

СИЛАБУС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ «ТЕХНОЛОГІЇ BIG DATA»

Лектор курсу			Щербина Ірина Сергіївна , кандидат технічних наук, доцент.		Контактна інформація лектора (e-mail), сторінка курсу в Moodle	e-mail: sherbina@duikt.ua сторінка курсу в Classroom - https://classroom.google.com/u/1/c/NzQ5NjE4NzUyNTM4 код доступу - dwn3lmc	
Галузь знань					Рівень вищої освіти		
Спеціальність					Семестр		
Освітня програма					Тип дисципліни	Вибіркова	
Обсяг:	Кредитів ECTS	Годин	За видами занять:				
			Лекцій	Семінарських занять	Практичних занять	Лабораторних занять	Самостійна підготовка
	3	90	18	-	18		54

АНОТАЦІЯ КУРСУ

Взаємозв'язок у структурно-логічній схемі

Освітні компоненти, які передують вивченню	
Освітні компоненти для яких є базовою	

Мета курсу:	Отримання теоретичних знань та практичних навичок про основи статистичної та інтелектуальної обробки даних, а також засвоєння практичних навичок роботи з програмними продуктами для обробки інформації та візуалізації даних.
--------------------	--

Компетентності відповідно до освітньої програми

Soft- skills / Загальні компетентності (ЗК)	Hard-skills / Спеціальні компетентності (СК)
Здатність до пошуку, оброблення та аналізу інформації з різних джерел. Здатність розв'язувати комплексні проблеми комп'ютерних наук на основі системного наукового світогляду та загального культурного кругозору із дотриманням принципів професійної етики та академічної доброчесності. Здатність застосовувати сучасні інформаційні та комунікаційні технології у науковій діяльності.	Здатність застосовувати сучасні методології, методи та інструменти експериментальних і теоретичних досліджень у сфері комп'ютерних наук, сучасні цифрові технології, бази даних та інші електронні ресурси у науковій та освітній діяльності. Здатність застосовувати сучасні концепції побудови високопродуктивних надійних комп'ютерних платформ для корпоративної ІТ-інфраструктури в умовах невизначеності, принципи розробки та функціонування паралельних, розподілених і гетерогенних обчислювальних систем, методи та засоби організації високопродуктивних обчислень у науковій діяльності та/або професійній практиці.

Програмні результати навчання (ПР)

Формулювати і перевіряти гіпотези; використовувати для обґрунтування висновків належні докази, зокрема, результати теоретичного аналізу, експериментальних досліджень і математичного та/або комп'ютерного моделювання, наявні літературні дані.

Застосовувати сучасні інструменти і технології пошуку, оброблення та аналізу інформації, зокрема, статистичні методи аналізу даних великого обсягу та/або складної структури, спеціалізовані бази даних та інформаційні системи.

Застосовувати сучасні концепції побудови інтелектуальних інформаційних технологій, організації високопродуктивних обчислень у науковій діяльності та/або професійній практиці для створення інноваційних продуктів у комп'ютерних науках та дотичних міждисциплінарних напрямках.

ОРГАНІЗАЦІЯ НАВЧАННЯ			
Тема, опис теми	Вид заняття	Оцінювання за тему	Форми і методи навчання/питання до самостійної роботи
Розділ 1. Сутність великих даних			
Тема 1. Сутність великих даних Знати: 1. Сутність поняття «великі дані» 2. Історію розвитку та перспективи використання 3. Етичні та правові аспекти використання великих даних Вміти: 1. Використовувати великі дані в розробці, проектуванні та науковій діяльності Рекомендовані джерела: 1–4			
Заняття 1.1. Історія та перспективи розвитку великих даних	Лекція 1 2 год		Пояснювально-ілюстративний, лекція-візуалізація, бліц опитування
Заняття 1.2. Практичні кейси використання великих даних у фінансах та медицині	Практичне заняття 1 2 год	4 бали	Усне опитування, навчальна дискусія, ситуаційне завдання, круглий стіл, мозковий штурм, тестування
Заняття 1.3. Використання великих даних – виклики та можливості	Лекція 2 2 год		Пояснювально-ілюстративний, лекція-візуалізація, бліц опитування
Заняття 1.4. Практичні кейси використання великих даних у цифровому маркетингу	Практичне заняття 2 2 год	4 бали	Усне опитування, навчальна дискусія, ситуаційне завдання, круглий стіл, мозковий штурм, тестування
Заняття 1.5. Етичні та правові аспекти використання великих даних	Лекція 3 2 год		Пояснювально-ілюстративний, лекція-візуалізація, бліц опитування
Заняття 1.6. Практики збору та обробки великих даних – потокові та неструктуровані дані	Практичне заняття 3 2 год	4 бали	Усне опитування, навчальна дискусія, ситуаційне завдання, круглий стіл, мозковий штурм, тестування
Тема 1. Сутність великих даних	Самостійна робота		
	6 год	2 бали	Збір та аналіз сенсорних даних з IoT
	6 год	2 бали	Використання великих даних в комп'ютерних науках
	6 год	2 бали	Регулювання та законодавство в сфері великих даних
Розділ 2. Архітектура великих даних			
Тема 2. Архітектура великих даних Знати: 1. Системи зберігання даних 2. Системи обробки даних			

3. Системи аналізу даних

Вміти: 1. Користуватися технологіями HADOOP, NoSQL, Apache, Spark

Рекомендовані джерела: 1–4

Заняття 2.1. Системи зберігання даних	Лекція 4 2 год		Пояснювально-ілюстративний, лекція-візуалізація, бліц опитування
Заняття 2.2. Розподілена файлова система HADOOP (HDFS) та нереляційні бази даних NoSQL	Практичне заняття 4 2 год	5 балів	Навчальна дискусія, ситуаційне завдання, круглий стіл, мозковий штурм, тестування
Заняття 2.3. Системи обробки даних (MapReduce, Spark)	Лекція 5 2 год		Пояснювально-ілюстративний, лекція-візуалізація, бліц опитування
Заняття 2.4. Створення потокового потоку даних з використанням Apache Kafka.	Практичне заняття 5 2 год	5 балів	Усне опитування, навчальна дискусія, ситуаційне завдання, круглий стіл, мозковий штурм, тестування
Заняття 2.5. Системи аналізу даних	Лекція 6 2 год		Пояснювально-ілюстративний, лекція-візуалізація, бліц опитування
Заняття 2.6. Реалізація моделі машинного навчання на платформі Spark MLlib.	Практичне заняття 6 2 год	5 балів	Усне опитування, навчальна дискусія, ситуаційне завдання, круглий стіл, мозковий штурм, тестування
Тема 2. Архітектура великих даних	Самостійна робота		
	6 год	2 бали	1. Безпекові заходи зберігання великих даних
	6 год	2 бали	2. Методи резервного копіювання та відновлення при зберіганні великих даних
	6 год	2 бали	3. Розподілені системи зберігання великих даних
Розділ 3. Методи аналізу та візуалізації великих даних			
Тема 3. Методи аналізу та візуалізації великих даних			
Знати: 1. Алгоритми машинного та глибинного навчання			
2. Методи візуалізації			
3. Сутність графової аналітики			
Вміти: 1. Користуватися програмними продуктами аналізу великих даних			
Рекомендовані джерела: 1–4			
Заняття 3.1. Машинне та глибинне навчання	Лекція 7 2 год		Пояснювально-ілюстративний, лекція-візуалізація, бліц опитування
Заняття 3.2. Використання Apache Spark для аналізу та машинного навчання	Практичне заняття 7 2 год	5 балів	Усне опитування, навчальна дискусія, ситуаційне завдання, круглий стіл, мозковий штурм, тестування
Заняття 3.3. Візуалізація великих даних	Лекція 8 2 год		Пояснювально-ілюстративний, лекція-візуалізація, бліц опитування

Заняття 3.4. Технології візуалізації великих даних за допомогою Tableau	Практичне заняття 8 2 год	5 балів	Усне опитування, навчальна дискусія, ситуаційне завдання, круглий стіл, мозковий штурм, тестування
Заняття 3.5. Графова аналітика в великих даних	Лекція 9 2 год		Пояснювально-ілюстративний, лекція-візуалізація, бліц опитування
Заняття 3.6. Використання графових баз даних для аналізу зв'язків у складних даних, таких як соціальні мережі або транспортні системи.	Практичне заняття 9 2 год	5 балів	Усне опитування, навчальна дискусія, ситуаційне завдання, круглий стіл, мозковий штурм, тестування
Тема 3. Методи аналізу та візуалізації великих даних	Самостійна робота		
	6 год	2 бали	1. Робота з Hadoop та MapReduce
	6 год	2 бали	2. Базы даних онлайнової аналітичної обробки (OLAP)
	6 год	2 бали	3. Робота з розподіленими базами даних NoSQL, такими як MongoDB або Cassandra
МАТЕРІАЛЬНО-ТЕХНІЧНЕ ЗАБЕЗПЕЧЕННЯ ДИСЦИПЛІНИ			
- Мультимедійний проектор; - Комп'ютерний клас для проведення практичних занять з доступом до інтернет.			
ІНФОРМАЦІЙНЕ ЗАБЕЗПЕЧЕННЯ ДИСЦИПЛІНИ			
1. Foster Provost, Tom Fawcett. Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking. – O’Railly, 413p. 2. Joi Raise, Matt Housley: Fundamentals of Data Engineering. – O’Railly, 2022, 446с.			
ПОЛІТИКА КУРСУ («ПРАВИЛА ГРИ»)			
- Курс передбачає роботу в колективі. - Середовище в аудиторії є дружнім, творчим, відкритим до конструктивної критики. - Освоєння дисципліни передбачає обов’язкове відвідування лекцій і практичних занять, а також самостійну роботу. - 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 відсутній з поважної причини, він презентує виконані завдання під час самостійної підготовки та консультації викладача. - 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. - Студент, який спізнився, вважається таким, що пропустив заняття з неповажної причини з виставленням 0 балів за заняття, і при цьому має право бути присутнім на занятті. - За використання телефонів і комп’ютерних засобів без дозволу викладача, порушення дисципліни студент видаляється з заняття, за заняття отримує 0 балів.			
КРИТЕРІЙ ТА МЕТОДИ ОЦІНЮВАННЯ			
Умовою допуску до підсумкового контролю є виконання всіх практичних робіт і виконання самостійних завдань, які передбачені структурою освітньої компоненти Технології BIG DATA. Якщо аспіранта не допущено до складання заліку, як такого, що не виконав індивідуальний план, йому надається час до перескладання для виконання всіх вимог допуску. Аспірант має право на два перескладання. При повторному перескладанні заліку його у аспіранта може приймати комісія, яка створюється директором ННІТ. Оцінка комісії є остаточною. У випадку отримання аспірантом 0 балів (неприйнятно), що тягне відрахування за невиконання навчального плану. Оцінювання аспірантів здійснюється за накопичувальною 100-бальною системою і складається із двох основних оцінкових блоків і розподіляється в певних пропорціях 60 (бали напрацьовані під час вивчення дисципліни – Поточний контроль), 40 (підсумкове оцінювання - Залік):			
Форми	Види навчальної роботи		Оцінювання

контролю		
ПОТОЧНИЙ КОНТРОЛЬ	• Виконання практичних робіт	42 бали
	• Самостійна робота	18 балів
ПІДСУМКОВЕ ОЦІНЮВАННЯ залік	Залік проходить у письмовій формі.	40 балів
Додаткова оцінка		
Види навчальної роботи		Оцінювання
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
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.
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
75-81	Студент в загальному добре володіє матеріалом, знає основні положення матеріалу, що	Достатній
		Добре /

	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 завдань та поясненні прийнятих рішень, в межах дисципліни, що вивчається.	Конкретний рівень, за вивченим матеріалом робочої програми дисципліни. Додаткові питання про можливість використання теоретичних положень для практичного використання викликають утруднення.	Зараховано (C)
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 розглядалися 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 відсутня.	Низький Не забезпечує практичної реалізації задач, що формуються при вивченні дисципліни	Незадовільно з можливістю повторного складання) / Не зараховано (FX) <i>В залікову книжку не проставляється</i>
1-34	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) <i>В залікову книжку не проставляється</i>

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.