

СИЛАБУС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ «ОСНОВИ DATA SCIENCE»

Лектор курсу			Корецька Вікторія Олександрівна, кандидат педагогічних наук, доцент, професор кафедри ІТ		Контактна інформація лектора (e-mail), сторінка курсу в Google Classroom		e-mail: o.koretskyi@duikt.edu.ua сторінка курсу в Google Classroom – https://classroom.google.com/c/NzIxMjIwMDMwNTA3?cjc=vhfljp7	
Галузь знань					Рівень вищої освіти		Доктор філософії	
Спеціальність					Семестр			
Освітня програма					Тип дисципліни		Цикл вибірових компонент	
Обсяг:	Кредитів ECTS	Годин	За видами занять:					
			Лекцій	Семінарських занять	Практичних занять	Лабораторних занять	Самостійна підготовка	
	3	90	18	-	-	36	36	
АНОТАЦІЯ КУРСУ								
Взаємозв’язок у структурно-логічній схемі								
Освітні компоненти, які передують вивченню			-					
Освітні компоненти для яких є базовою			-					
Мета курсу:	Вивчення методів машинної (комп’ютерної) підготовки та обробки розрізнених і неупорядкованих даних задля пошуку та виявлення латентних закономірностей їх вмісту, походження, цінності.							
Компетентності відповідно до освітньої програми								
Soft- skills / Загальні компетентності (ЗК)					Hard-skills / Спеціальні компетентності (СК)			
Здатність до абстрактного мислення, аналізу та синтезу.					Здатність використовувати системний аналіз як сучасну міждисциплінарну методологію, що базується на прикладних математичних методах і орієнтована на вирішення задач аналізу і синтезу технічних, економічних, соціальних, екологічних та інших складних систем. Здатність використовувати сучасні інформаційні технології для комп’ютерної реалізації математичних моделей та прогнозування поведінки конкретних систем, а саме: об’єктно-орієнтований підхід при проектуванні складних систем різної природи, прикладні математичні пакети, застосування баз даних і знань.			
Програмні результати навчання (ПР)								
Проектувати, реалізовувати, тестувати, впроваджувати, супроводжувати, експлуатувати програмні засоби роботи з даними і знаннями в комп’ютерних системах та мережах.								

ОРГАНІЗАЦІЯ НАВЧАННЯ

Тема, опис теми	Вид заняття	Оцінювання за тему	Форми і методи навчання/питання до самостійної роботи
Тема 1 Вступ. Типи даних. Структуровані дані, неструктуровані дані. Дані природньою мовою, машинні дані. Графові (мережні) дані, аудіо, відео й графіка, потокові дані. <u>Рекомендовані джерела:</u> 1,2,5			
Заняття 1.1 Вступ. Типи даних. Структуровані дані, неструктуровані дані. Дані природньою мовою, машинні дані. Графові (мережні) дані, аудіо, відео й графіка, потокові дані	Лекція 1 2 год	2,5 бали	Пояснювально-ілюстративний, лекція-візуалізація.
Заняття 1.2. Типи даних. Структуровані дані, неструктуровані дані. Дані природньою мовою, машинні дані.	Лабораторн е заняття 1 2 год	1,5 бали	Виконання завдань на практичне застосування знань і вмінь.
Заняття 1.3. Графові (мережні) дані, аудіо, відео й графіка, потокові дані	Лабораторн е заняття 2 2 год	1,5 бали	Виконання завдань на практичне застосування знань і вмінь.
Тема 2. Область визначення великих даних і data science. Розподілені файлові системи. Інфраструктура розподіленого програмування. Інфраструктура інтеграції даних. Інфраструктури машинного навчання бази даних NoSQL. Інструменти планування. <u>Рекомендовані джерела:</u> 1,2,3,5			
Заняття 2.1. Область визначення великих даних і data science. Розподілені файлові системи. Інфраструктура розподіленого програмування. Інфраструктура інтеграції даних. Інфраструктури машинного навчання бази даних NoSQL. Інструменти планування.	Лекція 2,3 4 год	2,5 бали	Інтерактивна лекція, бліц опитування.
Заняття 2.2. Область визначення великих даних і data science. Розподілені файлові системи.	Лабораторн е заняття 3 2 год	1,5 бали	Виконання завдань на практичне застосування знань і вмінь.
Заняття 2.3. Інфраструктура розподіленого програмування. Інфраструктура інтеграції даних. Інфраструктури машинного навчання бази даних NoSQL. Інструменти планування.	Лабораторн е заняття 4 2 год	1,5 бали	Виконання завдань на практичне застосування знань і вмінь, розробка архітектури системи.

Заняття 2.4. Інструменти порівняльного аналізу. Розгортання системи. Програмування служб. Безпека. Приклад використання Nadoop.	Лабораторн е заняття 5,6 4 год	3 бали	Виконання завдань на практичне застосування знань і вмінь.
Тема 3. Визначення цілей дослідження й створення проектного завдання. Збір даних. Очищення, інтеграція й перетворення даних. Дослідницький аналіз даних. Побудова моделей. Представлення результатів і побудова додатків на їхній основі. <u>Рекомендовані джерела:</u> 1,2,3,5			
Заняття 3.1. Визначення цілей дослідження й створення проектного завдання. Збір даних. Очищення, інтеграція й перетворення даних. Дослідницький аналіз даних. Побудова моделей. Представлення результатів і побудова додатків на їхній основі.	Лекція 4,5 4 год	2,5 бали	Пояснювально-ілюстративний, лекція-візуалізація, бліц опитування.
Заняття 3.2. Визначення цілей дослідження й створення проектного завдання. Збір даних.	Лабораторн е заняття 7,8 4 год	3 бали	Виконання завдань на практичне застосування знань і вмінь.
Заняття 3.3. Очищення, інтеграція й перетворення даних. Дослідницький аналіз даних.	Лабораторн е заняття 9 2 год	1,5 бали	Виконання завдань на практичне застосування знань і вмінь.
Заняття 3.4. Побудова моделей. Представлення результатів і побудова додатків на їхній основі	Лабораторн е заняття 10 2 год	1,5 бали	Виконання завдань на практичне застосування знань і вмінь.
Тема 4. Застосування машинного навчання в процесі data science. Інструменти Python, використовувані в машинному навчанні. Створення нових показників і вибір моделі, тренування моделі, перевірка адекватності моделі, прогнозування нових спостережень. <u>Рекомендовані джерела:</u> 1,2,3,5			
Заняття 4.1. Застосування машинного навчання в процесі data science. Інструменти Python, використовувані в машинному навчанні. Створення нових показників і вибір моделі, тренування моделі, перевірка адекватності моделі, прогнозування нових спостережень.	Лекція 6,7 4 год	2,5 балів	Пояснювально-ілюстративний, лекція-візуалізація, бліц опитування.
Заняття 4.2. Застосування машинного навчання в процесі data science.	Лабораторн е заняття 11 2 год	2,5 балів	Виконання завдань на практичне застосування знань і вмінь.
Заняття 4.3. Інструменти Python, використовувані в машинному навчанні.	Лабораторн е заняття 12 2 год	3 бали	Виконання завдань на практичне застосування знань і вмінь.

Заняття 4.4. Створення нових показників і вибір моделі, тренування моделі, перевірка адекватності моделі, прогнозування нових спостережень. Типи машинного навчання	Лабораторн е заняття 13,14 4 год	3 бали	Виконання завдань на практичне застосування знань і вмінь.
Тема 5. Проблеми при роботі з великими об'ємами даних і загальні методи обробки великих об'ємів даних. Правильний вибір алгоритму, структури даних, інструментів. <u>Рекомендовані джерела:</u> 1,2,3,5			
Заняття 5.1. Проблеми при роботі з великими об'ємами даних і загальні методи обробки великих об'ємів даних. Правильний вибір алгоритму, структури даних, інструментів.	Лекція 8,9 4 год	2,5 балів	Інтерактивна лекція.
Заняття 5.2. Проблеми при роботі з великими об'ємами даних і загальні методи обробки великих об'ємів даних.	Лабораторн е заняття 15,16 4 год	4,5 балів	Виконання завдань на практичне застосування знань і вмінь.
Заняття 5.3. Правильний вибір алгоритму, структури даних, інструментів.	Лабораторн е заняття 17,18 4 год	4,5 балів	Виконання завдань на практичне застосування знань і вмінь.
Самостійна робота			
Тема 1. Прогнозування шкідливих адрес URL. Побудова рекомендаційної системи. Прогнозування шкідливих адрес URL. Побудова рекомендаційної системи.		10 балів	Підготовка до лабораторних робіт. Самостійне вивчення теоретичного матеріалу.
Тема 2. Перевірка гіпотез.		5 балів	Підготовка до лабораторних робіт. Самостійне вивчення теоретичного матеріалу.
МАТЕРІАЛЬНО-ТЕХНІЧНЕ ЗАБЕЗПЕЧЕННЯ ДИСЦИПЛІНИ			
~ Мультимедійний проектор. ~ Комп'ютерний клас для проведення практичних занять з встановленим програмним забезпеченням Huawei, Cisco Packet Tracer. ~ Мережне обладнання Huawei, Cisco. ~ Середовище розробки JetBrains WebStorm; Текстовий редактор Microsoft Visual Studio Code; Система контролю версій Git; Програмне середовище Node.js; Веб-сервіс для зберігання Google Диск; Веб-сервіс для хостінга ІТ-проектів GitHub.			
ІНФОРМАЦІЙНЕ ЗАБЕЗПЕЧЕННЯ ДИСЦИПЛІНИ			
1. Таран В.І., Гордієнко Ю.Г., Стіренко С.Г. Технології BigData. Практикум. Київ : КПІ ім. І. Сікорського, 2021 . - 56 с . - https://ela.kpi.ua/handle/123456789/50014 2. Peng, R.D. R Programming for Data Science [Електронний ресурс] / R.D. Peng. — Lean Publishing. — 140 p. 3. Наука про дані (https://www.coursera.org/specializations/jhu-data-science). 4. The Data Mine. URL: http://www.the-data-mine.com . 5. Олещенко Л.М. Технології оброблення великих даних. Комп'ютерний практикум. Київ: вид-во КПІ ім. І.Сікорського, 2021. - 85 с. - https://ela.kpi.ua/handle/123456789/45890 .			

ПОЛІТИКА КУРСУ («ПРАВИЛА ГРИ»)

- Курс передбачає роботу в колективі.
- Середовище в аудиторії є дружнім, творчим, відкритим до конструктивної критики. Спілкуючись з учасниками навчального процесу, студенти мають дотримуватися етичних норм, утримуватися від гучних проявів емоцій, бути політично коректними й толерантними, поважати звичай й традиції різних етнічних, культурних, соціальних груп і релігійних конфесій.
- Освоєння дисципліни передбачає обов'язкове відвідування лекцій, практичних і лабораторних занять, а також самостійну роботу.
- Студенти зобов'язані відвідувати заняття за обраним і затвердженим індивідуальним навчальним планом та вчасно інформувати викладача про неможливість із поважних причин відвідувати заняття, бути присутніми на заліку.
- Самостійна робота включає в себе теоретичне вивчення питань, що стосуються тем лекційних занять, які не ввійшли в теоретичний курс, або ж були розглянуті коротко, їх поглиблена проробка за рекомендованою літературою.
- Усі завдання, передбачені програмою, мають бути виконані у встановлений термін.
- Якщо студент із поважних причин був відсутній на практичному чи лабораторному занятті, він має право його відпрацювати. Відпрацювання полягає у виконанні індивідуального завдання за прикладом, наданим викладачем. Якщо для виконання завдання необхідно використання обладнання лабораторій кафедри, тоді час відпрацювання оговорується з викладачем індивідуально і погоджується з завідувачем відповідної лабораторії, де розміщено обладнання.
- Під час роботи над завданнями не допустимо порушення академічної доброчесності: при використанні Інтернет ресурсів та інших джерел інформації студент повинен вказати джерело, використане в ході виконання завдання. У разі виявлення факту плагіату студент отримує за завдання 0 балів.
- За порушення дисципліни студент видаляється з заняття, за заняття отримує 0 балів.

КРИТЕРІЇ ТА МЕТОДИ ОЦІНЮВАННЯ

Умовою допуску до підсумкового контролю є виконання всіх практичних та лабораторних робіт і виконання самостійних завдань, які передбачені структурою освітньої компоненти Основи Data Science.

Якщо студента не допущено до складання заліку, як такого, що не виконав індивідуальний план, йому надається час до перескладання для виконання всіх вимог допуску. Студент має право на два перескладання. При повторному перескладанні заліку його у студента може приймати комісія, яка створюється директором ННІТ. Оцінка комісії є остаточною. У випадку отримання студентом 0 балів (неприйнятно), що тягне відрахування за невиконання навчального плану.

Оцінювання студентів здійснюється за накопичувальною 100-бальною системою.

Для отримання додаткових балів, студент повинен надати копію друкованої публікації чи письмове повідомлення видавця, про прийняття до друку публікації. Тематика публікації повинна відповідати змісту дисципліни Інформаційні мережі і тільки в цьому випадку додаткові бали будуть зараховані. При пред'явленні публікації студент звільняється від виконання практичної роботи, тема якої відповідає тематиці публікації, при цьому студенту зараховується додаткові бали замість балів за виконання суміжних за тематикою практичних робіт. Максимальна кількість додаткових балів, що можуть бути зараховані за дисципліну – 10 балів.

Форми контролю	Види навчальної роботи	Оцінювання
ПОТОЧНИЙ КІНТРОЛЬ	Виконання практичних, лабораторних робіт	45 балів
	Самостійна робота	15 балів
ПІДСУМКОВЕ ОЦІНЮВАННЯ <i>Залік/Екзамен</i>		40 балів

Додаткова оцінка

Участь у наукових конференціях, підготовка наукових публікацій за тематикою освітньої компоненти:

- Тези доповіді на фаховій конференції.3 бали

- Стаття у фаховому виданні. 5 балів
 - Стаття в іноземному рецензованому виданні. 10 балів
- Максимальна кількість додаткових балів, які можуть бути зараховані здобувачу освіти - 10 балів.

ПІДСУМКОВА ОЦІНКА ЗА ДИСЦИПЛІНУ

бали	Критерії оцінювання	Рівень компетентності	Оцінка /запис в екзаменаційній відомості
90-100	Студент демонструє повні й міцні знання навчального матеріалу в обсязі, що відповідає робочій програмі дисципліни, правильно й обґрунтовано приймає необхідні рішення в різних нестандартних ситуаціях. Вміє реалізувати теоретичні положення дисципліни в практичних розрахунках, аналізувати та співставляти дані об'єктів діяльності фахівця на основі набутих з даної та суміжних дисциплін знань та умінь. Знає сучасні технології та методи розрахунків з даної дисципліни. За час навчання при проведенні практичних занять, при виконанні індивідуальних/контрольних завдань проявив вміння самостійно вирішувати поставлені завдання, активно включатись в дискусії, може відстоювати власну позицію в питаннях та рішеннях, що розглядаються. Зменшення 100-бальної оцінки може бути пов'язане з недостатнім розкриттям питань, що стосується дисципліни, яка вивчається, але виходить за рамки об'єму матеріалу, передбаченого робочою програмою, або студент проявляє невпевненість в тлумаченні теоретичних положень чи складних практичних завдань.	Високий Повністю забезпечує вимоги до знань, умінь і навичок, що викладені в робочій програмі дисципліни. Власні пропозиції студента в оцінках і вирішенні практичних задач підвищує його вміння використовувати знання, які він отримав при вивченні інших дисциплін, а також знання, набуті при самостійному поглибленому вивченні питань, що відносяться до дисципліни, яка вивчається.	Відмінно / Зараховано (А)
82-89	Студент демонструє гарні знання, добре володіє матеріалом, що відповідає робочій програмі дисципліни, робить на їх основі аналіз можливих ситуацій та вміє застосовувати теоретичні положення при вирішенні практичних задач, але допускає окремі неточності. Вміє самостійно виправляти допущені помилки, кількість яких є незначною. Знає сучасні технології та методи розрахунків з даної дисципліни. За час навчання при проведенні практичних занять, при виконанні індивідуальних / контрольних завдань та поясненні прийнятих рішень, дає вичерпні пояснення.	Достатній Забезпечує студенту самостійне вирішення основних практичних задач в умовах, коли вихідні дані в них змінюються порівняно з прикладами, що розглянуті при вивченні дисципліни.	Добре / Зараховано (В)
75-81	Студент в загальному добре володіє матеріалом, знає основні положення матеріалу, що відповідає робочій програмі дисципліни, робить на їх основі аналіз можливих ситуацій та вміє застосовувати при вирішенні типових практичних завдань, але допускає окремі неточності. Вміє пояснити основні положення виконаних завдань та дати правильні відповіді при зміні результату при заданій зміні вихідних параметрів. Помилки у відповідях/ рішеннях/ розрахунках не є системними. Знає характеристики основних положень, що мають визначальне значення при проведенні практичних занять, при виконанні індивідуальних / контрольних	Достатній Конкретний рівень, за вивченим матеріалом робочої програми дисципліни. Додаткові питання про можливість використання теоретичних положень для практичного використання викликають утруднення.	Добре / Зараховано (С)

	завдань та поясненні прийнятих рішень, в межах дисципліни, що вивчається.		
67-74	Студент засвоїв основний теоретичний матеріал, передбачений робочою програмою дисципліни, та розуміє постанову стандартних практичних завдань, має пропозиції щодо напрямку їх вирішень. Розуміє основні положення, що є визначальними в курсі, може вирішувати подібні завдання тим, що розглядались з викладачем, але допускає значну кількість неточностей і грубих помилок, які може усувати за допомогою викладача.	Середній Забезпечує достатньо надійний рівень відтворення основних положень дисципліни.	Задовільно / Зараховано (D)
60-66	Студент має певні знання, передбачені в робочій програмі дисципліни, володіє основними положеннями, що вивчаються на рівні, який визначається як мінімально допустимий. З використанням основних теоретичних положень, студент з труднощами пояснює правила вирішення практичних/розрахункових завдань дисципліни. Виконання практичних / індивідуальних / контрольних завдань значно формалізовано: є відповідність алгоритму, але відсутнє глибоке розуміння роботи та взаємозв'язків з іншими дисциплінами.	Середній Є мінімально допустимим у всіх складових навчальної програми з дисципліни.	Задовільно / Зараховано (E)
35-59	Студент може відтворити окремі фрагменти з курсу. Незважаючи на те, що програму навчальної дисципліни студент виконав, працював він пасивно, його відповіді під час практичних робіт в більшості є невірними, необґрунтованими. Цілісність розуміння матеріалу з дисципліни у студента відсутні.	Низький Не забезпечує практичної реалізації задач, що формуються при вивченні дисципліни.	Незадовільно з можливістю повторного складання) / Не зараховано (FX) В залікову книжку не представляється
0-34	Студент повністю не виконав вимог робочої програми навчальної дисципліни. Його знання на підсумкових етапах навчання є фрагментарними. Студент не допущений до здачі екзамену/заліку.	Незадовільний Студент не підготовлений до самостійного вирішення задач, які окреслює мета та завдання дисципліни.	Незадовільно з обов'язковим повторним вивченням / Не допущений (F) В залікову книжку не представляється

ПОЛІТИКА ДОБРОЧЕСНОСТІ

Здобувач вищої освіти виконуючи самостійну або індивідуальну роботу повинен дотримуватись політики доброчесності. У разі наявності плагіату в будь-яких видах робіт Здобувача, він отримує незадовільну оцінку і повинен повторно виконати завдання, які передбачені у Силабусі.