

**Міністерство освіти і науки України**  
**Львівський національний університет імені Івана Франка**  
**Кафедра радіофізики та комп'ютерних технологій**

**“ЗАТВЕРДЖУЮ”**

Проректор з науково-педагогічної роботи

\_\_\_\_\_ Височанський В.С.

“ \_\_\_\_\_ ” \_\_\_\_\_ 2017 р.

**РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ**

**Інтелектуальний аналіз даних**

Галузі знань 12 " Інформаційні технології ", спеціальності 122 " Комп'ютерні науки ",  
факультету електроніки та комп'ютерних технологій

**Кредитно-модульна система  
організації навчального процесу**

**Львів – 2017**

**Інтелектуальний аналіз даних.** Робоча програма навчальної дисципліни для студентів галузі знань 12 "Інформаційні технології" спеціальності 122 "Комп'ютерні науки" факультету електроніки та комп'ютерних технологій, - Львів: ЛНУ імені Івана Франка, 2016. - 11с.

**Розробник:**

**Любунь З.М.** канд. технічних наук, доцент кафедри радіофізики

Робоча програма затверджена на засіданні кафедри радіофізики

Протокол № \_\_\_ від. " \_\_\_ " \_\_\_\_\_ 2016 р.

Завідувач кафедри радіофізики

Болеста І. М.

" \_\_\_ " \_\_\_\_\_ 2016р

Схвалено методичною радою факультету електроніки

Протокол № \_\_\_ від. " \_\_\_ " \_\_\_\_\_ 20\_\_ р.

Голова методичної ради

Шувар Р. Я.

" \_\_\_ " \_\_\_\_\_ 2016 р

## 1. Опис навчальної дисципліни

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

<b>Найменування показників</b>	<b>Галузь знань, напрям підготовки, освітньо-кваліфікаційний рівень</b>	<b>Характеристика навчальної дисципліни</b>	
		<i>денна форма навчання</i>	
Кількість кредитів, – 3	<b>Галузь знань 12 "Інформаційні технології"</b>	Дисципліна професійно-практичної підготовки	
Модулів – 2	<b>Спеціальність 122 " Комп'ютерні науки "</b>	<i>Рік підготовки: 3-й</i>	
Змістових модулів – 3		<i>Семестр 5-й</i>	
Курсова робота –			
Загальна кількість годин – 90			<i>Лекції 32год.</i>
Тижневих годин для денної форми навчання: <i>Аудиторних: 5 семестр – 4 Самостійної роботи студента: 5 семестр – 1.6.</i>			<i>Практичні -</i>
			<i>Лабораторні 32 год.</i>
	<i>Самостійна робота 26 год.</i>		
	<b>ІНДЗ:</b>		
	<i>Вид контролю: іспит</i>		

## 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 набути навички практичної роботи з програмними засобами для інтелектуального аналізу даних.

**Завдання:** Після вивчення даної 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.

**Повинен 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.

**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 рішення.**

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	4	2				2
Тема 2. Методи первісної обробки даних. Методи дослідження структури даних: візуалізація та автоматичне групування даних	5	2		4		1
Тема 3. Статистичні методи аналізу даних. Визначення математичних очікувань та дисперсій для одновимірних послідовностей випадкових величин. Кореляційний і регресійний аналіз даних.	10	2		2		2
Тема 4. Кластерний аналіз. Ієрархічна та секційна кластеризації.	10	2		4		2
<i>Разом – змістовий модуль 1</i>	25	8		10		7
<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)	4	4		8		2
Тема 6 Асоціаційні правила. Послідовне відображення шаблонів даних. Метод Аргіогі, побудова FP-дерев пошуку шаблонів даних. Min-max асоціації у базах даних.	4	2				2
<i>Разом – змістовий модуль 2</i>	18	6		8		4

Усього годин	43	14		18		11
<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.	6	4		2		2
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.	14	2				2
Тема 9. Виявлення закономірностей у багатомірному потоку даних за допомогою нейромереж. Навчання без вчителя. Пониження розмірності даних. Карті Кохонена. Кластеризація даних. Приклади використання нейромереж для вирішення прикладних задач ІАД.	16	4		8		2
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.	8	2				4
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.	6	2				4
Тема 12. Генетичні алгоритми.	16	4		4		1

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>84</i>	<i>18</i>		<i>14</i>		<i>15</i>
<b>Усього годин</b>	<b>144</b>	<b>32</b>		<b>32</b>		<b>26</b>

### *5. Темі 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 одної нейронної мережі	4
13	Лабораторна робота № 12. Вирішення задачі кластеризації за допомогою нейронної мережі	2
14	Заключне заняття	2



## 6. Самостійна робота

№ з/п	Назва теми	Кількість годин
1	Загальні поняття про ІАД.	6
2	Детерміновані технології.	8
3	Імовірнісні технології.	2
4	СУБД та ІАД.	4
5	Data Mining.	8
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.	8

## 7. Методи контролю

**Звітність – іспит.** Поточний контроль знань студентів здійснюється шляхом усного та письмового контролю за матеріалами лекцій. Кожна лабораторна робота розпочинається із перевірки теоретичного матеріалу що відноситься до теми лабораторної роботи та аналізу моделюючої програми, яку студент розробив самостійно. Підсумковим контролем знань є **з-іспит**.

### *11. Розподіл балів, що присвоюється студентам для іспиту*

Підсумкова кількість балів виставляється по наступним критеріям:

- обов'язкове виконання 13 лабораторних робіт;
- іспит.

Бали виставляються по наступним критеріям.

1. *Виконання лабораторних робіт.* Максимальна кількість балів – 50 бали. Включає в себе:
  - максимальна кількість балів за кожну лабораторну роботу – 2-4 бали,
  - при оцінюванні лабораторної роботи враховується
    - підготовка до виконання лабораторної роботи (0.5),
    - своєчасна здача і якість оформлення звіту (1)
    - захист звіту про виконану лабораторну роботу (0.5-2.5),
    -
2. Іспит проводиться в письмовій формі з наступною при необхідності усною співбесідою Максимальна кількість балів – 50 балів

Лабораторні роботи	Іспит	Сума
50	50	100

## Шкала оцінювання: Університету, національна та ECTS

<i>Оцінка в балах</i>	<i>Оцінка ECTS</i>	<i>Визначення</i>	<i>За національною шкалою</i>
			<i>Залік</i>
90–100	<b>A</b>	<i>Відмінно</i>	<i>Зараховано</i>
81-89	<b>B</b>	<i>Дуже добре</i>	
71-80	<b>C</b>	<i>Добре</i>	
61-70	<b>D</b>	<i>Задовільно</i>	
51-60	<b>E</b>	<i>Достатньо</i>	
26–50	<b>FX</b>	<i>Недостатньо</i>	<i>Незараховано</i>
0-25	<b>F</b>	<i>Незадовільно</i>	

### 8. Рекомендована література

#### Базова

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..
3. Дюк В. Data Mining: учебный курс (+CD)/Дюк В., Самойленко А. .. — СПб: Изд. Питер 2001. — 368 с.
4. Акіменко В.В., Загородній Ю.В. Проектування СППР на основі нечіткої логіки. Навчально-методичний посібник. – К.: Вид-во КНУ, 2007. – 94с.
5. Глибовец Н.Н., Медведь С.А. Генетические алгоритмы и их использование для решения задачи составления расписаний //Кибернетика и системный анализ. 2003. – № 1. – С. 95–108.
6. Нікольський Ю., Щербина Ю., Якимечко Р. Древа прийняття рішень та їхнє застосування для прогнозування діагнозу у медицині // Вісник Львівського університету. Серія прикладна математика та інформатика. Випуск 6, 2003.- С. 191-211
7. Ф. Уоссермен. Нейрокомпьютерная техника: Теория и практика Перевод на русский язык, Ю.А. Зуев, В.А. Точенов, 1992.193с.
8. С. Хайкин. Нейронные сети: полный курс, 2-е издание. :Пер. с англ.-М. : Издательский дом «Вильямс», 2006.-1104с.

#### Допоміжна

1. Ф. Уоссермен. Нейрокомпьютерная техника: Теория и практика Перевод на русский язык, Ю.А. Зуев, В.А. Точенов, 1992.193с.
2. Горбань А.Н. Возможности нейронных сетей / Нейроинформатика Новосибирск: Наука, Сибирская издательская фирма РАН, 1998.
3. Ежов А.А., Шумский С.А. Нейрокомпьютинг и его применения в экономике и бизнесе (серия "Учебники экономико-аналитического института МИФИ" под ред. проф. В.В. Харитонов). М.: МИФИ, 1998. - 224 с.

4. *Мандзій В.М.* Розробка нової модифікації нейромережі Гопфілда для некорельованих зображень. ISSN 0474-8662. Відбір і обробка інформ.2004. Вип. 21(97). 100-105с.