

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ВСП «ЕКОНОМІКО-ПРАВНИЧИЙ ФАХОВИЙ КОЛЕДЖ
ЗАПОРІЗЬКОГО НАЦІОНАЛЬНОГО УНІВЕРСИТЕТУ»



ЗАТВЕРДЖУЮ

Директор

[Handwritten signature]
09

О.Є.Грибанова

2022

ВВЕДЕННЯ ДО ТЕОРІЇ
ХМАРНИХ ОБЧИСЛЕНЬ

(вибіркова дисципліна)

підготовки фахових молодших бакалаврів

галузі знань: 12 Інформаційні технології

спеціальності: 121 Інженерія програмного забезпечення

освітньо-професійної програми: розробка програмного забезпечення

Укладач програми: Юлія БОРИСОВСЬКА

Обговорено та ухвалено
на засіданні циклової комісії економічних,
математичних дисциплін, менеджменту та
туризму

Протокол № 1 від "29" серпня 2022

Голова ЦК *[Handwritten signature]* Т.М. Смолянкова
(підпис) (ініціали, прізвище)

Ухвалено методичною радою коледжу

Протокол № 1 від "22" вересня 2022

Заступник директора з НМР

[Handwritten signature] А.В. Ходаковська
(підпис) (ініціали, прізвище)

• 2022 рік

1. Опис навчальної дисципліни

Найменування показників	Галузь знань, спеціальність, освітня програма, рівень вищої освіти	Характеристика навчальної дисципліни	
		денна форма навчання	заочна форма навчання
Кількість кредитів – 3	Галузь знань <i>12 інформаційні технології</i>	Нормативна	
Розділів - 2	Спеціальність: <i>121 інженерія програмного забезпечення</i>	Рік підготовки: 4	
Загальна кількість годин – 90		3-й	4-й
Тижневих годин для денної форми навчання: 1,5 год	Освітньо-кваліфікаційний рівень: фаховий молодший бакалавр	Лекції	
		10	
		Практичні, семінарські	
		20 год	
		Лабораторні	
		Самостійна робота	
		60 год	
Вид підсумкового контролю: залік			

2. Мета та завдання навчальної дисципліни

Метою викладання навчальної дисципліни «Введення до теорії хмарних обчислень» є формування компетенцій основаних на вивчення підходів, методів і механізмів функціонування та використання хмарних інфраструктур для обробки великого обсягу даних, що потребують використання великої кількості обчислювальних ресурсів.

У разі успішного завершення курсу здобувач освіти набуває програмні компетентності

- здатність алгоритмічно та логічно мислити;
- здатність розробляти модулі і компоненти програмного забезпечення за допомогою типових алгоритмів та інструментів;
- здатність розробляти архітектури, модулі та компоненти програмного забезпечення;
- здатність реалізовувати високопродуктивні обчислення на основі хмарних сервісів і технологій при розробці й експлуатації розподілених систем паралельної обробки інформації.

Програмні результати навчання, які досягаються за допомогою вивчення дисципліни:

- вміти застосовувати методи обчислення та структури даних для вирішення задач аналізу та синтезу алгоритмів;

- знати основні методи оптимізації алгоритмів, вміти розробляти ефективні алгоритми розв'язування завдань та на їх основі створювати програмний код;
- знати основні поняття та області застосувань хмарних технологій;
- вміти оцінювати і вибирати необхідні методи та технології для хмарних обчислень та вирішення поставленої задачі.

Міждисциплінарні зв'язки.

Навчальна дисципліна «Введення до теорії хмарних обчислень» базується на знаннях, отриманих під час вивчення дисциплін «Основи програмування», «Вступ до спеціальності», «Об'єктно-орієнтоване програмування», «Організація комп'ютерних мереж».

Знання, отримані після опанування даної навчальної дисципліни, можуть бути використані при вивченні дисципліни «Введення в інженерію великих даних» та безпосередньо при розв'язанні складних спеціалізованих задач та практичних проблем у певній галузі професійної діяльності або навчанні.

3. Програма навчальної дисципліни

Розділ 1. Введення до теорії хмарних обчислень

Тема 1. Джерела великих даних. Інтернет Речей.

Інтернет Речей та зростання даних. Платформа Kaggle. DrivenData. Визначення великих даних. Приклади великих даних у реальному світі. Відкриті дані. Приватність даних. Структуровані та неструктуровані дані. Хмарні та туманні обчислення. Дані в спокої та дані в русі. Інфраструктура великих даних. Розподілені дані та їх обробка.

Тема 2. Розроблення програмного забезпечення для аналізу веб-сайтів, які надають відкриті. Відкриті дані, їх формати та засоби оброблення.

Можливості інструментів аналізу даних. Роль Python в аналізі даних. Традиційна аналітика великих даних та аналітика нового покоління. Життєвий цикл аналізу даних. Відкриті дані, їх формати та засоби обробки. Веб-скрепінг. Витягування, перетворення та завантаження даних.

Тема 3. Форматування даних про час та дату, читання та запис файлів в Python. Взаємодія із зовнішніми додатками.

Форматування даних про час та дату у Python. Зчитування та запис файлів в Python. Взаємодія із зовнішніми додатками.

Тема 4. Процедура імпорту даних із файлів у Pandas. Імпорт даних з мережі Інтернет. Засоби для кореляційного аналізу в Pandas.

Статистичні підходи до аналітики великих даних. Використання Pandas. Імпорт даних з файлів. Імпорт даних з мережі Інтернет. Описова статистика в Pandas. Засоби для кореляційного аналізу в Pandas.

Розділ 2. Технології і типи хмарних обчислень

Тема 5. Регресійний аналіз даних в Python.

Методи та типи аналізу машинного навчання. Регресійний аналіз. Типи регресійного аналізу. Застосування регресійного аналізу.

Тема 6. Помилки в аналізі даних та прогнозній аналітиці. Оцінка помилок регресії засобами Python. Призначення бібліотеки scikit-learn.

Помилки в аналізі даних та прогнозній аналітиці. Оцінка помилок регресії засобами Python. Призначення бібліотеки scikit-learn.

Тема 7. Аналіз даних в R. Фактори, списки, фрейми та дії над ними.

Історія розвитку мови R. Можливості мови R. Об'єкти, пакети, функції. Вектори, матриці та операції над ними в R. Фактори, списки, фрейми та дії над ними.

Тема 8. Експорт, імпорт та оброблення даних в R.

Експорт та імпорт даних в R. Використання R для аналізу часових рядів. Оброблення даних в R.

4. Структура навчальної дисципліни

Назви тематичних розділів і тем	Кількість годин				
	денна форма				
	усього	у тому числі			сам.роб. інд.завд.
л		с/п	лаб.		
1	2	3	4	5	6
Розділ 1. Введення до теорії хмарних обчислень					
Тема 1 Джерела великих даних. Інтернет Речей.	10	2	4		4
Тема 2. Розроблення програмного забезпечення для аналізу веб-сайтів, які надають відкриті. Відкриті дані, їх формати та засоби оброблення.	13	1	2		10
Тема 3. Форматування даних про час та дату, читання та запис файлів в Python. Взаємодія із зовнішніми додатками.	13	1	2		10
Тема 4. Процедура імпорту даних із файлів у Pandas. Імпорт даних з мережі Інтернет. Засоби для кореляційного аналізу в Pandas.	9	1	2		6
Разом за розділом 1	45	5	10		30
Розділ 2. Технології і типи хмарних обчислень					
Тема 5. Регресійний аналіз даних в Python.	10	2	4		4
Тема 7. Аналіз даних в R. Фактори, списки, фрейми та дії над ними.	13	1	2		10
Тема 7. Аналіз даних в R. Фактори, списки, фрейми та дії над ними.	13	1	2		10
Тема 8. Експорт, імпорт та оброблення даних в R.	9	1	2		6
Разом за розділом 2	45	5	10		30
Усього годин	90	10	20		60

5. Теми лекційних занять

№ з/п	Назва теми	Кількість годин
1	Тема 1. Джерела великих даних. Інтернет Речей. Розроблення програмного забезпечення для аналізу веб-сайтів, які надають відкриті. Відкриті дані, їх формати та засоби оброблення.	2
2	Тема 2. Форматування даних про час та дату, читання та запис файлів в Python. Взаємодія із зовнішніми додатками. Процедура імпорту даних із файлів у Pandas. Імпорт даних з мережі Інтернет. Засоби для кореляційного аналізу в Pandas.	2
3	Тема 3. Регресійний аналіз даних в Python. Помилки в аналізі даних та	2

	прогнозній аналітиці. Оцінка помилок регресії засобами Python. Призначення бібліотеки scikit-learn.	
4	Тема 4. Аналіз даних в R. Фактори, списки, фрейми та дії над ними.	2
5	Тема 5. Експорт, імпорт та оброблення даних в R.	2
Разом		10

6. Теми лабораторних занять

№ з/п	Назва теми	Кількість годин
1	Тема 1. Дослідження джерел відкритих даних. завантаження датасету та збереження даних в форматі csv.	4
2	Тема 2. Аналіз та візуалізація даних у Python.	4
3	Тема 3. Кореляційний аналіз у Python.	2
4	Тема 4. Побудова лінійної регресії в Python.	4
5	Тема 5. Аналіз та візуалізація даних в R.	2
6	Тема 6. Розподілені обчислення даних з використанням Spark-кластера та мови R.	4
Разом		20

7. Самостійна робота

№ з/п	Назва теми	Кількість годин
1	Тема 1. Архітектурні моделі Big Data. Технології віртуалізації. Гіпервізори. Контейнерна технологія виконання програмного коду на сервері. SaaS, PaaS і IaaS.	15
2	Тема 2. Технології Hadoop Big Data. Розподілена обробка MapReduce. HDFS.	10
3	Тема 3. Розподілена потокова платформа Kafka. Переваги Cassandra.	10
4	Тема 4. Платформа Apache Spark.	10
5	Тема 5. Lambda та Карра архітектури оброблення великих даних.	15
Разом		60

Індивідуальне завдання

Максимізувати ефективність та продуктивність в роботі, використовуючи Google Cloud та інші послуги Google. Для цього розробити інфографіку для ефективного використанні ресурсів Google, які використовуються для забезпечення росту продуктивності, автоматизації рутинних завдань та полегшенні робочих процесів.

8. Види контролю і система накопичення балів

Поточний контроль знань		Самостійна робота інд. завд.	Залік	Сума
Модульна атестація № 1 (30 балів)	Модульна атестація № 2 (30 балів)			
30	30	20	20	100

	Вид контролю	Кількість балів
Розділ 1	Захист лабораторної роботи №1	7
	Захист лабораторної роботи №2	7
	Захист лабораторної роботи №3	7
	Модульний контроль №1	9
Разом за розділом 1		30
Розділ 2	Захист лабораторної роботи №4	7
	Захист лабораторної роботи №5	7
	Захист лабораторної роботи №6	7
	Модульний контроль №2	9
Разом за розділом 2		30
Індивідуальне завдання		20
Залік		20
Всього за семестр		100

Критерії оцінювання кожного з проведених видів контролю:

1) Захист лабораторної роботи відбувається після виконання завдання та завантаження відповідного звіту до СДН Moodle.

2) Залік складається з 50 питань, кожне з яких оцінюється в залежності від рівня складності запитань.

Шкала оцінювання: національна та ECTS

ЗА ШКАЛОЮ ECTS	За шкалою університету	За національною шкалою	
		Екзамен	Залік
A	90 – 100 (відмінно)	5 (відмінно)	Зараховано
B	85 – 89 (дуже добре)	4 (добре)	
C	75 – 84 (добре)		
D	70 – 74 (задовільно)	3 (задовільно)	
E	60 – 69 (достатньо)		
FX	35 – 59 (незадовільно – з можливістю повторного складання)	2 (незадовільно)	Не зараховано
F	1 – 34 (незадовільно – з обов'язковим повторним курсом)		

9. Рекомендована література

Основна:

1. Зінченко О.В., Іщеряков С.М., Прокопов С.В., Сєрих С.О., Василенко В.В. Хмарні технології. Навчальний посібник. К: ФОП Гуляєва В.М., 2020. 74 с.
2. S. Bhowmik, Cloud Computing // Cambridge University Press, 2017. 408 p.
3. Davy Cielen, Arno D. B. Meysman, and Mohamed Ali Introducing Data Science: Big data, machine learning, and more, using Python tools. – Manning, 2019. 320 p.
4. Олексюк В. Основи хмарних технологій / В. Олексюк. Тернопіль: Тернопільський обласний інститут післядипломної педагогічної освіти, 2018. 156 с.
5. Murugesan S., Vojanova I.(eds.), Encyclopedia of Cloud Computing // John Wiley & Sons, Ltd, 2020. 725 p.

Додаткова:

1. S. Akhter, J. Roberts. Multi-Core Programming. Intel Press, 2019. 344p.
2. Richard Gerber, Aart J.C. Bik, Kevin B. Smith, and Xinmin Tian The Software Optimization Cookbook, Second Edition. Intel Press, 2019. 404p.
3. Czarnul P. Parallel Programming for Modern High Performance Computing Systems// CRC Press, 2018. 304p.
4. Kurgalin S., Borzunov S. A Practical Approach to High-Performance Computing// Springer, 2019. 206 p.

Інформаційні ресурси:

1. <https://epkmoodle.znu.edu.ua/course/view.php?id=112>.
2. Byte Size Infographic: Visualising data. URL : <https://www.redcentricplc.com/resources/infographics/byte-size/>
3. IoT Fundamentals: Big Data & Analytics. URL : <https://www.netacad.com/courses/iot/big-data-analytics>
4. Kaggle. URL : <https://www.kaggle.com/>
5. DrivenData. URL : <https://www.drivendata.org/>
6. Big Data: the 3 VS explained. URL : <https://bigdataldn.com/intelligence/big-data-the-3-vs-explained/>
7. Computing. URL : <https://home.cern/science/computing>
8. Open Knowledge Foundation. URL : <https://okfn.org>
9. Gapminder. URL : <https://www.gapminder.org>
10. Портал відкритих даних України. URL : <https://data.gov.ua>

