

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ОДЕСЬКА НАЦІОНАЛЬНА АКАДЕМІЯ ЗВ'ЯЗКУ ім. О.С. ПОПОВА

(повне найменування вищого навчального закладу)

Кафедра інформаційних технологій

“ЗАТВЕРДЖУЮ”

Зав. каф. інформаційних технологій

_____ 2020 року
“ ____ ” _____

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Алгоритми та технології побудови рекомендаційних систем

спеціальність **121 Інженерія програмного забезпечення**

факультет, відділення **Інфокомунікацій та програмної інженерії**

Одеса
2020 рік

Робоча програма дисципліни Алгоритми та технології побудови рекомендаційних систем для студентів

(назва навчальної дисципліни)

за спеціальністю 121 Інженерія програмного забезпечення

Розробник програми: к.т.н., доц. Єгошина Г.А.

Робоча програма затверджена на засіданні кафедри Інформаційних технологій

Протокол від “_____” _____ 20__ року № _____

Завідувач кафедри Інформаційних технологій

_____ (підпис) (_____)
(прізвище та ініціали)

“_____” _____ 20__ року

1. Опис навчальної дисципліни

Найменування показників	Галузь знань, напрям підготовки, освітньо-кваліфікаційний рівень	Характеристика навчальної дисципліни	
		денна форма навчання	
Кількість кредитів ЄКТС – 6	Галузь знань <u>12 Інформаційні технології</u> (шифр і назва)	Обов'язкова	
	Спеціальність <u>122 Інженерія програмного забезпечення</u> (шифр і назва)		
Семестрів – 1	Спеціальність (професійне спрямування): _____	Рік підготовки:	
Змістових модулів – 2		1-ий	
Індивідуальне науково-дослідне завдання _____		Семестр	
Загальна кількість годин – 180		1.2	
Тижневих годин для денної форми навчання: аудиторних – 4 самостійної роботи студента – 5	Ступінь вищої освіти: <i>магістр</i>	Лекції	
		20 год.	
		Практичні, семінарські	
		20 год.	
		Лабораторні	
		40 год.	
		Самостійна робота	
		100 год.	
		Індивідуальні завдання	
		КР	–
Вид контролю: залік, КР			

2. Анотація дисципліни

Програма вивчення навчальної дисципліни “Алгоритми та технології побудови рекомендаційних систем” складена відповідно до освітньо-професійної програми підготовки магістрів спеціальності 121 Інженерія програмного забезпечення.

Курс призначений для вивчення сучасних алгоритмів та технології побудови рекомендаційних систем.

Сьогодні спостерігається стрімке вдосконалення технологій рекомендаційних систем. Вперше випробувані в середовищі електронної комерції, на даний момент вони активно використовуються в різних сферах. Рекомендаційні системи допомагають користувачам, які зіткнулися з проблемою величезного вибору, визначити, що саме найкраще відповідає їхнім потребам. Сучасні системи такого класу вивчають переваги кожного користувача і надають персоніфіковані рекомендації. У процесі роботи ці системи збирають дані про користувачів, використовуючи поєднання явних і неявних методів.

Програма навчальної дисципліни складається з таких змістових модулів:

1. Формування рекомендацій.
2. Виявлення груп. Кластеризація. Оптимізація.
3. Вдосконалення методик Collaborative Filtering.
4. Моделі класу Neighborhood.

3. Мета та завдання навчальної дисципліни

Метою викладання навчальної дисципліни є підготовка студента до ефективного використання сучасних методів розробки рекомендаційних систем різного профілю. Ознайомлення з основними поняттями колективного розуму та машинного навчання, методами колаборативної фільтрації та існуючими метриками оцінки подібності, сучасними програмними засоби для проектування і розробки рекомендаційних систем.

Курс передбачає теоретичні та практичні заняття.

В межах теоретичної частини забезпечуються **знання**:

- основні поняття та визначення колективного розуму та машинного навчання;
- методи колаборативної фільтрації;
- існуючі метрики оцінки подібності в рекомендаційних системах;
- сучасні програмні засоби та технології для проектування і розробки рекомендаційних систем.

Проходження практичної частини курсу формує **вміння**:

- обґрунтовувати й аналізувати вибір конкретного типу моделі та методу колаборативної фільтрації при вирішенні практичних задач;
- використовувати сучасні програмні засоби для проектування та дослідження рекомендаційних систем;
- створювати програми для формування рекомендацій при розв’язку конкретних практичних задач;

- аналізувати результати побудови та використання рекомендаційних систем при вирішенні прикладних задач

Вивчення навчальної дисципліни передбачає формування та розвиток у студентів **компетентностей**:

загальних:

- ЗК-1. Здатність до абстрактного мислення, аналізу та синтезу.
- ЗК-2. Здатність спілкуватися іноземною мовою як усно, так і письмово.
- ЗК-3. Здатність проведення теоретичних та прикладних досліджень на відповідному рівні.
- ЗК-4. Здатність мотивувати людей та рухатися до спільної мети, працювати в команді співробітників.
- ЗК-5. Здатність спілкуватися з представниками інших професійних груп різного рівня (з експертами з інших галузей знань/видів економічної діяльності).
- ЗК-6. Здатність удосконалювати свої навички на основі аналізу попереднього досвіду.
- ЗК-7. Здатність генерувати нові ідеї (креативність).
- ЗК-11. Здатність управляти своїм часом та розуміти важливість дедлайнів

фахових:

- СК-1. Здатність аналізувати предметні області, формувати, аналізувати та моделювати вимоги до програмного забезпечення.
- СК-2. Здатність ідентифікувати, класифікувати та описувати проектні завдання, знаходити раціональні методи й підходи до їх розв'язання.
- СК-3. Здатність проектувати програмне забезпечення, включаючи проведення моделювання його архітектури, поведінки та процесів функціонування окремих підсистем і модулів.
- СК-4. Здатність розвивати і реалізовувати нові конкурентоспроможні ідеї в інженерії програмного забезпечення.
- СК-5. Здатність оцінювати ступінь обґрунтованості застосування специфікацій, стандартів, правил і рекомендацій в професійній галузі та дотримуватися їх при реалізації процесів життєвого циклу програмного забезпечення.
- СК-7. Здатність систематизувати професійні знання щодо створення і супроводження програмного забезпечення.
- СК-12. Здатність використовувати методи машинного навчання та роботи з Big Data, засоби штучного інтелекту для дослідження та аналізу процесів життєвого циклу програмного забезпечення.
- СК-14. Здатність використовувати знання сучасних комп'ютерних та інформаційних технологій та інструментів інженерних і наукових досліджень, розрахунків, обробки та аналізу даних, моделювання та оптимізації.
- СК-16. Здатність демонструвати та застосовувати на практиці знання методів моделювання програмних систем, оцінки їх ефективності та якості.

Результати навчання даної дисципліни деталізують такі **програмні результати навчання**:

- Знати і системно застосовувати методи аналізу та моделювання прикладної області, виявлення інформаційних потреб і збору вихідних даних для проектування програмного забезпечення
- Обґрунтовувати вибір методів формування вимог до програмної системи, розробляти, аналізувати та систематизувати вимоги.
- Розробляти і оцінювати стратегії проектування програмних засобів, в тому числі з урахуванням впливу факторів різновекторного спрямування; обґрунтовувати, аналізувати і оцінювати прийняті проектні рішення з точки зору якості кінцевого програмного продукту.
- Аналізувати, оцінювати і вибирати методи, сучасні програмно-апаратні інструментальні та обчислювальні засоби, технології, алгоритмічні та програмні рішення для ефективного виконання конкретних виробничих задач з програмної інженерії.
- Обґрунтовано вибирати парадигми і мови програмування для вирішення прикладних завдань; застосовувати на практиці системні та спеціалізовані засоби, компонентні технології (платформи) та інтегровані середовища розробки програмного забезпечення.
- Знати і застосовувати сучасні професійні стандарти і інші нормативно-правові документи з інженерії програмного забезпечення.
- Вміти приймати організаційно-управлінські рішення в умовах невизначеності.
- Набувати нові наукові і професійні знання, вдосконалювати навички, прогнозувати розвиток програмних систем та інформаційних технологій.
- Формулювати, експериментально підтверджувати, обґрунтовувати і застосовувати на практиці в процесі розробки програмного забезпечення конкурентоспроможні ідеї, методи, технології вирішення професійних, науково-технічних завдань в умовах невизначеності.
- Здобувати необхідну інформацію з іншомовної літератури, аналізувати та вибирати необхідні для вирішення фахових наукових і прикладних задач інформаційно-довідкові та науково-технічні ресурси і джерела знань з урахуванням сучасних досягнень науки і техніки, здійснювати різні види комунікації під час спілкування
- Проектувати та створювати програмні системи зберігання та обробки великих масивів даних, розробляти високонавантажені системи обробки даних, використовувати методи штучного інтелекту та машинного навчання у вирішенні практичних завдань
- Організовувати командну роботу, управляти проектами, підбирати команду проекту, ефективно працювати в групі, визначати та розподіляти завдання з метою вирішення різноманітних дослідницьких та практичних завдань

На вивчення навчальної дисципліни відводиться 180 годин / 6 кредитів ECTS.

4. Інформаційний обсяг навчальної дисципліни

Семестр 1.2

Змістовий модуль 1. Формування рекомендацій

Тема 1. Поняття колективного розуму. Сутність web 2.0. Машинне навчання. Основні поняття рекомендаційних систем. Відомі реалізації.

Тема 2. Збір інформації про вподобання. Колаборативна фільтрація. Основні підходи. Пошук схожих користувачів. Оцінка схожості.

Тема 3. Розрахунок оцінок подібності користувачів: Евклідова відстань, коефіцієнт кореляції Пірсона, коефіцієнт Жаккара, відстань Хемінга.

Тема 4. Розрахунок оцінок подібності користувачів: манхетенівська відстань, косинусна міра подібності, модифікована косинусна міра подібності.

Тема 5. Рекомендування та підбір об'єктів.

Тема 6. Фільтрація на основі подібності зразків. Порівняння методів фільтрації на основі подібності.

Змістовий модуль 2. Виявлення груп. Кластеризація. Оптимізація

Тема 7. Вектори слів. Систематизація блогерів. Розрахунок кількості слів у RSS каналі.

Тема 8. Ієрархічна кластеризація. Кластеризація методом K-середніх.

Тема 9. Кластери вподобань.

Тема 10. Групові мандрівки. Представлення рішень Цільова функція.

5. Структура навчальної дисципліни

Назви змістових модулів і тем	Кількість годин					
	денна форма					
	усього	у тому числі				
л		п	лаб	інд	ср	
<i>1</i>	2	3	4	5	6	7
<i>Семестр 1.2</i>						
Змістовний модуль 1. Формування рекомендацій						
Тема 1. Поняття колективного розуму. Сутність web 2.0. Машинне навчання. Основні поняття рекомендаційних систем. Відомі реалізації.	13	1	2	2	-	8
Тема 2. Збір інформації про вподобання. Колаборативна фільтрація. Основні підходи. Пошук схожих користувачів. Оцінка схожості.	15	1	-	4	-	10
Тема 3. Розрахунок оцінок подібності користувачів: Евклідова відстань, коефіцієнт кореляції Пірсона, коефіцієнт Жаккара, відстань Хемінга.	18	2	2	4	-	10
Тема 4. Розрахунок оцінок подібності користувачів: манхетенівська відстань, косинусна міра подібності, модифікована косинусна міра подібності.	18	2	2	4	-	10
Тема 5. Рекомендування та підбір об'єктів.	18	2	2	4	-	10
Тема 6. Фільтрація на основі подібності зразків. Порівняння методів фільтрації на основі подібності.	18	2	2	4	-	10
<i>Разом за змістовним модулем 1</i>	<i>100</i>	<i>10</i>	<i>10</i>	<i>22</i>	<i>-</i>	<i>58</i>
Змістовний модуль 2. Виявлення груп. Кластеризація. Оптимізація						
Тема 7. Вектори слів. Систематизація блогерів. Розрахунок кількості слів у RSS каналі.	18	2	2	4	-	10
Тема 8. Ієрархічна кластеризація. Кластеризація методом К-середніх.	20	2	2	4	-	12
Тема 9. Кластери вподобань.	24	4	4	6	-	10
Тема 10. Групові мандрівки. Представлення рішень Цільова функція.	18	2	2	4	-	10
<i>Разом за змістовним модулем 2</i>	<i>80</i>	<i>10</i>	<i>10</i>	<i>18</i>	<i>-</i>	<i>42</i>
Усього годин	180	20	20	40		100

6. Теми практичних занять

№	Назва теми	Години
<i>Семестр 1.2</i>		
1	Установка дистрибутиву Anaconda та робота з Jupyter Notebook	2
2	Формування рекомендацій: пошук схожих користувачів на основі коефіцієнта кореляції Пірсона, відстані Хемінга, коефіцієнта Жакара	12
3	Рекомендування та підбір об'єктів	8
4	Розрахунок кількості слів у RSS каналі на основі Universal Feed Parser	4
5	Формування ієрархії блогів за допомогою кластеризації	4
6	Підготовка презентації курсового проекту	2
Усього		20

7. Теми лабораторних занять

№	Назва теми	Години
1	Списки, кортежі в Python. Словники в Python	6
2	Формування рекомендацій: пошук схожих користувачів на основі коефіцієнта кореляції Спірмена, косинусної міри подібності, модифікованої косинусної міри подібності	10
3	Фільтрація на основі подібності зразків	4
4	Розрахунок кількості слів у RSS каналі	4
5	Кластеризація методом К-середніх	10
6	Групові мандрівки. Представлення рішень Цільова функція	6
Усього		40

8. Самостійна робота

№	Види роботи	Години
1	Опрацювання лекцій	10
2	Вивчення додаткового матеріалу до лекцій	10
3	Підготовка до практичних занять	20
4	Підготовка до лабораторних робіт	20
5	Робота над КР	40
Усього годин		100

9. Методи навчання

Процес навчання дисципліни базується на використанні сучасних методів навчання: мультимедійні лекції, демонстрування, викладання-пояснення, практичні заняття та лабораторні роботи, метод проблемно-орієнтовного навчання, самостійне спостереження, запитання-бесіда, ілюстрування.

10. Методи контролю

- контроль рівня виконавських умінь для перевірки практичної підготовки;
- усне та письмове опитування;
- захист індивідуальних робіт.

Оцінювання проводиться за шкалою ЄКТС, національною та за шкалою ОНАЗ ім. О.С. Попова (100 бал.)

11.Розподіл балів, які отримують студенти

Частина 1 – Семестр 1.2

Поточне тестування та самостійна робота		Контрольні завдання (після кожного модулю)	Сума
ЗМ1	ЗМ2		
30	30	40	100

Шкала оцінювання: національна та ЄКТС

Сума балів за всі види навчальної діяльності	Оцінка ЄКТС	Оцінка за національною шкалою	
		для екзамену, курсової роботи, практики	для заліку
90-100	A	відмінно	зараховано
82-89	B	добре	
74-81	C		
64-73	D	задовільно	
60-63	E		
35-59	FX	незадовільно з можливістю повторного складання	не зараховано з можливістю повторного складання
0-34	F	незадовільно з обов'язковим повторним вивченням дисципліни	не зараховано з обов'язковим повторним вивченням дисципліни

12.Методичне, технічне й програмне забезпечення /обладнання

Для реалізації програмних проектів використовується мова програмування Python 3.X. В якості програмних засобів рекомендується використовувати дистрибутив Anaconda (www.anaconda.com), бібліотеки NumPy (numpy.org), Pandas (pandas.pydata.org), Scikit-Learn (scikit-learn.org), Surprise (surpriselib.com). Також можна використовувати засоби Google Colaboratory (colab.research.google.com) або IDE PyCharm (www.jetbrains.com/pycharm)-додатково вимагає інсталяції всіх необхідних Python бібліотек окремо.

Методичне забезпечення курсу: конспект лекцій, завдання для практичних та лабораторних робіт, презентаційні матеріали.

13.Рекомендована література

1. Барсегян А. А. Методы и модели анализа данных: OLAP и Data Mining / А. А. Барсегян, М. С. Куприянов, В. В. Степаненко та ін. – 2-е изд., перераб. и доп. – СПб. : БХВ-Петербург, 2004. – 336 с.
2. Плас Дж. Вандер. Python для сложных задач: наука о данных и машинное обучение. — СПб.: Питер, 2018. — 576 с.: ил.
3. Recommender Systems Handbook. F. Ricci, L. Rokach, B. Shapira, P. В. Kantor, 2015 (second edition).
4. Charu С. Aggarwal Recommender Systems: The Textbook 1st ed. 2016 Edition
5. Data clustering : algorithms and applications / [edited by] Charu С. Aggarwal, Chandan К. Reddy.// <http://charuaggarwal.net/clusterbook.pdf>
6. С. Aggarwal and P. Yu. Privacy-preserving data mining: models and algorithms, Springer, 2008. // [https://doc.lagout.org/Others/Data%20Mining/Privacy-Preserving %20Data %20Mining_ %20Models %20and %20Algorithms %20%5BAggarwal %20%26%20Yu%202008-07-07%5D.pdf](https://doc.lagout.org/Others/Data%20Mining/Privacy-Preserving%20Data%20Mining_%20Models%20and%20Algorithms%20%5BAggarwal%20%26%20Yu%202008-07-07%5D.pdf)
7. Рассел С. Искусственный интеллект: современный подход / Рассел С., Норвинг П. – М. : Издательский дом "Вильямс", 2006. – 1408 с.

Інформаційні ресурси

- <https://ru.coursera.org/specializations/recommender-systems>
- <https://medium.com/ai-society/a-concise-recommender-systems-tutorial-fa40d5a9c0fa>
- <https://www.datacamp.com/community/tutorials/recommender-systems-python>
- <https://www.analyticsvidhya.com/blog/2018/06/comprehensive-guide-recommendation-engine-python/>
- <https://cloud.google.com/solutions/machine-learning/recommendation-system-tensorflow-overview>
- <https://surprise.readthedocs.io/en/stable/index.html>
- <http://www.foursquare.com>
- <http://grouplens.org>
- <http://grouplens.org/datasets/movielens/>
- <http://eigentaste.berkeley.edu/user/index.php>
- <http://www.netflixprize.com/community/viewtopic.php?id=828>
- <http://blog.netflix.com/2010/03/this-is-neil-hunt-chief-product-officer.html>
- <http://www.kddcup2012.org/workshop>
- S. Agarwal. Ranking methods in machine learning. Tutorial at SIAM Conference on Data Mining, 2010. Slides available at: <http://www.siam.org/meetings/sdm10/tutorial1.pdf>