

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**

**ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ імені ІВАНА ФРАНКА**

Кафедра (циклова комісія) \_\_\_\_\_ радіофізики та комп'ютерних технологій \_\_\_\_\_

**“ЗАТВЕРДЖУЮ”**

**Декан факультету**

**доц. Юрій ФУРГАЛА**

“ \_\_\_\_\_ ” \_\_\_\_\_ 2022 року

**РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ**

**ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ДАНИХ**

(шифр і назва навчальної дисципліни)

спеціальність \_\_\_\_\_ **122 – Комп'ютерні науки** \_\_\_\_\_

(шифр і назва спеціальності)

спеціалізація \_\_\_\_\_

(назва спеціалізації)

факультет \_\_\_\_\_ **електроніки та комп'ютерних технологій** \_\_\_\_\_

(назва інституту, факультету, відділення)

2022 – 2023 навчальний рік

Робоча програма “Інтелектуальний аналіз даних” для студентів

(назва навчальної дисципліни)

галузі знань “12 – Інформаційні технології”

за спеціальністю “122 Комп’ютерні науки”

Розробники: Зіновій ЛЮБУНЬ (кандидат технічних наук, доцент,

доцент кафедри радіофізики та комп’ютерних технологій)

(вказати авторів, їхні посади, наукові ступені та вчені звання)

Робочу програму схвалено на засіданні кафедри (циклової комісії) \_\_\_\_\_

радіофізики та комп’ютерних технологій

Протокол від “\_\_\_\_” \_\_\_\_\_ 2022 року № \_\_\_\_

Завідувач кафедри радіофізики та комп’ютерних технологій

( Іван КАРБОВНИК )

(підпис)

(прізвище та ініціали)

Ухвалено Вченою радою \_\_\_\_\_ факультету електроніки та комп’ютерних технологій

Протокол від “\_\_\_\_” \_\_\_\_\_ 2022 року № \_\_\_\_

© Любунь З., 2022 р

© ЛНУ ім. І. Франка, 2022 рік

## 1. Опис навчальної дисципліни

(Витяг з робочої програми навчальної дисципліни “ Інтелектуальний аналіз даних ”)

Найменування показників	Галузь знань, напрям підготовки, освітньо-кваліфікаційний рівень	Характеристика навчальної дисципліни	
		<i>денна форма навчання</i>	
Кількість кредитів, – 5	Галузь знань <b>12</b> <b>"Інформаційні технології"</b>	Дисципліна професійно-практичної підготовки	
Модулів – 2	Спеціальність <b>122</b> <b>" Комп'ютерні науки "</b>	<i>Рік підготовки:</i> 3-й	
Змістових модулів – 3	Освітньо-кваліфікаційний рівень: <b>бакалавр</b>	<i>Семестр</i> 5-й	
Курсова робота –		<i>Лекції</i> 32год.	
Загальна кількість годин – 150			
Тижневих годин для денної форми навчання: <i>Аудиторних:</i> 5 семестр – 4 <i>Самостійної роботи студента:</i> 5.4 5 семестр – 1.6.			<i>Практичні -</i>
			<i>Лабораторні</i> 32 год.
	<i>Самостійна робота</i> 86 год.		
	ІНДЗ:		
	<i>Вид контролю: залік</i>		

Примітка.

Співвідношення кількості годин аудиторних занять до самостійної і індивідуальної роботи становить

для денної форми навчання – 0,74

для заочної форми навчання – немає

## 2. Мета та завдання навчальної дисципліни

Курс Інтелектуальний аналіз даних (ІАД) є дисципліною професійно-практичної підготовки студента.

**Мета:** Засвоєння студентами основних теоретичних відомостей та практичних вмінь з курсу. Підготувати студента до ефективного використання як класичних так і сучасних методів інтелектуального аналізу даних та обробки інформації з використанням класичних методів обчислень та нейромереж задля створення елементів автоматизованих комп'ютерних систем та їх використання у подальшій професійній діяльності; допомогти набутти навички практичної роботи з програмними засобами для інтелектуального аналізу даних.

. **Цілі:** Формування у студентів розуміння побудови програм для вирішення задач аналізу даних.

У результаті вивчення навчальної дисципліни студент повинен

**знати:** основні поняття та визначення інтелектуального аналізу даних, моделі та методи побудови моделей та аналізу залежностей у даних, сучасні програмні засоби для проектування і розробки систем інтелектуального аналізу даних, критерії порівняння моделей і методів інтелектуального аналізу даних.

**вміти:** обґрунтовувати й аналізувати вибір конкретного типу моделі та методу інтелектуального аналізу даних при вирішенні практичних задач, використовувати сучасні програмні засоби для проектування та дослідження систем інтелектуального аналізу даних, створювати програми для інтелектуального аналізу даних при розв'язку конкретних практичних задач, аналізувати результати побудови та використання систем інтелектуального аналізу даних при вирішенні прикладних задач.

Після вивчення курсу «Інтелектуальний аналіз даних» здобувачі набувають таких Загальних та Фахових компетентностей та Програмних результатів навчання:

ЗК1. Здатність до абстрактного мислення, аналізу та синтезу.

ЗК2. Здатність застосовувати знання у практичних ситуаціях.

ЗК5. Здатність вчитися і оволодівати сучасними знаннями.

ЗК6. Здатність до пошуку, оброблення та аналізу інформації з різних джерел.

С11. Здатність до інтелектуального аналізу даних на основі методів обчислювального інтелекту включно з великими та погано структурованими даними, їхньої оперативної обробки та візуалізації результатів аналізу.

ФК29. Здатність до алгоритмічного та логічного мислення.

ПР1. Застосовувати знання основних форм і законів абстрактно-логічного мислення, основ методології наукового пізнання, форм і методів аналізу, обробки та синтезу інформації в предметній області комп'ютерних наук.

**Місце в структурно-логічній схемі спеціальності.** Для вивчення дисципліни необхідні знання з таких предметів: “Вища математика”, та “Алгоритмізація та програмування”, “Об’єктно орієнтоване програмування”, СУБД.

### 3. Програма навчальної дисципліни

#### МОДУЛЬ 1

##### *Змістовий модуль 1. Традиційні технології ІАД.*

##### **Тема 1. Загальні поняття про ІАД.**

Означення ІАД. Задачі ІАД. Етапи вирішення задач ІАД. Методи ІАД.

**Data Mining.** Представлення нової технології інтелектуального аналізу даних. Визначення Data Mining. Можливості інтелектуального аналізу. Недоліки технології інтелектуального аналізу даних. Data Mining і OLAP. Сфера застосування технологій інтелектуальних обчислень. Бізнес-застосування Data Mining

##### **Тема 2. Основи ІАД.**

Методи первісної обробки даних. Методи дослідження структури даних: візуалізація та автоматичне групування даних

##### **Тема 3. Методи використання навчальної інформації. Імовірнісні технології.**

Статистичні методи аналізу даних. Визначення математичних очікувань та дисперсій для одновимірних послідовностей випадкових величин. Кореляційний і регресійний аналіз даних.

**Тема 4. Методи багатомірного розвідувального аналізу**  
Кластерний аналіз. Ієрархічна та секційна кластеризації.

**Змістовий модуль 2. Нові технології інтелектуальних обчислень.**

**Тема 5. Методи класифікації та прогнозування.**

Прогнозування часових послідовностей (наївні моделі прогнозу).

Прогнозування часових послідовностей (Метод Хольта-Брауна)

Прогнозування часових послідовностей (лінійний та параболічний тренди)

**Тема 6. Методи пошуку шаблонів даних.**

Асоціаційні правила. Послідовне відображення шаблонів даних. Метод Аргіогі, побудова FP-дерев пошуку шаблонів даних. Min-max асоціації у базах даних.

**Змістовий модуль 3. Методи інтелектуальних обчислень.**

**Тема 7. Детерміновані технології.**

Аналітичні обчислення на основі відомих закономірностей. Методи апроксимації та екстраполяції залежностей заданих аналітично. Метод найменших квадратів з точки зору виявлення закономірностей. Загальні поняття про методи дослідження операцій.

Градентні методи знаходження екстремумів функцій заданих аналітично.

**Тема 8. Процес знаходження нового знання.**

Визначення проблеми (постановка задачі). Збір та підготовка даних. Оцінка даних. Об'єднання й очищення даних. Відбір даних. Перетворення. Побудова моделі. Оцінка й інтерпретація. Зовнішня перевірка. Використання моделі. Спостереження за моделлю.

**Тема 9. Нейронні мережі.**

Виявлення закономірностей у багатомірному потоку даних за допомогою нейромереж.

Навчання без вчителя. Пониження розмірності даних. Карти Кохонена. Кластеризація даних.

Приклади використання нейромереж для вирішення прикладних задач ІАД.

**Тема 10. Деревя рішень.**

Технологія використання дерев прийняття рішень. Типи задач прийняття рішень, формальні означення дерева рішень і системи прийняття рішень. Приклади дерев рішень та їх використання для вирішення прикладних задач ІАД. Алгоритм ID3 побудови дерева рішень та його застосування. Методика проектування бази знань на основі використання дерев рішень із застосуванням алгоритму ID3.

**Тема 11. Нечітка логіка .**

Основні моделі та методи синтезу нейро-нечітких систем. Нейро-нечіткий апроксиматор Мамдані .Принципи побудови нейро-нечітких мереж за допомогою програмних засобів. Використання нейро-нечітких моделей для інтелектуального аналізу даних. Приклади прикладних задач.

**Тема 12. Генетичні алгоритми.**

Історія появи еволюційних алгоритмів. Генетичні алгоритми і традиційні методи оптимізації

Основні поняття генетичних алгоритмів. Класичний генетичний алгоритм. Модифікації класичного ГА. Приклади використання генетичних алгоритмів.

#### 4. Структура навчальної дисципліни

Назви змістових модулів і тем	Кількість годин					
	Денна форма					
	Усього	у тому числі				
		лк	пр	лаб	інд	ср
1	2	3	4	5	6	7
<b>МОДУЛЬ 1</b>						
<b>Змістовий модуль 1. Традиційні технології ІАД.</b>						
Тема 1. Означення ІАД. Задачі ІАД. Етапи вирішення задач ІАД. Методи ІАД. <b>Data Mining.</b> Представлення нової технології інтелектуального аналізу даних. Визначення Data Mining. Можливості інтелектуального аналізу. Недоліки технології інтелектуального аналізу даних. Data Mining і OLAP. Сфера застосування технологій інтелектуальних обчислень. Бізнес-застосування Data Mining	6	2				4
Тема 2. Методи первісної обробки даних. Методи дослідження структури даних: візуалізація та автоматичне групування даних	8	2		4		6
Тема 3. Статистичні методи аналізу даних. Визначення математичних очікувань та дисперсій для одновимірних послідовностей випадкових величин. Кореляційний і регресійний аналіз даних.	10	2		2		6
Тема 4. Кластерний аналіз. Ієрархічна та секційна кластеризації.	18	2		4		12
<i>Разом – змістовий модуль 1</i>	46	8		10		28
<b>Змістовий модуль 2. Нові технології інтелектуальних обчислень.</b>						
Тема 5. Прогнозування часових послідовностей (наївні моделі прогнозу). Прогнозування часових послідовностей (Метод Хольта-Брауна) Прогнозування часових послідовностей (лінійний та параболічний тренди)	24	4		8		12
Тема 6 Асоціаційні правила. Послідовне відображення шаблонів даних. Метод Аргіогі, побудова FP-дерев пошуку шаблонів даних. Min-тах асоціації у базах даних.	4	2				2
<i>Разом – змістовий модуль 2</i>	28	6		8		14

<b>Усього годин</b>	<b>43</b>	<b>14</b>		<b>18</b>		<b>42</b>
<b>МОДУЛЬ 2</b>						
<b>Змістовий модуль 3. Методи інтелектуальних обчислень.</b>						
Тема 7. Аналітичні обчислення на основі відомих закономірностей. Методи апроксимації та екстраполяції залежностей заданих аналітично. Метод найменших квадратів з точки зору виявлення закономірностей. Загальні поняття про методи дослідження операцій. Градієнтні методи знаходження екстремумів функцій заданих аналітично.	10	4		2		6
Тема 8. Визначення проблеми (постановка задачі). Збір та підготовка даних. Оцінка даних. Об'єднання й очищення даних. Відбір даних. Перетворення. Побудова моделі. Оцінка й інтерпретація. Зовнішня перевірка. Використання моделі. Спостереження за моделлю.	2	2				2
Тема 9. Виявлення закономірностей у багатомірному потоку даних за допомогою нейромереж. Навчання без вчителя. Пониження розмірності даних. Карти Кохонена. Кластеризація даних. Приклади використання нейромереж для вирішення прикладних задач ІАД.	24	4		8		12
Тема 10 Технологія використання дерев прийняття рішень. Типи задач прийняття рішень, формальні означення дерева рішень і системи прийняття рішень. Приклади дерев рішень та їх використання для вирішення прикладних задач ІАД. Алгоритм ID3 побудови дерева рішень та його застосування. Методика проектування бази знань на основі використання дерев рішень із застосуванням алгоритму ID3.	2	2				2
Тема 11 Основні моделі та методи синтезу нейро-нечітких систем. Нейро-нечіткий апроксиматор Мамдані .Принципи побудови нейро-нечітких мереж за допомогою програмних засобів. Використання нейро-нечітких моделей для інтелектуального аналізу даних. Приклади прикладних задач.	10	2				8
Тема 12. Генетичні алгоритми.	22	4		4		14

Історія появи еволюційних алгоритмів. Генетичні алгоритми і традиційні методи оптимізації Основні поняття генетичних алгоритмів. Класичний генетичний алгоритм. Модифікації класичного ГА. Приклади використання генетичних алгоритмів.						
<i>Разом – змістовний модуль 3</i>	<i>76</i>	<i>18</i>		<i>14</i>		<i>44</i>
<b>Усього годин</b>	<b>150</b>	<b>32</b>		<b>32</b>		<b>86</b>

## 5. Теми семінарських занять

## 6. Теми практичних занять

## 7. Теми лабораторних занять

№ з/п	Назва теми	Кількість годин
1.	Вступне заняття. Елементарні задачі аналізу даних.	2
2.	Методи первісної обробки даних 5	2
3.	Використання методу найменших квадратів для згладжування даних	2
4.	. Візуалізація даних. Автоматичне групування даних	2
5.	Визначення числових характеристик статистичного розподілу одновимірної послідовності випадкових величин	2
6.	Кластеризація даних	4
7.	Прогнозування часових послідовностей (наївні моделі прогнозу)	2
8	Прогнозування часових послідовностей (Метод Хольта-Брауна)	2
10	Прогнозування часових послідовностей (лінійний та параболічний тренди)	4
11	Методи одно та багатовимірної оптимізації	2
12	Вирішення задачі класифікації за допомогою однеї нейронної мережі	4
13	Вирішення задачі кластеризації за допомогою нейронної мережі	2
14	Заключне заняття	2
	Разом	32

## 6. Самостійна робота

№ з/п	Назва теми	Кількість годин
1	Загальні поняття про ІАД.	6



2	Детерміновані технології.	10
3	Імовірнісні технології.	12
4	СУБД та ІАД.	19
5	Data Mining.	18
6	Основні моделі технологій інтелектуальних обчислень.	6
7	Процес знаходження нового знання.	4
8	Нейронні мережі.	4
9	Дерева рішень.	6
10	Системи міркування на основі аналогічних випадків.	6
11	Алгоритми виявлення асоціацій.	4
12	Нечітка логіка .	6
13	Генетичні алгоритми.	14
	Разом	86

## 9. Індивідуальні завдання

### 10. Методи навчання

Інформаційні методи (лекція, бесіда, ілюстрація, демонстрація); дедуктивні методи на основі узагальнень; евристичні методи (проблемна лекція); інтерактивні методи (дискусія).

### 11. Методи контролю

Поточний контроль здійснюється шляхом проведення усного опитування та написання письмових звітів по виконаних лабораторних роботах. У кінці курсу проводиться залік.

### 12. Розподіл балів, які отримують студенти

Оцінювання проводиться упродовж семестру та під час екзаменаційної сесії за 100-бальною шкалою. Бали нараховуються за такими видами робіт із таким співвідношенням:

- лабораторні роботи: 100% семестрової оцінки; максимальна кількість балів 100. Загалом 100 балів.

#### Шкала оцінювання: національна та ЄКТС

Оцінка ЄКТС	Сума балів за всі види навчальної діяльності	Оцінка за національною шкалою	
		для екзамену, курсового проекту (роботи), практики	для заліку
A	90 – 100	відмінно	зараховано
B	81-89	добре	
C	71-80		
D	61-70		
E	51-60	задовільно	
FX	21-50	незадовільно з можливістю повторного складання	не зараховано з можливістю повторного складання
F	0-20	незадовільно з обов'язковим повторним вивченням дисципліни	не зараховано з обов'язковим повторним вивченням дисципліни

\* кількість балів для оцінок «незадовільно» (FX і F) визначається Вченими радами факультетів (педагогічними радами коледжів).

### 13. Методичне забезпечення

1. Любунь З. М., Рабик В. Г., Карбовник І. Д. Інтелектуальний аналіз даних. Методичні вказівки до виконання лабораторних робіт для студентів за напрямом підготовки 6.050101 “Комп’ютерні науки”. – Львів: Видавничий центр ЛНУ імені Івана Франка, 2015. –70 с.
2. Любунь З. М. Основи теорії нейромереж / З. М. Любунь /: Текст лекцій. – Львів: Видавничий центр ЛНУ імені Івана Франка, 2007. –142 с.

### 14. Рекомендована література

#### Основна

1. Черняк О.І. Інтелектуальний аналіз даних: підручник. – К: Знання, 2014. –599с.
2. Олійник А. О. Інтелектуальний аналіз даних : навчальний посібник. Запоріжжя : ЗНТУ, 2012. 278 с .
3. Нікольський Ю., Щербина Ю., Якимечко Р. Дерева прийняття рішень та їхнє застосування для прогнозування діагнозу у медицині // Вісник Львівського університету. Серія прикладна математика та інформатика. Випуск 6, 2003.- С. 191-211
4. І. А. Терейковський, Д. А. Бушуєв, Л. О. Терейковська. Штучні нейронні мережі: базові положення: навчальний посібник – Київ : КПІ ім. Ігоря Сікорського, 2022. – 123 с.
5. Хайкін С. Нейронні мережі. Повний курс. Друге видання. - М .: Вільямс, 2006. - 1104 с.

#### Додаткова

1. Liubun Z. Hover Signal-Profile Detection / Liubun, V. Mandziy, H. Klein, O. Karpin, V. Rabyk // Proceedings of the XV International Scientific and Technical Conference “Computer Science and Information Technologies” – 2020. P. 7 – 10. (Scopus)
2. Karpin O. Method of Neural Network Training with Integer Weights / O. Karpin, Mandziy, Z. Liubun, V. Rabyk // Proceedings of the XI<sup>th</sup> International Scientific and Practical Conference “Electronics and Information Technologies” (ELIT – 2019), September 16 – 18, 2019, Lviv, Ukraine. P. 168 – 172. doi: [10.1109/ELIT.2019.8893349](https://doi.org/10.1109/ELIT.2019.8893349).
3. З. Любунь, Р. Лісовський, І. Поплавський, Б. Рачій. Прогнозування властивостей вуглецевих матеріалів з використанням нейронних мереж. // Електроніка та інформаційні технології. – 2019. – Випуск 12. – С. 64–72.

### 15. Інформаційні ресурси

1. Internet джерела.
2. Наукова бібліотека Львівського національного університету імені Івана Франка (<https://www.lnulibrary.lviv.ua/to-users-2/paid-services/internet/>).
3. Львівська національна наукова бібліотека України імені Василя Стефаника (<https://www.lsl.lviv.ua/index.php/uk/elektronni-resursy1/>).