

<b>Назва курсу</b>	Алгоритми та технології побудови рекомендаційних систем
<b>Викладачі</b>	к.т.н., доц. Єгошина Ганна Анатолівна
<b>Контактний тел.</b>	705-04-42
<b>E-mail:</b>	yegoshyna@onat.edu.ua
<b>Консультації</b>	ауд. 245, Чт. 15.00-16.00

### 1. Анотація до курсу

Курс призначений для вивчення сучасних алгоритмів та технологій 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 персоналізовані рекомендації. У процесі роботи ці системи збирають дані про користувачів, використовуючи поєднання явних і неявних методів.

### 2. Мета та цілі курсу

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.

### 3. Формат курсу

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

#### **4. Результати навчання**

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 високонавантаженні 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

## 5. Обсяг курсу

Семестр 1.2				
Вид заняття	лекції	лабораторні заняття	практичні заняття	самостійна робота (+ КР)
К-сть годин (180)	20	40	20	100

## 6. Ознаки курсу:

Рік викладання	семестр	спеціальність	Курс, (рік навчання)	Нормативний\вибірковий
2020	1.2	Інженерія програмного забезпечення	1	н

## 7. 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](http://www.anaconda.com)), бібліотеки NumPy ([numpy.org](http://numpy.org)), Pandas ([pandas.pydata.org](http://pandas.pydata.org)), Scikit-Learn ([scikit-learn.org](http://scikit-learn.org)), Surprise ([surpriselib.com](http://surpriselib.com)). Також можна використовувати засоби Google Colaboratory ([colab.research.google.com](http://colab.research.google.com)) або IDE PyCharm ([www.jetbrains.com/pycharm](http://www.jetbrains.com/pycharm))- додатково вимагає інсталяції всіх необхідних Python бібліотек окремо.

## 8. Політики курсу

Для успішного проходження курсу та складання контрольних заходів необхідним є виконання наступних обов'язків:

- не запізнюватися на заняття;
- не пропускати заняття (як лабораторні, практичні так і лекційні), в разі хвороби мати довідку або її ксерокопію;
- самостійно опрацювати весь лекційний матеріал та ресурси для самостійної роботи;

- конструктивно підтримувати зворотній зв'язок з викладачем на всіх етапах проходження курсу (особливо під час виконання курсового проекту);
- своєчасно і самостійно виконувати всі передбачені програмою лабораторні та практичні завдання;
- не користуватися мобільним телефоном під час аудиторних занять;
- брати очну участь у контрольних заходах;
- будь-яке відтворення результатів чужої праці (виключаючи практичну роботу над командним проектом), в тому числі використання завантажених з Інтернету матеріалів, як власних результатів, кваліфікується, як порушення норм і правил академічної доброчесності, та передбачає притягнення до відповідальності у порядку, визначеному чинним законодавством.

## 9. Схема курсу

Тиж. / год.	Тема, план, короткі тези	Форма діяльності (заняття) / Формат	Матеріали	Завдання ( в тому числі тематика лабораторної роботи), год	Вага оцінки	Термін виконання
Тиж. 1-10 30 акад. год.	<b>Змістовий модуль 1. Формування рекомендацій</b> <i>Тема 1.</i> Поняття колективного розуму. Сутність web 2.0. Машинне навчання. Основні поняття рекомендаційних систем. Відомі реалізації.	Лекція, практика, лабораторна F2F	Презентація	Передивитись презентацію, Установка дистрибутиву Anaconda та робота з Jupyter Notebook, Списки, кортежі в Python 4 год	5	1 тиж.
	<i>Тема 2.</i> Збір інформації про вподобання. Колаборативна фільтрація. Основні підходи. Пошук схожих користувачів. Оцінка схожості.	Лекція, лабораторна F2F	Презентація	Передивитись презентацію, Словники в Python 4 год	5	1 тиж.
	<i>Тема 3.</i> Розрахунок оцінок подібності користувачів: Евклідова відстань, коефіцієнт кореляції Пірсона, коефіцієнт Жаккара, відстань Хемінга.	Лекція, практика, лабораторна F2F	Презентація	Передивитись презентацію, Формування рекомендацій: пошук схожих користувачів на основі коефіцієнта кореляції Пірсона, відстані Хемінга, коефіцієнта Жакара	5	1 тиж.

Тиж. / год.	Тема, план, короткі тези	Форма діяльності (заняття) / Формат	Матеріали	Завдання ( в тому числі тематика лабораторної роботи), год	Вага оцінки	Термін виконання
				6 год		
	Тема 4. Розрахунок оцінок подібності користувачів: манхетенівська відстань, косинусна міра подібності, модифікована косинусна міра подібності.	Лекція, практика, лабораторна F2F	Презентація	Передивитись презентацію, Формування рекомендацій: пошук схожих користувачів на основі коефіцієнта кореляції Спірмена, косинусної міри подібності, модифікованої косинусної міри подібності 6 год	5	1 тиж.
	Тема 5. Рекомендування та підбір об'єктів.	Лекція, практика, лабораторна F2F	Презентація	Передивитись презентацію, Рекомендування та підбір об'єктів 6 год	5	1 тиж.
	Тема 6. Фільтрація на основі подібності зразків. Порівняння методів фільтрації на основі подібності	Лекція, практика, лабораторна F2F	Презентація	Передивитись презентацію, Фільтрація на основі подібності зразків 6 год Контрольні завдання	5 10	1 тиж.
Тиж.11-20 30 акад. год.	<b>Змістовий модуль 2. Виявлення груп. Кластеризація. Оптимізація</b>	Лекція, практика, лабораторна F2F	Презентація	Передивитись презентацію, Розрахунок кількості слів у RSS каналі на основі Universal Feed Parser	10	1 тиж.

Тиж. / год.	Тема, план, короткі тези	Форма діяльності (заняття) / Формат	Матеріали	Завдання ( в тому числі тематика лабораторної роботи), год	Вага оцінки	Термін виконання
	<i>Тема 7.</i> Вектори слів. Систематизація блогерів. Розрахунок кількості слів у RSS каналі.			Робота над завданнями курсового проекту <i>6 год</i>		
	<i>Тема 8.</i> Ієрархічна кластеризація. Кластеризація методом К-середніх.	Лекція, практика, лабораторна <i>F2F</i>	Презентація	Передивитись презентацію, Кластеризація методом К-середніх Робота над завданнями курсового проекту <i>6 год</i>	10	1 тиж.
	<i>Тема 9.</i> Кластери вподобань.	Лекція, практика, лабораторна <i>F2F</i>	Презентація	Передивитись презентацію, Формування ієрархії блогів за допомогою кластеризації Робота над завданнями курсового проекту <i>10 год</i>	10	1 тиж.
	<i>Тема 10.</i> Групові мандрівки. Представлення рішень Цільова функція	Лекція, практика, лабораторна <i>F2F</i>	Презентація	Передивитись презентацію, Підготовка пояснювальної записки до курсового проекту <i>6 год</i>  Контрольні завдання	10	1 тиж.

## 10. Система оцінювання та вимоги

### 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 ECTS

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



## 11. 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. B. Kantor, 2015 (second edition).
4. Charu C. Aggarwal Recommender Systems: The Textbook 1st ed. 2016 Edition
5. Data clustering : algorithms and applications / [edited by] Charu C. Aggarwal, Chandan K. Reddy.// <http://charuaggarwal.net/clusterbook.pdf>
6. C. 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 / Рас- сел С., Норвинг П. – М. : Издательский дом "Вильямс", 2006. – 1408 с.

## 12 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>