



АНАЛІЗ ДАНИХ В ІНФОРМАЦІЙНИХ СИСТЕМАХ

Робоча програма навчальної дисципліни (Силабус)

- Реквізити навчальної дисципліни

Рівень вищої освіти	Перший (бакалаврський)
Галузь знань	Інформаційні технології
Спеціальність	121 Інженерія програмного забезпечення
Освітня програма	Інженерія програмного забезпечення інформаційних систем
Статус дисципліни	нормативна
Форма навчання	очна (денна)
Рік підготовки, семестр	2 курс, весняний семестр
Обсяг дисципліни	4 кредити / 120 годин (36 лекції, 18 лабораторні, 18 практичні, 48 СРС)
Семестровий контроль/ контрольні заходи	залік / захист лабораторних робіт, МКР
Розклад занять	Лекції: 1,5 пари на тиждень Практичні заняття: одна пара на два тижні Лабораторні заняття: одна пара на два тижні Розклад: https://schedule.kpi.ua
Мова викладання	Українська
Інформація про керівника курсу / викладачів	Лектори: доц. Ліхоузова Т.А., доц. Олійник Ю.О. Практичні заняття: доц. Ліхоузова Т.А., likhouzova.tetiana@edu.kpi.ua Лабораторні заняття: доц. Олійник Ю.О., yurii.oliinyk-fiote@edu.kpi.ua
Розміщення курсу	https://classroom.google.com/c/Nzc2MDY5Mzc4ODMz?cjc=phvmqfm7_sikorsky-distance.org доступ до курсу за запрошенням викладача

- Програма навчальної дисципліни

1. Опис навчальної дисципліни, її мета, предмет вивчення та результати навчання

Мета вивчення дисципліни – набуття ключових фахових компетентностей, теоретичних знань і практичних навичок аналізу даних для подальшого застосування у різних сферах професійної діяльності.

Предметом вивчення дисципліни є технології, методи та засоби інтелектуального аналізу даних.

Завдання вивчення дисципліни: оволодіння основними моделями та алгоритмами обробки та аналізу даних.

Навчальна дисципліна покликана допомогти студенту отримати:

знання:

- основних типових задач, що вирішуються при аналізі даних;
- методів підготовки даних до аналізу;
- різновидів моделей для опису випадкових даних;
- принципи вибору моделей та їх верифікації.

вміння:

- підготувати дані для аналізу, оцінити їх повноту та якість;
- вибрати моделі, що якнайкраще підійдуть для дослідження;
- побудувати моделі та оцінити їх якість;
- проаналізувати отримані результати, зробити висновки щодо можливості використання моделей.

досвід:

- вибору та використання методів обробки та аналізу даних або результатів досліджень.

Компетентності

Інтегральна компетентність Здатність розв'язувати складні спеціалізовані задачі та практичні проблеми у галузі інженерії програмного забезпечення, що характеризується комплексністю та невизначеністю умов із застосування теорій та методів інформаційних технологій.

Спеціальні (фахові, предметні) компетентності

ФК 11. Здатність реалізовувати фази та ітерації життєвого циклу програмних систем та інформаційних технологій на основі відповідних моделей і підходів розробки програмного забезпечення.

ФК 16. Здатність розробляти інтелектуальні інформаційні системи, технології генерації та аналізу знань, алгоритми штучного інтелекту для вирішення прикладних задач і підтримки прийняття рішень в різних прикладних областях життєдіяльності людини.

Програмні результати навчання

ПРН 29. Вміти використовувати методи та засоби аналізу даних при розв'язанні прикладних проблем у спеціалізованих сферах професійної діяльності, в тому числі і з використанням відповідного програмного забезпечення.

2. Пререквізити та постреквізити дисципліни (місце в структурно-логічній схемі навчання за відповідною освітньою програмою)

Пререквізити – Математичний аналіз, Теорія ймовірностей, Комп'ютерна дискретна математика, Алгоритми та структури даних, Основи програмування, Проектування алгоритмів, Базы даних.

Постреквізити – Програмування інтелектуальних ІС.



3. Зміст навчальної дисципліни

<i>Розділ 1. Сховища даних</i>	1.1. Задачі аналізу. Емпірична вибірка та її абстрактні моделі. Способи збору даних 1.2. Що таке Data mining. Методи і стадії Data mining. Завдання Data mining. 1.3. ETL- процес. SQL Window functions
<i>Розділ 2. Описова статистика</i>	2.1. Первинний аналіз даних. Візуалізація даних 2.2. Підготовка даних до аналізу. Заповнення пропусків 2.3. Business Intelligence та дашборди
<i>Розділ 3. Вивідна статистика</i>	3.1. Точкові та інтервальні оцінки параметрів законів розподілу ймовірностей та методи їх визначення 3.2. Перевірка статистичних гіпотез 3.3. Дисперсійний та кореляційний аналіз
<i>Розділ 4. Побудова та верифікація моделей</i>	4.1. Вибір моделі. Верифікація та оцінка якості моделі 4.2. Регресійні моделі 4.3. Кластеризація та класифікація 4.4. Дерева рішень. Випадковий ліс 4.5. Методи перевірки моделі в умовах недостатнього обсягу даних 4.6. Методи зниження розмірності задачі. Факторний аналіз 4.7. Аналіз часових послідовностей 4.8. Аналіз текстів

4. Навчальні матеріали та ресурси

Основна література

1. Конспект лекцій в класрумі

Додаткова література

2. Ліхоузова Т.А. Теорія ймовірностей та математична статистика [Електронний ресурс]: підручник / Електронні текстові данні (1 файл: 5,2 Мбайт). – Київ : НТУУ «КПІ ім. Ігоря Сікорського», 2018. – Доступ : <http://ela.kpi.ua/handle/123456789/22404>
3. Бахрушин В. Є. Методи аналізу даних : навч. посіб. для студентів / В. Є. Бахрушин. – Запоріжжя : КПУ, 2011. – 268 с.
4. Пасічник В.В., Шаховська Н.Б. Сховища даних: Навчальний посібник. - Львів: “Магнолія 2006”, 2008.-496 С. ISBN 978-966-2025-18-7
5. Inmon W. H. Building the Data Warehouse. – 4th ed. – Hoboken (NJ) : Wiley Publishing, Inc., 2005. – ISBN 978-0-7645-9944-6.
6. Kimball R., Ross M., Thornthwaite W., Mundy J., Becker B. The Data Warehouse Lifecycle Toolkit. – 2nd ed. – Indianapolis (IN) : Wiley Publishing, Inc., 2008. – 672 p.
7. Shmueli G., Bruce P. C., Gedeck P., Patel N. R., Yahav I. Data Mining for Business Analytics: Concepts, Techniques, and Applications in R. – Hoboken (NJ) : Wiley, 2020. – 608 p.
8. James G., Witten D., Hastie T., Tibshirani R. An Introduction to Statistical Learning: With Applications in R. – 2nd ed. – New York : Springer, 2021. – 607 p. – (Springer Texts in Statistics).
9. Hyndman R. J., Athanasopoulos G. Forecasting: Principles and Practice. – 2nd ed. – Melbourne : OTexts, 2018. – Режим доступу: <https://OTexts.com/fpp2>.
- 10.

Інформаційні ресурси

<https://classroom.google.com>

5. Методика опанування навчальної дисципліни (освітнього компонента)

5.1. Тематика лекцій

№ з/п	Назва теми заняття та перелік основних питань
<i>Лекція 1</i>	Що таке аналіз даних Сучасна екосистема даних. Корпоративне середовище даних. Ключові гравці в екосистемі даних. Різновиди аналізу даних. Роль аналітика даних. Які бувають дані. Як можна зберігати дані. Джерела даних. Мови для обробки даних.
<i>Лекція 2</i>	Основи сховищ даних Способи збору та збереження даних. Основні характеристики сховищ даних. Відмінності OLTP та OLAP. Багатовимірний аналіз (OLAP). Архітектура сховищ даних. Slowly Changing Dimensions (SCD). Вітрини даних (Data Marts).
<i>Лекція 3</i>	ETL- процес Що таке Data mining. Методи і стадії Data mining. Завдання Data mining. ETL- процес. Первинний аналіз даних. Підготовка даних до аналізу. Заповнення пропусків.
<i>Лекція 4</i>	SQL Window functions Призначення віконних функцій. Різновиди віконних функцій.
<i>Лекція 5</i>	Описова статистика Статистичний розподіл. Міри центральної тенденції. Міри варіативності. Графічне представлення даних. Комунікація та візуалізація. Популярні інструменти візуалізації.
<i>Лекція 6</i>	Створення дашбордів у Power BI Візуалізація даних. Business Intelligence та дашборди. Завантаження даних. Побудова візуалізації. Вибір правильної візуалізації. Форматування та оптимізація.
<i>Лекція 7</i>	Покрокове керівництво для основних візуалів Power BI Стовпчасті та лінійчасті діаграми. Кругові діаграми. Трімап (Treemap). Лінійні графіки. Картки KPI. Зрізи та фільтри. Як зібрати все на дашборді.
<i>Лекція 8</i>	Вивідна статистика. Статистичні оцінки Вимоги до статистичних оцінок. Точкові та інтервальні оцінки параметрів та методи їх визначення.
<i>Лекція 9</i>	Перевірка статистичних гіпотез Типи статистичних гіпотез. Похибки при перевірці гіпотез. Критерії перевірки гіпотез. Критична область та потужність критерію. Алгоритм перевірки гіпотез.
<i>Лекція 10</i>	Аналіз впливу факторів Методи аналізу факторів впливу. Дисперсійний аналіз. Кореляційний аналіз.
<i>Лекція 11</i>	Побудова та верифікація моделей Класифікація методів машинного навчання. Мета побудови моделі. Методи побудови моделей. Вибір моделі. Верифікація та оцінка якості моделі. Компроміс між точністю передбачення та інтерпретації.
<i>Лекція 12</i>	Регресійні моделі Загальна характеристика регресійного аналізу. Види регресійних моделей. Основні методи регресійного аналізу. Перевірка адекватності моделі. Порядок виконання регресійного аналізу.
<i>Лекція 13</i>	Кластеризація та класифікація Задача класифікації. Типи класів. Параметричні методи класифікації без навчання. Кластерний аналіз. Класифікація з навчанням. Вибір методів класифікації.

Лекція 14	Методи перевірки моделі в умовах недостатнього обсягу даних Методи повторної вибірки. Перехресна перевірка (Cross-Validation). Покращення лінійних моделей. Оцінка якості моделі.
Лекція 15	Дерева в задачах регресії та класифікації Загальна характеристика методів. Дерева в задачі регресії. Дерева в задачі класифікації. Дерево чи лінійна модель? Ансамблеві методи: Bagging, Random Forests, Boosting. Переваги та недоліки дерев.
Лекція 16	Методи зниження розмірності задачі. Факторний аналіз Мета та завдання факторного аналізу. Основні методи факторного аналізу. Критерії вибору та оцінки моделі. Практичні вимоги до даних та обмеження.
Лекція 17	Аналіз часових послідовностей Загальний огляд задач прогнозування. Особливості аналізу часових рядів. Прогнозування на основі регресії. Використання автокореляції та ARIMA моделей. Методи згладжування.
Лекція 18	Аналіз текстів Основи обробки природної мови (NLP). Підходи до NLP. Компоненти NLP. Процес роботи з текстом. Методи представлення тексту. Інструменти та бібліотеки для NLP.
Лекція 19	Моделі на основі нейромереж. Моделі для рекомендаційних систем (резервне заняття)

5.2. Тематика комп'ютерних практикумів/лабораторних робіт

Робота студентів складається з:

- виконання практичного завдання за наданим прикладом (за вибором студента на практичному занятті або в час для самостійної роботи),
- виконання завдання лабораторної роботи та оформлення звіту (за вибором студента на практичному занятті, в час для самостійної роботи або під час лабораторного заняття),
- захист лабораторної роботи, звіт має бути завантажено в класрум не менше ніж за добу до захисту (під час лабораторного заняття).

№ з/п	Назва теми заняття та перелік відповідних лабораторних робіт
практ.1	Знайомство з середовищем
практ.1 + лаб. 1	Підключення до джерел даних (АД1)
практ.2 + лаб. 2	Створення сховища даних (АД2)
практ.3 + лаб. 3	Створення ВІ рішення (АД3)
практ.4 + лаб. 4	Обробка даних. Усунення суперечностей та заповнення пропусків (Python1)
практ.4 + лаб. 4	Пошуковий аналіз даних. Виявлення предикторів (Python2)
практ.5 + лаб. 5	Прогнозування кількісної ознаки. Регресійні моделі (Python3, Python4)
практ.6 + лаб. 6	Прогнозування якісної ознаки. Класифікація та кластеризація (Python5)
практ.7 + лаб. 7	Аналіз часових послідовностей (Python6)
практ.7 + лаб. 8	Аналіз текстів (Python7)
практ.8 + 9	Захист лабораторних робіт
лаб. 9	Захист лабораторних робіт

6. Самостійна робота студента

Самостійна робота студентів складається з:

- підготовки до аудиторних занять (8 годин),
- підготовки до виконання практичних робіт (16 годин),

- виконання та підготовки до захисту лабораторних робіт (20 годин),
- написання тестів (4 години).

Кількість годин СРС на кожен тему залежить від студента і може коливатись в межах від 0.5 до 2 годин.

- Політика та контроль

7. Політика навчальної дисципліни (освітнього компонента)

Форми організації освітнього процесу, види навчальних занять і оцінювання результатів навчання регламентуються Положенням про організацію освітнього процесу в Національному технічному університеті України «Київському політехнічному інституті імені Ігоря Сікорського».

Політика виставлення оцінок: кожна оцінка виставляється відповідно до розроблених викладачем та заздалегідь оголошених студентам критеріїв, а також мотивується в індивідуальному порядку на вимогу студента; у випадку не виконання студентом усіх передбачених навчальним планом видів занять (лабораторних робіт, тесту) до заліку він не допускається; пропущені лабораторні заняття обов'язково мають бути відпрацьовані.

Відвідування не є обов'язковим. Якщо студент не може бути присутнім на заняттях, він все одно несе відповідальність за опрацювання теоретичного матеріалу та виконання завдань, що розглядалися на лабораторному занятті. Відеозаписи лекційних занять, а також конспект лекцій доступні студентам в класрумі.

Порядок зарахування пропущених занять. Відпрацювання пропущеного заняття з лекційного курсу здійснюється шляхом перегляду відеозапису або конспекту лекції. Вікторини для студентів, що пропустили лекцію, додатково не проводяться - ці бали вони втрачають. Відпрацювання пропущеного лабораторного заняття здійснюється шляхом самостійного виконання завдання і подачі звіту в класрум. Звіт має бути зданий мінімум за добу до захисту роботи. Захист виконується або на наступному занятті, або відповідно до графіка консультацій викладача.

Політика дедлайнів та перескладань. Дедлайни виконання тестів жорсткі, пізніше бали не нараховуються. Звіти з практичних робіт здавати не потрібно. Вчасність виконання цих робіт впливає лише на можливість приступити до лабораторної роботи з тієї ж теми. Дедлайни виконання лабораторних гнучкі, перший з них вказаний в класрумі для кожної роботи. Не здані в частині захисту лабораторні роботи можна захищати пізніше дедлайну, але не пізніше останнього практичного заняття в семестрі і не більше трьох робіт за раз. Якщо студент не набрав достатню кількість балів або не виконав мінімально необхідну кількість лабораторних робіт, то він не допущений до залікової сесії. Перенесення вивчення дисципліни на наступний семестр як додаткової послуги неможливе.

Політика академічної поведінки та доброчесності: конфліктні ситуації мають відкрито обговорюватись в академічних групах з викладачем, необхідно бути взаємно толерантним, поважати думку іншого. Плагіат та інші форми нечесної роботи неприпустимі. Всі індивідуальні завдання та курсову роботу студент має виконати самостійно із використанням рекомендованої літератури й отриманих знань та навичок. Цитування в письмових роботах допускається тільки із відповідним посиланням на авторський текст. Недопустимі підказки і списування у ході захисту лабораторних робіт, на контрольних роботах, на іспиті.

Норми академічної етики: дисциплінованість; дотримання субординації; чесність; відповідальність; робота в аудиторії з відключеними мобільними телефонами. Повага один до одного дає можливість ефективніше досягати поставлених командних результатів. При виконанні лабораторних робіт студент може користуватися ноутбуком. Якщо ви використовуєте свій ноутбук чи телефон для аудіо- чи відеозапису, необхідно заздалегідь отримати дозвіл викладача.

Дотримання академічної доброчесності студентів й викладачів регламентується кодексом честі Національного технічного університету України «Київський політехнічний інститут», положення про організацію освітнього процесу в КПІ ім. Ігоря Сікорського

8. Види контролю та рейтингова система оцінювання результатів навчання (PCO)

Розподіл балів, які отримують студенти на заняттях

Види контролю	бали
Практичні та лабораторні роботи (10 робіт)	5*10
Тести (4 роботи)	10+12+8+10

Вікторини (10 робіт + 2 додаткові)	1*10 (+2)
------------------------------------	-----------

$$R=50+40+10=100$$

Практична робота: за вибором студента проводиться або під час практичного заняття, або в час, відведений на самостійну роботу. Під час практичного заняття можна отримати консультацію викладача щодо завдань, які викликають труднощі.

Лабораторна робота: за вибором студента виконується під час практичного заняття або в час, відведений на самостійну роботу, а захищається під час лабораторного заняття. Спочатку необхідно здати звіт в класрумі, потім захистити роботу. Оцінка включає 1 бал за звіт з основного завдання, 1 бал за звіт з додаткового завдання, за відповіді на питання студент може отримати ще 3 бали:

- повна відповідь 3 бали;
- часткова відповідь 1-2 бали;
- незадовільна відповідь 0 балів.

Гранична дата прийому робіт - 10 травня.

МКР у формі тестів: проводиться в час, відведений на самостійну роботу. Завдання включає 4 тести з теоретичними питаннями, по одному тесту на кожний розділ. За кожне питання тесту студент може отримати 1 бал:

- правильна відповідь 1 бал;
- незадовільна відповідь 0 балів.

Після дедлайну бали за тест не нараховуються.

Вікторини: проводяться на лекційному занятті, включають теоретичні питання на тему лекції. Студент може отримати:

- 50-99% правильних відповідей 1 бал;
- 0-49% правильних відповідей 0 балів.

Заохочувальні бали нараховуються за всі правильні відповіді на вікторині (2 бали замість 1), а також є дві додаткові вікторини.

Календарний контроль: провадиться двічі на семестр як моніторинг поточного стану виконання вимог силабусу.

За результатами навчальної роботи за перші 7 тижнів максимально можлива кількість балів – 36 балів. На першій атестації (7-й тиждень) студент отримує “зараховано”, якщо його поточний рейтинг не менше 20 балів.

За результатами 12 тижнів навчання максимально можлива кількість балів – 80 балів. На другій атестації (12-й тиждень) студент отримує “зараховано”, якщо його поточний рейтинг не менше 40 балів.

Семестровий контроль: залік

Таблиця відповідності рейтингових балів оцінкам за університетською шкалою:

Кількість балів	Оцінка
100-95	Відмінно
94-85	Дуже добре
84-75	Добре
74-65	Задовільно
64-60	Достатньо
менше 60	Незадовільно
Не виконані умови допуску (менше 7 лаб, менше 40 балів)	Не допущено

9. Додаткова інформація з дисципліни (освітнього компонента)

По кожній темі дисципліни в конспекті лекцій надано перелік додаткової літератури, якою можна скористатись для більш глибокого опанування теми.

Зарахування сертифікатів проходження дистанційних чи онлайн курсів за відповідною тематикою не передбачено.

Перелік питань, які виносяться на семестровий контроль, відповідає темам і наведено в класрумі.

Робочу програму навчальної дисципліни (силабус):

Складено доцент Ліхоузова Т.А., доцент Олійник Ю.О.

Ухвалено кафедрою ІПІ (протокол № 16 від 23.06.2025р.)

Погоджено Методичною комісією факультету (протокол № 3 від 17.10.2025р.)