**Задание №1**

Есть данные по клиентам широкополосного доступа в сеть Internet. Данные разбиты по 4 таблицам.

**Тарифы.xlsx:**

*account* - номер клиентского договора

*connected* - дата подключения клиента

*tariff*- текущий тариф клиента

*IP* - наличие (1) или отсутствие (0) на договоре услуги выделенный IP

*lease* - наличие (1) или отсутствие (0) на договоре услуги аренда роутера

**Оплата. xlsx:**

*account* - номер клиентского договора

*datetime* - дата оплаты

*amount* - сумма

*method* - метод оплаты

**Изменение тарифов. xlsx:**

*account* - номер клиентского договора

*datetime* - дата изменения тарифа

*new\_tariff* - тариф, НА который перешёл клиент

*old\_tariff* - тариф, С которого перешёл клиент

**Обращения в КЦ. xlsx:**

*date* - дата обращения в контактный центр

*time* - время обращения в контактный центр

*duration* – продолжительность разговора в секундах

*account* - номер клиентского договора, если

*region* - район клиента

*theme* - тема обращения

*type* - тип обращения

КЦ – контактный центр

Задание в прилагаемом Jupiter notebook - tasks.ipynb.

**Задание №2**

Задание по NLP (Natural Language Processing). Одной из типичных задач NLP является определение тональности текста. В нашем случае даны тексты диалогов клиентов компании с сотрудниками контактного центра. В файле sentiment\_train.xlsx на листе Positive представлены позитивные диалоги, на листе Negative, соответственно, диалоги, имеющие негативную тональность. Каждый диалог состоит из фраз либо клиента (visitor), либо сотрудника контактного центра (agent). Фразы сотрудника контактного центра предполагаются априори позитивными. Конечная цель построить классификатор, который, обучившись на представленных данных, будет способен классифицировать (на позитивные и негативные) новые диалоги.

В данном задании необходимо:

1. Предложить и кратко охарактеризовать основные шаги по построению классификатора на основе представленных данных.
2. Попытаться реализовать предложенные шаги (в идеале). В качестве метрике используем обычную accuracy\_score.