

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА**

**ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ
Кафедра інтелектуальних та інформаційних систем**



«ЗАТВЕРДЖУЮ»

Заступник декана
з навчальної роботи

М.О. Мостопалова

« 20 » 03 2018 року

**РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ
ІНТЕЛЕКТУАЛЬНІ ОБЧИСЛЕННЯ
ТА АНАЛІЗ ДАНИХ**

для здобувачів освітньо-наукового рівня «доктор філософії»

галузь знань	для усіх галузей, за якими здійснюється навчання в університеті
спеціальність	для усіх спеціальностей, за якими здійснюється навчання в університеті
освітній рівень	третій (освітньо-науковий)
вид дисципліни	вибіркова

Форма навчання	денна
Навчальний рік	2018/2019
Рік навчання	2
Кількість кредитів ECTS	4
Мова викладання, навчання та оцінювання	українська
Форма заключного контролю	екзамен

Викладач: доктор технічних наук, професор Снитюк Віталій Євгенович

Пролонговано: на 20 19/20 н.р. (В.С. Снитюк) «20» травня 2019 р.
на 20 /20 н.р. () « » 20 р.

КИЇВ – 2018

Розробник: **Снитюк Віталій Євгенович**, д.т.н., проф., професор кафедри інтелектуальних та інформаційних систем

ЗАТВЕРДЖЕНО

Завідувач кафедри інтелектуальних та інформаційних систем

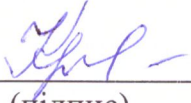


Снитюк В.Є.

Протокол № 4 від 27 грудня 2017 р.

Схвалено науково-методичною комісією факультету інформаційних технологій

Протокол № 5 від 19 березня 2018 року

Голова науково-методичної комісії  доцент, к.т.н. Красовська Г.В.
(підпис)

1. Мета дисципліни формування теоретичних знань та практичних умінь з інтелектуальних обчислень та аналізу даних, що необхідно для побудови сучасних інформаційно-аналітичних систем, систем підтримки прийняття рішень та застосування інформаційних технологій у науковій діяльності, що передбачає глибоке переосмислення наявних та створення нових цілісних знань з технологій інформаційної аналітики.

2. Попередні вимоги до опанування або вибору навчальної дисципліни:

- 1. Знати:** основні принципи, моделі та математичні методи аналізу даних, а також способи їх реалізації в сучасних пакетах статистичного аналізу.
- 2. Вміти:** здійснювати вибір, розробляти алгоритми аналізу даних та використовувати для цього сучасні мови програмування і програмні системи.

3. Анотація навчальної дисципліни:

Дисципліна «Інтелектуальні обчислення та аналіз даних» належить до переліку дисциплін вільного вибору. Вона забезпечує професійний розвиток, спрямована на формування концептуальних та методологічних знань у галузі інформаційних технологій, вміння критично аналізувати, оцінювати і синтезувати нові та комплексні ідеї, формулювати задачі та здійснювати їх формалізацію, здійснювати вибір або розробку алгоритмів аналізу даних, аналізувати їх ефективність на основі різних критеріїв. В рамках дисципліни вивчаються основні методи структурної та параметричної ідентифікації невідомих залежностей на основі інтелектуальних обчислень.

4. Завдання (навчальні цілі): набуття знань, умінь та навичок (компетентностей) на рівні новітніх досягнень у комп'ютерних науках, відповідно науково-освітньої кваліфікації «Доктор філософії». Зокрема, розвивати: здатність застосовувати теоретичні та практичні основи методології та технології аналізу даних, реалізовувати алгоритми інтелектуальних обчислень для ідентифікації та оптимізації залежностей в умовах невизначеності, проводити експерименти, аналізуючи якість розв'язків та здійснюючи вибір кращих алгоритмів.

5. Результати навчання за дисципліною:

Результат навчання (РН) (1. знати; 2. вміти; 3. комунікація; 4. автономність та відповідальність)		Форми (та/або методи і технології) викладання і навчання	Методи оцінювання та пороговий критерій оцінювання (за необхідності)	Відсоток у підсумковій оцінці з дисципліни
Код	Результат навчання			
РН 1.1	Знати основні проблеми та алгоритми попередньої обробки даних	Лекція, практичне заняття	Контрольна робота 1 (60% правильних відповідей), екзамен, активна робота на лекції, усні відповіді	20%
РН 1.2	Знати основні алгоритми кластеризації та відновлення даних			
РН 1.3	Знати методи структурної ідентифікації невідомих залежностей та методи їх параметричної оптимізації			20%
РН 1.4	Знати основи сучасних гібридних технологій аналізу даних та прогнозування			
РН 2.1	Вміти здійснювати вибір та розробку у разі потреби алгоритмів ідентифікації та оптимізації	Лекція, практичне заняття, самостійна робота	Контрольна робота 2 (60% правильних відповідей), екзамен, захист проекту, виконання завдань, винесених на самостійну роботу	20%
РН 2.2	Вміти оцінювати ефективність існуючих та розроблених алгоритмів обробки даних за різними критеріями.			

PH 2.3	Вміти застосовувати технології аналізу даних до розв'язання практичних задач.	<i>Практичне заняття, самостійна робота</i>	<i>Захист проекту</i>	5%
PH3.1	Обґрунтовувати власний погляд на вибір та розробку алгоритмів обробки даних, спілкуватися з колегами з питань проектування та розробки інформаційно-аналітичних систем, скласти письмові звіти			5%
PH4.1	Демонстрація авторитетності, інноваційності, високий ступінь самостійності, академічна та професійна доброчесність, послідовна відданість розвитку нових ідей або процесів у передових контекстах професійної та наукової діяльності.			5%
PH4.2	Відповідально ставитися до виконуваних робіт, нести відповідальність за їх якість			5%

6. Схема формування оцінки.

6.1. Форми оцінювання здобувачів освітньо-наукового ступеня:

- оцінювання впродовж навчального періоду:

1. Активна робота на лекції, усні відповіді: PH1.1, PH1.2, PH1.3, PH1.4 – 5 балів/3 бали;
2. Виконання завдань, винесених на самостійну роботу: PH2.1, PH2.2 – 5 балів/3 бали;
3. Контрольна робота 1: PH1.1, PH1.2, PH2.1, PH2.2 – 15 балів/9 балів;
4. Контрольна робота 2: PH1.3, PH1.4, PH2.1, PH2.2 – 15 балів/9 балів;
6. Захист проекту: PH2.1, PH2.2, PH2.3, PH3.1, PH4.1, PH4.2, – 20 балів/12 балів;

- підсумкове оцінювання: екзамен.

- максимальна кількість балів, які можуть бути отримані: 40 балів;
- рубіжний рівень оцінки на іспиті є 24 бали;
- результати навчання, які будуть оцінюватись: PH1.1, PH1.2, PH1.3, PH1.4;
- форма проведення і види завдань: письмова робота

Для здобувачів освітньо-наукового ступеня, які набрали сумарно меншу кількість балів ніж критично-розрахунковий мінімум – 20 балів для допуску до повторного складання іспиту необхідно здати контрольні роботи та захистити проект.

Рекомендований мінімум – 36 балів.

6.2. Організація оцінювання:

Обов'язковим є виконання завдань, винесених на самостійну роботу, та модульних контрольних робіт за графіком робочої програми.

У частину 1 входять теми 1 - 3, у частину 2 – теми 4 – 6 у частину 3 – теми 7 – 9. Обов'язковим для екзамену є виконання усіх контрольних робіт та захист проекту до вказаної викладачем дати, перед початком екзаменаційної сесії, згідно навчального плану. Переписування чи перескладання тем не практикується. Дозволяється здача окремих завдань модульних тем у проміжках між написанням модульних контрольних робіт (наприклад, перша тема здається до здачі наступної модульної контрольної роботи у будь-який зручний для викладача та студента час).

Терміни проведення форм оцінювання:

1. Контрольна робота: до 5 тижня навчального періоду.
2. Контрольна робота: до 13 тижня навчального періоду.
3. Захист проекту: до 10 тижня навчального періоду.

У випадку відсутності з поважних причин відпрацювання та перездачі контрольні роботи здійснюються у відповідності до „Положення про організацію освітнього процесу”.

6.3. Шкала відповідності оцінок

Відмінно / Excellent	90-100
Добре / Good	75-89
Задовільно / Satisfactory	60-74
Незадовільно / Fail	0-59

7. СТРУКТУРА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ. ТЕМАТИЧНИЙ ПЛАН ЛЕКЦІЙ І ПРАКТИЧНИХ ЗАНЯТЬ

№	Назва лекції	Кількість годин		
		Лекції	Практичні	Самостійна робота
Частина 1. „Методи попередньої обробки даних”				
1	<p>Тема 1. Препроцесінг даних. Ентропія та кількість інформації. Нормалізація та стандартизація даних. Алгоритми визначення інформативних ознак. Алгоритм «вибілювання входів». Методика «box-counting».</p> <p><i>Самостійна робота:</i> Для задач дисертаційного дослідження виконати препроцесінг даних, здійснивши їх нормалізацію, стандартизацію, «вибілення». Вибрати інформативні ознаки.</p>	2		8
2	<p>Тема 2. Кластеризація. Постановка задачі та її попередній аналіз. Характеристика методів кластерного аналізу. Алгоритми, що базуються на гіпотезі компактності та лямбда-компактності. Пірамідальні мережі.</p> <p><i>Самостійна робота:</i> Для задач дисертаційного дослідження виділити групи однорідних об'єктів чи експериментів. Зробити висновки про оптимальну кількість кластерів.</p>	2		12
3	<p>Тема 3. Відновлення даних. Постановка задачі відновлення пропусків у даних. Евристичні методи обробки некомп-лектних даних. Відновлення пропусків значень залежної змінної. Локальні методи відновлення пропусків. Ітераційних метод головних компонент для даних з пропусків. EM-алгоритм.</p> <p><i>Самостійна робота:</i> Для задач дисертаційного дослідження визначити, чи є наявними всі потрібні дані. Якщо є відсутні дані, то відновити їх та оцінити якість відновлення. Якщо відсутніх даних немає, то згенерувати пропуски та оцінити ефективність пропозованих алгоритмів.</p>	1	2	8
<i>Контрольна робота 1</i>		1		
Частина 2. „Структурна та параметрична ідентифікація невідомих залежностей”				
4	<p>Тема 4. Класичні методи ідентифікації. Метод найменших квадратів. Множинна лінійна регресія. Автокореляція. Мультиколінеарність. Гетероскедастичність. Множинна нелінійна регресія.</p> <p><i>Самостійна робота:</i></p>	2		12

	Для даних дисертаційного дослідження побудувати лінійні моделі, оцінити їх адекватність. Перевірити наявність мультиколінеарності та гетероскедастичності.			
5	Тема 5. Нейромережева ідентифікація Парадигми нейромережевого навчання. Ітераційні методи навчання нейромереж. Стохастичне навчання нейромереж. Пряме навчання нейромереж. Мережа зустрічного поширення. Мережі з оберненими зв'язками. <i>Самостійна робота:</i> Для даних дисертаційного дослідження виконати ідентифікацію невідомої залежності з використанням різних нейромереж та виконати порівняльний аналіз точності результату..	2	2	8
6	Тема 6. Еволюційне моделювання та методи самоорганізації Метод групового врахування аргументів. Алгоритми поділу початкової вибірки даних. Генетичний алгоритм. Еволюційні стратегії. Метод деформованих зірок. <i>Самостійна робота:</i> Для задачі дисертаційного дослідження побудувати цільову функції, виконати її оптимізацію та порівняти результати.	2		12
Частина 3. „ Слабко структуровані дані та методи їх обробки”				
7	Тема 7. Обробка нечітко заданих даних Нечіткі відношення та нечітке логічне виведення. Аналіз нечітких експертних висновків. Прийняття рішень в умовах нечіткості. <i>Самостійна робота:</i> Для даних дисертаційного дослідження побудувати моделі у вигляді нечітких продукційних правил та здійснити логічне виведення.	2		12
8	Тема 8. Гібридні методи аналізу даних. Нечіткі нейромережні парадигми. Навчання нечітких нейромереж. Еволюційно-параметрична оптимізація нейромереж. Композиційні методи зменшення невизначеності. <i>Самостійна робота:</i> Для задач дисертаційного дослідження побудувати відповідні нечіткі нейромережі та виконати порівняльний аналіз їх ефективності..	2		12
9	Тема 9. Сучасні методи ідентифікації та оптимізації Мурашині алгоритми та алгоритми рою. Програмування генетичних виразів. Нечіткі системи як універсальні апроксиматори. <i>Самостійна робота:</i>	1		12

Для задач дисертаційного дослідження виконати пошук оптимізованих розв'язків, або побудувати моделі у вигляді нечітких продукційних правил			
<i>Контрольна робота 2</i>	1		
ВСЬОГО	18	4	96

Загальний обсяг 120 годин, в тому числі:

Лекцій – **18 годин**,

Практичні – **4 години**.

Консультації - **2 години**.

Самостійна робота – **96 годин**.

8. Рекомендовані джерела

Основні:

1. Снитюк В.Є. Прогнозування. Моделі, методи, алгоритми. – К.: Маклаут, 2008. – 364 с.
2. Хайкин С. Нейронные сети: Полный курс. – М., СПб, К.: Вильямс, 2006.
3. Luke S. Essential of Metaheuristics. Режим доступу:
<https://cs.gmu.edu/~sean/book/metaheuristics/Essentials.pdf>

Додаткові:

1. Лук'яненко І., Краснікова Л. Економетрика. – К.: Знання, 1998. – 494 с.
2. Люгер Ф. Дж. Искусственный интеллект. Стратегии и методы решения сложных проблем. – М.: “Вильямс”, 2003. – 864 с.
3. Генетические алгоритмы, искусственные нейронные сети и проблемы виртуальной реальности. /Вороновский Г.К., Махотило К.В., Петрашев С.Н., Сергеев С.Н. – Харьков: Основа, 1997. – 112 с.
4. Зайченко Ю.П. Основи проектування інтелектуальних систем. – К.: “Слово”, 2004. – 352 с.
5. Факторный, дискриминантный и кластерный анализ: Пер. с англ. /Дж. – О. Ким, Ч.У. Мьюллер, У.Р. Клекка и др.: Под ред. И.С. Енюкова. – М.: Финансы и статистика, 1989. – 215 с.
6. Гладун В.П. Растущие пирамидальные сети // Новости искусственного интеллекта. – 2004. – № 1. – С. 30–40.
7. Злоба Е., Яцкив И. Статистические методы восстановления пропущенных значений // Computer Modelling & New Technologies. – 2002. – Vol. 6. – № 1. – Р. 51–61.
8. Литтл Р. Дж. А., Рубин Д.Б. Статистический анализ данных с пропусками. – М.: Финансы и статистика, 1991. – 336 с.
9. Рутковская Д., Пилиньский М., Рутковский М. Нейронные сети, генетические алгоритмы и нечеткие системы. – М.: Горячая линия–Телеком, 2006. – 452 с.
10. Круглов В.В., Длин М.И., Голунов Р.Ю. Нечеткая логика и искусственные нейронные сети. – М.: Физматлит, 2001. – 224 с.
11. Dorigo, M., and Gambardella, L.M., Ant Colony System: A cooperative learning approach to the traveling salesman problem. IEEE Transactions on Evolutionary Computation, 1(1):53–66, 1997.
12. Ferreira, C., 2001. Gene Expression Programming: A New Adaptive Algorithm for Solving Problems. Complex Systems, forthcoming.