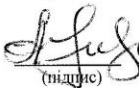


Міністерство освіти і науки України
Національний аерокосмічний університет ім. М. Є. Жуковського
«Харківський авіаційний інститут»

Кафедра біомедичних та радіоелектронних комп'ютеризованих засобів і
технологій (№ 502)

ЗАТВЕРДЖУЮ

Керівник проектної групи

 О.Й. Довнар
(підпис) (ініціали та прізвище)

«31» серпня 2021 р.

**РОБОЧА ПРОГРАМА ВИБІРКОВОЇ
НАВЧАЛЬНОЇ ДИСЦИПЛІНИ**

Сучасні технології медичних баз та сховищ даних
(назва навчальної дисципліни)

Галузь знань: 12 «Інформаційні технології»
(шифр і найменування галузі знань)

Спеціальність: 122 «Комп'ютерні науки»
(код і найменування спеціальності)

Освітня програма: «Комп'ютерні технології в біології та медицині»
(найменування освітньої програми)

Форма навчання: денна

Рівень вищої освіти: перший (бакалаврський)

Харків 2021 рік

Робоча програма дисципліни «Сучасні технології медичних баз та сховищ даних»

(назва дисципліни)

для студентів за спеціальністю 122 «Комп'ютерні науки»

освітньою програмою «Комп'ютерні технології в біології та медицині»

«31» серпня 2021 р. – 11 с.

Розробник: Панферова І.Ю. професор каф. 502, к.т.н., доцент
(прізвище та ініціали, посада, науковий ступінь і вчене звання)



(підпис)

Робочу програму розглянуто на засіданні кафедри біомедичних та радіоелектронних комп'ютеризованих засобів і технологій (№ 502)

(назва кафедри)

Протокол № 1 від «31» серпня 2021 р.

Завідувач кафедри д.т.н., проф.
(науковий ступінь і вчене звання)



(підпис)

О.В. Висоцька
(ініціали та прізвище)

1. ОПИС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Найменування показника	Галузь знань, спеціальність, освітня програма, рівень вищої освіти	Характеристика навчальної дисципліни (денна форма навчання)
Кількість кредитів – 5,0	Галузь знань <u>12 «Інформаційні технології»</u> (шифр і найменування) Спеціальність <u>122 «Комп’ютерні науки»</u> (код і найменування) Освітня програма <u>«Комп’ютерні технології в біології та медицині»</u> (найменування) Рівень вищої освіти: перший (бакалаврський)	Вибіркова
Кількість модулів – 2		Навчальний рік
Кількість змістовних модулів – 4		2021/2022
Індивідуальне завдання _____ (назва)		Семестр
Загальна кількість годин – 64/150		4-й
		Лекції*
		32
Кількість тижневих годин для денної форми навчання: аудиторних – 4,0 самостійної роботи студента – 5,4		Практичні, семінарські*
		Лабораторні*
	32	
	Самостійна робота	
	86	
	Вид контролю	
	модульний контроль, іспит	

Співвідношення кількості годин аудиторних занять до самостійної роботи становить: 64/86 (кількість годин аудиторних занять/ кількість годин самостійної роботи).

*Аудиторне навантаження може бути зменшене або збільшене на одну годину залежно від розкладу занять.

2. МЕТА ТА ЗАВДАННЯ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Мета вивчення: навчання студентів сучасним технологіям баз даних, що пов'язані із створенням, підтримкою і використанням баз даних в інформаційних системах, отримання практичних навичок і теоретичних основ побудови структур даних сучасних систем управління. Отримання практичних навичок і теоретичних основ роботи з Big Data в сучасних систем управління.

Завдання

Згідно з вимогами освітньо-професійної програми студенти повинні досягти таких **компетентностей**:

- здатністю до абстрактного мислення, аналізу та синтезу (ЗК1);
- здатністю застосовувати знання у практичних ситуаціях (ЗК2);
- знання та розуміння предметної області та розуміння професійної діяльності (ЗК3)
- здатність вчитися й оволодівати сучасними знаннями (ЗК6).
- здатністю реалізувати багаторівневу обчислювальну модель на основі архітектури клієнт-сервер, включаючи бази даних, сховища даних і бази знань, для забезпечення обчислювальних потреб багатьох користувачів, обробки транзакцій, у тому числі на хмарних сервісах (ФК8) ;
- здатність реалізувати багаторівневу обчислювальну модель на основі архітектури клієнт-сервер, включаючи бази даних, знань і сховища даних, виконувати розподілену обробку великих наборів даних на кластерах стандартних серверів для забезпечення обчислювальних потреб користувачів, у тому числі на хмарних сервісах (ФК9);
- здатність до інтелектуального аналізу даних на основі методів обчислювального інтелекту включно з великими та погано структурованими даними, їхньої оперативної обробки та візуалізації результатів аналізу в процесі розв'язування прикладних задач медико-біологічної спрямованості (ФК11);
- здатність проектувати та впроваджувати медичні бази та сховища даних із дотриманням заданих вимог щодо їх надійності, ефективності, продуктивності, навантаженості та масштабування; оптимізувати роботу великих баз даних, забезпечувати реалізацію складних запитів(ФК18).

Програмні результати навчання:

- використовувати методи обчислювального інтелекту, машинного навчання, нейромережевої та нечіткої обробки даних, генетичного та еволюційного програмування для розв'язання задач розпізнавання, прогнозування, класифікації, ідентифікації об'єктів керування тощо (ПРН 4);
- використовувати інструментальні засоби розробки клієнт-серверних застосувань, проектувати концептуальні, логічні та фізичні моделі баз даних, розробляти та оптимізувати запити до них, створювати розподілені бази даних, сховища та вітрини даних, бази знань, у тому числі на хмарних сервісах, із застосуванням мов веб-програмування (ПРН 10);

- застосовувати методи та алгоритми обчислювального інтелекту та інтелектуального аналізу даних в задачах класифікації, прогнозування, кластерного аналізу, пошуку асоціативних правил з використанням програмних інструментів підтримки багатовимірного аналізу даних на основі технологій DataMining, TextMining, WebMining (ПРН 12);

- застосовувати знання методології та CASE-засобів проектування складних систем, методів структурного аналізу систем, об'єктно-орієнтованої методології проектування при розробці і дослідженні функціональних моделей організаційно-економічних і виробничо-технічних систем (ПРН 15);

- використовувати методи та технології аналізу великих даних для розв'язання прикладних задач медико-біологічної спрямованості, новітні інструменти та методи проектування, розробки, оптимізації та адміністрування медичних баз та сховищ даних, в тому числі розподілених (ПРН 19);

- застосовувати сучасні підходи для розв'язання нових проблем, що виникають в сфері комп'ютерних технологій біології та медицини, враховуючи соціальні, біоетичні, біофізичні та економічні аспекти, існуючі державні і закордонні стандарти; забезпечувати ефективне управління якістю інформаційного, алгоритмічного та програмного забезпечення медичних комп'ютерних систем на основі використання сучасних моделей, методів та інструментальних засобів (ПРН 20).

Міждисциплінарні зв'язки:

Дисципліна «Сучасні технології медичних баз та сховищ даних» базується на знаннях з дисциплін «Організація медичних баз даних та знань», «Оптимізація та адміністрування медичних баз даних»».

3 ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Модуль 1

Змістовий модуль 1. . Напрями розвитку технологій баз даних.

Тема 1. Поняття технології баз даних. Еволюція технологій баз даних. Основні досягнення. Напрямки перспективних досліджень.

Тема 2. Нові можливості SQL. Перспективи розвитку стандарту.

Тема 3. Засоби API у стандарті мови SQL(аналітичний SQL, динамічний SQL).

Тема 4. Бази даних і UML.

Змістовий модуль 2. DataWarehouse

Тема 1. Концепція сховищ даних. Системи підтримки прийняття рішень. Способи побудови СППР. Преваги та недоліки. Фактори, що впливають на вибір архітектури СППР

Тема 2. Створення сховищ даних. Архітектурні рішення.

Тема 3. Оптимізація сховища даних.

Тема 4. Управління зведеними даними.

Тема 5. Завантаження даних в сховище. ETL процес. Видобуток даних. Data mining. Визначення, область. Застосування, інструментарій.

Модуль 2

Змістовий модуль 3. Новітні технології баз даних

Тема 1. Database In-Memory

Тема 2. Екосистема NoSQL.

Тема 3. Технологія blockchain

Тема 4. Хмарні технології баз даних.

Змістовий модуль 4. BigData

Тема 1. BigData. Концептуальна архітектура кластера для великих даних.

Тема 2. Огляд системи HADOOP. Архітектура HADOOP. Робота з HDFS.

Тема 3. Парадигма Map-Reduce. Сучасні і перспективні навколо-HADOOP технології (YARN, SQOOP, SPARK)

4. СТРУКТУРА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Назва змістовного модуля і тем	Кількість годин				
	Усього	У тому числі			
		л	п	лаб.	с. р.
1	2	3	4	5	6
Модуль 1.					
Змістовий модуль 1. . Напрями розвитку технологій баз даних.					
Тема 1. Поняття технології баз даних. Еволюція технологій баз даних. Основні досягнення. Напрямки перспективних досліджень.	4	2			2
Тема 2. Нові можливості SQL. Перспективи розвитку стандарту.	4	2			2
Тема 3. Засоби API у стандарті мови SQL(аналітичний SQL, динамічний SQL).	20	2		8	10
Тема 4. Бази даних і UML.	12	2		4	6
<i>Разом за змістовним модулем 1</i>	<i>40</i>	<i>8</i>		<i>12</i>	<i>20</i>
Змістовний модуль 2 DataWarehouse					
Тема 1. Концепція сховищ даних. Системи підтримки прийняття рішень. Способи побудови СППР. Преваги та недоліки. Фактори, що впливають на вибір архітектури СППР	4	2			2
Тема 2. Створення сховищ даних. Архітектурні рішення.	10	2		4	6
Тема 3. Оптимізація сховища	2	1			1

даних.					
Тема 4. Управління зведеними даними.	12	2		4	6
Тема 5. Завантаження даних в сховище. ETL процес. Видобуток даних. Data mining. Визначення, область. Застосування, інструментарій.	12	2		4	6
Разом за змістовним модулем 2	42	9		12	21
Модульний контроль 1	3	1			2
Разом за модуль 1	85	18		24	43
Модуль 2.					
Змістовий модуль 3. Новітні технології баз даних					
Тема 1. Database In-Memory	4	2			2
Тема 2. Екосистема NoSQL.	12	2		4	6
Тема 3. Технологія blockchain	4	2			2
Тема 4. Хмарні технології баз даних.	12	2		4	6
Разом за змістовим модулем 3.	32	8		8	16
Змістовий модуль 4. BigData.					
Тема 1. BigData. Концептуальна архітектура кластера для великих даних.	15	3			4+8
Тема 2. Огляд системи HADOOP. Архітектура HADOOP. Робота з HDFS.	15	2			4+9
Разом за змістовим модулем 4	30	5			25
Модульний контроль 2	3	1			2
Разом за модуль 2	65	14		8	45
Усього годин	150	32		32	86

5. ТЕМИ СЕМІНАРСЬКИХ ЗАНЯТЬ

не передбачено навчальним планом

6. ТЕМИ ПРАКТИЧНИХ ЗАНЯТЬ

не передбачено навчальним планом

7. ТЕМИ ЛАБОРАТОРНИХ ЗАНЯТЬ

№ п/п	Назва теми	Кількість годин
1	Використання аналітичних функцій SQL	4
2	Динамічне формування SQL команд	4
3	Дослідження об'єктно-орієнтованих технологій баз даних	4
4	Створення сховищ даних	4

5	Експлуатація сховищ даних	4
6	Завантаження даних у сховище даних	4
7	Основи роботи з NoSQL базами даних	4
8	Використання хмарних технологій БД	4
Разом		32

8. Самостійна робота

№ п/п	Назва теми	Кількість годин
1	Вивчення теоретичного матеріалу з використанням конспекту і навчально-методичної літератури	37
2	Підготовка до лабораторних занять, оформлення результатів виконання завдань	32
3	Вивчення додаткових тем за літературними джерелами:	
	Shared nothing architecture (SNA)	8
	Map reduce	9
Разом		86

9. ІНДИВІДУАЛЬНІ ЗАВДАННЯ

не передбачено навчальним планом

10. МЕТОДИ НАВЧАННЯ

Словесний (лекція, дискусія) у сполученні з новітніми інформаційними технологіями та комп'ютерними засобами навчання (дистанційні, мультимедійні), лабораторні (виконання лабораторних занять), робота з навчально-методичною літературою (конспектування, самостійне опрацювання окремих розділів).

11. МЕТОДИ КОНТРОЛЮ

Усне та письмове опитування, захист лабораторних робіт, самооцінювання тощо, а також модульний контроль та іспит.

12. КРИТЕРІЇ ОЦІНЮВАННЯ ТА РОЗПОДІЛ БАЛІВ, ЯКІ ОТРИМУЮТЬ СТУДЕНТИ

12.1. Розподіл балів, які отримують студенти (кількісні критерії оцінювання)

Для оцінювання роботи студента протягом семестру використовується підсумкова рейтингова оцінка $P_{\pi} = O_{\text{сем}}$, де $O_{\text{сем}}$ – оцінка за семестра за 100-бальною шкалою, що обчислюється як сума оцінок за присутність на лекціях, виконані лабораторні завдання та контрольні заходи.

Складові навчальної роботи	Бали за одне заняття (завдання)	Кількість занять (завдань)	Сумарна кількість балів
Модуль 1			
Робота на лекціях	1...2	9	9...18
Виконання лабораторних робіт	2...3	6	12...18
Модульний контроль	16...22	1	16...22
Модуль 2			
Робота на лекціях	1...2	7	7...14
Виконання лабораторних робіт	2...3	2	4...6
Модульний контроль	12...22	1	12...22
Усього за семестр			60...100

Семестровий контроль (іспит) проводиться за наявності допуску до іспиту (виконання усіх лабораторних завдань та модульного контролю). Під час складання семестрового іспиту студент має можливість отримати максимум 100 балів.

Білет для іспиту складається з двох теоретичних та двох практичних завдань. Кожне теоретичне завдання оцінюється максимум в 30 балів, кожне практичне завдання – максимум 20 балів.

12.2. Якісні критерії оцінювання

Необхідний обсяг знань для одержання позитивної оцінки.

Об'єктно-орієнтовані технології баз даних

Динамічний SQL.

Аналітичні функції

Системи підтримки прийняття рішень

Концепція OLAP

Сховища даних

Завантаження даних у сховища

Поняття Data Mining

Архітектура HADOOP

Парадигма Map-Reduce

Екосистема NOSQL

Необхідний обсяг вмінь для одержання позитивної оцінки.

1. Динамічно будувати SQL-команди.

2. Вміти задати аналітичні функції.

3. Побудувати схему даних за допомогою UML.

4. Розробити схему сховища даних.

5.Реалізувати розроблену схему у середовищі обраної СУБД.

6.Завантажити дані у сховища.

7.Виконувати секціонування та індексування сховища даних. Виконувати менеджмент зведених даних.

12.3 Критерії оцінювання роботи студента протягом семестру

Задовільно (60-74). Студент знайом з основними поняттями навчального матеріалу, може самостійно відтворити значну частину навчального матеріалу і робити певні узагальнення, ознайомився з основною літературою, рекомендованою програмою, вміє виконувати навчальні завдання, передбачені програмою.

Добре (75-89). Студент вільно володіє навчальним матеріалом, вміє застосовувати вивчений матеріал у стандартних ситуаціях, узагальнювати та систематизувати навчальну інформацію, самостійно виконує передбачені програмою навчальні знання, самостійно знаходить і виправляє допущені помилки, може аргументовано обрати раціональний спосіб виконання навчального завдання.

Відмінно (90-100). Знання, вміння і навички студента повністю відповідають вимогам програми. Студент володіє глибокими, міцними знаннями, самостійно визначає проміжні цілі і вміє планувати особисту навчальну діяльність, оцінювати результати власної практичної роботи, вміє знаходити додаткову інформацію та самостійно використовує її для реалізації поставлених перед ним навчальних цілей, судження його логічні і достатньо обґрунтовані, засвоїв взаємозв'язок основних понять дисципліни, їх значення для подальшої професійної діяльності, вміє вільно використовувати сучасні програмні засоби для поповнення власних знань та розв'язування задач.

Шкала оцінювання: бальна і традиційна

Сума балів	Оцінка за традиційною шкалою	
	Іспит, диференційований залік	Залік
90 – 100	Відмінно	Зараховано
75 – 89	Добре	
60 – 74	Задовільно	
0 – 59	Незадовільно	Не зараховано

13. МЕТОДИЧНЕ ЗАБЕЗПЕЧЕННЯ

13.1 Базова література

1. Scott Urman, Michael McLaughlin, Ron Hardman, Ron Hartman Oracle Database 10g PL/SQL Programming , 2010. – 535 p.
2. Thomas Kyte, Darl Kuhn «Expert Oracle Database Architecture: Oracle Database Programming, Techniques and Solutions . – Apress, 2015. – 824с.
3. Thomas M. Connolly Database Systems: A Practical Approach to Design, Implementation, and Management. - DMCA CONTACT, 2017. - 1440p.

4. Peter Rob, Carlos M. Coronel Database Systems: Design, Implementation, and Management – Boston, MA 02210 USA, 2017. – 818 p.
5. Jinyu Wang Oracle Database 11g Building Oracle XML DB Applications, 2011
6. Drew Adams Oracle XML DB Developer's Guide, 11g Release 1/ - 2008, 972 p.
7. Jan L. Harrington Object-Oriented Database Design Clearly Explained - 2012 . - 312 c
8. Lilian Hobbs, Susan Hillson, Shilpa Lawande, Pete Smith Oracle 10g Data Warehousing Digital Press. 2011. - : 872 p

13.2 Методичні вказівки до різних видів занять

1. Конспект лекцій з дисципліни „ Сучасні технології медичних баз та сховищ даних ” для студентів денної форми навчання за напрямом 122 “Комп’ютерні науки ” / Упоряд.: І.Ю.Панфьорова. – Харків:, 2021 (електронний варіант).

2. Методичні вказівки до лабораторного практикуму з дисциплін „ медичних баз та сховищ даних ” / Упоряд.: І.Ю. Панфьорова. – Харків:, 2021. – 48с. (електронний варіант)

14 ІНФОРМАЦІЙНЕ ЗАБЕЗПЕЧЕННЯ

ORACLE XE,
ORACLE Client,
SQL*Plus,
SQL*Loader,
MS OFFICE.