

СИСТЕМА РЕКОМЕНДАЦИИ ВИДЕОКОНТЕНТА

Калоша А.Л.

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Стержанов М.В. – к.т.н., доцент

Аннотация. Цель данной работы заключается в создании системы для рекомендации видеоматериалов. В качестве видеоматериалов были выбраны фильмы и сериалы с сайта kinopoisk. В данной системе используется нейронная сеть, которая обучена на наборе метрик. В качестве набора метрик используется название, описание, список актеров, год выпуска. Обучение нейронной сети производилось на 10 000 фильмах. В результате обучения нейронная сеть способна рекомендовать довольно интересные подборки фильмов и сериалов.

Идея рекомендательных систем видеоматериалов не нова в современном мире. Как правило, в данных системах производится поиск людей с похожими интересами. Однако это не всегда работает, так как многим нравятся смежные жанры. Для решения данной проблемы отлично подойдет система на основе нейронной сети.

Данная система обучена для рекомендации фильмов, однако она может быть обучена для рекомендации видеоматериалов с любого веб-ресурса. Сбор данных с веб-сайтов возможен при помощи веб-краулеров. В данном исследовании для этих целей использовался Selenium WebDriver.

Система рекомендации обучается на основе оценок уже просмотренных видеоматериалов, а также по метрикам самих видео.

Для обучения нейронной сети была выбрана библиотека TensorFlow как один из лучших инструментов машинного обучения. TensorFlow — это библиотека программного обеспечения с открытым исходным кодом для численного расчета с использованием графиков потока данных [1].

Для работы системы был создан пользовательский интерфейс, на котором можно выставлять оценки и просматривать рекомендованные системой видео. Список фильмов отсортирован по рекомендациям нейронной сети, то есть фильмы на первых страницах предпочтительнее фильмам на последних. Пример пользовательского интерфейса приведен на рисунке 1.

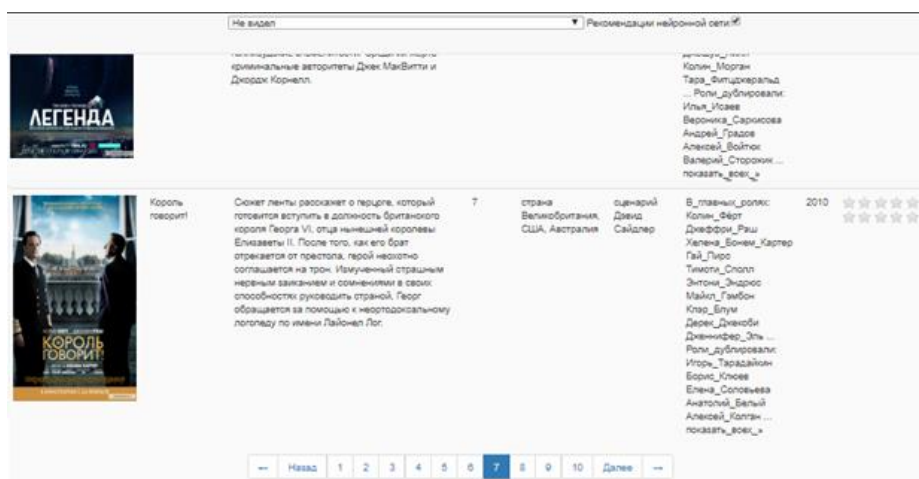


Рисунок 1 – Пример пользовательского интерфейса

Точность рекомендаций системы хорошая. Многие интересные фильмы, не оцененные в данном сервисе, были порекомендованы, а также много новых. Некоторые категории фильмов можно наблюдать в конце списка. К таким, например, относятся аниме или старые фильмы в плохом качестве. Это также свидетельствует о высокой точности обучения системы, так как при обучении фильмам из данных категорий выставлялись минимальные оценки.

Список использованных источников:

1. Library for numerical computation using data flow graphs [Электронный ресурс]. – Режим доступа: <https://www.tensorflow.org/>. – Дата доступа: 15.04.2020г.