



СИЛАБУС ДИСЦИПЛІНИ «ДЕЙТАМАЙНІНГ»

Ступінь вищої освіти	третій (освітньо-науковий)
Спеціальність	<u>071 Облік і оподаткування</u>
Освітня програма	Облік і оподаткування
Рік навчання	<u>2 семестр 3</u>
Форма навчання	денна, вечірня, заочна
Кількість кредитів ЄКТС	<u>5</u>
Мова викладання	українська

Лектор курсу

Доктор економічних наук, професор кафедри статистики та економічного аналізу Савчук Василь Кирилович

Контактна інформація
лектора (e-mail)

Телефон службовий Тел. 044 527 82 36
savchukvk@ukr.net

Сторінка курсу в eLearn

: <https://elearn.nubip.edu.ua/course/view.php?id=2900>

ОПИС ДИСЦИПЛІНИ

Метою навчальної дисципліни «Дейтамайнінг» є вивчення методів сучасної обробки даних – інтелектуального аналізу даних (Data Mining, Knowledge Discovery in Data), аналітичного дослідження великих масивів інформації з метою виявлення нових раніше невідомих, практично корисних знань і закономірностей, необхідних для прийняття рішень; огляд методів, програмних продуктів і різних інструментальних засобів, які використовуються Data Mining; розгляд практичних прикладів застосування Data Mining; підготовка аспірантів до самостійної роботи з вирішення задач засобами Data Mining і розробки інтелектуальних систем.

Завданнями вивчення дисципліни «Дейтамайнінг» є:

- здійснення пошуку, систематизація, аналіз та критичне осмислення інформації, отриманої з різних джерел;
- генерування і продукування власних науково-теоретичних та практично-спрямованих ідей;
 - опанування базовими принципами побудови інтелектуальних моделей даних;
 - ознайомлення з концепцією Knowledge Discovery in Data (виявлення знань в даних) і Data Mining («видобування» знань);
 - навчитися ефективно використовувати методи здобуття знань з великих масивів даних;
 - отримати практичні навички з використання інструментальних засобів інтелектуального аналізу даних при вирішенні прикладних задач та навчитися інтерпретувати отримані результати;
 - проводити емпіричні дослідження та використовувати економіко-математичні методи для встановлення тенденцій розвитку об'єктів дослідження;
 - застосовувати інформаційні технології у науковій діяльності, сучасні методи наукових досліджень із використанням новітніх прикладних пакетів і програмних продуктів для наукового обґрунтування та підтвердження / спростування висунутих гіпотез;
 - ідентифікувати наукові та практичні проблеми, здійснювати апробацію результатів наукових досліджень, висновків і практичних рекомендацій.

СТРУКТУРА КУРСУ

Тема	Години (лекції/ практичні)	Результати навчання	Завдання	Оцінювання, бали
Тема 1. Концепція дейтамайнінгу	2/2	Знати сутність: Data Mining та області застосування, задачі, моделі та методи Data Mining; поняття Business Intelligence; циклів одержання, попередньої обробки, аналізу даних, інтерпретації результатів та їхнього використання. Уміння обґрунтовувати етапи процесу Data Mining, пов'язані з побудовою, перевіркою, оцінкою, вибором і корекцією моделей. Володіти методами первісної обробки даних та методами дослідження їх структури: візуалізацією та автоматичним групуванням даних.	Підготовка реферату з оглядом дискусії щодо можливостей і областей застосування дейтамайнінгу	10
Тема 2. Алгоритми Data Mining: класифікація і регресія	2/2	Вміти ставити задачі класифікації, оцінювати та представляти її результати. Володіти правилами класифікації та методами побудови математичних функцій. Знати особливості класифікації об'єктів у випадку невідомих розподілень даних. Володіти методами оцінювання помилок класифікації та методами вирішення задач регресії.	Підбір задач і методів класифікації та вирішення задач регресії	10
Тема 3. Інтелектуальний аналіз часових рядів	4/4	Знати ознаки та вміти оцінювати часові ряди; володіти методами моделювання, аналізу та прогнозування їх поведінки	Підбір методів побудови, аналізу та прогнозування часових рядів	10
Тема 4. Алгоритми Data Mining: кластеризація	4/4	Вміти ставити задачі кластеризації, представляти та інтерпретувати її результати. Володіти типологією кластерів. Знати базові алгоритми кластеризації та адаптивні методи її проведення.	Підбір задач, алгоритмів та методів кластеризації.	10
Тема 5. Технології нейронних мереж та генетичні алгоритми	4/6	Знати сутність та еволюцію нейронних мереж. Вміти підбирати математичну модель штучного нейрона; програмне та апаратне забезпечення нейромереж; визначати діапазон їх застосування, оцінювати переваги та недоліки.	Розуміння сутності нейронних мереж та методик їх побудови та	10

			застосуван ня.	
Тема 6. Дерев рішень як метод дейтамайнінгу	2/6	Розуміти сутність дерев рішень, особливості їх використання та побудови.. Вміти описувати і класифікувати альтернативні їх варіанти та методи її проведення і способи побудови класифікаційних дерев рішень.	Оцінка дерева рішень та його побудова стосовно свого об'єкту дослідженн я	10
Тема 7. Сховища даних та оперативний аналіз даних (OLAP)	2/6	Визначати сховища даних, оцінювати їх архітектуру та використання. Розуміти ETL-процеси (добування, перетворення й завантаження даних) та особливості архітектури OLAP-систем: MOLAP, ROLAP, HOLAP.	Обґрунтув ання вибору архітектур и даних, методика і практика вибору та використан ня.	10
Разом	20/30			70
Екзамен				30
Всього за курс				100

ПОЛІТИКА ОЦІНЮВАННЯ

Політика щодо дедлайнів та перескладання:	Роботи, які здаються із порушенням термінів без поважних причин, оцінюються на нижчу оцінку. Перескладання відбувається із дозволу лектора за наявності поважних причин (наприклад, лікарняний).
Політика щодо академічної добросовісності:	Списування під час контрольної атестації заборонене (в т. ч. із використанням мобільних девайсів). Реферати, презентації та наукові доповіді повинні мати коректні текстові посилання на використані літературні джерела.
Політика щодо відвідування:	Відвідування занять є обов'язковим. За об'єктивних причин (наприклад, хвороба, закордоне стажування) навчання може відбуватись індивідуально (в онлайн формі за погодженням із деканом факультету)

ШКАЛА ОЦІНЮВАННЯ ЗДОБУВАЧІВ СТУПЕНЯ ДОКТОРА ФІЛОСОФІЇ

Рейтинг здобувача вищої освіти, бали	Оцінка національна за результати складання екзаменів/ заліків	
	екзаменів	заліків
90-100	відмінно	зараховано
74-89	добре	
60-73	задовільно	
0-59	незадовільно	не зараховано

