**Тестовое задание: *отличия ответчиков на противоопухолевую терапию от неответчиков***

**Данные**

Вам дан *tsv*-файл, содержащий проценты крупных лейкоцитарных популяций в крови у 100 пациентов с различными онкологиями. Проценты популяций крови подсчитаны относительно общего числа лейкоцитов в крови каждого пациента. Пациенты пронумерованы от 1 до 100, номер пациента записан в колонке *donor\_num*.

Каждый пациент подвергался лечению новым методом противоопухолевой терапии. Забор крови у пациентов проводился два раза: до и после терапии. Таким образом, для каждого пациента доступна информация о *двух временных точках*. Номер временной точки (1 - до терапии, 2 - после терапии) приведен в колонке *Time point*.

По результатам терапии пациенты разделены на 2 группы: ответчики и неответчики на терапию. Статус ответа на лечение (0 - неответчик, 1 - ответчик) приведен в колонке *response*.

**Основное задание**

Ваша задача проанализировать данные и предоставить параметры, по которым можно достоверно различить ответчиков от неответчиков. Помимо самих параметров, нужно предоставить доказательства того, что эти параметры действительно достоверно различают две группы пациентов. В качестве параметров можно использовать как отдельные популяции клеток, так и совокупности нескольких популяций или совокупность всех популяций. Можно анализировать отличия в конкретной временной точке, а также наблюдать *динамику изменений*. В данном задании продемонстрируйте свое умение выдвигать гипотезы и проводить их проверку. Отрицательный результат при проверке осмысленной гипотезы в данном задании будет также засчитываться в вашу пользу.

**Дополнительное задание**

На основании полученных вами параметров создайте и обучите классификатор, который сможет различать респондеров и нереспондеров по предоставленным вам данным. Протестируете различные типы моделей, подберите параметры моделей на кросс-валидации, оцените качество моделей (обоснуйте ваш выбор метрики для оценки качества) и выберете лучшую. Опишите, каким образом вы подходили к решению задачи для маленького объема выборки, как проводили отбор признаков для обучения. Сделайте финальный вывод о том, получилось ли у вас создать модель, достоверно отделяющую ответчиков от неответчиков и что могло бы помочь в дальнейшем улучшении модели.

**Требования**

Весь анализ проведите на языке Python в Jupyter Notebook.

В качестве ответа на задание прикрепите ваш Jupyter Notebook, содержащий все необходимые шаги анализа, визуализацию результатов и комментарии, иллюстрирующие ход ваших мыслей и рассуждения при решении.