский государственный технический университет им. Н.Э. Баумана

Факультет «Информатика и системы управления» Кафедра ИУ5 «Системы обработки информации и управления»

Курс «Технологии машинного обучения» Отчет по рубежному контролю №2

Вариант 11

Выполнил: Студент группы ИУ5-63Б Лобанов Иван

Руководители: Гапанюк Ю.Е.

Москва, 2024 г.

Задание. Для заданного набора данных (по Вашему варианту) постройте

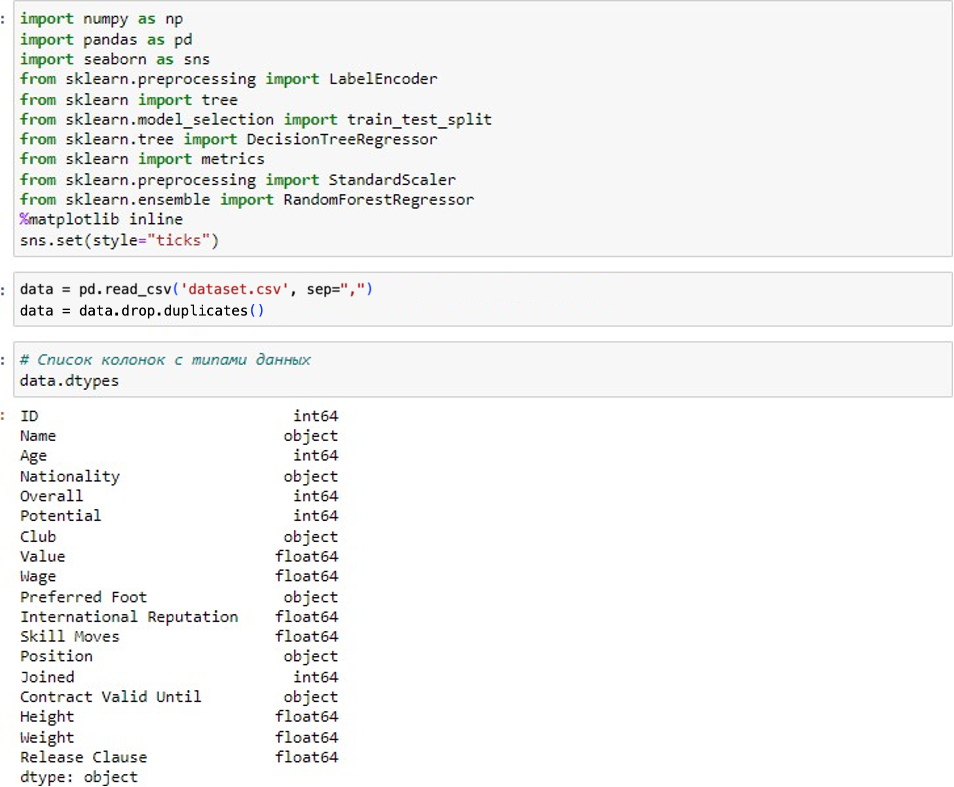
модели классификации или регрессии (в зависимости от конкретной задачи, рассматриваемой в наборе данных). Для построения моделей используйте методы 1 и 2 (по варианту для Вашей группы). Оцените качество моделей на основе подходящих метрик качества (не менее двух метрик). Какие метрики качества Вы использовали и почему? Какие выводы Вы можете сделать о качестве построенных моделей? Для построения моделей необходимо выполнить требуемую предобработку данных: заполнение пропусков,

кодирование категориальных признаков, и т.д. Метод 1 – дерево решений

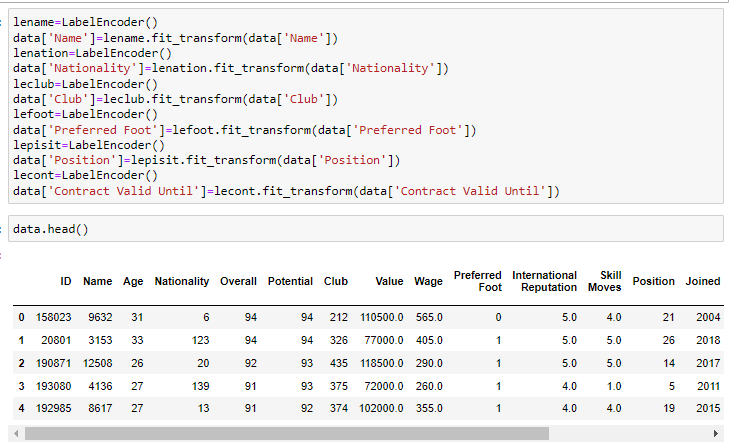
Метод 2 – случайный лес

Датасет - <https://www.kaggle.com/datasets/winterbreeze/fifa19eda> Выполнение:

Загрузка датасет



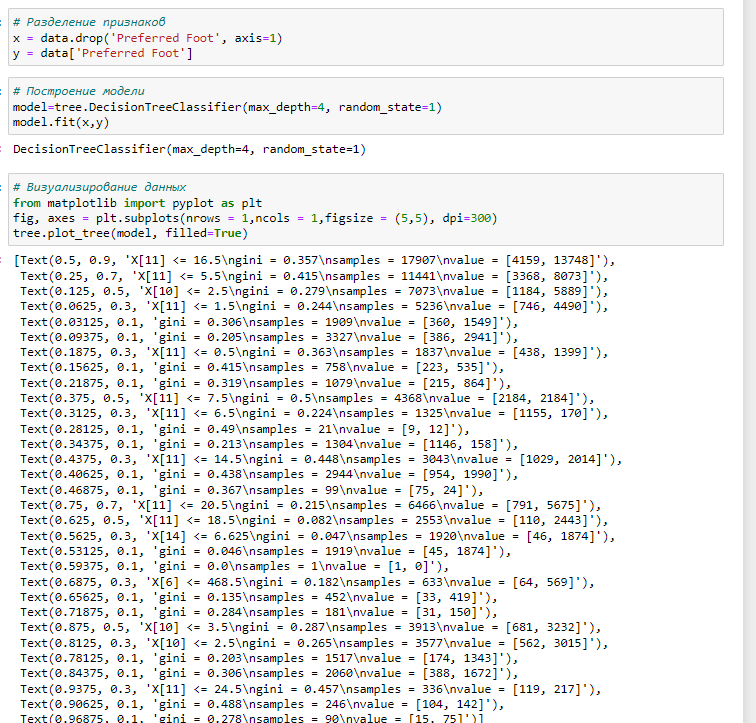
Кодирование категориальных признаков

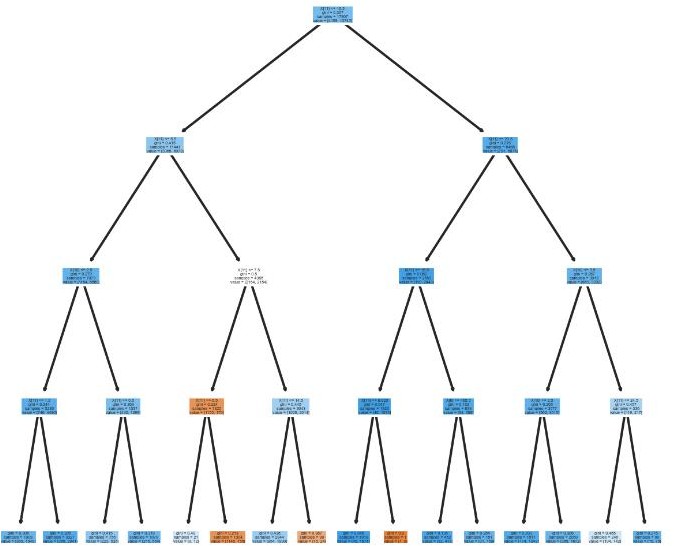


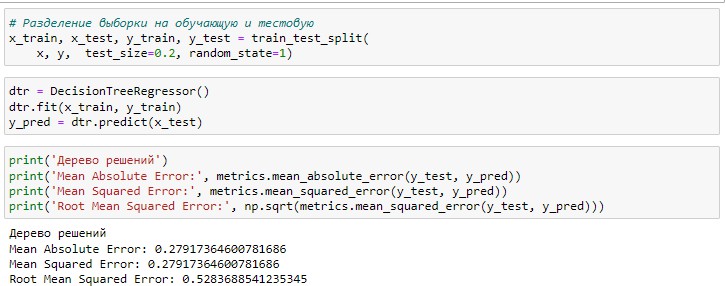
Удаление строк с пустыми значениями



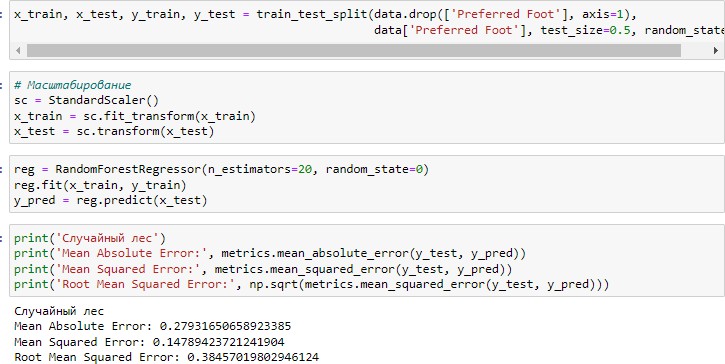
Дерево решений







Случайный лес



Случайный лес оказался более качественной моделью.