

**ДЕРЖАВНИЙ ВИЩИЙ НАВЧАЛЬНИЙ ЗАКЛАД
«УЖГОРОДСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ»
ФАКУЛЬТЕТ МАТЕМАТИКИ ТА ЦИФРОВИХ ТЕХНОЛОГІЙ
Кафедра кібернетики і прикладної математики**

«ЗАТВЕРДЖУЮ»
Декан факультету математики та
цифрових технологій
_____ /Маляр М. М./
« ____ » _____ 20__ року

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Моделі та методи Data Mining

Рівень вищої освіти	бакалавр
Галузь знань	11 Математика та статистика
Спеціальність	113 Прикладна математика
Предметна спеціальність (Спеціалізація) <i>(за наявності)</i>	
Освітня програма	Системи штучного інтелекту
Статус дисципліни	нормативна
Мова навчання	українська

Ужгород 2021

Робоча програма навчальної дисципліни «**Моделі та методи Data Mining**» для здобувачів вищої освіти галузі знань **11 Математика та статистика** спеціальності **113 Прикладна математика** освітньої програми **Системи штучного інтелекту**.

Розробник: Кондрук Н. Е., доцент, к.т.н, доцент кафедри кібернетики і прикладної математики

Робочу програму розглянуто та затверджено на засіданні

Кафедри кібернетики і прикладної математики

протокол № 17 від «29» червня 2021 р.

Завідувач кафедри _____ Гече Ф.Е.

Схвалено науково-методичною комісією факультету математики та цифрових технологій

протокол № 8 від «30» червня 2021 р.

Голова науково-методичної комісії _____ Мулеса О.Ю.

© Кондрук Н.Е., 2021 р.

© ДВНЗ «Ужгородський національний університет», 2021 р.

1. ОПИС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Найменування показників	Розподіл годин за навчальним планом	
	Денна форма навчання	Заочна форма навчання
Кількість кредитів ЄКТС – 5	Рік підготовки:	
Загальна кількість годин – 150	2-й	-й
Кількість модулів – 2	Семестр:	
Тижневих годин для денної форми навчання: аудиторних – 74 самостійної роботи студента – 76	4-й	-й
	Лекції:	
	32	
	Практичні (семінарські):	
Вид підсумкового контролю: екзамен	Лабораторні:	
	42	
Форма підсумкового контролю: екзамен	Самостійна робота:	
	76	

2. МЕТА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Метою вивчення навчальної дисципліни Моделі та методи Data Mining є вивчення методів інтелектуального аналізу даних, спрямованих на аналітичне дослідження масивів інформації з метою виявлення нових раніше невідомих, практично корисних знань і закономірностей, необхідних для прийняття рішень; огляд методів та алгоритмів, які використовуються в Data Mining; розгляд практичних прикладів застосування Data Mining; підготовка студентів до самостійної роботи з вирішення задач засобами Data Mining.

Відповідно до освітньої програми, вивчення дисципліни сприяє формуванню у здобувачів вищої освіти таких компетентностей:

- ЗК01. Здатність учитися і оволодівати сучасними знаннями.
- ЗК02. Здатність застосовувати знання у практичних ситуаціях.
- ЗК03. Здатність генерувати нові ідеї (креативність).
- ЗК04. Здатність бути критичним і самокритичним.
- ЗК05. Здатність проведення досліджень на відповідному рівні.
- ЗК06. Здатність до абстрактного мислення, аналізу та синтезу.
- ЗК08. Знання та розуміння предметної області та розуміння професійної діяльності.
- ЗК10. Навички у використанні інформаційних і комунікаційних технологій.
- ЗК11. Здатність працювати в міжнародному контексті.
- ЗК16. Здатність до планування та розподілу часу.
- ЗК17. Здатність приймати обґрунтовані рішення.
- ЗК18. Здатність до розроблення та управління проектами.
- ФК03. Здатність обирати та застосовувати математичні методи для розв'язання прикладних задач, моделювання, аналізу, проектування, керування, прогнозування, прийняття рішень.
- ФК09. Здатність до проведення математичного і комп'ютерного моделювання, аналізу та обробки даних, обчислювального експерименту, розв'язання формалізованих задач за допомогою спеціалізованих програмних засобів.
- ФК13. Здатність зрозуміти постановку завдання, сформульовану мовою певної предметної галузі, здійснювати пошук та збір необхідних вихідних даних.
- ФК17. Здатність вибирати адекватні методи машинного навчання, включаючи методи глибокого навчання, та використовувати їх для вирішення конкретних задач інтелектуального аналізу даних
- ФК19. Здатність проводити маркетингові дослідження та готувати бізнес-аналітику різних процесів в економіці, техніці та фінансовій сфері.

3. ПЕРЕДУМОВИ ДЛЯ ВИВЧЕННЯ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Передумовами вивчення навчальної дисципліни «Інтелектуальні технології Data Mining» є опанування таких навчальних дисциплін (НД) освітньої програми (ОП):

- ОК 9. Теорія ймовірностей
- ОК 11. Функціональний аналіз
- ОК 10. Математична статистика
- ОК 15. Вступ до програмування
- ОК 19. Об'єктно-орієнтоване до програмування

4. ОЧІКУВАНІ РЕЗУЛЬТАТИ НАВЧАННЯ

Відповідно до освітньої програми «Системи штучного інтелекту», вивчення навчальної дисципліни повинно забезпечити досягнення здобувачами вищої освіти таких програмних результатів навчання (ПРН):

Програмні результати навчання	Шифр ПРН
Демонструвати знання й розуміння основних концепцій, принципів, теорій прикладної математики і використовувати їх на практиці.	РН01
Формалізувати задачі, сформульовані мовою певної предметної галузі; формулювати їх математичну постановку та обирати раціональний метод вирішення; розв'язувати отримані задачі аналітичними та чисельними методами, оцінювати точність та достовірність отриманих результатів	РН03
Вміти проводити практичні дослідження та знаходити розв'язок некоректних задач.	РН07
Володіти методиками вибору раціональних методів та алгоритмів розв'язання математичних задач оптимізації, дослідження операцій, оптимального керування і прийняття рішень, аналізу даних.	РН10
Використовувати в практичній роботі спеціалізовані програмні продукти та програмні системи комп'ютерної математики.	РН13
Виявляти здатність до самонавчання та продовження професійного розвитку.	РН14
Уміти організувати власну діяльність та одержувати результат у рамках обмеженого часу.	РН15
Вміти застосовувати методи та алгоритми Data Mining для задач класифікації, прогнозування, кластерного аналізу, пошуку асоціативних правил з використанням програмних інструментів підтримки багатовимірного аналізу даних.	РН21

Очікувані результати навчання, які повинні бути досягнуті здобувачами освіти після опанування навчальної дисципліни «Моделі і методи Data Mining»:

Очікувані результати навчання з дисципліни	Шифр ПРН
Розуміння основних класів задач інтелектуального аналізу даних	РН01, РН14
Застосування відповідних методів інтелектуального аналізу даних до певних класів прикладних задач (за наявності комп'ютерної техніки та відповідного ПЗ)	РН01, РН03, РН10, РН13, РН07, РН15, РН14, РН15
Аналіз та оцінка отриманих результатів прикладних досліджень	РН03, РН07
Застосовувати можливості табличного редактора Excel для розв'язання окремих класів задач інтелектуального аналізу даних	РН13, РН21

5. ЗАСОБИ ДІАГНОСТИКИ ТА КРИТЕРІЇ ОЦІНЮВАННЯ РЕЗУЛЬТАТІВ НАВЧАННЯ

Засоби оцінювання та методи демонстрування результатів навчання

Засобами оцінювання та методами демонстрування результатів навчання з навчальної дисципліни є:

командні проекти; розрахункові та розрахунково-графічні роботи; завдання на комп'ютерному обладнанні.

Форми контролю та критерії оцінювання результатів навчання

Форми поточного контролю: виконання та захист лабораторних завдань;

Форма модульного контролю: модульна контрольна робота;

Форма підсумкового семестрового контролю: екзамен.

Розподіл балів, які отримують здобувачі вищої освіти (модуль 1)

Поточне оцінювання та самостійна робота			Модульна контрольна робота	Сума
T1	T2	T3	30	100
0	30	40		

Розподіл балів, які отримують здобувачі вищої освіти (модуль 2)

Поточне оцінювання та самостійна робота			Модульна контрольна робота	Сума
T1	T2	T3	30	100
30	20	20		

Оцінювання окремих видів навчальної роботи з дисципліни

Вид діяльності здобувача вищої освіти	Модуль 1		Модуль 2	
	Кількість	Максимальна кількість балів (сумарна)	Кількість	Максимальна кількість балів (сумарна)
Лабораторні заняття		70		70
Модульна контрольна робота		30		30
Разом		100		100

Критерії оцінювання модульної контрольної роботи

Модульна контрольна робота складається із 3-ох теоретичних питань, кожне з яких оцінюється в 10 балів.

Критерії оцінювання підсумкового семестрового контролю

Сума балів за всі види навчальної діяльності	Оцінка ECTS	Оцінка за національною шкалою	
		для екзамену, курсового проєкту (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

6. ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

6.1. Зміст навчальної дисципліни

Модуль 1

Тема 1. Суть, мета та сфера застосування технології Data Mining. Класи систем інтелектуального аналізу даних. Класи задач Data Mining. Види атрибутів векторів ознак.

Тема 2. Задача та методи класифікації (kNN, Наївний Баєсовський класифікатор). Види відстаней між об'єктами. Методи нормалізації даних.

Тема 3. Дерева прийняття рішень (Decision trees). Алгоритм ID3. Дерево регресії CART.

Модуль 2

Тема 1. Задача кластеризації. Центроїдні методи кластерного аналізу (k-means). Ієрархічна кластеризація. Види відстаней між кластерами. Просторова кластеризація на основі щільності DBScan.

Тема 2. Методи прогнозування. Методи оцінки точності прогнозів. Визначення тренду.

Тема 3. Асоціативні правила. Метод Аргіогі. Послідовна асоціація. Метод GSP.

6.2. Структура навчальної дисципліни

Назви змістових модулів і тем	Кількість годин											
	Денна форма						Заочна форма					
	Усього	у тому числі					Усього	у тому числі				
		л	п	лаб	інд	ср		л	п	лаб	інд	ср
1	2	3	4	5	6	7	8	9	10	11	12	13
2-й семестр												
Модуль 1.												
Тема 1.		3				4						
Тема 2.		4		8		10						
Тема 3.		5		8		10						
Модульна контрольна робота		2				6						
<i>Разом за модуль 1</i>		<i>14</i>		<i>16</i>		<i>30</i>						
Модуль 2.												
Тема 1.		6		10		16						
Тема 2.		6		8		10						
Тема 3.		4		8		10						
Модульна контрольна робота		2				10						
<i>Разом за модуль 2</i>		<i>18</i>		<i>26</i>		<i>46</i>						
Разом за семестр		32		42		76						

6.3. Теми лабораторних занять

№ з/п	Назва теми	Кількість годин	
		денна	заочна
1	Метод класифікації kNN	6	
2	Наївний Баєсовський класифікатор	4	

3	Алгоритм ID3	4	
4	Дерево регресії CART	4	
5	Методи кластеризації	10	
6	Методи прогнозування	6	
7	Метод Apriori	4	
8	Метод GSP	4	
	Разом	42	

7. ІНСТРУМЕНТИ, ОБЛАДНАННЯ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ, ВИКОРИСТАННЯ ЯКИХ ПЕРЕДБАЧАЄ НАВЧАЛЬНА ДИСЦИПЛІНА

Технічні засоби – комп’ютер.

Програмне забезпечення: стандартні офісні додатки, обране студентом програмне середовище.

8. РЕКОМЕНДОВАНІ ДЖЕРЕЛА ІНФОРМАЦІЇ

Основна література

1. Ситник В. Ф. Інтелектуальний аналіз даних (дейтамайнінг): Навч. Посібник/ В. Ф. Ситник, М.Т. Краснюк - К: КНЕУ, 2007. - 376 с.
2. Черняк О.І.. Інтелектуальний аналіз даних: Підручник / О.І. Черняк, П.В. Захарченко ; Київ. нац. ун-т ім. Т. Шевченка. — К. : Знання, 2014. — 599 с.
3. Барсегян, А. А. Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP / А. А. Барсегян, М. С. Куприянов, В. В. Степаненко, И. И. Холод - 2-е изд., перераб. и доп. СПб.: БХВ-Петербург, 2010. -384 с.
4. Дюк В. Data Mining: учебный курс / В. Дюк, А. Самойленко - СПб: Питер, 2001. -368 с.

Допоміжна література

1. Корнеев В.В. Базы данных. Интеллектуальная обработка информации / В.В. Корнеев., А.Ф. Гареев, С.В. Васютин, В.В. Райх – М.: «Нолидж», 2000. – 352 с.
2. Кацко И.А. Практикум по анализу данных на компьютере / И.А.Кацко, Н.Б. Паклин – М.: Колос, 2009. – 278 с.
3. Барсегян А. А. Анализ данных и процессов: учеб. пособие / А.А. Барсегян, М. С. Куприянов, И. И. Холод, М. Д. Тесс, С. И. Елизаров. – 3-е изд., перераб. доп. – СПб.: БХВ-Петербург, 2009. – 512 с.

Інформаційні ресурси в мережі Інтернет

1. Бізнес-аналітика: від даних до знань. Доступно з електронного джерела: <http://padaread.com/?book=52105>.
2. Візуалізація кластеризації. Доступно з електронного джерела: <https://www.naftaliharris.com/blog/visualizing-k-means-clustering/>

**Результати перегляду
робочої програми навчальної дисципліни**

Робоча програма перезатверджена на 20__ / 20__ н.р. без змін; зі змінами (Додаток __).
(потрібно підкреслити)

протокол № __ від «__» _____ 20__ р. Завідувач кафедри _____
(підпис) (Прізвище ініціали)

Робоча програма перезатверджена на 20__ / 20__ н.р. без змін; зі змінами (Додаток __).
(потрібно підкреслити)

протокол № __ від «__» _____ 20__ р. Завідувач кафедри _____
(підпис) (Прізвище ініціали)

Робоча програма перезатверджена на 20__ / 20__ н.р. без змін; зі змінами (Додаток __).
(потрібно підкреслити)

протокол № __ від «__» _____ 20__ р. Завідувач кафедри _____
(підпис) (Прізвище ініціали)

Робоча програма перезатверджена на 20__ / 20__ н.р. без змін; зі змінами (Додаток __).
(потрібно підкреслити)

протокол № __ від «__» _____ 20__ р. Завідувач кафедри _____
(підпис) (Прізвище ініціали)