**Задание 2023\_Е73. Построение прогноза рейтинга телевизионного канала (машинное обучение, нейронные сети)**

Задача состоит в построении прогноза рейтинга телевизионного канала с использованием признаков, найденных в открытых источниках. Необходимо выделить отдельные элементы программы (фильмы, ТВ-шоу, сериалы), найти по ним оценки, комментарии, список актеров. Также можно использовать дополнительные признаки (прогноз погоды, курс доллара, …), которые можно получить автоматически из открытых источников в Интернет. Использовать найденные и построенные признаки для формирования прогнозной модели. Проверить полученный результат с данными Mediascope.

Этапы решения задачи:

1. Парсинг телепрограммы за текущий месяц (например, <https://tv.yandex.ru/> ), формирование перечня программ
2. Парсинг открытых источников
   1. <https://dbpedia.org/sparql>
   2. IMDB
   3. Youtube API
   4. … (список неограничен)
3. Формирование прогнозной модели рейтинга канала на месяц.
4. Сравнение с данными <https://mediascope.net/data/> . (Регион: Россия 100+, Неделя: текущая, Целевая аудитория: население 18-54, Отчет «100 наиболее популярных программ среди россиян в возрасте от 18 до 54 лет»)

Решение требуется оформить в виде **Jupyter Notebook** <https://docs.jupyter.org/en/latest/start/index.html> .

Литература:

* <https://www.researchgate.net/publication/349594007_Forecast_Model_of_TV_Show_Rating_Based_on_Convolutional_Neural_Network>
* <https://developers.google.com/youtube/v3/quickstart/python>
* MAP@K метрика https://medium.com/@misty.mok/how-mean-average-precision-at-k-map-k-can-be-more-useful-than-other-evaluation-metrics-6881e0ee21a9

**Требования к структуре оформления решения с указанием максимального количества баллов за каждую часть решения:**

1. Введение (описание решения).

2. Основная часть:

2.1. Решение задачи парсинга телепрограммы.

2.2. Парсинг открытых источников.

2.3. Анализ методов машинного обучения и обоснование выбора для решения поставленной задачи.

2.4. Описание реализации модели машинного обучения для решения задачи, а также обоснование выбора, используемого ПО и библиотек.

2.5. Программная реализация прогнозной модели рейтинга канала на месяц (приложить файл (ы) с решением)).

2.6. Демонстрация решения.

2.7. Результаты сравнения полученного решения с данными <https://mediascope.net/data/> и их анализ.