

СИЛАБУС

Назва дисципліни: Інтелектуальний аналіз даних				
<p>Мета дисципліни: забезпечити теоретичну та практичну підготовку щодо методів сучасної обробки даних (Data Mining), а саме: ознайомлення та практичне застосування методів та алгоритмів Data Mining, Text Mining, WebMining; пошук у необроблених масивах даних знань та закономірностей, необхідних для прийняття обґрунтованих рішень.</p> <p>Основні компетентності, що формуються:</p> <p>ІК-1. Здатність розв'язувати складні спеціалізовані задачі та практичні проблеми у галузі комп'ютерних наук або у процесі навчання, що передбачає застосування певних теорій та методів і має комплексний характер.</p> <p>ЗК1. Здатність до абстрактного мислення, аналізу та синтезу.</p> <p>ЗК2. Здатність застосовувати знання у практичних ситуаціях.</p> <p>ЗК3. Знання та розуміння предметної області та розуміння професійної діяльності.</p> <p>ЗК6. Здатність вчитися й оволодівати сучасними знаннями.</p> <p>ЗК7. Здатність до пошуку, оброблення та аналізу інформації з різних джерел.</p> <p>ЗК8. Здатність генерувати нові ідеї (креативність).</p> <p>ЗК10. Здатність бути критичним і самокритичним</p> <p>ЗК11. Здатність приймати обґрунтовані рішення.</p> <p>ЗК12. Здатність оцінювати та забезпечувати якість виконуваних робіт.</p> <p>ЗК13. Здатність діяти на основі етичних міркувань.</p> <p>СК1.2 Здатність до обґрунтування вибору методів і підходів для розв'язування теоретичних і прикладних задач у галузі комп'ютерних наук, аналізу та інтерпретування.</p> <p>СК2. Здатність до виявлення статистичних закономірностей недетермінованих явищ, застосування методів обчислювального інтелекту, зокрема статистичної, нейромережевої та нечіткої обробки даних, методів машинного навчання та генетичного програмування тощо.</p> <p>СК7. Здатність проводити обчислювальні експерименти з обробкою й аналізом результатів.</p> <p>СК11. 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.</p>				
Мова викладання	Семестр	Кредити ECTS / Тип дисципліни (обов'язкова, вибіркова)	Викладач	Навчальне навантаження
Укр.	5	5 / обов'язкова	Барташевська Ю.М., к.е.н., доц., Rizun N., Ph.D., Associate Professor	150 год. (28 год. лекцій, 28 год. лабораторних занять, 15 год. інд. роботи, 79 год. самостійної роботи)
Результати навчання По закінченню вивчення дисципліни здобувачі будуть здатні		Методи викладання, навчання		Форми оцінювання (поточний та підсумковий контроль)
РН-1. застосовувати знання основних форм і законів абстрактно-логічного мислення, основ методології наукового пізнання, форм і методів вилучення, аналізу, обробки та синтезу інформації в предметній області комп'ютерних наук		Лекція, семінар-діалог, розбір/ аналіз ситуаційних задач		Участь в семінарі, відповіді на запитання, вирішення ситуаційних задач
РН-3. використовувати знання закономірностей		Лекція, семінар-діалог, розбір практичних case-study, проблемно-пошуковий метод з використанням мережі		Усні відповіді на запитання, вирішення/ пояснення

<p>випадкових явищ, їх властивостей та операцій над ними, моделей випадкових процесів та сучасних програмних середовищ для розв'язування задач статистичної обробки даних і побудови прогнозних моделей.</p> <p>РН-4. використовувати методи обчислювального інтелекту, нейромережевої та нечіткої обробки даних, генетичного та еволюційного програмування для розв'язання задач розпізнавання, прогнозування, класифікації, ідентифікації об'єктів керування тощо.</p> <p>РН-12. 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.</p>	<p>Інтернет</p> <p>Проблемна лекція, вирішення практичних case-study, обговорення вирішення завдань в мікрогрупах, проблемно-пошуковий метод з використанням мережі Інтернет, самостійна робота</p> <p>Проблемна лекція, вирішення практичних case-study, обговорення вирішення завдань в мікрогрупах, проблемно-пошуковий метод з використанням мережі Інтернет, самостійна робота</p>	<p>практичних case-study</p> <p>Усні відповіді на запитання, вирішення/ пояснення практичних case-study, презентація вирішення завдань в мікрогрупах, оцінювання практичних навичок, презентація результатів самостійної роботи</p> <p>Усні відповіді на запитання, вирішення/ пояснення практичних завдань, оцінювання практичних навичок, презентація результатів самостійної роботи</p>
---	---	--

Оцінка

Підсумкова оцінка в результаті 100% постійного оцінювання:

60% виконання індивідуальних практичних завдань
40% підсумковий контроль (екзамен: тести, есе, міні-case)

Критерії оцінювання:

1 модуль (max – 30 балів). Бали за 1 модуль здобувач отримує за умови виконання 5 поточних робіт, чотири з яких оцінюються по 5 балів кожна, одна – в 10 балів:

Критерії оцінювання (4 роботи, кожна по 5 балів відповідно):

5 балів – здобувач вірно виконав роботу, демонструє глибоке розуміння матеріалу. Вірно обрано алгоритм реалізації, якісне представлення результатів. Обґрунтовані висновки.

4 бали – здобувач виконав роботу, однак є незначні неточності, що не здатні вплинути на кінцевий результат. Зроблено висновки і якісне подання результатів.

3 бали – здобувач демонструє недостатнє розуміння матеріалу. Є помилки у виборі та реалізації алгоритму рішення. Відсутні висновки і здобувач не може якісно пояснити отриманий результат. Завдання виконано частково або в загальному вигляді.

2 бали – здобувач демонструє переважне незнання матеріалу, невірно обрано алгоритм реалізації і отриманий результат не є кінцевим, містить істотні помилки.

1 бал – здобувач демонструє незнання матеріалу. Виконання завдання не доведено до кінця, а наявне рішення містить грубі помилки.

0 балів – завдання не виконано здобувачем.

Критерії оцінювання (1 робота, 10 балів):

9-10 балів – здобувач вірно виконав роботу, демонструє глибоке розуміння матеріалу. Вірно обрано алгоритм реалізації, якісне представлення результатів. Обґрунтовані висновки.

7-8 бали – здобувач виконав роботу, однак є незначні неточності, що не здатні вплинути на кінцевий результат. Зроблено висновки і якісне подання результатів.

5-6 бали – здобувач демонструє недостатнє розуміння матеріалу. Є помилки у виборі та реалізації алгоритму рішення. Відсутні висновки і здобувач не може якісно пояснити отриманий результат. Завдання виконано частково або в загальному вигляді.

3-4 бали – здобувач демонструє переважне незнання матеріалу, невірно обрано алгоритм реалізації і отриманий результат не є кінцевим, містить істотні помилки.

1-2 бали – здобувач демонструє незнання матеріалу. Виконання завдання не доведено до кінця, а наявне рішення містить грубі помилки.

0 балів – завдання не виконано здобувачем

2 модуль (max – 30 балів). Бали за 2 модуль здобувач отримує, за умови виконання 3 поточних робіт, кожна відповідно оцінюється у 5, 15 і 10 балів:

Критерії оцінювання першої роботи (максимально 5 балів):

5 балів – здобувач вірно виконав роботу, демонструє глибоке розуміння матеріалу. Вірно обрано алгоритм реалізації, якісне представлення результатів. Обґрунтовані висновки.

4 бали – здобувач виконав роботу, однак є незначні неточності, що не здатні вплинути на кінцевий результат. Зроблено висновки і якісне подання результатів.

3 бали – здобувач демонструє недостатнє розуміння матеріалу. Однак є помилки у виборі та реалізації алгоритму рішення. Відсутні висновки і здобувач не може якісно пояснити отриманий результат. Завдання виконано частково або в загальному вигляді.

2 бали – здобувач демонструє незнання матеріалу, невірно обрано алгоритм реалізації і отриманий результат не є кінцевим, містить істотні помилки.

1 бал – здобувач демонструє незнання матеріалу. Виконання завдання не доведено до кінця, а наявне рішення містить грубі помилки.

0 балів – завдання не виконано здобувачем

Критерії оцінювання другої роботи (максимально 15 балів):

13-15 балів – здобувач вірно виконав роботу, демонструє глибоке розуміння матеріалу. Вірно обрано алгоритм реалізації, якісне представлення результатів. Обґрунтовані висновки.

10-12 бали – здобувач виконав роботу, однак є незначні неточності, що не здатні вплинути на кінцевий результат. Зроблено висновки і якісне подання результатів.

7-9 бали – здобувач демонструє недостатнє розуміння матеріалу. Однак є помилки у виборі та реалізації алгоритму рішення. Відсутні висновки і здобувач не може якісно пояснити отриманий результат. Завдання виконано частково або в загальному вигляді.

5-6 бали – здобувач демонструє незнання матеріалу, невірно обрано алгоритм реалізації і отриманий результат не є кінцевим, містить істотні помилки.

2-4 бали – здобувач демонструє незнання матеріалу. Виконання завдання не доведено до кінця, а наявне рішення містить грубі помилки.

0 балів – завдання не виконано здобувачем

Критерії оцінювання третьої роботи (максимально 10 балів):

9-10 балів – здобувач вірно виконав роботу, демонструє глибоке розуміння матеріалу. Вірно обрано алгоритм реалізації, якісне представлення результатів. Обґрунтовані висновки.

7-8 бали – здобувач виконав роботу, однак є незначні неточності, що не здатні вплинути на кінцевий результат. Зроблено висновки і якісне подання результатів.

5-6 бали – здобувач демонструє недостатнє розуміння матеріалу. Однак є помилки у виборі та реалізації алгоритму рішення. Відсутні висновки і здобувач не може якісно пояснити отриманий результат. Завдання виконано частково або в загальному вигляді.

3-4 бали – здобувач демонструє незнання матеріалу, невірно обрано алгоритм реалізації і отриманий результат не є кінцевим, містить істотні помилки.

1-2 бали – здобувач демонструє незнання матеріалу. Виконання завдання не доведено до кінця, а наявне рішення містить грубі помилки.

0 балів – завдання не виконано здобувачем.

По закінченні курсу здобувачі складають екзамен, максимальна кількість балів за який – 40. Кількість балів за кожне завдання (тест, есе, міні-case) та критерії оцінювання наведені в екзаменаційному білеті.

Зміст

Змістовий модуль 1. Візуальний аналіз даних. Tableau

Тема 1. Аналіз даних: поняття, види, застосування

Тема 1.1. Забезпечення дотримання академічної доброчесності викладачами та здобувачами вищої освіти, у тому числі створення і забезпечення функціонування ефективної системи запобігання та виявлення академічного плагіату.

Тема 2. Знайомство з Tableau

Тема 3. Робота з даними та функціями у Tableau

Тема 4. Візуалізація даних

Тема 5. Дашборди та карти

Змістовий модуль 2. Мова програмування R

Тема 6. Основи R. Статистика. Описова статистика

Тема 7. Основи R. Матриці та вектори

Тема 8. Візуалізація

Тема 9. Аналітична статистика. Порівняння груп виборок.

Змістовий модуль 3. Text Mining

Тема 10. Завдання аналізу текстів (етапи аналізу текстів, попередня обробка тексту, завдання Text Mining)

Тема 11. Вилучення ключових понять із тексту

Тема 12. Класифікація текстових документів

Тема 13. Методи кластеризації текстових документів. Аналіз настрою. Аналіз тональності

Тема 14. Кількісний аналіз тексту

Література

Основна

1. Литвин В.В., Пасічник В.В., Нікольський Ю.В. Аналіз даних та знань: підручник. - Л.: Магнолія, 2020.- 276с. (базовий підручник).
2. Фісун М.Т., Кравець І.О., Казмірчук П.П. Інтелектуальний аналіз даних: практикум. - Л.: Новий світ 2000, 2021. - 162 с.
2. Сидорова А. В., Біленко Д. В., Буркіна Н. В. Бізнес-аналітика: навчально-методичний посібник. - Вінниця: ДонНУ імені Василя Стуса. 2019. - 104 с.
3. Hadley Wickham, Garrett Grolemund R for Data Science: Import, Tidy, Transform, Visualize, and Model Data. - Sebastopol: O'Reilly Media, 2017. - 520 p.
4. О'Ніл К. BIG DATA. Зброя математичного знищення. - К.: Форс, 2019. - 336 с.
5. Ланде Д.В., Субач І.Ю., Бояринова Ю.Є. Основи теорії і практики інтелектуального аналізу даних у сфері кібербезпеки: навчальний посібник. - К.: ІСЗІ КПІ ім. Ігоря Сікорського», 2018. - 300 с.
6. Акіменко В.В. Прикладні задачі інтелектуального аналізу даних (DATA MINING). - К.: КНУ ім. Тараса Шевченка, 2018. - 152 с.

Додаткова

1. Олещенко Л.М. Технології оброблення великих даних: конспект лекцій. - К.: КПІ ім. Ігоря Сікорського, 2021. - 227 с.
2. Brian Steele, John Chandler, Swarna Reddy Algorithms for Data Science. - Switzerland: Springer, 2016. - 438 p.
3. Сергєєв-Горчинський О.О., Іщенко Г.В. Інтелектуальний аналіз даних: комп'ютерний практикум. - Київ : КПІ ім. Ігоря Сікорського, 2018. - 75 с.
4. <https://www.tutorialspoint.com/r/index.htm>
5. Kantarjic M. Data Mining. Concepts, Models, Methods and Algorithms: 3rd Ed. - Publisher: Wiley, 2019. - 672 p.
6. Ethical Dimensions of Renewable Energy and Sustainability Systems: 2.1 Falsification, fabrication, plagiarism. The John A. Dutton e-Education Institute of The Pennsylvania State University. URL:

<https://www.e-education.psu.edu/bioet533/node/654>

7. Електронний курс: <https://www.coursera.org/learn/r-programming>

Політика курсу

Політика щодо відвідування занять: Здобувачі мають відвідувати заняття регулярно. У випадку ситуацій, коли здобувач пропускає заняття, він несе особисту відповідальність за опрацювання матеріалів лекції, розміщених у Google Classroom. Частина матеріалу, який виноситься на іспит у вигляді есе та тесту, базується на лекціях. Пропущені заняття здобувач має відпрацювати, захистивши виконані практичні завдання під час чергової консультації викладача.

Здобувачі з особливими освітніми потребами: Мають право на індивідуальне визначення способів проходження поточного модульного та підсумкового контролю за письмовою заявою, яка подається до загального деканату на початку викладання курсу. Можливе навчання за індивідуальним графіком, який оформлюється відповідно до п. 3.4 Положення про організацію освітнього процесу.

Академічна доброчесність: Здобувач має усвідомити, що академічна недоброчесність є неприпустимою. Викриття будь-якого порушення академічної доброчесності під час виконання будь-якого завдання призведе до його нульової оцінки. Порушення академічної доброчесності на екзамені призведе до негативної оцінки за весь курс та можливого виключення з програми. Під час екзамену здобувачам забороняється користуватися жодним електронним пристроєм (окрім ПК для виконання завдання), навчальними та додатковими матеріалами. Всі суперечливі питання, у разі їх виникнення, можуть бути врегульовані шляхом звернення до Комісії з академічної доброчесності та етики, відповідно до п. 4.9 Положення про організацію освітнього процесу.

Політика щодо використання телефонів та інших електронних пристроїв: Під час проведення навчальних занять електронні пристрої та телефони мають перебувати в безшумному режимі роботи і можуть використовуватися для доступу до навчальних матеріалів у Google Classroom. У разі невиконання даної вимоги, викладач може запропонувати здобувачу залишити аудиторію.

Політика щодо скарг здобувачів. Здобувач може обговорити проблемне питання з викладачем після заняття. Якщо питання залишається невирішеним, здобувач має право звернутися до завідувача кафедри інформаційних технологій.

Політика щодо підвищення оцінки з дисципліни: Здобувач має право підвищити оцінку з дисципліни відповідно до пп. 2.4.5. Положення про організацію освітнього процесу. Заява на підвищення оцінки має бути оформлена у загальному деканаті.

Пропозиції від здобувачів вищої освіти: Протягом вивчення курсу здобувачі можуть звернутися до викладача з пропозиціями щодо вдосконалення курсу (доповнення тем, зміни методів викладання, форм оцінювання та ін.). Дані пропозиції можуть бути висловлені усно або письмово (електронною поштою, коментарі у Google Classroom). Для вирішення будь-якого питання, яке пов'язане із вивченням даної дисципліни, здобувач може звернутися до викладача (усно – в ауд. 2302 або письмово (bartashevaska@duan.edu.ua) або до гаранта ОПП (bartashevaska@duan.edu.ua)).