

Міністерство освіти і науки України
Національний технічний університет
«Дніпровська політехніка»

Кафедра системного аналізу та управління



«ЗАТВЕРДЖЕНО»

завідувач кафедри

Желдак Т.А.

« 31 » серпня 2023 року

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

«Аналіз даних та знань»

Галузь знань	12 Інформаційні технології
Освітній рівень.....	перший (бакалаврський)
Освітня програма	Освітньо-професійна
Статус	Вибіркова
Загальний обсяг	4 кредити ECTS (120 годин)
Форма підсумкового контролю	Диференційований залік
Термін викладання	5-й семестр
Мова викладання	українська

Викладачі: PhD, асистент Хабарлак Костянтин Сергійович

Пролонговано: на 20__/20__ н.р. _____ (_____) «__»__ 20__ р.
(підпис, ПІБ, дата)

на 20__/20__ н.р. _____ (_____) «__»__ 20__ р.
(підпис, ПІБ, дата)

Дніпро
НТУ «ДП»
2023

Робоча програма навчальної дисципліни «Аналіз даних та знань» для бакалаврів галузі 12 Інформаційні технології / Нац. техн. ун-т. «Дніпровська політехніка», каф. Системного аналізу та управління. Д. : НТУ «ДП», 2023. 12 с.

Розробник: Хабарлак К.С., PhD, асистент каф. системного аналізу та управління

Робоча програма регламентує:

- мету дисципліни;
- дисциплінарні результати навчання, сформовані на основі трансформації очікуваних результатів навчання освітньої програми;
- базові дисципліни;
- обсяг і розподіл за формами організації освітнього процесу та видами навчальних занять;
- програму дисципліни (тематичний план за видами навчальних занять);
- алгоритм оцінювання рівня досягнення дисциплінарних результатів навчання (шкали, засоби, процедури та критерії оцінювання);
- інструменти, обладнання та програмне забезпечення;
- рекомендовані джерела інформації.

Робоча програма призначена для реалізації компетентнісного підходу під час планування освітнього процесу, викладання дисципліни, підготовки студентів до контрольних заходів, контролю провадження освітньої діяльності, внутрішнього та зовнішнього контролю забезпечення якості вищої освіти, акредитації освітніх програм у межах спеціальності.

Робоча програма буде в пригоді для формування змісту підвищення кваліфікації науково-педагогічних працівників кафедр університету.

Погоджено рішенням методичної комісії спеціальності 124 «Системний аналіз» (протокол № 3 від 31.08.2023).

ЗМІСТ

1 МЕТА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ	4
2 ОЧІКУВАНІ ДИСЦИПЛІНАРНІ РЕЗУЛЬТАТИ НАВЧАННЯ.....	4
3 БАЗОВІ ДИСЦИПЛІНИ.....	4
4 ОБСЯГ І РОЗПОДІЛ ЗА ФОРМАМИ ОРГАНІЗАЦІЇ ОСВІТНЬОГО ПРОЦЕСУ ТА ВИДАМИ НАВЧАЛЬНИХ ЗАНЯТЬ	4
5 ПРОГРАМА ДИСЦИПЛІНИ ЗА ВИДАМИ НАВЧАЛЬНИХ ЗАНЯТЬ ..	5
6 ОЦІНЮВАННЯ РЕЗУЛЬТАТІВ НАВЧАННЯ.....	6
6.1 Шкали	6
6.2 Засоби та процедури	7
6.3 Критерії	8
7 ІНСТРУМЕНТИ, ОБЛАДНАННЯ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ	7
8 РЕКОМЕНДОВАНІ ДЖЕРЕЛА ІНФОРМАЦІЇ.....	11

1 МЕТА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Мета дисципліни – сформувати у бакалаврів навички обробки та аналізу даних за допомогою мови програмування Python та сучасних бібліотек машинного навчання та обробки даних: scikit-learn, SciPy, Pandas, NumPy, Matplotlib, що буде корисним для подальшого працевлаштування здобувача.

2 ОЧІКУВАНІ ДИСЦИПЛІНАРНІ РЕЗУЛЬТАТИ НАВЧАННЯ

Дисциплінарні результати навчання (ДРН)	
Шифр ДРН	зміст
ДРН-1	Знати як працювати та візуалізувати дані з електронних таблиць за допомогою мови програмування Python та бібліотеки Pandas.
ДРН-2	Отримати навички з роботою бібліотеки scikit-learn (sklearn) для розв'язання задач методами машинного навчання.
ДРН-3	Володіти навичками проведення А/В тестування для визначення подальшого напрямку розвитку програмного продукту компанії із перевіркою статистичної значущості результатів.
ДРН-4	Вміти будувати регресію та класифікувати дані. Оцінювати результати на тестовій вибірці або за допомогою крос-валідації.
ДРН-5	Отримати базові навички обробки зображень і текстів за допомогою вивчених методів.

3 БАЗОВІ ДИСЦИПЛІНИ

Назва дисципліни	Здобуті результати навчання
Б1 Алгебра та геометрія	Знати і вміти застосовувати на практиці розв'язувати системи лінійних рівнянь, диференціальне та інтегральне числення, ряди та інтеграл Фур'є, аналітичну геометрію, лінійну алгебру та векторний аналіз, функціональний аналіз та дискретну математику в обсязі необхідному для вирішення типових завдань системного аналізу
Б2 Алгоритми та структури даних	Розуміти основи побудови програм, поняття класу, функції, умовного оператора циклу.

4 ОБСЯГ І РОЗПОДІЛ ЗА ФОРМАМИ ОРГАНІЗАЦІЇ ОСВІТНЬОГО ПРОЦЕСУ ТА ВИДАМИ НАВЧАЛЬНИХ ЗАНЯТЬ

Вид навчальних занять	Розподіл за формами навчання, години							
	денна			вечірня		заочна		
	обсяг	аудиторні заняття	самостійна робота	аудиторні заняття	самостійна робота	обсяг	аудиторні заняття	самостійна робота
лекційні	64	26	38	-	-	64	6	58
практичні	50	13	37	-	-	50	4	46
лабораторні	-	-	-	-	-	-	-	-
семінари	-	-	-	-	-	-	-	-
контрольні заходи	6	6	-	-	-	-	-	-
РАЗОМ	120	45	75	-	-	120	10	104

5 ПРОГРАМА ДИСЦИПЛІНИ ЗА ВИДАМИ НАВЧАЛЬНИХ ЗАНЯТЬ

Шифри ДРН	Види та тематика навчальних занять	Обсяг складових, години
	ЛЕКЦІЇ	78
ДРН-1	1. Первинний аналіз даних Знайомство із мовою програмування Python для аналізу даних. Використання бібліотеки Pandas для табличних даних та розрахунку статистичних оцінок. Типи даних. Бібліотека SciPy.	6
ДРН-1	2. Візуалізація даних Основні види графіків, діаграм. Їх побудова, доречність використання для наявних даних, особливості.	6
ДРН-1 ДРН-2	3. Бутстрап Поняття популяції та вибірки даних, зміщеної вибірки, повторного та статифікованого відбору. Алгоритм бутстрапу, розрахунок бутстрапівських оцінок.	8
ДРН-1 ДРН-2 ДРН-4	4. Побудова та підгонка лінійної регресії Знайомство із бібліотекою sklearn. Інтерфейс бібліотеки. Побудова лінійної регресії за даними. Створення та використання нелінійних ознак.	8
ДРН-1 ДРН-2 ДРН-4	5. Передбачення за допомогою регресії. Перехресний контроль Проблеми передбачення. Оцінка довірчого інтервалу за допомогою бутстрапу. Проведення експериментів. Поняття тренувальної та тестової вибірок, перехресного контролю (крос-валідації)	8
ДРН-1 ДРН-2 ДРН-3	6. А/В тестування Знайомство із поняттям А/В тестування. Коли доречно та як необхідно проводити А/В тестування. Вклад випадковості в результати тестування. Перестановочних тест. Перевірка статистичної значущості	8
ДРН-1 ДРН-2 ДРН-3	7. Дисперсійний аналіз та багатурукий бандит Проблема множинного тестування в А/В тесті. Дисперсійний аналіз. Вибір найкращого варіанта за допомогою алгоритму багатурукого бандита.	8
ДРН-1 ДРН-2 ДРН-4	8. Лінійні, метричні та ймовірнісні методи класифікації Поняття задачі класифікації. Основні ідеї та методи: логістична регресія, K найближчих сусідів, наївний Баєс	10
ДРН-1 ДРН-2	9. Методи зменшення розмірності даних Проблема обробки даних із великою розмірністю та методи її зменшення. Метод головних компонент.	8

Шифри ДРН	Види та тематика навчальних занять	Обсяг складових, години
ДРН-2 ДРН-5	10. Основи аналізу текстів та зображень випадковим лісом та градієнтним бустінгом Як працювати із зображеннями та текстом за допомогою розглянутих методів. Бустінг як основа методів класифікації випадкового лісу та градієнтного бустінгу	8
	ПРАКТИЧНІ ЗАНЯТТЯ	36
ДРН-1	Лабораторна робота №1: Первинний аналіз даних Мета: закріпити теоретичні знання і розвинути практичні навички роботи з таблицями в Pandas, розрахунку статистик та візуалізації даних.	6
ДРН-1 ДРН-2 ДРН-4	Лабораторна робота №2: Регресійний аналіз Мета: закріплення навичок збору даних, роботи із бібліотекою sklearn для побудови лінійної регресії, створення нелінійних ознак. Застосування на практиці підходу пересхесної-валідації.	6
ДРН-1 ДРН-2 ДРН-4	Лабораторна робота №3: Класифікація за допомогою логістичної регресії та наївного Баєса Мета: закріплення навичок роботи із категорійними даними, їх попередньої обробки та класифікація даних.	8
ДРН-1 ДРН-2 ДРН-3	Лабораторна робота №4: А/В тест та багаторукі бандити Мета: проведення власного експерименту щодо оцінки різного дизайну веб-сторінок, закріплення навичок проведення А/В тесту, алгоритму ϵ -жадібного бандиту для довільної кількості варіантів дизайну веб-сторінки та підтвердження статистичної значущості результатів експерименту.	8
ДРН-2 ДРН-5	Лабораторна робота №5: Основи аналізу текстів та зображень Мета: закріпити навички класифікації текстів та зображень за допомогою методів випадкового лісу та градієнтного бустінгу.	8
КОНТРОЛЬНІ ЗАХОДИ		6
РАЗОМ		120

6 ОЦІНЮВАННЯ РЕЗУЛЬТАТІВ НАВЧАННЯ

Сертифікація досягнень студентів здійснюється за допомогою прозорих процедур, що ґрунтуються на об'єктивних критеріях відповідно до Положення університету «Про оцінювання результатів навчання здобувачів вищої освіти».

Досягнутий рівень компетентностей відносно очікуваних, що ідентифікований під час контрольних заходів, відображає реальний результат навчання студента за дисципліною.

6.1 Шкали

Оцінювання навчальних досягнень студентів НТУ «ДП» здійснюється за рейтинговою (100-бальною) та конвертаційною шкалами. Остання необхідна (за офіційною відсутністю національної шкали) для конвертації (переведення) оцінок здобувачів вищої освіти різних закладів.

Шкали оцінювання навчальних досягнень студентів НТУ «ДП»

Рейтингова	Конвертаційна
90...100	відмінно / Excellent
74...89	добре / Good
60...73	задовільно / Satisfactory
0...59	незадовільно / Fail

Кредити навчальної дисципліни зараховується, якщо студент отримав підсумкову оцінку не менше 60-ти балів. Нижча оцінка вважається академічною заборгованістю, що підлягає ліквідації відповідно до Положення про організацію освітнього процесу НТУ «ДП».

6.2 Засоби та процедури

Зміст засобів діагностики спрямовано на контроль рівня сформованості знань, умінь, комунікації, автономності та відповідальності студента за вимогами НРК до 6-го кваліфікаційного рівня під час демонстрації регламентованих робочою програмою результатів навчання.

Студент на контрольних заходах має виконувати завдання, орієнтовані виключно на демонстрацію дисциплінарних результатів навчання (розділ 2).

Засоби діагностики, що надаються студентам на контрольних заходах у вигляді завдань для поточного та підсумкового контролю, формуються шляхом конкретизації вихідних даних та способу демонстрації дисциплінарних результатів навчання.

Засоби діагностики (контрольні завдання) для поточного та підсумкового контролю дисципліни затверджуються кафедрою.

Види засобів діагностики та процедур оцінювання для поточного та підсумкового контролю дисципліни подано нижче.

Засоби діагностики та процедури оцінювання

ПОТОЧНИЙ КОНТРОЛЬ			ПІДСУМКОВИЙ КОНТРОЛЬ	
навчальне заняття	засоби діагностики	процедури	засоби діагностики	процедури
лекції	контрольні завдання за кожною темою	виконання завдання під час лекцій	комплексна контрольна робота (ККР)	визначення середньозваженого результату поточних контролів;
практичні	контрольні завдання за кожною темою	виконання завдань під час практичних занять		виконання ККР під час екзамену за бажанням студента
	або індивідуальне завдання	виконання завдань під час самостійної роботи		

Під час поточного контролю лекційні заняття оцінюються шляхом визначення якості виконання контрольних конкретизованих завдань. Практичні

заняття оцінюються якістю виконання контрольного або індивідуального завдання.

Якщо зміст певного виду занять підпорядковано декільком дескрипторам, то інтегральне значення оцінки може визначатися з урахуванням вагових коефіцієнтів, що встановлюються викладачем.

За наявності рівня результатів поточних контролів з усіх видів навчальних занять не менше 60 балів, підсумковий контроль здійснюється без участі студента шляхом визначення середньозваженого значення поточних оцінок.

Незалежно від результатів поточного контролю кожен студент під час екзамену має право виконувати ККР, яка містить завдання, що охоплюють ключові дисциплінарні результати навчання.

Кількість конкретизованих завдань ККР повинна відповідати відведеному часу на виконання. Кількість варіантів ККР має забезпечити індивідуалізацію завдання.

Значення оцінки за виконання ККР визначається середньою оцінкою складових (конкретизованих завдань) і є остаточним.

Інтегральне значення оцінки виконання ККР може визначатися з урахуванням вагових коефіцієнтів, що встановлюється кафедрою для кожного дескриптора НРК.

6.3 Критерії

Реальні результати навчання студента ідентифікуються та вимірюються відносно очікуваних під час контрольних заходів за допомогою критеріїв, що описують дії студента для демонстрації досягнення результатів навчання.

Для оцінювання виконання контрольних завдань під час поточного контролю лекційних і практичних занять в якості критерія використовується коефіцієнт засвоєння, що автоматично адаптує показник оцінки до рейтингової шкали:

$$O_i = 100 a/m,$$

де a – число правильних відповідей або виконаних суттєвих операцій відповідно до еталону рішення; m – загальна кількість запитань або суттєвих операцій еталону.

Індивідуальні завдання та комплексні контрольні роботи оцінюються експертно за допомогою критеріїв, що характеризують співвідношення вимог до рівня компетентностей і показників оцінки за рейтинговою шкалою.

Зміст критеріїв спирається на компетентнісні характеристики, визначені НРК для бакалаврського рівня вищої освіти (подано нижче).

**Загальні критерії досягнення результатів навчання
для 7-го кваліфікаційного рівня за НРК**

Інтегральна компетентність – здатність розв’язувати складні спеціалізовані задачі та практичні проблеми у професійної діяльності або в процесі навчання, що передбачають застосування теоретичних положень та методів системного аналізу і характеризується комплексністю та невизначеністю умов.

Дескриптори НРК	Вимоги до знань, умінь, комунікації, автономності та відповідальності	Показник оцінки
Знання		
<ul style="list-style-type: none"> ◆ спеціалізовані концептуальні знання, набуті у процесі навчання та/або професійної діяльності, ◆ критичне осмислення проблем у навчанні та /або професійній діяльності та на межі предметних галузей 	Відповідь відмінна – правильна, обґрунтована, осмислена. Характеризує наявність: <ul style="list-style-type: none"> - спеціалізованих концептуальних знань на рівні новітніх досягнень; - критичне осмислення проблем у навчанні та/або професійній діяльності та на межі предметних галузей 	95-100
	Відповідь містить негрубі помилки або описки	90-94
	Відповідь правильна, але має певні неточності	85-89
	Відповідь правильна, але має певні неточності й недостатньо обґрунтована	80-84
	Відповідь правильна, але має певні неточності, недостатньо обґрунтована та осмислена	74-79
	Відповідь фрагментарна	70-73
	Відповідь демонструє нечіткі уявлення студента про об’єкт вивчення	65-69
	Рівень знань мінімально задовільний	60-64
	Рівень знань незадовільний	<60
Уміння		
<ul style="list-style-type: none"> • розв’язання складних непередбачуваних задач і проблем у спеціалізованих сферах професійної діяльності та/або навчання, що передбачає збирання та інтерпретацію інформації (даних), вибір методів та інструментальних засобів, застосування інноваційних підходів 	Відповідь характеризує уміння: <ul style="list-style-type: none"> - виявляти проблеми; - формулювати гіпотези; - розв’язувати проблеми; - оновлювати знання; - інтегрувати знання; - провадити інноваційну діяльність; - провадити наукову діяльність 	95-100
	Відповідь характеризує уміння застосовувати знання в практичній діяльності з негрубими помилками	90-94
	Відповідь характеризує уміння застосовувати знання в практичній діяльності, але має певні неточності при реалізації однієї вимоги	85-89
	Відповідь характеризує уміння застосовувати знання в практичній діяльності, але має певні неточності при реалізації двох вимог	80-84
	Відповідь характеризує уміння застосовувати знання в практичній діяльності, але має певні неточності при реалізації трьох вимог	74-79
	Відповідь характеризує уміння застосовувати знання в практичній діяльності, але має певні неточності при реалізації чотирьох вимог	70-73

Дескриптори НРК	Вимоги до знань, умінь, комунікації, автономності та відповідальності	Показник оцінки
	Відповідь характеризує уміння застосовувати знання в практичній діяльності при виконанні завдань за зразком	65-69
	Відповідь характеризує уміння застосовувати знання при виконанні завдань за зразком, але з неточностями	60-64
	Рівень умінь незадовільний	<60
Комунікація		
<ul style="list-style-type: none"> ◆ донесення до фахівців і нефахівців інформації, ідей, проблем, рішень та власного досвіду в галузі професійної діяльності ◆ здатність ефективно формувати комунікаційну стратегію 	<p>Зрозумілість відповіді (доповіді). Мова:</p> <ul style="list-style-type: none"> - правильна; - чиста; - ясна; - точна; - логічна; - виразна; - лаконічна. <p>Комунікаційна стратегія:</p> <ul style="list-style-type: none"> - послідовний і несуперечливий розвиток думки; - наявність логічних власних суджень; - доречна аргументації та її відповідність відстоюваним положенням; - правильна структура відповіді (доповіді); - правильність відповідей на запитання; - доречна техніка відповідей на запитання; - здатність робити висновки та формулювати пропозиції; - використання іноземних мов у професійній діяльності 	95-100
	Достатня зрозумілість відповіді (доповіді) та доречна комунікаційна стратегія з незначними хибами	90-94
	Добра зрозумілість відповіді (доповіді) та доречна комунікаційна стратегія (сумарно не реалізовано три вимоги)	85-89
	Добра зрозумілість відповіді (доповіді) та доречна комунікаційна стратегія (сумарно не реалізовано чотири вимоги)	80-84
	Добра зрозумілість відповіді (доповіді) та доречна комунікаційна стратегія (сумарно не реалізовано п'ять вимог)	74-79
	Задовільна зрозумілість відповіді (доповіді) та доречна комунікаційна стратегія (сумарно не реалізовано сім вимог)	70-73
	Задовільна зрозумілість відповіді (доповіді) та комунікаційна стратегія з хибами (сумарно не реалізовано дев'ять вимог)	65-69
	Задовільна зрозумілість відповіді (доповіді) та комунікаційна стратегія з хибами (сумарно не реалізовано 10 вимог)	60-64
	Рівень комунікації незадовільний	<60
Автономність та відповідальність		
<ul style="list-style-type: none"> ◆ управління комплексними діями або проектами, відповідальність за прийняття рішень у непередбачуваних умовах ◆ відповідальність за професійний розвиток окремих осіб та/або груп осіб 	<p>Відмінне володіння компетенціями:</p> <ul style="list-style-type: none"> - використання принципів та методів організації діяльності команди; - ефективний розподіл повноважень в структурі команди; - підтримка врівноважених стосунків з членами команди (відповідальність за взаємовідносини); - стресовитривалість; - саморегуляція; - трудова активність в екстремальних ситуаціях; - високий рівень особистого ставлення до справи; - володіння всіма видами навчальної діяльності; 	95-100

Дескриптори НРК	Вимоги до знань, умінь, комунікації, автономності та відповідальності	Показник оцінки
♦ здатність до подальшого навчання з високим рівнем автономності	- належний рівень фундаментальних знань; - належний рівень сформованості загальнонавчальних умінь і навичок	
	Упевнене володіння компетенціями автономності та відповідальності з незначними хибами	90-94
	Добре володіння компетенціями автономності та відповідальності (не реалізовано дві вимоги)	85-89
	Добре володіння компетенціями автономності та відповідальності (не реалізовано три вимоги)	80-84
	Добре володіння компетенціями автономності та відповідальності (не реалізовано чотири вимоги)	74-79
	Задовільне володіння компетенціями автономності та відповідальності (не реалізовано п'ять вимог)	70-73
	Задовільне володіння компетенціями автономності та відповідальності (не реалізовано шість вимог)	65-69
	Задовільне володіння компетенціями автономності та відповідальності (рівень фрагментарний)	60-64
	Рівень автономності та відповідальності незадовільний	<60

7 ІНСТРУМЕНТИ, ОБЛАДНАННЯ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ

Технічні засоби навчання: мультимедійні та комп'ютерні пристрої.

Засоби дистанційної освіти: Moodle, MS Teams.

Пакети прикладних програм: Python, бібліотеки: scikit-learn, SciPy Pandas, NumPy, Matplotlib (безкоштовні).

8 РЕКОМЕНДОВАНІ ДЖЕРЕЛА ІНФОРМАЦІЇ

1. Математичні методи інтелектуального аналізу даних: [навчальний посібник для здобувачів першого рівня вищої освіти спеціальності 124 Системний аналіз] / Т. Шабельник, О. Дяченко. – Маріуполь: МДУ, 2021. – 163 с
2. Кононова К. Ю. Машинне навчання: методи та моделі / К. Ю. Кононова. – Харків: ХНУ імені В. Н. Каразіна, 2020. – 301 с.
3. Документація бібліотеки машинного навчання scikit-learn. URL: <https://scikit-learn.org> (дата звернення: 02.11.2023).
4. Документація бібліотеки аналізу даних в Python: pandas. URL: <https://pandas.pydata.org/> (дата звернення: 02.11.2023).
5. Practical Statistics for Data Scientists / P. Bruce, A. Bruce, P. Gedeck. – O'Reilly Media, 2020.

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

«Аналіз даних та знань»

для бакалаврів галузі
12 Інформаційні технології

Розробник:
Костянтин ХАБАРЛАК

В редакції автора

Підготовлено до виходу в світ
у Національному технічному університеті
«Дніпровська політехніка».

Свідоцтво про внесення до Державного реєстру ДК № 1842
4960050, м. Дніпро, просп. Д. Яворницького, 19