



Politechnika  
Wrocławska

# Systemy rekomendacyjne

Projekt: semestr letni 2020/21

Mateusz Nurek



# Cele przedmiotu

1. Zrozumienie celów i zadań systemów rekomendacyjnych i personalizacji
2. Nabycie umiejętności zaprojektowania systemu rekomendacyjnego
3. Nabycie umiejętności pracy w grupie



# Sposób pracy

Cel:

**zaprojektowanie i implementacja proof of concept systemu  
rekomendacyjnego na koniec semestru**

- praca w 3-4 os. grupach
- ok. 60h w semestrze pracy własnej, tj. ok. 180-240h pracy na standardową grupę 3 osobową
- ok. 4h pracy własnej tygodniowo, 12-16h per grupę



# Ocenianie

W ramach zaliczenia, oceniane będą:

1. pierwsze dwie prezentacje związane z metodami rekomendacji
2. wykonany PoC systemu rekomendacyjnego
  - a. jakość projektu:
    - i. jasne zdefiniowanie potrzeb użytkowników,
    - ii. wybór adekwatnych zmiennych wejściowych do potrzeb użytkowników,
    - iii. dobór odpowiedniego modelu rekomendującego
  - b. repo githubowe z kodem, (opcjonalnie dvc jako pipeline do przetwarzania danych)
  - c. chodzący PoC - pozwalający wgrać dane i uzyskać rekomendacje, np. z użyciem streamlit
3. zdolność do pracy w grupie - realizacja celu nr 3

Każda z grup będzie oceniać w tym zakresie pozostałe i samą siebie. Na podstawie tych ocen prowadzący wystawia ocenę ostateczną, uwzględniając swoją ocenę.

Możliwa **jedna nieusprawiedliwiona nieobecność**, każda kolejna to pół oceny w dół.



# Plan kursu (TN)

1. Wprowadzenie **(3.03)**
2. Zapoznanie się z podstawowymi metodami tworzenia systemów rekomendacyjnych **(10.03)**
3. Zapoznanie się z zaawansowanymi metodami tworzenia systemów rekomendacyjnych **(24.03)**
4. Wybór rodzaju systemu rekomendacyjnego jaki chcecie zbudować, określenie grupy odbiorców, definiowanie ich punktów widzenia. Wybór na które z nich będzie odpowiadał wasz system, wybór zbiorów danych do wykorzystania lub określenie strategii pozyskania danych. **(7.04)**
5. Zbieranie i eksploracja danych, ewentualna redefinicja założeń z pkt. 3 **(21.04)**
6. Prototypowanie interfejsu użytkownika PoC, ewentualna redefinicja założeń z pkt. 3 **(5.05)**

— redefinicja nie jest możliwa od tego punktu —

7. Przetwarzanie danych i trenowanie modelu **(19.05)**
8. Ostatnie korekty, wpięcie modelu do PoC, prezentacja i ocena **(9.06)**



# Plan kursu (TP)

1. Wprowadzenie **(3.03)**
2. Zapoznanie się z podstawowymi metodami tworzenia systemów rekomendacyjnych **(17.03)**
3. Zapoznanie się z zaawansowanymi metodami tworzenia systemów rekomendacyjnych **(31.03)**
4. Wybór rodzaju systemu rekomendacyjnego jaki chcecie zbudować, określenie grupy odbiorców, definiowanie ich punktów widzenia. Wybór na które z nich będzie odpowiadał wasz system, wybór zbiorów danych do wykorzystania lub określenie strategii pozyskania danych. **(14.04)**
5. Zbieranie i eksploracja danych, ewentualna redefinicja założeń z pkt. 3 **(28.04)**
6. Prototypowanie interfejsu użytkownika PoC, ewentualna redefinicja założeń z pkt. 3 **(12.05)**

— redefinicja nie jest możliwa od tego punktu —

7. Przetwarzanie danych i trenowanie modelu **(26.05)**
8. Ostatnie korekty, wpięcie modelu do PoC, prezentacja i ocena **(16.06)**



# Z1: Podstawowe zagadnienia systemów rekomendacyjnych

Tematyka najbliższego tygodnia:

- collaborative filtering
- content-based filtering
- demographic recommendation
- hybrid approaches
- K-nearest neighbours classification
- porównanie miar podobieństwa

Wymagany czas: ok. 4h / os. pracy własnej:

1. czytanie 1h
2. przykłady kodu, zapoznanie się z istniejącymi projektami 1h
3. spotkanie/call grupowy - wymiana wiedzy 20 min
4. implementacja 1h
5. spotkanie / call grupowy - przygotowanie prezentacji 40 minut



# Lektura [1h]

1. [Jakie są rodzaje systemów rekomendacyjnych?](#)
2. [Jaki algorytmów się używa w systemach rekomendacyjnych?](#)
3. [Skąd się bierze rekomendacja?](#)
4. [Inny tekst o metodach rekomendacji](#)
5. [Przegląd hybrydowych systemów rekomendacyjnych](#)
6. [Wprowadzenie do systemów rekomendacyjnych](#)

Podzielcie się tekstami, poczytajcie i przygotujcie się do opowiedzenia innym członkom grupy o tym czego nauczyliście się z tekstu.





# Przykłady kodu [1h]

- <https://github.com/topics/content-based-recommendation>
- <https://github.com/topics/collaborative-filtering>
- <https://realpython.com/build-recommendation-engine-collaborative-filtering/>
- <https://github.com/topics/recommender-system>

Przejrzyjcie repozytoria, wiele z nich ma przykłady gotowych systemów rekomendacyjnych, poczytajcie kod.



# Spotkanie / call grupowy [20min]

Ustalcie termin spotkania i porozmawiajcie o tekstach, opowiedzcie sobie nawzajem:

- czego się nauczyliście nowego, czego nie wiedzieliście,
- co uważacie za podstawową wiedzę z tego co przeczytaliście, a co było zaskakujące,
- jakie repozytoria zwróciły wasza uwagę i co w nich było ciekawego.

Niech za każdym razem jedna osoba notuje na slajdach ważne punkty.



# Implementacja [1h]

**Korzystając z przykładów kodu** zaimplementujcie i uruchomcie prosty system rekomendacyjny na tym samym zbiorze danych, w oparciu o minimum 3 z poniższych technik (po 1 na os.):

- KNN i miary podobieństwa
- collaborative filtering
- content-based filtering
- demographic recommendation
- podejścia hybrydowe



# Spotkanie / call grupowy [40min]

Na ostatnim spotkaniu:

- porównajcie uzyskane rekomendacje, jak to się ma do waszej intuicji,
- jak się mają uzyskane rekomendacje do potencjalnych potrzeb użytkowników?
- omówcie: **wady i zalety** poszczególnych podejść, **jak je walidować**, do czego się nadają,
- ponownie niech ktoś notuje najważniejsze punkty,
- zebrane slajdy z notatkami z tego i poprzedniego spotkania ułóżcie w krótką 5-7 minutową prezentację na zajęcia

Wybierzcie 1 osobę, która prezentuje.



Politechnika  
Wrocławska

# Ocena prezentacji

Składowe:

- złożoność podjętego tematu
- uzyskane wyniki
- analiza wad i zalet prezentowanego rozwiązania
- analiza walidacji rozwiązania
- projekt na githubie

Każda grupa ocenia siebie i inne grupy.



## Z2: Zaawansowane zagadnienia systemów rekomendacyjnych

Rekomendacja z użyciem:

- partitioning clustering, np. K-means
- hierarchical clustering
- reguły asocjacyjne (frequent items set), zależność między produktami
- złożoność akcji użytkowników
  - streaming - kolejność oglądania
  - różne typy działań - oglądał, kupił, włożył do koszyka, ocenił
- sekwencyjność, sequential patterns, sezonowość
- embedding
- deep learning
- context based recommendation

Wymagany czas: ok. 4h / os. pracy własnej:

1. czytanie 1h
2. przykłady kodu, zapoznanie się z istniejącymi projektami 1h
3. spotkanie/call grupowy - wymiana wiedzy 20 min
4. implementacja 1h
5. spotkanie / call grupowy - przygotowanie prezentacji 40 minut



# Prezentacja końcowa

10-12 min na prezentację, 5-7min na dyskusję

Proponowane slajdy:

- motywacja i cel
- **krótkie** przedstawienie danych
- omówienie wybranych modeli
- omówienie zalet i wad rozwiązania
- wnioski



Politechnika  
Wrocławska

# Ocena projektu

Składowe:

- złożoność podjętego tematu
- uzyskane wyniki
- analiza wad i zalet zastosowanych rozwiązań
- analiza walidacji modelu/modeli
- projekt na githubie / strona www

Każda grupa ocenia siebie i inne grupy.





Politechnika  
Wrocławska

# Ocena końcowa

$$Z1 * 20\% + Z2 * 20\% + \text{Projekt} * 60\%$$



Politechnika  
Wrocławska

**Dziękuję, do zobaczenia!**