|  |  |  |  |
| --- | --- | --- | --- |
| agh_nzw_s_pl_1w_wbr_rgb_150ppi  **NAZWA WYDZIAŁ ELEKTROTECHNIKI, AUTOMATYKI, INFORMATYKI I INŻYNIERII BIOMEDYCZNEJ**  Informatyka w sterowaniu i zarządzaniu  Optymalizacja wielokryterialna  *Aplikacja do rekomendacji filmów z wykorzystaniem metod optymalizacji wielokryterialnej* | | | |
| ***L.p.*** | **Członek** | **Numer albumu** | **Adres e-mail** |
| *1* | Artur Mzyk | 400658 | arturmzyk@student.agh.edu.pl |
| *2* | Joanna Nużka | 400561 | [joannanuzka@student.agh.edu.pl](mailto:joannanuzka@student.agh.edu.pl) |
| *3* | Adrian Poniatowski | 401346 | [adrianponiat@student.agh.edu.pl](mailto:adrianponiat@student.agh.edu.pl) |

Spis treści

[1. Wstęp 3](#_Toc153811445)

[1.1. Cel projektu 3](#_Toc153811446)

[1.2. Akwizycja danych 3](#_Toc153811447)

[2. Badany problem 3](#_Toc153811448)

[2.1. Opis problemu 3](#_Toc153811449)

[3. Propozycja rozwiązania 4](#_Toc153811450)

[3.1. TOPSIS 4](#_Toc153811451)

[3.2. UTA 4](#_Toc153811452)

[3.3. Metoda zbiorów referencyjnych 4](#_Toc153811453)

[4. Aplikacja 4](#_Toc153811454)

[5. Eksperymenty 5](#_Toc153811455)

[6. Podsumowanie/wnioski 5](#_Toc153811456)

[6.1. Zrealizowane punkty 5](#_Toc153811457)

[6.2. Napotkane problemy 5](#_Toc153811458)

[6.3. Kroki dalszego rozwoju 5](#_Toc153811459)

[7. Spis literatury 5](#_Toc153811460)

[8. Podział pracy 6](#_Toc153811461)

# Wstęp

## Cel projektu

Celem projektu było zbudowanie aplikacji do rekomendacji użytkownikowi filmów na podstawie jego preferencji. Wykorzystywane są poznane na zajęciach metody optymalizacji wielokryterialnej, głównie metody rankingowe, takie jak Topsis, UTA czy metoda zbiorów referencyjnych. Użytkownikowi proponowane jest dziesięć filmów najbardziej zbliżonych do jego wymagań, na które składają się:

* gatunek,
* rok produkcji,
* obsada,
* ogólna ocena użytkowników portalu.

## Akwizycja danych

Dane zaczerpnięte zostały z ogólnodostępnej bazy filmów IMDB za pośrednictwem pakietu Pythona IMDbPy. Oprócz zbioru filmów zawiera ona kluczowe informacje na ich temat i udostępnia szeroki zakres funkcji ułatwiających filtrowanie.

Dane zostały na początku poddane preprocessingowi i zweryfikowane pod względem kompletności. Interfejs graficzny pozwala na filtrowanie filmów ze względu na gatunek.

# Badany problem

## Opis problemu

Ważne jest zdefiniowanie funkcji dopasowania określającej, jak dobre jest dopasowanie danego filmu do podanych przez użytkownika preferencji. Pozwala ona ocenić film na podstawie trzech kryteriów, z których każde ma określoną skalę:

* liczba pokrywających się z preferencjami aktorów i reżyserów: 0-10,
* ocena na portalu IMDB: 0-10,
* odległość od średniego roku produkcji obejrzanych przez użytkownika filmów: 0-3 (danemu interwałowi przypisana jest odpowiednia ocena ze skali).

# Propozycja rozwiązania

## TOPSIS

Metoda TOPSIS to metoda optymalizacji wielokryterialnej bazująca na wyznaczaniu odległości badanych punktów od danego punktu idealnego i antyidealnego dla badanego problemu. Na podstawie tych odległości wyliczany jest współczynnik scoringowy służący do uszeregowania wyników. Największa wartość tego współczynnika odpowiada najlepszemu obiektowi dla danego problemu.

Zaimplementowana przez nas metoda TOPSIS składa się z opisanych dalej etapów   
i elementów. Pierwszym krokiem jest zbadanie kierunków optymalizacji dla każdego kryterium. Dla kryterium maksymalizowanego następuje odwrócenie wszystkich wartości tak, aby otrzymać problem minimalizacji. Następnie dokonywane jest odfiltrowanie punktów zdominowanych a dalsza analiza występuje tylko dla punktów niezdominowanych. Punkty niezdominowane zapisywane są w macierzy ewaluacji. Macierz ta jest normalizowana przy użyciu normy euklidesowej oraz skalowana za pomocą przemnożenia przez wektor wag. Wagi są wprowadzane przez użytkownika przy pomocy interfejsu. Następnie wyznaczany jest punkt idealny (wektor najmniejszych wartości dla każdego kryterium) oraz antyidealny (wektor największych wartości) przeskalowanej macierzy. W kolejnym kroku dla każdego   
z punktów niezdominowanych mierzone są odległości od punktu idealnego i antyidealnego przy użyciu metryki euklidesowej. Na podstawie otrzymanych odległości obliczany jest współczynnik topsis z zależności:

, gdzie:

s – współczynnik topsis

– odległość od punktu antyidealnego

– odległość od punktu idealnego.

Ostatnim krokiem jest uszeregowanie punktów niezdominowanych według malejących wartości współczynnika oraz zwrócenie wyników.

## UTA

## Metoda zbiorów referencyjnych

# Aplikacja

Wykorzystany został język wysokiego poziomu Python, gdyż pozwala na proste tworzenie interfejsów graficznych i łatwy dostęp do bazy danych filmów IMDB.

Dane pobrano, wykorzystując REST API z dedykowanej do filmów biblioteki Pythona IMDbPy, a następnie poddane je parsowaniu, co zostało przedstawione na *Rys. 1.*

Zbudowany został interfejs graficzny pozwalający na intuicyjne modyfikowanie preferencji w celu znalezienia odpowiedniego filmu. Ponadto, umożliwia on pobranie listy filmów na podstawie preferowanego gatunku. Szkic interfejsu znajduje się na *Rys. 2.*

Obraz zawierający tekst, zrzut ekranu, oprogramowanie

Opis wygenerowany automatycznie

Rys. 1. Skrypt do pobrania i parsowania danych

Obraz zawierający tekst, zrzut ekranu, Prostokąt, design

Opis wygenerowany automatycznie

Rys. 2. Interfejs graficzny

# Eksperymenty

# Podsumowanie/wnioski

## 6.1. Zrealizowane punkty

## 6.2. Napotkane problemy

## 6.3. Kroki dalszego rozwoju

# Spis literatury

<https://www.sciencedirect.com/science/article/pii/S187705091631273X>

<https://www.researchgate.net/publication/236153959_Decision_Support_Systems_Based_on_Reference_Sets>

<https://imdbpy.readthedocs.io/en/latest/>

# Podział pracy

|  |  |  |  |
| --- | --- | --- | --- |
| **Funkcjonalność** | **Artur [wkład w %]** | **Joanna [wkład w %]** | **Adrian [wkład w %]** |
| Preprocessing danych | - | 50 | 50 |
| Wstępne GUI do pobierania filmów | 100 | - | - |