@eel.expose

Эта строка кода указывает, что функция movie\_rec будет доступна для вызова из JavaScript с использованием библиотеки Eel.

def movie\_rec(movie\_name):

Это определение функции movie\_rec, которая принимает один аргумент movie\_name, это название фильма.

global similarity, movies\_data

Эта строка объявляет переменные similarity и movies\_data как глобальные, что позволяет им использоваться внутри функции без необходимости передавать их в качестве аргументов.

list\_of\_all\_titles = movies\_data['title'].tolist()

Эта строка создает список list\_of\_all\_titles, содержащий названия всех фильмов из столбца 'title' в данных movies\_data.

find\_close\_match = difflib.get\_close\_matches(movie\_name, list\_of\_all\_titles)

Эта строка использует модуль difflib для поиска наиболее близкого совпадения movie\_name среди названий фильмов в списке list\_of\_all\_titles.

close\_match = find\_close\_match[0]

Эта строка присваивает наиболее близкое совпадение из найденных в переменную close\_match.

index\_of\_the\_movie = movies\_data[movies\_data.title == close\_match]['index'].values[0]

Эта строка находит индекс фильма в базе данных movies\_data, соответствующий найденному наиболее близкому совпадению close\_match.

similarity\_score = list(enumerate(similarity[index\_of\_the\_movie]))

Эта строка создает список кортежей, содержащий индексы и оценки сходства данного фильма с другими фильмами.

sorted\_similar\_movies = sorted(similarity\_score, key=lambda x: x[1], reverse=True)

Эта строка сортирует список фильмов по убыванию оценки сходства.

ans = []

i = 1

Эти строки инициализируют пустой список ans и счетчик i для формирования списка рекомендаций.

for movie in sorted\_similar\_movies:

Этот оператор запускает цикл по отсортированному списку фильмов с высокой оценкой сходства.

index = movie[0]

title\_from\_index = movies\_data[movies\_data.index == index]['title'].values[0]

Эти строки извлекают название фильма из базы данных movies\_data на основе индекса.

if (i < 11):

ans.append(title\_from\_index)

i += 1

Этот блок кода добавляет название фильма в список рекомендаций ans, если количество рекомендаций не превышает 10.

return ans

Эта строка возвращает список рекомендаций пользователю.

movies\_data = pd.read\_csv('movies.csv')

Эта строка кода читает данные о фильмах из файла CSV с именем "movies.csv" и загружает их в переменную movies\_data в виде объекта DataFrame библиотеки pandas.

selected\_features = ['genres','keywords','tagline','cast','director']

Эта строка определяет список выбранных признаков, которые будут использоваться для анализа и рекомендации фильмов. В данном случае выбраны 'genres' (жанры), 'keywords' (ключевые слова), 'tagline' (слоган), 'cast' (актерский состав) и 'director' (режиссер).

for feature in selected\_features:

movies\_data[feature] = movies\_data[feature].fillna('')

Этот блок кода проходит по каждому выбранному признаку из списка selected\_features и заменяет отсутствующие значения в столбцах DataFrame movies\_data на пустые строки. Это делается с помощью метода fillna('') библиотеки pandas.

all\_features = movies\_data['genres'] + ' ' + movies\_data['keywords'] + ' ' + movies\_data['tagline'] + ' ' + movies\_data['cast'] + ' ' + movies\_data['director']

Эта строка объединяет значения выбранных признаков для каждого фильма в одну строку all\_features, разделяя каждый признак пробелом. Таким образом, для каждого фильма создается одна строка, содержащая информацию о его жанрах, ключевых словах, слогане, актерском составе и режиссере.

vectorizer = TfidfVectorizer()

feature\_vectors = vectorizer.fit\_transform(all\_features)

Этот блок кода использует класс TfidfVectorizer из библиотеки scikit-learn для преобразования текстовых данных all\_features в векторы TF-IDF. Метод fit\_transform применяется к all\_features, чтобы вычислить векторы признаков и одновременно подготовить векторайзер к кодированию новых данных. Результат сохраняется в переменной feature\_vectors.

similarity = cosine\_similarity(feature\_vectors)

Косинусное сходство

eel.init('web')

eel.start("gui.html", size = (700,700))

Подключение графического интерфейса.