

KARTA OPISU MODUŁU KSZTAŁCENIA		
Nazwa modułu/przedmiotu Systemy rekomendacyjne		Kod 1010331571010337135
Kierunek studiów Informatyka	Profil kształcenia (ogólnoakademicki, praktyczny) (brak)	Rok / Semestr 4 / 7
Ścieżka obieralności/specjalność Technologie informatyczne	Przedmiot oferowany w języku: polski	Kurs (obligatoryjny/obieralny) obligatoryjny
Stopień studiów: I stopień	Forma studiów (stacjonarna/niestacjonarna) stacjonarna	
Godziny Wykłady: 15 Ćwiczenia: - Laboratoria: - Projekty/seminaria: 15		Liczba punktów 3
Status przedmiotu w programie studiów (podstawowy, kierunkowy, inny) (brak)		(ogólnouczelniany, z innego kierunku) (brak)
Obszar(y) kształcenia i dziedzina(y) nauki i sztuki nauki techniczne nauki techniczne		Podział ECTS (liczba i %) 3 100% 3 100%
Odpowiedzialny za przedmiot / wykładowca: dr inż. Andrzej Szwabe email: Andrzej.Szwabe@put.poznan.pl tel. 61 665 3958 Wydział Elektryczny ul. Piotrowo 3A 60-965 Poznań		
Wymagania wstępne w zakresie wiedzy, umiejętności, kompetencji społecznych:		
1	Wiedza:	K_W07: ma uporządkowaną i podbudowaną teoretycznie wiedzę w zakresie technologii sieciowych K_W05: ma uporządkowaną i podbudowaną teoretycznie wiedzę w zakresie podst. konstrukcji programistycznych, implementacji algorytmów, paradygmatów i stylów programowania, metod weryfikacji poprawności programów, języków formalnych, kompilatorów, platform K_W08: ma uporządkowaną i podbudowaną teoretycznie wiedzę w zakresie baz danych oraz hurtowni danych
2	Umiejętności:	K_U01: potrafi pozyskiwać informacje z literatury, baz danych i innych źródeł; potrafi integrować uzyskane informacje, dokonywać ich interpretacji, a także wyciągać wnioski oraz formułować i uzasadniać opinie K_U03: potrafi opracować dokumentację dotyczącą realizacji zadania inżynierskiego i przygotować tekst zawierający omówienie wyników realizacji tego zadania
3	Kompetencje społeczne	K_K02: ma świadomość ważności i rozumie pozatechniczne aspekty i skutki działalności inżyniera-informatyka i związaną z tym odpowiedzialność za podejmowane decyzje
Cel przedmiotu: Głównym celem przedmiotu jest zapoznanie z najważniejszymi technologiami współcześnie stosowanych systemów rekomendacyjnych. W celu ułatwienia zaprojektowania i implementacji własnego systemu rekomendacyjnego, system taki jest przedstawiony studentowi jako szczególny typ systemu uczącego się (ang. Machine Learning system). Prezentowane jest, w jaki sposób biblioteki programistyczne i pakiety oprogramowania wykorzystywane do budowy systemów uczących się mogą być wykorzystane do sprawnego stworzenia nowoczesnego, efektywnego i skalowalnego systemu rekomendacyjnego. W celu dostarczenia umiejętności obiektywnej oceny systemu rekomendacyjnego i przyszłego prowadzenia przez studenta własnych badań, przedstawiona jest standardowa metoda oceny (powszechnie stosowana w środowisku akademickim i w przemyśle) wraz z najważniejszymi miarami jakości rekomendacji i ich przykładowymi implementacjami przygotowanymi z użyciem wiodących bibliotek oprogramowania. Ważnym celem kursu jest przedstawienie studentom tych zagadnień związanych z systemami rekomendacyjnymi, które mają kluczowe znaczenie z perspektywy udanych komercyjnie wdrożeń - w szczególności w inteligentnych internetowych systemach reklamowych i w systemach automatycznej oceny bankowych wniosków kredytowych.		
Efekty kształcenia i odniesienie do kierunkowych efektów kształcenia		
Wiedza:		

1. ma uporządkowaną i podbudowaną teoretycznie wiedzę w zakresie sztucznej inteligencji, systemów eksperckich i agentowych - [K_W09]
2. ma uporządkowaną i podbudowaną teoretycznie wiedzę w zakresie technologii internetowych - [K_W11]
3. ma uporządkowaną i podbudowaną teoretycznie wiedzę w zakresie podstaw teleinformatyki oraz protokołów i usług w sieciach telekomunikacyjnych - [K_W15]
Umiejętności:
1. potrafi opracować dokumentację dotyczącą realizacji zadania inżynierskiego i przygotować tekst zawierający omówienie wyników realizacji tego zadania - [K_U03]
2. potrafi realizować podstawowe zadania dotyczące portali i usług internetowych - [K_U15]
3. potrafi zaprojektować oraz zrealizować prosty system ekspercki lub agentowy - [K_U13]
Kompetencje społeczne:
1. ma świadomość ważności i rozumie pozatechniczne aspekty i skutki działalności inżyniera-informatyka i związaną z tym odpowiedzialność za podejmowane decyzje - [K_K02]

Sposoby sprawdzenia efektów kształcenia
Wykład: egzamin pisemny (sprawdzenie wiedzy teoretycznej i umiejętności opracowania koncepcji systemu rekomendacyjnego dla zadanego scenariusza aplikacyjnego). Projekt: ocena jakości wykonania systemu rekomendacyjnego z perspektywy zadanego scenariusza aplikacyjnego; ocena końcowa średnią ocen cząstkowych przyznawanych w trakcie realizacji projektu na podstawie oceny postępu w przygotowaniu projektu, prezentacji implementacji i dokumentacji przygotowywanego systemu.
Treści programowe

Wykład

Kluczowe zagadnienia prezentowane na wykładzie to: algorytmy systemów uczących się (ang. machine learning) mające zastosowanie w budowie systemów rekomendacyjnych, filtracja kolaboratywna, sposoby reprezentacji i przetwarzania danych w wielowymiarowych przestrzeniach wektorowych, rekomendacja hybrydowa, modelowanie kontekstu, rekomendacja spersonalizowana, nadzorowane uczenie się (w szczególności w oparciu o metodę gradient boosting, regresję logistyczną i sieci neuronowe), komercyjne systemy rekomendacyjne, systemy rekomendacyjne używane jako komponenty systemów reklamy internetowej (systemów optymalizacji funkcji spersonalizowanego i kontekstowego kierowania (ang. targeting) reklam internetowych) i systemów automatycznej oceny bankowych wniosków kredytowych, wiodące biblioteki programistyczne i pakiety oprogramowania, które mogą być stosowane do szybkiego przygotowania i przetestowania wydajnego i skalowalnego systemu rekomendacyjnego (w szczególności Scikit-learn, Numpy, Scipy, Wowpal Wabbit, LibFFM). Najnowsze osiągnięcia w badaniach nad systemami rekomendacyjnymi przedstawiane są ze szczególnym uwzględnieniem rozwiązań Big Data, które umożliwiają zastosowanie metod przetwarzania danych tekstowych w języku naturalnym i modelowania semantycznego oraz rozwiązań umożliwiających skalowalne przetwarzanie w trybie offline niekompletnych danych heterogenicznych i skuteczne generowanie rekomendacji stosowalne w systemach działających w czasie rzeczywistym.

Metody kształcenia

Stosowane są następujące metody kształcenia:

- prezentacja teorii z częstymi odniesieniami do praktycznych przykładów implementacji,
- wykład z prezentacjami multimedialnymi i prezentacjami przykładów kodu źródłowego w języku programowania Python: jego modyfikacji i wyników uruchomienia,
- dyskusje inicjowane w trakcie wykładu.

Projekt

Projekt studencki polega na implementacji nieskomplikowanego systemu rekomendacyjnego z użyciem publicznie dostępnego oprogramowania.

Projekty skupiają się na wykorzystaniu publicznie dostępnych zbiorów testowych, np. MovieLens, HetRec, S3, OAEI, zbiory służące ocenie predykcji tzw. współczynnika klikalności (ang. Click-Through Rate, CTR) reklam internetowych i zbiory służące automatycznej ocenie wniosków kredytowych pochodzące z serwisu Kaggle, komponenty oprogramowania opracowane w kilku projektach badawczych Politechniki Poznańskiej, w szczególności w projekcie "Teorioinformacyjne wnioskowanie abdukcyjne dla rekomendacji kontekstowej" i projekcie poświęconym zastosowaniu metod systemów uczących się do oceny (ratingu) bankowych wniosków kredytowych finansowanym przez Talax S.A. Praca w projekcie studenckim polega m.in. na zastosowaniu standardowej metodyki oceny jakości rekomendacji, tj. jakości predykcji uzyskanych z użyciem systemu rekomendacyjnego (powszechnie stosowanej w środowisku akademickim i przemysłowym), która została początkowo opracowana przez naukowców zajmujących się systemami uczącymi się i systemami wyszukiwania informacji, ale obecnie są łatwo implementowalne z użyciem wiodących bibliotek oprogramowania (Scikit-learn, Numpy, Scipy) .

Metody kształcenia

Stosowane są następujące metody kształcenia:

- indywidualna praca studenta (wykonywana głównie poza zajęciami),
- praca z narzędziami o otwartym kodzie źródłowym (ang. open source) i składnikami oprogramowania opracowanymi w projektach badawczych Politechniki Poznańskiej udostępnianymi studentom przez prowadzącego zajęcia,
- okresowe (raz na kilka tygodni) krótkie przedstawienie przez studenta postępów w realizacji projektu (każdorazowo uzupełnione podpowiedziami wykładowcy, rozmowami z innymi studentami oraz jawną oceną wyników dokonaną przez wykładowcę),
- krótka prezentacja końcowych wyników pracy studenta w projekcie (wraz z rozmowami z udziałem innych studentów i z jawną oceną wyników dokonaną przez wykładowcę).

Aktualizacja w 2017 r.

Dokonano poważnej modyfikacji całego opisu kursu i bibliografii. W szczególności wprowadzono kilka nowych tematów, w tym: systemy rekomendacyjne dla systemów spersonalizowanego i kontekstowego kierowania reklam internetowych oraz systemów automatycznej oceny bankowych wniosków kredytowych z użyciem technologii systemów uczących się (ang. Machine Learning system), system rekomendacyjny ujęty jako szczególny przypadek systemu uczącego się, uzyskane w ostatnich latach postępy w badaniach nad systemami rekomendacyjnymi, w tym nad rozwiązaniami klasy Big Data umożliwiającymi zastosowanie metod przetwarzania danych tekstowych w języku naturalnym i metod modelowania semantycznego, skalowalne przetwarzanie heterogenicznych i niepełnych danych w trybie offline i efektywne generowanie rekomendacji stosowalne w systemach czasu rzeczywistego, biblioteki programistyczne i pakiety oprogramowania umożliwiające przygotowanie i ocenę systemów rekomendacyjnych, stosowana w środowisku akademickim i w przemyśle metodyka oceny systemów rekomendacyjnych.

Literatura podstawowa:

1. Kim Falk, Practical Recommender Systems, Manning Publications, www.manning.com/books/practical-recommender-systems, 2017.
2. Francesco Ricci, Lior Rokach, Bracha Shapira, Recommender Systems Handbook (2nd ed.). Springer-Verlag New York, Inc., New York, NY, USA, 2015.
3. Trevor Hastie, Robert Tibshirani, Jerome Friedman, The elements of statistical learning : data mining, inference, and prediction, 2nd ed., Springer Series in Statistics, ISBN 978-0-387-84857-0, 978-0-387-84858-7 (e-ISBN), New York, 2016.
4. Szwabe A., Misiorek P., Ciesielczyk M., "Logistic Regression Setup for RTB CTR Estimation", in: Proceedings of the 9th International Conference on Machine Learning and Computing, ICMLC 2017, Singapore, Singapore, pp. 61-70, ACM, New York, USA, DOI = 10.1145/3055635.3056584, 2017, <http://dl.acm.org/citation.cfm?id=3056584>, http://ncn6788.cie.put.poznan.pl/images/ncn6788_icmlc2017.pdf

Literatura uzupełniająca:

1. Christopher D. Manning, Prabhakar Raghavan, Hinrich Schütze, An Introduction to Information Retrieval, online edition, <https://nlp.stanford.edu/IR-book/>, <https://nlp.stanford.edu/IR-book/pdf/irbookprint.pdf>, Cambridge University Press, England, 2009.
2. Szwabe, A., Misiorek, P., Walkowiak, P., Multi-Relational Learning for Recommendation of Matches between Semantic Structures, in: Grana, M., Toro, C., Howlett, R.J., Jain, L.C. (Eds.), Knowledge Engineering, Machine Learning and Lattice Computing with Applications, LNCS/LNAI Volume 7828, 2013, pp. 98-107, http://link.springer.com/chapter/10.1007/978-3-642-37343-5_11, http://ncn6788.cie.put.poznan.pl/images/ncn6788_lnai.pdf
3. Szwabe A., Ciesielczyk M., Misiorek P., Blinkiewicz M., "Application of the tensor-based recommendation engine to semantic service matchmaking", Proceedings of The Ninth International Conference on Advances in Semantic Processing, pp. 116-125, ISBN: 978-1-61208-420-6, July 2015, Nice, France, http://www.thinkmind.org/index.php?view=article&articleid=semapro_2015_5_40_30093, <http://ncn6788.cie.put.poznan.pl/images/ncn6788-semapro2015.pdf>

Bilans nakładu pracy przeciętnego studenta

Czynność	Czas (godz.)
1. Udział w wykładach	15
2. Udział w zajęciach projektowych	15
3. Konsultacje i egzamin	5
4. Przygotowanie do zajęć projektowych	25
5. Opracowanie sprawozdań	10
6. Przygotowanie do egzaminu	10

Obciążenie pracą studenta

forma aktywności	godzin	ECTS
Łączny nakład pracy	75	3
Zajęcia wymagające bezpośredniego kontaktu z nauczycielem	35	1
Zajęcia o charakterze praktycznym	50	2