



Силабус курсу

Інтелектуальний аналіз даних

Освітньо-професійна програма «Системний аналіз»
Ступінь вищої освіти – бакалавр
Спеціальність: 124 «Системний аналіз»

Рік навчання: 4 Семестр: 7

Кількість кредитів: 6 Мова викладання: українська

Керівник курсу

ПШ

к.е.н., доцент Гончар Людмила Іванівна

Контактна інформація

l.honchar@wunu.edu.ua

Опис

Предметом дисципліни є вивчення теоретичних і практичних основ інтелектуального аналізу даних, що охоплюють процеси консолідації та підготовки інформаційних ресурсів (у тому числі створення й використання сховищ даних, методів ETL та OLAP), базові методи побудови аналітичних моделей (класифікація, регресія, пошук асоціативних правил, кластеризація), візуальний аналіз та обробку текстових даних (Text Mining), проектування й оцінювання ансамблів моделей, застосування інструментів аналізу в реальному часі й у веб-середовищі, дослідження бізнес-процесів (Process Mining), а також ознайомлення зі стандартами та галузевими технологіями для автоматизації процесів знань.

Мета та цілі курсу

Передбачено вивчення формальних методів застосування прикладних інформаційних технологій для розв'язання прикладних задач, зокрема щодо аналітики даних та процесів та їх автоматизації.

Завдання дисципліни. В результаті вивчення курсу "Інтелектуальний аналіз даних" студенти повинні:

- знати основні поняття лінійного та квадратичного аналізу вибірок, машин опорних векторів, стохастичного градієнтного спуску, методу найближчих сусідів, сумішей нормальних розподілів, наївних Байєсівських класифікаторів, дерев рішень, ансамблевих методів, напівконтрольованого навчання;
- вміти здійснювати лінійний та квадратичний аналіз вибірок, реалізовувати кластеризацію вибірок за допомогою машин опорних векторів, стохастичного градієнтного спуску, найближчих сусідів, наївних Байєсівських класифікаторів, дерев рішень, будувати прогнози на основі сумішей нормальних розподілів, ансамблевих методів, напівконтрольованого навчання та ізотонічної регресії.

Перелік тем

1. Технології аналізу даних
2. Консолідація даних
3. OLAP системи
4. Класифікація і регресія
5. Пошук асоціативних правил

6. Кластеризація
7. Візуальний аналіз даних
8. Аналіз текстової інформації (Text Mining)
9. Ансамблі моделей
10. Інтелектуальний аналіз даних в реальному часі
11. Здобування даних через Web
12. Методи аналізу бізнес-процесів
13. Стандарти аналізу даних

Рекомендовані джерела інформації

Основна література

1. Болюбаш Н.М. Інтелектуальний аналіз даних : навч. посіб. / Н.М. Болюбаш. – Миколаїв : Вид-во ЧНУ ім. Петра Могили, 2023. – 320 с. // https://www.researchgate.net/publication/379994907_Bolubas_N_M_Intelektualnij_analiz_daniv
2. Гороховатський В.О., Творошенко І.С. Методи інтелектуального аналізу та оброблення даних: навч. посібник. – Харків: ХНУРЕ, 2021. 92 с.
3. Лупан І. В. Інтелектуальний аналіз даних Data Mining: навчально-методичний посібник / Ірина Володимирівна Лупан. – Кропивницький: ФОП Піскова М. А., 2022. – 112 с // <https://dspace.cusu.edu.ua/items/f2aabfff-9146-43cc-8525-761437210a00>
4. Гончар Л.І. Методичні вказівки до виконання лабораторних робіт з дисципліни «Інтелектуальний аналіз даних» для студентів спеціальності 124 «Системний аналіз» денної та заочної форм навчання Тернопіль: ФОП Шпак. 2022. 54 с.
5. Гончар Л.І. Методичні вказівки до виконання самостійної роботи з дисципліни «Інтелектуальний аналіз даних» для студентів спеціальності 124 «Системний аналіз» денної та заочної форм навчання Тернопіль: ФОП Шпак. 2022. 22 с.
6. Гончар Л.І. Методичні вказівки до вивчення дисципліни (конспект лекцій) «Інтелектуальний аналіз даних» для студентів спеціальності 124 «Системний аналіз» денної та заочної форм навчання Тернопіль: ФОП Шпак. 2022. 50 с.

Додаткова література:

1. Honchar. L. Information Technology of Motor Vehicle Databases Use to Prevent Terrorist Emergencies. *Advanced Computer Information Technologies (ACIT'2020)*: Proc. of the 9th Intern. Conf. Germany 2020. P.339-343.
2. Honchar L, A. Pukas, A. Melnyk, I.Voytyuk. A. Yushko, M. Romanyuk. Transactional Business Application Based on Microservice Architecture , *2021 11 th International Conference on Advanced Computer Information Technologies (ACIT)*. 2021. P.564-567.
3. Honchar L, M. Dyvak, A. Melnyk, L. Dostalek, V. Ostroverkhov and I. Romanets, Repository of Interval Models of Dynamics of Concentrations of Harmful Emissions of Motor Transport, *2022 12th International Conference on Advanced Computer Information Technologies (ACIT)*. 2022. P. 89-94.
4. Honchar L. Classification of Information Objects with Fuzzy Parameters in E-learning Systems / V. Mukhin, V. Zavgorodnii, V. Liskin and others// *The 12th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, 7-9 September, 2023. Dortmund, Germany. P. 1189-1193.
5. Honchar L Ontology as a Software Superstructure to the System for Mathematical Modeling based on Interval Data. *13th International Conference on Advanced Computer Information Technologies (ACIT)*. 2023. P. 543-548.
6. Decision Tree <https://www.geeksforgeeks.org/decision-tree/>
7. Kaixin Wang. Introduction to Gaussian process regression, Part 1: The basics // <https://medium.com/data-science-at-microsoft/introduction-to-gaussian-process-regression-part-1-the-basics-3cb79d9f155f>

8. Kaixin Wang. Introduction to Gaussian process regression, Part 1: Application to predicting concrete strength // <https://medium.com/data-science-at-microsoft/introduction-to-gaussian-process-regression-part-2-application-to-predicting-concrete-strength-3facef04f639>
9. Naive Bayes Classifiers // <https://www.geeksforgeeks.org/naive-bayes-classifiers/>
10. Scikit-learn. User Guide // https://scikit-learn.org/1.5/user_guide.html
11. Scikit-learn. Stochastic Gradient Descent // <https://scikit-learn.org/1.5/modules/sgd.html>
12. Scikit-learn. SGD: Maximum margin separating hyperplane // https://scikit-learn.org/dev/auto_examples/linear_model/plot_sgd_separating_hyperplane.html
13. Scikit-learn. SGD: Weighted samples // https://scikit-learn.org/dev/auto_examples/linear_model/plot_sgd_weighted_samples.html
14. Scikit-learn. Nearest Neighbors. // <https://scikit-learn.org/1.5/modules/neighbors.html>

Оцінювання

Підсумковий бал (за 100-бальною шкалою) з дисципліни «Інтелектуальний аналіз» визначається як середньозважена величина, в залежності від питомої ваги кожної складової залікового кредиту:

Модуль 1		Модуль 2		Модуль 3	Модуль 4	Модуль 5
10%	10%	10%	10%	15%	15%	40%
Поточне оцінювання	Модульний контроль 1	Поточне оцінювання	Модульний контроль 2	Тренінг	Самостійна робота	Екзамен
Оцінка за даний модуль визначається як середнє арифметичне за роботу на практичних заняттях (по темах №1-6)	Підсумкова письмова робота за темами №1-6	Оцінка за даний модуль визначається як середнє арифметичне за роботу на практичних заняттях (по темах №7-13)	Підсумкова письмова робота за темами №7-13	Визначається як середнє арифметичне за виконання завдань за темами №1-13 тренінгу	Оцінка за виконаний і представлений звіт із оброблення отриманих даних	1. Теоретична частина: 3 запитання по 20 балів (60 балів) 2. Практична частина: задача – 40 балів

Шкала оцінювання:

За шкалою ЗУНУ	За національною шкалою	За шкалою ECTS
90–100	Відмінно	A (відмінно)
85–89	Добре	B (дуже добре)
75-84		C (добре)
65-74	Задовільно	D (задовільно)
60-64		E (достатньо)
35-59	Незадовільно	FX (незадовільно з можливістю повторного складання)
1-34		F (незадовільно з обов'язковим повторним курсом)

Політики курсу

Академічна доброчесність. Дотримання академічної доброчесності студентами передбачає:

- самостійне виконання навчальних завдань, завдань поточного та підсумкового контролю результатів навчання (для осіб з особливими освітніми потребами ця вимога застосовується з урахуванням їхніх індивідуальних потреб і можливостей);
- посилення на джерела інформації у разі використання ідей, розробок, тверджень, відомостей;
- дотримання норм законодавства про авторське право і суміжні права;
- надання достовірної інформації про результати власної навчальної (наукової, творчої) діяльності, використані методики досліджень і джерела інформації.

Порушенням академічної доброчесності вважається:

академічний плагіат - оприлюднення (частково або повністю) наукових (творчих) результатів, отриманих іншими особами, як результатів власного дослідження (творчості) та/або відтворення опублікованих текстів (оприлюднених творів мистецтва) інших авторів без зазначення авторства; самоплагіат - оприлюднення (частково або повністю) власних раніше опублікованих наукових результатів як нових наукових результатів;

фабрикація - вигадкування даних чи фактів, що використовуються в освітньому процесі або наукових дослідженнях;

фальсифікація - свідомо зміна чи модифікація вже наявних даних, що стосуються освітнього процесу чи наукових досліджень;

списування - виконання письмових робіт із залученням зовнішніх джерел інформації, крім дозволених для використання, зокрема під час оцінювання результатів навчання.

За порушення академічної доброчесності здобувачі освіти можуть бути притягнені до такої академічної відповідальності:

- повторне проходження оцінювання (контрольна робота, іспит, залік тощо);
- повторне проходження відповідного освітнього компонента освітньої програми.

Політика запізнення. За несвоєчасно виконані завдання буде накладено штраф 10 відсотків від загальної кількості балів за це завдання. Примітка. Виключення можуть бути зроблені до невчасно зданих завдань з поважних причин.