



АНАЛІЗ ДАНИХ В ІНФОРМАЦІЙНИХ СИСТЕМАХ

Робоча програма навчальної дисципліни (Силабус)

– Реквізити навчальної дисципліни

Рівень вищої освіти	Перший (бакалаврський)
Галузь знань	12 Інформаційні технології
Спеціальність	121 Інженерія програмного забезпечення
Освітня програма	Інженерія програмного забезпечення інформаційних систем
Статус дисципліни	нормативна
Форма навчання	очна (денна)/заочна/дистанційна
Рік підготовки, семестр	2 курс, весняний семестр
Обсяг дисципліни	4 кредити
Семестровий контроль/ контрольні заходи	залік
Розклад занять	Лекції: одна пара 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@iit.kpi.ua Лабораторні заняття: доц. Олійник Ю.О., oliyura@gmail.com
Розміщення курсу	https://classroom.google.com , 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, Бази даних.

Постреквізити – Програмування інтелектуальних ІС.



3. Зміст навчальної дисципліни

Розділ 1. Сховища даних	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
Розділ 2. Описова статистика	2.1. Первинний аналіз даних. Візуалізація даних 2.2. Підготовка даних до аналізу. Заповнення пропусків 2.3. Business Inteligence та дашборди
Розділ 3. Вивідна статистика	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
Розділ 4. Побудова та верифікація моделей	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. *Ліхоузова Т.А.* Теорія імовірностей 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>
2. *Бахрушин В.С.* Методи аналізу даних: навчальний посібник для студентів / Запоріжжя: КПУ, 2011 — 268с

Додаткова література

3. W. H. Inmon, Building the Data Warehouse, Fourth Edition: Wiley Publishing, Inc. ISBN-13: 978-0-7645-9944-6, ISBN-10: 0-7645-9944-5
4. Jaspersoft OLAP Ultimate Guide. Навчальний посібник
5. Ralph Kimball, Margy Ross, Warren Thornthwaite, Joy Mundy, Bob Becker. The Data Warehouse Lifecycle Toolkit, 2nd Edition. Wiley Publishing, Inc, 2008.672с
6. Data Mining for Business Analytics: Concepts, Techniques, and Applications in R
7. *James, Witten, Hastie, Tibshirani* An Introduction to Statistical Learning, with Applications in R
8. *Hyndman, R.J., & Athanasopoulos, G.* (2018) Forecasting: principles and practice, 2nd edition, OTexts: Melbourne, Australia. <https://OTexts.com/fpp2>

Інформаційні ресурси

<https://classroom.google.com>

– Навчальний контент

5. Методика опанування навчальної дисципліни (освітнього компонента)

5.1. Тематика лекцій

№ з/п	Назва теми заняття та перелік основних питань
Лекція 1	Сховища даних Задачі аналізу. Способи збору та збереження даних
Лекція 2	Що таке Data mining. Методи і стадії Data mining. Завдання Data mining. ETL- процес. SQL Window functions.

<i>Лекція 3</i>	Описова статистика Первинний аналіз даних. Візуалізація даних. Business Intelligence та дашборди
<i>Лекція 4</i>	Підготовка даних до аналізу. Заповнення пропусків
<i>Лекція 5</i>	Вивідна статистика Точкові та інтервальні оцінки параметрів та методи їх визначення
<i>Лекція 6</i>	Перевірка статистичних гіпотез
<i>Лекція 7</i>	Дисперсійний аналіз
<i>Лекція 8</i>	Кореляційний аналіз
<i>Лекція 8</i>	Побудова та верифікація моделей Вибір моделі. Верифікація та оцінка якості моделі
<i>Лекція 10</i>	Регресійні моделі
<i>Лекція 11</i>	Кластеризація та класифікація
<i>Лекція 12</i>	Дерева рішень. Випадковий ліс
<i>Лекція 13</i>	Методи перевірки моделі в умовах недостатнього обсягу даних
<i>Лекція 14</i>	Методи зниження розмірності задачі. Факторний аналіз
<i>Лекція 15</i>	Аналіз часових послідовностей
<i>Лекція 16</i>	Аналіз текстів
<i>Лекція 17</i>	Моделі для рекомендаційних систем (резервне заняття)
<i>Лекція 18</i>	Залік

5.2. Тематика комп'ютерних практикумів/лабораторних робіт

№ з/п	Назва теми заняття та перелік основних питань
<i>практ.1 / лаб. 1</i>	Знайомство з середовищем
<i>практ.2 / лаб. 2</i>	Підключення до джерел даних
<i>практ.3 / лаб. 3</i>	Створення сховища даних
<i>практ.4 / лаб. 4</i>	Створення ВІ рішення
<i>практ.5 / лаб. 5</i>	Обробка даних. Усунення суперечностей та заповнення пропусків
<i>практ.6 / лаб. 6</i>	Пошуковий аналіз даних. Виявлення предикторів
<i>практ.7 / лаб. 7</i>	Прогнозування кількісної ознаки. Регресійні моделі
<i>практ.8 / лаб. 8</i>	Прогнозування якісної ознаки. Класифікація та кластеризація
<i>практ.9 / лаб. 9</i>	Аналіз часових послідовностей
<i>практ.10 / лаб. 10</i>	Аналіз текстів

6. Самостійна робота студента/аспіранта

Самостійна робота студентів складається з:

- підготовки до аудиторних занять (лекцій та лабораторних робіт),
- виконання практичних робіт (<https://classroom.google.com>),
- виконання лабораторних робіт (<https://classroom.google.com>),
- написання тестів (<https://classroom.google.com>).

– 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 розглядались 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 його захисту відповідно до графіку консультацій викладача.

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
Лабораторні роботи (10 робіт)	5*10
Тести (4 роботи)	10+12+8+10
Вікторини (10 робіт + 2 додаткові)	1*10 (+2)

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

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 (8-й та 9-й тиждень) студент отримує “зараховано”, якщо його поточний рейтинг не менше 20 балів.

За результатами 13 тижнів навчання максимально можлива кількість балів – 80 балів. На другій атестації (14-й тиждень) студент отримує “зараховано”, якщо його поточний рейтинг не менше 40 балів.

Семестровий контроль: залік

Таблиця відповідності рейтингових балів оцінкам за університетською шкалою:

<i>Кількість балів</i>	<i>Оцінка</i>
100-95	Відмінно
94-85	Дуже добре
84-75	Добре
74-65	Задовільно
64-60	Достатньо
менше 60	Незадовільно
Не виконані умови допуску (менше 6 лаб, менше 30 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.

Робочу програму навчальної дисципліни (силабус):

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

Ухвалено кафедрою ІПІ (протокол № 16 від 29.05.2024р.)

Погоджено Методичною комісією факультету (протокол № 10 від 21.06.2024р.)