

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА

ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ
Кафедра інтелектуальних та інформаційних систем



**РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ
АВТОМАТИЗОВАНІ СИСТЕМИ ОБРОБКИ ДАНИХ І
ЗНАНЬ**

для здобувачів освітньо-наукового рівня «доктор філософії»

галузь знань	12 «Інформаційні технології»	
спеціальність	121 – “Інженерія програмного забезпечення”	
освітній рівень	третій (освітньо-науковий)	
освітньо-наукова програма	«Комп’ютерні науки»	
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	

Викладачі: Снитюк Віталій Євгенович, д.т.н., професор

Пролонговано: на 20 / 20 н.р. () « 01 » 07 2019 р.
на 20 / 20 н.р. () « » 20 р.

КИЇВ – 2018

1. Мета дисципліни формування теоретичних знань та практичних умінь з аналізу задач системи обробки даних та знань, вибір відповідних моделей та технологій для розробки автоматизованих систем обробки даних та знань, що передбачає знання особливостей обробки добре структурованих та слабо-структурованих даних, виявлення прихованих знань, знання відповідних методів та засобів для роботи з даними, створення автоматизованих системи обробки даних і знань в різноманітних галузях.

2. Попередні вимоги до опанування або вибору навчальної дисципліни:

Знати: поняття, сутність інформаційних систем (ІС), ресурсів та інформаційних технологій, які використовуються на підприємствах; системне і прикладне програмне забезпечення інформаційних технологій в ІС; основні поняття та засоби проектування баз даних, основи мережевих технологій та їх вплив на вирішення аналітичних процедур; організацію інформаційного забезпечення управління підприємством.

Вміти: використовувати методи і прийоми організації інформаційних систем в залежності від їх видів, рівнів та функціонального призначення; удосконалювати, поглиблювати і розробляти форми і конфігурації інформаційних систем в напрямку максимального охоплення і рішення задач управління з даної предметної області; проектувати бази даних та знань, розробляти постановки і моделювати окремі задачі та комплекси задач з обліку.

3. Анотація навчальної дисципліни:

Дисципліна «Автоматизовані системи обробки даних і знань» 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, аналізувати складність існуючих методів і моделей обробки даних та знань, вміння проектувати системи автоматизовані системи обробки даних і знань для різних галузей та напрямків діяльності, вміння працювати та обробляти дані, представлені в різних форматах, виявляти приховані знання, використовуючи відповіді технологій для конкретної задачі.

4. Завдання (навчальні цілі): набуття знань, умінь та навичок (компетентностей) на рівні новітніх досягнень у програмуванні, відповідно науково-освітньої кваліфікації «Доктор філософії». Зокрема, розвивати: здатність застосовувати теоретичні та практичні основи методології та технології обробки даних та знань, використовувати тензорну методологію для проектування систем обробки знань, вміти розробляти автоматизовані системи проектування обробки даних і знань для різних галузей, вміти працювати з даними різних форматів, вміти застосовувати для побудови даних систем технології Hadoop.

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 функції та структуру автоматизованих систем обробки даних і знань	<i>Лекція, практичне заняття</i>	<i>Контрольна робота 1 (60% правильних відповідей),</i>	15%
РН 1.2	Знати основні поняття тензорної алгебри, методи представлення та виявлення прихованих знань			
РН 1.3	Знати архітектуру систем обробки даних та знань.			15%

PH 1.4	Знати основні формати документів автоматизованих систем обробки даних та знань, особливості роботи з ними		екзамен, активна робота на лекції, усні відповіді	
PH 2.1	Вміти проектувати автоматизовані системи обробки даних і знань відповідно до поставлених задач, галузі та форматів даних, які має обробляти система	Лекція, практичне заняття, самостійна робота	Контрольна робота 2 (60% правильних відповідей), екзамен, захист проекту, виконання завдань, винесених на самостійну роботу	15%
PH 2.2	Вміти використовувати тензорну методологію для проектування систем обробки даних та знань			15%
PH 2.3	Вміти застосовувати MapReduce для обробки даних, вміти працювати з технологіями Hadoop, пакетом MATLAB (Parallel Computing Toolbox, Database Toolbox, Distributed Computing Server)	Практичне заняття, самостійна робота	Захист проекту	15%
PH3.1	Обґрунтовувати власний погляд на задачу, спілкуватися з колегами з питань проектування та розробки програм, скласти письмові звіти			15%
PH4.1	Демонстрація авторитетності, інноваційний підхід, високий ступінь самостійності, академічна та професійна добросесність, 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. Співвідношення результатів навчання дисципліни із програмними результатами навчання

Результати навчання дисципліни	Р	Р	Р	Р	Р	Р	Р	Р	Р	Р
	Н	Н	Н	Н	Н	Н	Н	Н	Н	Н
Програмні результати навчання	1	1	1	1	2	2	2	3	4	4
	·	·	·	·	·	·	·	·	·	·
	1	2	3	4	1	2	3	1	1	2
<i>(з опису освітньої програми)</i>										
ПРН-7. Знати, розуміти і самостійно застосовувати методи аналізу предметної області, виявлення інформаційних потреб і збір даних для проектування.	+	+	+	+	+	+	+	+	+	+

7. Схема формування оцінки.

7.1. 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;
- результати навчання, які будуть оцінюватись: PH1.1, PH1.2, PH1.3, PH1.4;
- форма проведення і види завдань: письмова робота.

- умови допуску до підсумкового оцінювання:

Рекомендований мінімум для допуску до іспиту – **36 балів**, критично розрахунковий мінімум – **20 балів**.

При цьому обов'язковим є виконання всіх передбачуваних робочою програмою видів робіт не менше ніж на 60%, а також отримання позитивної оцінки з контрольних робіт.

Для здобувачів, які не набрали рекомендований мінімум (60% від максимально можливої кількості балів) обов'язковим є написання **комплексної контрольної роботи**, яка включає увесь пройдений за навчальний період матеріал і максимальна оцінка за яку не може перевищувати 40% підсумкової оцінки (до 40 балів за 100-бальною шкалою).

7.2. 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 - 2, у частину 2 – теми 3 – 4, у частину 3 – теми 5 – 6, у частину 4- 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.*

Умовою отримання позитивної результуючої оцінки з дисципліни є досягнення не менш як 60% від максимально можливої кількості балів, при цьому, оцінка за результати навчання, передбачені пунктами 2, 3, 4 не може бути меншою ніж 50% від максимального рівня.

Максимальна кількість балів, яку може отримати здобувач за роботу протягом навчального періоду становить **60 балів** за 100-бальною шкалою.

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

7.3. Шкала відповідності оцінок

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

При визначені оцінки визначальною є робота протягом навчального періоду

Після завершення розгляду тем проводяться письмові контрольні роботи та теоретичне опитування.

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

№	Назва лекції	Кількість годин		
		Лекції	Практичні	Самостійна робота
Частина 1. «Засоби систем обробки даних та знань . Основні визначення, поняття.»				
1	<p>Тема 1. Системи обробки даних та знань. Вступ, основні визначення та поняття. Класифікації автоматизованих систем. Автоматизовані системи управління. Автоматизовані системи обробки даних. Функціональні підсистеми різних галузей. Визначення основних функцій.</p> <p><i>Самостійна робота:</i> На прикладі задач по темі дисертаційної роботи провести аналіз та розробити модель інформаційної системи обробки даних.</p>	2		8
2	<p>Тема 2. Системи обробки даних та знань. Поняття добре структурованих даних, слабко структурованих даних. Застосування тензорної методології в розробці систем обробки даних і знань. Тензори: визначення та основні поняття. Тензорна алгебра</p> <p><i>Самостійна робота:</i> На прикладі виконати базові операції над заданим тензором, виконати розклад тензора за векторами. Створити фрагмент програми для реалізації даних завдань</p>	2		8
Частина 2 «Архітектура систем обробки даних та знань»				
3	<p>Тема 3. Розподілені інформаційні системи, класифікація Багаторівнева архітектура “клієнт-сервер”. Побудова систем на основі архітектури розподілених компонент. Сервіс-орієнтована архітектура(SOA), стандарти SOA.</p> <p><i>Самостійна робота:</i> Визначити тип архітектури, який буде відповідати потребам системи обробки даних і знань, розробленої для вирішення задач дисертаційного дослідження</p>	1		10
	<p>Тема 4. Засоби вилучення знань, присутніх в великих масивах даних в неявному вигляді. Системи для роботи зі знаннями (Knowledge-level Systems). Використання MATLAB для побудови систем обробки знань, вилучення знань, представлених в неявному вигляді. Робота з Parallel Computing Toolbox, Database Toolbox, Distributed Computing Server</p> <p><i>Самостійна робота:</i></p>	2		12

	Реалізувати обробку даних та знань, використовуючи засоби MATLAB Parallel Computing Toolbox, Database Toolbox.			
<i>Контрольна робота 1</i>		1		
Частина 3. «Характеристика компонент систем обробки даних та знань»				
	Тема 5. Компоненти систем обробки даних та знань: функціональні; компоненти системи опрацювання даних; організаційні компоненти. Характеристика компонентів системи опрацювання даних. БД, СУБД. Робота з великими даними. <i>Самостійна робота:</i> Проектування автоматизованої системи обробки даних та знань, визначення основних компонентів опрацювання даних, сховищ даних відповідно до тематики та задач дисертаційного дослідження	2		8
4	Тема 6. Робота з великими даними в парадигмі MapReduce. Стадії обробки даних. <i>Map – Shuffle – Reduce</i> . Технології Hadoop: основні компоненти, принципи роботи. Застосування технологій Hadoop при побудові автоматизованих систем обробки даних і знань <i>Самостійна робота:</i> Застосування принципів MapReduce для обробки даних на прикладі задач дисертаційного дослідження.	2	2	14
Частина 4. «Формати документів автоматизованих систем обробки даних та знань»				
7	Тема 7. Формат CSV, переваги та недоліки. Засоби роботи з CSV- форматами. Формати JSON. Призначення, структура та способи отримання даних в форматі JSON. Формати YAML. <i>Самостійна робота:</i> На прикладі задач дисертаційного дослідження виконати обробку даних, представлених в форматах CSV та JSON. Описати особливості обробки даних, представлених в форматі JSON	2	2	12
	Тема 8. Формати даних для роботи з великими даними: AVRO, Sequence, Parquet, ORC, RCFile. Колонкові та лінійні формати. <i>Самостійна робота:</i> Робота з форматами даних AVRO, Sequence, Parquet, ORC, RCFile. На прикладі задач дисертаційного дослідження опрацювати дані в одному з перерахованих форматів.	2		12
	Тема 9. Структура XML-документа, способи створення та редагування. Переваги та недоліки роботи з XML-документами. Відображення, зв'язування XML-документів. <i>Самостійна робота:</i> На прикладі задач дисертаційного дослідження опрацювати створення та редагування XML-	1		12

документів. Застосувати різні способи створення XML-документів.			
<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**.

9. Рекомендовані джерела

Основні:

1. Шаховська Н.Б., Пасічник В.В. Сховища та простори даних: Монографія. - Львів: Видавництво Національного університету "Львівська політехніка", 2009. - 244 с.
2. Гайдамакин Н. А. Автоматизированные информационные системы, базы и банки данных. Вводный курс: Учебное пособие. — М.: Гелиос АРВ, 2002. — 368 с., ил.
3. Том Уайт. Hadoop. Подробное руководство. СПб.: Питер, 2013. — 672 с.
4. Таненбаум Э. Распределенные системы. Принципы и парадигмы / Э.Таненбаум, М. ван Стеен. – СПб.: Питер, 2003. – 877 с.
5. Димитриенко Ю.И. Тензорное исчисление / Ю.И. Димитриенко.— М.: Высш. шк., 2001. — 575 с
6. George T. Heineman, William T. Councill (2001). Component-Based Software Engineering: Putting the Pieces Together. Addison-Wesley Professional, Reading, 2001.

Додаткові:

7. Снитюк В.Е., Юрченко К.Н. Элементы знаниеориентированных систем профессиональной подготовки адаптивного типа // Вестник Херсонского национального технического университета. – 2010. – № 2 (38). – С. 180-186
7. Pattern-Oriented Software Architecture. Volume 1: A System of Patterns / F. Buschmann, R. Meunier, H. Rohnert, P. Sommerlad, M.Stal. – N.Y.: John Wiley, 1996. – 476 p.
8. Pattern-Oriented Software Architecture. Volume 2: Patterns For Concurrent And Networked Objects / D. Schmidt, M. Stal, H. Rohnert, F. Buschmann. – N. Y.: John Wiley, 2000. – 633 p.
9. Tomas B. Co Methods of Applied Mathematics for Engineers and Scientists - y Cambridge University Press, 2013, 582 p.
10. Skillicorn D. Data Mining and Knowledge Discovery Series. Understanding Complex Datasets. Data Mining with Matrix Decompositions. Chapman & Hall/CRC.- 2007.- 257 pp.
11. V. Kreinovich; A.V. Lakeyev; J. Rohn; P.T. Kahl Computational Complexity and Feasibility of Data Processing and Interval Computations: Springer US, June 2013 – 447 p.
12. Fedra Pavlou . Big Data, Hidden Knowledge <https://thepathologist.com/inside-the-lab/big-data-hidden-knowledge>
13. Гарсия Х., Голдшмидт Г., «Разработка составных бизнес-сервисов на базе сервисно-ориентированной архитектуры». <http://www.ibm.com/developerworks/ru/library/ws-soa-composite/> 12.08.2014.
14. Проектування інформаційних систем : Навч. посібник/ Ред. Володимир Пономаренко, -К.: Академія, 2002. -486 с.
15. <https://www.mathworks.com/services/consulting/proven-solutions/matlab-with-hadoop-and-spark.html>