



Course of study (code) / Назва дисципліни (шифр)	Великі дані (Big Data) EI0711CBDAN
Academic year / Навчальний рік - Семестр	2022/2023 2 семестр
Course of study / Назва спеціальності	071 Облік і оподаткування
Educational program / Освітня програма Education - ECTS / Рівень – Кредити Status / Статус Learning language / Мова навчання	«Облік і оподаткування» Початковий рівень (короткий цикл) – 3 ECTS Вибіркова Українська
Author / Укладач	Шокотько Людмила Миколаївна, викладач Державний університет економіки і технологій e-mail: shokotko_lm@kneu.dp.ua , http://orcid.org/0000-0001-7294-2003 моб. +380679556883
Консультації	середа, 14.00-15.00

A. OBJECTIVE OF THE SUBJECT / МЕТА ТА ЗАВДАННЯ ДИСЦИПЛІНИ

Вивчення дисципліни «Великі дані (Big Data)» дозволяє сформувати у здобувачів освіти компетенції, необхідні для розв'язання практичних задач професійної та наукової діяльності, пов'язаної з підготовкою та аналізом великих даних.

Метою вивчення дисципліни «Великі дані (Big Data)» є формування у студентів здатностей застосовувати програмні методи та засоби обробки даних для аналізу великих даних та прийняття управлінських рішень в різних галузях науки та бізнесу

B. SUBJECT PROGRAM / ПРОГРАМА ДИСЦИПЛІНИ

Тема 1. Джерела великих даних.

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. Використання сучасних інформаційно-комунікаційних технологій для формування інформаційної бази (спектру) моделей.

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

Тема 3. Класифікація методів аналізу великих даних.

Особливості класифікації методів і технологій аналітики Великих даних, групи методів і технологій аналітики Великих даних, які класифікуються з урахуванням функціональних зв'язків та формальної моделі цієї інформаційної технології.

Проблеми класифікації. Алгоритми класифікації. Задача визначення концептів онтології аналітики Великих даних.

Тема 4. Аналіз та візуалізація великих даних.

Візуалізація великих даних, візуалізація текстів, візуалізація кластерів, візуалізація асоціацій, ландшафтна візуалізація, візуалізація гіпотез, візуалізація дерев рішень, багатовимірна візуалізація. Сіткова візуалізація. Класифікація. Gephі.

Тема 5. Регресійний аналіз даних в Python.

Методи та типи аналізу великих даних. Регресійний аналіз. Типи регресійного аналізу.

Застосування регресійного аналізу. Помилки в аналізі даних та прогнозній аналітиці. Оцінка помилок регресії засобами Python.



C. LIST OF COMPETENCIES AND STUDIES TARGETED RESULTS / ПЕРЕЛІК КОМПЕТЕНТНОСТЕЙ ТА ПРОГРАМНИХ РЕЗУЛЬТАТІВ НАВЧАННЯ

Загальні компетентності (ЗК)	ЗК01. Здатність реалізувати свої права і обов'язки як члена суспільства, усвідомлювати цінності громадянського (вільного демократичного) суспільства та необхідність його сталого розвитку, верховенства права, прав і свобод людини і громадянина в Україні.
Спеціальні (фахові) компетентності (ФК)	ЗК02. Здатність зберігати та примножувати моральні, культурні, наукові цінності і досягнення суспільства на основі розуміння історії та закономірностей розвитку предметної області, її місця у загальній системі знань про природу і суспільство та у розвитку суспільства, техніки і технологій. ЗК04. Здатність застосовувати методи аналізу та синтезу, вміти мислити абстрактно і критично. ЗК10. Здатність до пошуку, оброблення та аналізу інформації з різних джерел.
Програмні результати навчання (ПР)	ФК09. Здатність застосовувати знання податкового законодавства у професійній діяльності для розрахунків і сплати податків та зборів суб'єктів господарювання. ФК10. Здатність здійснювати зовнішній та внутрішній контроль діяльності підприємства з метою забезпечення його ефективного функціонування, оцінки результатів роботи та дотримання ним законодавства з бухгалтерського обліку і оподаткування. ПР05. Ідентифікувати та оцінювати ризики господарської діяльності підприємств. ПР08. Застосовувати спеціалізовані інформаційні системи і комп'ютерні технології для обліку, аналізу, контролю та оподаткування. ПР10. Вміти застосовувати знання дисциплін циклу загальної і професійної підготовки в обраній професії. ПР11. Володіти загальнонауковими та спеціальними методами дослідження соціально-економічних явищ і господарських процесів на підприємстві. ПР15. Демонструвати розуміння особливостей практики здійснення обліку, аналізу, контролю та оподаткування діяльності підприємств, установ та організацій різних форм власності та видів економічної діяльності в умовах цифрової економіки.

D. SEMESTER PLAN / СЕМЕСТРОВИЙ ПЛАН

D. SEMESTER PLAN / СЕМЕСТРОВИЙ ПЛАН

Тиждень/ Дата	Тема, план/короткі тези	Форма діяльності (заняття), години, формат	Завдання для СРС (література, ресурси в інтернеті, презентація, відеокурси)
за розкладом	Тема 1. Джерела великих даних. Інтернет Речей. 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	лекція (2 год) практичні заняття (2 год)	Опрацювання літератури: основна 1, 7, 8 додаткова 1
за розкладом	Тема 2. Використання сучасних інформаційно-комунікаційних технологій для формування інформаційної бази (спектру) моделей. Можливості інструментів аналізу даних. Традиційна аналітика великих даних та аналітика нового покоління. Відкриті дані, їх формати та засоби обробки. Веб-скрепінг. Витягування, перетворення та завантаження даних	лекція (4 год), практичні заняття (4 год)	Опрацювання літератури: основна 1 - 5 додаткова 2, 3
за розкладом	Тема 3. Класифікація методів аналізу великих даних. Особливості класифікації методів і технологій аналітики Великих даних, групи методів і технологій аналітики Великих даних, які класифікуються з урахуванням функціональних зв'язків та формальної моделі цієї інформаційної технології. Проблеми класифікації. Алгоритми класифікації. Задача визначення концептів онтології аналітики Великих даних.	лекція (4 год), практичні заняття (4 год)	Опрацювання літератури: основна 1, 6, 7 додаткова 1



за розкладом	Тема 4. Аналіз та візуалізація великих даних. Візуалізація великих даних, візуалізація текстів, візуалізація кластерів, візуалізація асоціацій, ландшафтна візуалізація, візуалізація гіпотез, візуалізація дерев рішень, багатовимірна візуалізація. Сіткова візуалізація. Класифікація. Gephi.	лекція (4 год), практичні заняття (4 год)	Опрацювання літератури: основна 3-7 додаткова 2,3
за розкладом	Тема 5. Регресійний аналіз даних в Python. Методи та типи аналізу машинного навчання. Регресійний аналіз. Типи регресійного аналізу. Застосування регресійного аналізу. Помилки в аналізі даних та прогнозній аналітиці. Оцінка помилок регресії засобами Python.	лекція (4 год) практичні заняття (4 год)	Опрацювання літератури: основна 1-3,6 додаткова 2,3

Методичні вказівки до виконання лабораторних робіт, індивідуальної та самостійної роботи містяться в системі MOODLE в методичному комплексі з вивчення дисципліни «Великі дані (Big Data)» для здобувачів вищої освіти на початковому (короткий цикл) рівні спеціальності 071 «Облік і оподаткування» денної та заочної форм навчання

E. BASIC LITERATURE (OBLIGATORY TEXTBOOKS) / ОСНОВНА ЛІТЕРАТУРА (ОБОВ'ЯЗКОВІ ПІДРУЧНИКИ)

- Олещенко Л. М. Технології оброблення великих даних: конспект лекцій з дисципліни «Технології оброблення великих даних» [Електронний ресурс] : навч. посіб. для студ. спеціальності 121 «Інженерія програмного забезпечення» / Л.М. Олещенко; КПІ ім. Ігоря Сікорського. – 2021. – 225 с.
- Олещенко Л. М. Технології оброблення великих даних: комп'ютерний практикум з дисципліни «Технології оброблення великих даних» [Електронний ресурс] : навч. посіб. для студ. спеціальності 121 «Інженерія програмного забезпечення» / Л.М. Олещенко; КПІ ім. Ігоря Сікорського. – 2021. – 85 с.
- Programming with Databases – Python // Електронний ресурс. Режим доступу: <https://swcarpentry.github.io/sql-novice-survey/10-prog/index.html>
- Силен Д., Мейсман А., Али М. Основы Data Science и Big Data. Python и наука о данных. - СПб.: Питер, 2017. - 336 с.
- Радченко И.А., Николаев И.Н. Технологии и инфраструктура Big Data. Учебное пособие. — СПб.: Университет ИТМО, 2018. — 52 с.
- Вайгенд Андреас. BIG DATA. Вся технология в одной книге. М.: Эксмо, 2018. — 384 с
- Raheem N. Big Data: A Tutorial-Based Approach. Taylor & Francis Group LLC, CRC Press, 2019. — 203 p 5.
- Волкова Светлана (ред.). Просто BIG DATA. СПб.: Страта, 2019. — 148 с.

F. COMPLEMENTARY LITERATURE / ДОДАТКОВА ЛІТЕРАТУРА

- Орлов А. Машинное обучение для Больших Данных. //Открытые системы. СУБД. - №01. – М., 2016. – Режим доступа: <https://www.osp.ru/os/2016/01/13048651>
- MapReduce Tutorial // Електронний ресурс. Режим доступу: https://hadoop.apache.org/docs/r1.2.1/mapred_tutorial.html
- Pandas // Електронний ресурс. Режим доступу: https://www.w3schools.com/python/pandas/pandas_intro.asp
- Мова програмування R // Електронний ресурс. Режим доступу: <https://coderlessons.com/tutorials/mashinnoe-obuchenie/r-programmirovaniye/r-programmirovaniye>

G. THE MOST IMPORTANT PUBLICATIONS OF THE AUTHOR(S) CONCERNING PROPOSED CLASSES / ОСНОВНІ ПУБЛІКАЦІЇ АВТОРА, ЩО ПОВ'ЯЗАНІ З ТЕМАТИКОЮ ЗАПЛАНОВАНИХ ЗАНЯТЬ

- Шокотько Л. М. Методи розподіленої обробки великих даних: концепція MAPREDUCE. // Innovations and prospects of world science. Proceedings of the 5th International scientific and practical conference. Perfect Publishing. Vancouver, Canada. 2021. Pp. 360-363. URL: <https://sci-conf.com.ua/v-mezhdunarodnaya-nauchno-prakticheskaya-konferentsiya-innovations-and-prospects-of-world-science-29-31-dekabrya-2021-goda-vankuver-kanada-arhiv>
- Шокотько Л. М. Порівняння провідних хмарних платформ як сервісів. The 7th International scientific and practical conference “Modern scientific research: achievements, innovations and development prospects” (December 19-21, 2021) MDPC Publishing, Berlin, Germany. 2021. 689 p. P. 222-227
- Шокотько Л. М. Технології інтернет-обчислень. Платформа Boinc. The 6 th International scientific and practical conference –Topical issues of modern science, society and educationII (December 26-28, 2021) SPC —Sci-conf.com.uall, Kharkiv, Ukraine. 2021. 1889 p. P. 495-499.
- Шокотько Л. М. Квантові індикатори-передвісники кризових явищ на корельованих фінансових ринках/ Л. М. Шокотько // Моніторинг, моделювання та менеджмент емерджентної економіки: зб. наук. праць Сьомої Міжнародної наук.-практ. конф., Одеса – Черкаси, 23–25 травня 2018 р. – Черкаси: видавець Вовчок О.Ю., 2018. – 284 с. <http://m3e2.ccjournals.eu/2019/m3e2-2018.pdf>
- Шокотько Л. М. Формалізація впливу інформаційного поля на виникнення квантової когерентності фінансових ринків / Л. М. Шокотько // зб. наук. праць VI Міжнародної науково-практич. конференції: Черкаси-Одеса, 24-26 травня, 2017 р. – Черкаси: видавець О.Вовчок, 2017. – 358 с.
- Шокотько Л. М. Використання теорії випадкових матриць для дослідження кореляційних ефектів складних систем / Л. М. Шокотько // Науковий вісник Міжнародного гуманітарного університету. Серія: Економіка і менеджмент.–Вип. 9.127-132– Одеса, – 2015.

7. Shokotko L. Study of correlated complex systems by network methods / L. Shokotko // збірник наукових праць за матеріалами IV Міжнародної науково-практичн. конференції, 2-4 жовтня 2014р–Хмельницький, 2014.–350с
8. Шокотько Л. М. Використання спектральних характеристик квантових графів для моделювання фондових ринків / Л. М. Шокотько // збірник наукових праць за матеріалами IX Міжнародної наук.-практ. конф. аспір., мол. учених та науковців): 4 квітня, 2014 р.– У 2-х томах.–Кременчук: КІ ДУЕП, 2014, – с. 268

H. PREREQUISITE AND POSTREQUISITE / ПРЕРЕКВІЗИТИ ТА ПОСТРЕКВІЗИТИ

Теоретико-методологічною базою для вивчення дисципліни є статистика, теорія ймовірності та математична статистика, вища математика, алгоритмізація та програмування.

I. SCOPE AND TYPE / КІЛЬКІСТЬ ВІДВЕДЕНИХ ГОДИН ТА ФОРМА ПРОВЕДЕННЯ ЗАНЯТЬ

	Денна	Заочна
Лекції	18	
Практичні (лабораторні)	18	
Самостійна робота студента (СРС)	54	
Індивідуально-консультативна робота (ІКР)		
Курсова робота	-	
Разом годин	90	

J. CURRENT AND FINAL EVALUATION / ПОТОЧНЕ ТА ПІДСУМКОВЕ ОЦІНЮВАННЯ

	Денна	Заочна
Поточний контроль, в т.ч.:	100	
оцінювання під час аудиторних занять	50	
виконання контрольних (модульних) робіт	20	
виконання і захист завдань самостійної роботи	30	
науково-дослідницька робота		
Підсумковий контроль (екзамен)		
Разом	100	

Шкала балів	Оцінка за 4-бальною шкалою	Шкала ECTS
90 – 100	Відмінно	A
80 – 89	Добре	B
70 – 79		C
66 – 69		D
60 – 65	Задовільно	E
21 – 59	незадовільно з можливістю повторного складання екзамену (заліку)	FX
0 – 20	незадовільно з можливістю вивчення дисципліни за індивідуальним графіком у формі додаткової індивідуально-консультативної роботи	F

K. CODE OF CONDUCT OF THE COURSE / КОДЕКС ПОВЕДІНКИ ПІД ЧАС ВИВЧЕННЯ КУРСУ

Для успішного проходження курсу та складання контрольних заходів необхідним є виконання наступних обов'язків:

- ❖ не запізнюватися на заняття;
- ❖ не пропускати заняття (як лекційні, так і практичні), в разі хвороби мати довідку або її ксерокопію;
- ❖ самостійно опрацювати весь лекційний матеріал та ресурси для самостійної роботи;
- ❖ при підготовці есе, есе-рефлексії на статтю, що запропонована вам для читання (див. семестровий план), оцінюється якість та оригінальність наведених вами аргументів. Есе повинно бути надіслано до 16:00 у попередній день перед семінаром. Усі повинні обов'язково підготувати есе, а його відсутність з будь-яких причин не може бути виправданням.
- ❖ конструктивно підтримувати зворотній зв'язок з викладачем на всіх етапах проходження курсу (особливо під час виконання індивідуальних проектів/курсів проекту);
- ❖ своєчасно і самостійно виконувати всі передбачені програмою лабораторні та практичні завдання;
- ❖ брати очну участь у контрольних заходах;

будь-яке відтворення результатів чужої праці (виключаючи практичну роботу над командним проектом), в тому числі використання завантажених з Інтернету матеріалів, як власних результатів, кваліфікується, як порушення норм і правил академічної доброчесності, та передбачає притягнення до відповідальності у порядку, визначеному чинним законодавством.



L. METHODS OF CONDUCTING / МЕТОДИ НАВЧАННЯ

Для формувань умінь та навичок застосовуються такі методи навчання:

- вербальні/словесні (лекція, пояснення, розповідь, бесіда, інструктаж);
- наочні (спостереження, ілюстрація, демонстрація);
- практичні (різні види вправлення, виконання графічних робіт, проведення експерименту, практики);
- пояснювально-ілюстративний або інформаційно-рецептивний, який передбачає пред'явлення готової інформації викладачем та її засвоєння студентами;
- репродуктивний, в основу якого покладено виконання різного роду завдань за зразком;
- метод проблемного викладу.
- частково-пошуковий або евристичний.
- Дослідницький.

M. TOOLS, EQUIPMENT AND SOFTWARE / ІНСТРУМЕНТИ, ОБЛАДНАННЯ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ

Zoom – хмарна платформа для відео і аудіо конференцій та вебінарів.

Telegram – програма месенджер.

viber – програма для відео та голосового зв'язку.

ZELIS - система призначена для тестування знань студентів в двох режимах: автоматизований контроль знань та тестування по бланкам.

N. STUDENT RESOURCES, MOOC PLATFORMS / ЦИФРОВІ РЕСУРСИ ДЛЯ СТУДЕНТІВ ТА ВІДКРИТІ ДИСТАНЦІЙНІ ОНЛАЙН КУРСИ

Студентам пропонується доступ до навчальних матеріалів дисципліни - moodle.kneu.dp.ua:

[Coursera](#) – безкоштовні онлайн-курси з різних дисциплін, у разі успішного закінчення яких користувач отримує сертифікат про проходження курсу.

[EdX](#) – онлайн-курси від закладів вищої освіти.

[Prometheus](#) — український громадський проект масових відкритих онлайн-курсів.

O. FEEDBACK/ ЗВОРОТНІЙ ЗВ'ЯЗОК

Електронні листи є найкращим способом зв'язатися з керівником курсу, і, будь ласка, додайте шифр групи в темі листа. Якщо ви надішлете мені електронне повідомлення, надайте мені, принаймні, 24 години, щоб відповісти. Якщо ви не отримуєте відповідь, відправте листа повторно.

P. ACADEMIC HONESTY/ АКАДЕМІЧНА ДОБРОЧЕСНІСТЬ

Державний університет економіки і технологій очікує від студентів розуміння та підтримання високих стандартів академічної чесності. Приклади академічної не доброчесності включають такі: плагіат, зловживання інформацією із застарілих джерел мережі. Очікується, що вся робота, виконана відповідно до вимог курсу, є власною роботою студента. Під час підготовки роботи, яка відповідає вимогам курсу, студенти повинні відрізнити власні ідеї від інформації, отриманої з інших джерел. Без попереднього письмового схвалення викладачем, студенти можуть не подавати один і той же звіт двічі.

Обов'язково вказати на положення про доброчесність й зробити гіперпосилання на сайт Університету (Положення про академічну доброчесність у Державному університеті економіки і технологій. Затверджено Вченою радою Державного університету економіки і технологій, Протокол № 5 від 25 листопада 2021 р.)

https://www.duet.edu.ua/uploads/normbase/243/pol_AD.pdf

APPROVED / ЗАТВЕРДЖЕНО

Рішенням кафедри «Економіки та цифрового бізнесу» Державного університету економіки і технологій - протокол № 1 від 05 жовтня 2022 року

Укладач

ЗАТВЕРДЖЕНО:

Кафедрою економіки та цифрового бізнесу
Протокол № 1 від 05 жовтня 2022 року
В.о. завідувача кафедри

Науково-методичною радою Державного університету
економіки і технологій
Протокол № 4 від 30 листопада 2022 року
Голова науково-методичної ради

Людмила ШОКОТЬКО

Вікторія СОЛОВІЙОВА

Валентин ОРЛОВ