

Zastosowanie spersonalizowanych systemów rekomendacyjnych jako potencjał dla badań socjologicznych – na przykładzie portalu filmowego

W swoim referacie postaram się zaprezentować koncepcję wykorzystania prototypowego portalu filmowego jako narzędzia badawczego. To, co wyróżnia taką metodę od konwencjonalnych metod ilościowych, to wielkość uzyskanej próby.

„W stronę praktyki” – to hasło przewodnie tegorocznej konferencji. Kwestia algorytmów rekomendacyjnych i wątek uczenia się maszyn to dziedziny, które są nierozłącznie związane z pragmatycznym wymiarem nauki. Wśród wielu współczesnych zastosowań sztucznej inteligencji, takich jak: maszynowe tłumaczenie tekstów, technologie oparte na logice rozmytej, rozpoznawanie obrazów, mowy lub pisma, istnieją także systemy ekstrapolacji danych oraz mechanizmy uczenia się, które stanowią podstawę dla współczesnej koncepcji personalizacji wyszukiwania, opartej na systemach rekomendacji.

Stosowanie algorytmów rekomendujących staje się dziś praktyką coraz bardziej powszechną. Na przestrzeni ostatnich paru lat mogliśmy zaobserwować ich gwałtowną popularyzację. Są one dziś potrzebne wielu firmom po to, by w miarę możliwości z jak najwyższą trafnością określić gusta swoich użytkowników, którzy mają się stać częstymi bywalcami serwisu lub jego klientami.

Serwis Amazon.com był pionierem zautomatyzowanych rekomendacji, ale współcześnie stosują je już takie serwisy jak YouTube, Apple, eBay czy rodzime Allegro. W swoim wywodzie pominię koncepcję *social graph* zaproponowaną przez Marka Zuckerberga, gdyż jest to temat zbyt obszerny.

Od października 2009 roku personalizacji podlega także wyszukiwanie poprzez najpopularniejszą światową wyszukiwarkę Google. Dla zalogowanego użytkownika wyniki wyszukiwania są modyfikowane na podstawie historii zapytań z ostatnich 180 dni. Kolejnym uwzględnianym przy personalizacji wyników czynnikiem jest geolokalizacja użytkownika. Informacja o domniemanym bieżącym położeniu geograficznym wyszukującego może zostać wykorzystana przy pewnych zapytaniach, np.: poszukując gabinetu lekarskiego lub specjalistycznego sklepu, wyszukiwarka w pierwszej kolejności zaproponuje obiekty znajdujące się w najbliższej okolicy.

Algorytm wykorzystywany przez wyszukiwarkę Google, mimo licznych usprawnień, umożliwia poważne manipulacje prezentowanymi wynikami. Osoby zorientowane w jego działaniu generują ogromne ilości komercyjnego spamu, który wyświet-

lany jest często przed wartościowymi informacjami. W rezultacie wyszukiwanie staje się coraz mniej efektywne. Wprowadzenie personalizacji nie jest krokiem rewolucyjnym, ale i tak zdaje się krokiem w dobrym kierunku.

Mechanizm rekomendacji nadal jednak kojarzy się nam chyba w pierwszym rzędzie ze sklepami internetowymi, gdzie taki drobiazg wydaje się najbardziej oczywisty. Jak pisze Marcin Majda [2010] w artykule *Sposoby implementacji systemów rekomendacyjnych*: „Algorytmy rekomendacji dają bardzo dobre efekty biznesowe. Greg Linden, autor systemu rekomendacyjnego Amazon.com, podaje na swoim blogu, że aktualnie w sklepie 35% sprzedanych pozycji pochodzi z algorytmów rekomendacyjnych”. Skalę przemian obrazuje fakt, że jeśli wierzyć obliczeniom, Amazon.com posiada obecnie około 24 milionów produktów, podczas gdy w 1994 roku w całych Stanach Zjednoczonych dostępnych było około 500 tysięcy różnych produktów [Grossman 2010]. Wracając jednak do efektów biznesowych rekomendacji: „Jeszcze większe współczynniki podaje Netflix, największa wypożyczalnia filmów DVD w USA. Rekomendacje są dla nich kluczowym elementem, gdyż odpowiadają za 60% sprzedaży” [Majda 2010].

To właśnie kwestia rekomendacji filmowych zyskała w ostatnich latach spory rozgłos za sprawą serwisu Netflix. W 2009 roku ogłosił on konkurs, którego celem było polepszenie o 10% trafności opracowywanego przez nich algorytmu zwanego Cinematch. Nagroda za to osiągnięcie wynosiła milion dolarów [Romański 2008].

Ale serwis Netflix to nie jedyny portal starający się pozyskać użytkownika poprzez rekomendacje. Trzy najciekawsze przykłady stron filmowych zaopatrzone w systemy rekomendacyjne to:

Jinni.com – wyszukiwarka i system rekomendujący dla filmów, seriali telewizyjnych i krótkich filmików. Mechanika serwisu bazuje na projekcie zwanym *Movie Genome*, który opiera się na niestandardowym podejściu do indeksowania filmów, wyróżniającym takie atrybuty jak: nastrój, fabuła, struktura historii. Technologia Jinni.com opiera się na taksonomii ustalonej przez profesjonalistów, w której nowe tytuły są automatycznie indeksowane przez boty analizujące recenzje i metadane na ich temat. Dzięki temu rozwiązaniu Jinni.com umożliwia oprócz klasycznego także wyszukiwanie semantyczne.

MovieLens.org – system rekomendujący typu *collaborative filtering* bazujący na preferencjach użytkownika. System porównuje użytkowników o podobnych gustach. Oceny filmów zaczerpnięte od „sąsiadów” są używane w kreowaniu spersonalizowanych rekomendacji. Projekt *GroupLens Research* powstał jako platforma projektu akademickiego na Wydziale Informatyki i Inżynierii Uniwersytetu Minnesota w 1997 roku. Do dziś platforma stanowi środowisko dla wielu interdyscyplinarnych badań prowadzonych przez studentów i pracowników uniwersytetu.

Criticker.com – Ciekawym rozwiązaniem zastosowanym w tym portalu jest wprowadzanie skali ocen od 1 do 100 oraz sposób, w jaki oceny są ważone. Poszczególne zakresy ocen przypisywane są do 10 poziomów tzw. *tier*. Pozwala to na wartościowanie ocen użytkowników, którzy stosują własne kryteria, np. przyjmują skalę od 1 do 10 lub przejawiają tendencje do wystawiania wyłącznie wysokich lub niskich ocen. W teorii umożliwia to skuteczniejsze porównywanie ocen różnych użytkowników.

Inne portale filmowe zaopatrzone w algorytm rekomendacji to: IMDb, Rotten Tomatoes, Flixster, Netflix, a w Polsce – od września 2010 – istniejący od 1998 roku

Filmweb oraz znacznie od niego młodsze: Filmster, fdb.pl, Szufler.pl. Przykład Netflix (który gotów był poświęcić milion dolarów na to, by zoptymalizować algorytm o 10%) pokazuje, że optymalizacja algorytmu staje się coraz trudniejszym zadaniem.

Pytanie, dlaczego nasza ocena (choć nierzadko moglibyśmy przysiąc o jej obiektywności) podlega, jak wiemy z codziennego doświadczenia, tak dużym wahaniom? Chris Volinsky – jeden z programistów zespołu, który wygrał główną nagrodę w konkursie Netflix – stwierdza, że „bardzo wiele czynników losowych zamieszanych jest w proces rekomendacji”, i dodaje, że „jest pewien nieodłączny poziom błędu tkwiący we wszelkich próbach przewidzenia ludzkich zachowań” [Grossman 2010].

Z punktu widzenia neurologii „pobudzona wrażeniami estetycznymi uwaga poznawcza aktywuje układ dopaminergiczny” [Vetulani 2010], gdy coś zaspokoi nasze estetyczne potrzeby. Jednak sprawa jest bardziej skomplikowana, ponieważ jak twierdzi Włodzisław Duch: „Świadome odczucia są stanami mózgu, które mogą powstać na skutek zewnętrznych albo też czysto *wewnętrznych* wzajemnych pobudzeń struktur neuronowych mózgu” [Duch 2010], co implikuje, że ocena obiektywna obiektu jako taka w ogóle nie istnieje. Jest to twierdzenie, które od czasu wydania w 1764 roku przez Immanuela Kanta rozprawy *O uczuciu piękności i wzniosłości* nikogo nie powinno szokować, jednak nasza intuicja błędnie podpowiada nam, że obiektywizm jest wartością niezaprzeczalną. Profesor Jerzy Vetulani w swojej najnowszej książce *Mózg: fascynacje, problemy, tajemnice* tak podsumowuje ten temat: „Filozofowie od stuleci wiedli spór na ten temat. Od starożytności do czasów oświecenia uważano za Platona, że piękno jest zawarte w przedmiocie, niezależnie od tego, kto je ocenia” [Vetulani 2010].

Józef Bremer w jednym ze swoich artykułów zauważa niezwykle trafnie, że „Zależność pomiędzy emocjami a poznaniem widziano najczęściej poprzez przyzmat zaburzeń poznania (...)” [Bremer 2010], i być może to jest odpowiedź na pytanie, dlaczego tak długo temat naszej oceny aksjologicznej zepchnięty był na margines naukowego dyskursu. „W kognitywistyce odchodzi się od tego spojrzenia, pytając o pozytywne wpływy uczuć na procesy poznawcze oraz o wpływy procesów poznawczych na uczucia” – zauważa profesor Bremer. Dlatego właśnie ujęcie kognitywne okazuje się przydatne w kontekście oceny aksjologicznej [Bremer 2010].

Z punktu widzenia praktycznego nasz główny problem polega na tym, że nie jesteśmy dobrzy w wyrażaniu swoich preferencji w formie podlegającej skalowaniu. Inaczej oceniamy rzeczy po ciężkim dniu pracy, inaczej w sytuacji relaksu (Netflix). Poza tym różnimy się podejściem do samego oceniania. Niektórzy starają się być powściągliwi w swojej ocenie, inni z kolei zawyżają swą skalę ocen (Criticster.com). Kolejnym czynnikiem jest kwestia, czy dany film oceniamy zaraz po jego obejrzeniu, czy też może po paru dniach bądź miesiącach (Netflix). Jeszcze inaczej oceniamy będziemy ten sam film obejrzany w różnych warunkach. I na ten właśnie aspekt oceny chciałbym zwrócić uwagę, ponieważ w odróżnieniu od powyższych czynników jest on stosunkowo mierzalnym modyfikatorem. Wymaga jednak pozyskania danych, które niekoniecznie w sposób oczywisty wiążą się z ocenianym filmem.

W moim odczuciu problem optymalizacji systemu Netflix pokazuje, iż większe zróżnicowanie jakościowe danych dostępnych dla algorytmu jest kolejnym niezbędnym

nym krokiem w rozwijaniu mechanizmów rekomendacji. Proponowane przeze mnie badanie porusza w tym względzie cztery modyfikatory oceny. Są to:

- I – czynniki indywidualne,
- II – kontekst grupy,
- III – jakość oglądanej kopii,
- IV – stan wiedzy osoby oceniającej.

I – czynniki indywidualne. Po to by lepiej określić preferencje użytkownika, wymagane byłoby wprowadzenie obowiązkowego formularza uwzględniającego takie czynniki indywidualne jak: wiek, płeć, status społeczny, grupa zawodowa itd.

II – kontekst grupy. O ile pierwszy postulat nie jest żadną innowacją, o tyle wprowadzenie kontekstu sytuacyjnego jest koniecznością, aby całą bazę można uznać za aparaturę badawczą. Przez kontekst grupy rozumiem *warunki, w jakich dany film był przez użytkownika oceniany*. Proponuję w tym miejscu konstrukcję prostego interfejsu graficznego, który za pomocą pięciu czytelnych ikon pozwoliłby na szybkie i wygodne zdefiniowanie kontekstu. Ikona pierwsza reprezentowana przez jedną postać symbolizowałaby, że użytkownik: 1) obejrzał dany film w pojedynkę, analogicznie kolejne specjalnie zaprojektowane ikony symbolizowałyby: oglądanie filmu: 2) w parze mieszanej, 3) w grupie do trzech osób, 4) na imprezie (powyżej trzech osób), 5) w kinie.

Jest to koncepcja oparta na twierdzeniu, że jak to zwięźle ujął Philip Zimbardo: „Nasze oczekiwania prowadzić mogą do takich zachowań werbalnych i niewerbalnych, które prowokują innych do zachowań, jakich od nich oczekujemy, przy czym żadna ze stron nie zdaje sobie sprawy z zachodzącego procesu” [Zimbardo 1999].

Rozwiązanie to nie ma obecnie zastosowania w żadnym ze znanych portali filmowych. Warto jednak w tym miejscu zanotować, że niedawno Mariusz Kryński napisał testową aplikację dla portalu Filmaster, po podaniu listy znajomych generującą listę filmów, których użytkownicy wspólnie nie widzieli, a które powinny spodobać się całej grupie [blog.filmaster.pl]. Celem wprowadzenia powyższej innowacji z punktu widzenia możliwości badawczych jest perspektywa poszukiwania odpowiedzi na pytanie, jak ściśle procesy poznawcze jednostki są skorelowane z kontekstem sytuacji.

III – jakość oglądanej kopii. Modyfikator oparty na jakości ocenianego materiału nie jest może czymś nowym, ale podobnie jak powyższa koncepcja nie znajduje odzwierciedlenia w portalach filmowych. Kwestie takie jak jakość dźwięku, głośność dźwięku czy przekątna ekranu nie są tak łatwe do sprawdzenia w sposób automatyczny jak sama jakość oglądanego materiału. Tą ostatnią kwestią mógłby się zająć bot, działający podobnie jak scrobbler Last.fm, który po odtworzeniu utworu muzycznego zapisuje ten fakt na koncie użytkownika w internetowej bazie serwisu.

IV – stan wiedzy osoby oceniającej. Tę kwestię można by rozwinąć opcjonalnie. Należałoby dać użytkownikom możliwość obejrzenia krótkiej prelekcji na temat filmu tuż przed jego odtworzeniem. Statystyczne zestawienie wyników z grupą kontrolną, która nie miałaby okazji obejrzeć prelekcji, mogłoby dać wartościowe wyniki.

System wymienionych powyżej modyfikatorów należałoby sprzęgnąć z modelem oceny zaproponowanym przez Byrona Reevesa i Clifforda Nassa w książce *Media i ludzie* [2005, s. 157]. Model tradycyjny, wynikający w prostej linii z recenzji prasowej, zakłada pewną obiektywną ocenę recenzenta, wyrażoną w stałej skali, która

najczęściej wizualizowana jest przez gwiazdki i tym podobne graficzne elementy. Podejście takie wydaje się anachroniczne w dobie powszechnego dostępu do Internetu. Patrząc na problem oceny z punktu widzenia współczesnych możliwości, dochodzimy do wniosku, że nie ma żadnej konieczności, by użytkownik nie oceniał filmów w sposób subiektywny. Także sposób oceniania na skali rozpiętej w określonym przedziale stracił swoją aktualność.

Autorzy publikacji *Media i ludzie* zwracają uwagę, że:

Szybkie podjęcie decyzji co do dobrej bądź złej natury jakiegoś obcego lub środowiska stanowi niezbędne przygotowanie najbardziej podstawowej z reakcji – unikania. Chociaż w słowniku wyrazów bliskoznacznych istnieją setki słów, które mogą poszerzyć lub sprecyzować ów sąd, droga do nich prowadzi przez prostą ocenę – dobry lub zły. Ocena ta może potem wpłynąć na to, jak myślimy, na co zwracamy uwagę i co robimy. [...] Dwie półkule ludzkiego mózgu rozwinęły się w wyspecjalizowane procesory pozytywnych i negatywnych informacji. Doświadczenia pozytywne łączą się z lewą półkulą, a negatywne – z prawą. [...] Istnieją przekonujące dowody na to, że struktura mózgu warunkuje zasadnicze rozróżnienie pomiędzy tym, co dobre, a tym, co złe [Reeves, Nass 2005, s. 134–135].

Zdaniem Nassa i Reevesa: „Lepszy system oceny mógłby zawierać dwa obrazki: jeden mógłby uśmiechem lub skrzywieniem wykazywać, czy film jest dobry, czy zły. Drugi obrazek mógłby pokazywać poziom entuzjazmu, z jakim dokonujemy oceny [...]”. W ich rozumieniu „Ten ostatni rysunek ilustruje drugi aspekt emocji: pobudzenie” [Reeves, Nass 2005, s. 157]. Zatem „pobudzenie i wartościowanie są to dwie podstawowe cechy emocji” [Reeves, Nass 2005, s. 158].

Zarysowaną w artykule platformę należałoby rozwijać przez dłuższy czas jako *interdyscyplinarny międzyinstytutowy projekt akademicki*. Dane pozyskane przy wprowadzeniu zaproponowanych innowacji mogłyby się okazać pomocne przy wyprowadzaniu nowych wniosków w dziedzinie ludzkiego postrzegania.

Szacując okres, w jakim można by założyć dostatecznie duży przyrost użytkowników, aby móc opublikować pierwsze wyniki badań, możemy odnieść się do entuzjazmu, z jakim spotkał się polski portal Filmaster.pl. Jeśli nawet byłby to przyrost o połowę mniejszy niż w wypadku Filmastera, już po upływie roku użytkownicy stanowiliby dostatecznie reprezentacyjną grupę, by móc doszukiwać się określonych prototypowych wzorców. Opierając się na przykładzie portalu Filmaster.pl, możemy zaobserwować, że przyrost liczby ocen jest bardzo duży, nawet przy relatywnie niewielkiej liczbie użytkowników. Portal ten wystartował 18 stycznia 2009 roku, a wersja angielska serwisu została uruchomiona w maju 2009 roku.

Statystyki portalu Filmaster.pl [blog.filmaster.pl] przedstawiają się następująco: **W pierwszym miesiącu działalności:** profile filmów odwiedzono w lutym 33 tysiące razy; liczba wystawionych w tym czasie ocen wyniosła ponad 50 tysięcy.

Po pół roku działalności:

Tabela 1

Liczba użytkowników	2528
Liczba filmów	6053
Liczba ocen do uczenia	237321
Liczba ocen w zbiorze testowym	58415

Źródło: dane testowe z portalu Filmaster.pl za: Tłałka [2010].

Po roku działalności:

- w serwisie zarejestrowało się ponad 4,5 tysiąca użytkowników;
- użytkownicy ocenili pół miliona pozycji (średnio ponad 100 filmów na użytkownika);
- przez rok stronę odwiedziło 358 tysięcy osób, generując 1,7 mln odsłon stron.

Przy pięciu tysiącach użytkowników, z których średnio każdy ocenia ponad 100 filmów miesięcznie, zebrany materiał badawczy, jak nietrudno policzyć, będzie miał już objętość ponad pół miliona zwaloryzowanych danych. Badanie na tak dużej próbie zwiększa jego wartość, uwierzytelniając wyniki i pozwalając na odkrycie pewnych niezauważalnych wzorów. Taka skala nie jest często spotykana przy tradycyjnych metodach pomiarów w przypadku eksperymentów z zakresu psychologii społecznej.

Problemem związanym z działalnością proponowanej platformy badawczej projektu akademickiego może się okazać rozmiar zgromadzonych danych. Twórcy portalu Filmaster.pl szacują, że pół miliona ocen generuje w ich przypadku około 8,5 miliona rekomendacji. W przypadku zwiększonej komplikacji algorytmu, związanej z większą liczbą modyfikatorów, kompilacja algorytmu rekomendacyjnego może się okazać nie lada wyzwaniem.

Mam nadzieję, że tym wywodem wykazałem, iż po pierwsze, kwestia systemów rekomendacji stanowi istotny aspekt w praktycznym zastosowaniu algorytmów sztucznej inteligencji we współczesnych technologiach informatycznych. Po drugie, wyrażam nadzieję, że zaprezentowane przeze mnie koncepcje innowacji staną się inspiracją kierującą w jakimś stopniu uwarunkowanie środowiska akademickiego względem perspektyw roztaczających się przed nowymi technologiami.

BIBLIOGRAFIA

- Bremer J. (2010). *Kilka słów o kognitywistyce*, <http://www.kognitywistyka.iphils.uj.edu.pl/index.php?section=article2> (dostęp: 14.01.2010).
- Duch W. (2010), *Neuroestetyka i ewolucyjne podstawy przeżyć estetycznych*, <http://www.fizyka.umk.pl/ftp/papers/kmk/07-Neuroestetyka.pdf> (dostęp: 14.01.2010).

- Grossman L. (2010). *How Computers Know What We Want – before We Do*. „Time Magazine”, maj.
<http://blog.filmaster.pl/> (dostęp: 14.01.2010).
<http://movielens.org/> (dostęp: 14.01.2010).
<http://netflixprize.com/> (dostęp: 14.01.2010).
- Majda M. (2010). *Sposoby implementacji systemów rekomendacyjnych*. „Software Developer’s Journal” 11.
- Reeves B., Nass C. (2005). *Media i ludzie*. Tłum. H. Szczerkowska. Warszawa: Państwowy Instytut Wydawniczy.
- Romański B. (2008). *Porównanie metod przewidywania preferencji użytkowników na przykładzie konkursu Netflix Prize*. Praca licencjacka, Uniwersytet Warszawski, Wydział Matematyki, Informatyki i Mechaniki.
- Tłałka J. (2010). Transkrypcja prezentacji *Algorytmy rekomendacji – Filmaster*. Spotkanie miłośników wolnego oprogramowania, Jesień Linuksowa 2010, Ustroń Jaszowice, 25.09.2010.
- Vetulani J. (2010). *Mózg: fascynacje, problemy, tajemnice*. Kraków: Wydawnictwo Homini.
- Zimbardo Ph.G. (1999). *Psychologia i życie*. Tłum. J. Radzicki. Warszawa: Wydawnictwo Naukowe PWN.

Personalised Recommendation Systems as Potential for Sociological Research in Case of Movie Database

In my article I would try to present concept of using film portal as a research tool. What distinguishes this method from conventional quantitative methods is the size of the resulting sample. The problem of assessing the film is necessity for evaluating it using a specific scale, to which we, as a human beings, are not fittest. Our judgments may vary depending on the conditions under which the film is viewed. It is this aspect of the evaluation, I would like to draw attention. Proposed study will takes account of four assessment modifiers, such as: individual factors, II context of group III quality of the image, IV evaluators state of knowledge.