**Задание №2. Классификаторы**

Написать программу на Python, которая загружает набор данных и выполняет задачи построения бинарных классификаторов, подбора гиперпараметров моделей и анализа качества работы классифицирующих моделей.

1. Выполнить предварительную обработку набора данных.

2. Построить классифицирующие модели с использованием GridSearch для алгоритмов RandomForest, LogisticRegression и XGBoost. Провести эксперименты с регуляризацией. Выполнить ROC-анализ и визуализировать результаты анализа.

3. Визуализировать значения важности признаков для моделей, использующих деревья и логистической регрессии для сбалансированных и первоначальной выборок.

4. Выполнить K-блочную стратифицированную проверку для указанных алгоритмов с значениями гиперпараметров, полученных ранее.

5. Получить усреднённую важность признаков для наборов моделей, использующих деревья и логистической регрессии

6. Визуализировать результаты ROC-анализа. Рассчитать среднюю ROC и отклонение (standard deviation).

7. Напишите короткое заключение о наиболее интересных зависимостях, которые Вы обнаружили в результате выполнения работы.

Выборка данных для задания выбирается исходя из номера Вашей зачетной книжки:

• если последняя цифра 0, 5: Выборка – Классификатор типа ландшафта в природном парке (<https://archive.ics.uci.edu/ml/datasets/Covertype>) метка классов – Cover\_Type; Положительный класс – 1, Отрицательный класс – 2, 3, 4, 5, 6, 7;

• если последняя цифра 1, 6: Выборка – Классификатор типа ландшафта в природном парке (<https://archive.ics.uci.edu/ml/datasets/Covertype>); метка классов – Cover\_Type; Положительный класс – 2, Отрицательный класс – 1, 3, 4, 5, 6, 7;

• если последняя цифра 2, 7: Выборка – Классификатор типа ландшафта в природном парке (<https://archive.ics.uci.edu/ml/datasets/Covertype>); метка классов – Cover\_Type; Положительный класс – 3, Отрицательный класс – 1, 2, 4, 5, 6, 7;

• если последняя цифра 3, 8: Выборка – Классификатор типа ландшафта в природном парке (<https://archive.ics.uci.edu/ml/datasets/Covertype>); метка классов – Cover\_Type; Положительный класс – 4, Отрицательный класс – 1, 2, 3, 5, 6, 7;

• если последняя цифра 4, 9: Выборка – Классификатор типа ландшафта в природном парке (<https://archive.ics.uci.edu/ml/datasets/Covertype>); метка классов – Cover\_Type; Положительный класс – 5, Отрицательный класс – 1, 2, 3, 4, 6, 7.