

Український державний університет залізничного транспорту

Рекомендовано на засіданні
кафедри інформаційних технологій
протокол № 1 від 29 серпня 2024 р.

Силабус з дисципліни
ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ДАНИХ
2024-2025 навчальний рік, I семестр

Освітній рівень	Перший (бакалаврський) рівень вищої освіти
Галузь знань	12 Інформаційні технології
Спеціальність	126 Інформаційні системи та технології
Освітня програма	Технології штучного інтелекту

Лектор та керівник лабораторних занять:

Іванюк Олександр Ігорович (доцент кафедри ІТ, доктор філософії)

Контакти: +38 (057) 730-19-84, e-mail: ivaniuk@kart.edu.ua

Консультації за дисципліною – понеділок, 12:40-14:00.

Мета курсу – формування знань та навичок щодо побудови, розуміння принципів роботи і практичного застосування методів інтелектуального аналізу даних та машинного навчання.

Анотація

Дисципліна «Інтелектуальний аналіз даних» є основою сучасної науки про дані та штучний інтелект, надаючи студентам навички та знання для ефективного використання даних. В рамках цього курсу студенти вивчатимуть фундаментальні принципи та методології машинного навчання, включаючи регресію, класифікацію, кластеризацію та зменшення розмірності. Вони навчатимуться будувати, оцінювати та вдосконалювати прогнозні моделі через поєднання теоретичних концепцій та практичних застосувань, що дозволить їм впевнено вирішувати складні реальні проблеми.

Теми дисципліни охоплюють такі методи, як лінійна та поліноміальна регресія, логістична регресія, дерева рішень та ансамблеві методи, включаючи випадкові ліси. Крім того, розглядаються критичні теми, як перенавчання, регуляризація та виявлення аномалій. Лабораторні заняття надають практичний досвід роботи з Python, що зміцнює розуміння та застосування цих методів на практиці.

Ця дисципліна є важливою для студентів, які прагнуть зробити кар'єру в аналізі даних, штучному інтелекті та суміжних галузях. Вона забезпечує всебічне розуміння того, як отримувати користь з даних та створювати інтелектуальні системи, що є вирішальним для сучасних фахівців у цій сфері.

Розподіл часу

Кількість кредитів – 5.

Загальна кількість годин вивчення дисципліни – 150.

Кількість годин на проведення лекцій – 30.

Кількість годин на проведення лабораторних робіт – 45.

Кількість годин на самостійну роботу – 75.

Рік, курс та семестр навчання – 2024/2025 н. р., IV курс, I семестр.

Теми

Лекції:

1. Введення в аналіз даних та машинне навчання.
2. Регресія. Метод лінійної регресії.
3. Метод поліноміальної регресії.
4. Перенавчання. Регуляризація.
5. Класифікація. Метод логістичної регресії.
6. Метод наївної баєсової класифікації.
7. Метрики якості класифікації. Техніки валідації.

8. Метрична класифікація та регресія. Метод k найближчих сусідів.
9. Логічна класифікація та регресія. Метод дерев рішень.
10. Ансамблеві методи. Метод випадкового лісу.
11. Кластеризація. Метод ієрархічної кластеризації. Метод k-means.
12. Зниження розмірності. Метод головних компонент.
13. Виявлення аномалій. Метод локального рівня викиду. Метод ізолюючого лісу.
14. Основи рекомендаційних систем.

Лабораторні роботи:

1. Основи лінійної алгебри в Python.
2. Побудова та застосування методу лінійної регресії.
3. Побудова та застосування методу поліноміальної регресії.
4. Побудова та застосування методу логістичної регресії.
5. Побудова та застосування методу наївної баєсової класифікації.
6. Побудова та застосування методу k найближчих сусідів.
7. Дослідження та застосування методів дерев рішень і випадкового лісу.
8. Побудова та застосування методу ієрархічної кластеризації та методу k-means.
9. Застосування методів зниження розмірності.
10. Застосування методів виявлення аномалій.

Правила оцінювання

Порядок оцінювання результатів навчання визначається [Положенням про контроль та оцінювання якості знань студентів в УкрДУЗТ](#).

Оцінювання знань студента з дисципліни здійснюється за наступними складовими:

- виконання та захист лабораторних робіт;
- проходження модульного контролю;
- складання іспиту (обов'язкове при отриманні загальної модульної оцінки F, за вибором студента при отриманні загальної модульної оцінки D або B).

Лабораторні роботи можуть бути захищені за умови наявності самостійно та вірно виконаного завдання. Захист лабораторних робіт можливий під час проведення лабораторних занять та у години консультацій, але не більше однієї лабораторної роботи в день.

У випадку пропуску студентом більше ніж двох лекцій протягом одного модуля, поточна оцінка за відповідний модуль зменшується на три бали, за кожну пропущену лекцію понад дві. Штрафні бали можуть бути відпрацьовані, шляхом виконання додаткових індивідуальних або групових завдань,

проявлення активності під час лекційних та лабораторних занять, участі у студентських конференціях (за тематикою дисципліни) або змаганнях з аналізу даних.

Модуль	Вид контролю	Вид робіт		Максимальна кількість балів	Сума	
1	Поточний контроль	Виконання та захист лабораторних робіт	ЛР №1	12	60	100
			ЛР №2	12		
			ЛР №3	12		
			ЛР №4	12		
			ЛР №5	12		
Модульний контроль			40	40		
2	Поточний контроль	Виконання та захист лабораторних робіт	ЛР №6	12	60	100
			ЛР №7	12		
			ЛР №8	12		
			ЛР №9	12		
			ЛР №10	12		
Модульний контроль			40	40		

При заповненні заліково-екзаменаційної відомості та залікової книжки (індивідуального навчального плану) здобувача, оцінка, виставлена за 100-бальною шкалою, переводиться до національної шкали (5, 4, 3, 2) та шкали ЄКТС (A, B, C, D, E, FX, F).

За національною шкалою	Визначення назви за шкалою ЄКТС	За 100-бальною шкалою	За шкалою ЄКТС
ВІДМІННО – 5	Відмінно – відмінне виконання лише з незначною кількістю помилок	90-100	A
ДОБРЕ – 4	Дуже добре – вище середнього рівня з кількома помилками	82-89	B
	Добре – в загальному правильна робота з певною кількістю грубих помилок	75-81	C
ЗАДОВІЛЬНО – 3	Задовільно – непогано, але зі значною кількістю недоліків	69-74	D
	Достатньо – виконання задовольняє мінімальні критерії	60-68	E
НЕЗАДОВІЛЬНО – 2	Незадовільно – потрібно попрацювати перед тим як отримати залік (без повторного вивчення модуля)	35-59	FX
	Незадовільно – необхідна серйозна подальша робота (повторне вивчення модуля)	<35	F

Результати навчання

Дисципліна має на меті сформувати та розвинути наступні компетентності студентів, сформованих у вигляді програмних результатів навчання:

ПР 4. Проводити системний аналіз об'єктів проектування та обґрунтовувати вибір структури, алгоритмів та способів передачі інформації в інформаційних системах та технологіях.

ПР 9. Здійснювати системний аналіз архітектури підприємства та його ІТ інфраструктури, проводити розроблення та вдосконалення її елементної бази і структури.

Кодекс академічної доброчесності

Порушення [Кодексу академічної доброчесності УкрДУЗТ](#) є серйозним порушенням, навіть якщо воно є ненавмисним.

Зокрема, дотримання Кодексу академічної доброчесності УкрДУЗТ означає, що вся робота на іспитах та заліках має виконуватися індивідуально. Під час виконання самостійної роботи студенти можуть консультиватися з викладачами та з іншими студентами, але повинні самостійно розв'язувати завдання, керуючись власними знаннями, уміннями та навичками. Посилання на всі ресурси та джерела (наприклад, у звітах, самостійних роботах чи презентаціях) повинні бути чітко визначені та оформлені належним чином. У разі спільної роботи з іншими студентами над виконанням індивідуальних завдань, ви повинні зазначити ступінь їх залучення до роботи.

Інтеграція студентів із обмеженими можливостями

Вища освіта є провідним чинником підвищення соціального статусу, досягнення духовної, матеріальної незалежності і соціалізації молоді з обмеженими функціональними можливостями й відображає стан розвитку демократичних процесів і гуманізації суспільства.

Для інтеграції студентів із обмеженими можливостями в освітній процес Українського державного університету залізничного транспорту створена система дистанційного навчання на основі сучасних педагогічних, інформаційних, телекомунікаційних технологій.

Доступ до матеріалів дистанційного навчання з цієї дисципліни можна знайти на [платформі дистанційного навчання УкрДУЗТ](#).

Викладачі

Лектор та керівник лабораторних занять:

[Олександр Ігорович Іванюк](#) – доктор філософії, доцент кафедри інформаційних технологій.

Контакти: +38 (057) 730-19-84, e-mail: ivaniuk@kart.edu.ua.

Ресурси

Інформаційні ресурси дисципліни розміщено на [платформі дистанційного навчання УкрДУЗТ](#).

Рекомендована література

Основна

1. Басюк Т.М., Литвин В.В., Захарія Л.М., Кунанець Н.Е. Машинне навчання: навч. посіб. Львів: Новий Світ-2000, 2019. 315 с.
2. Hastie T., Friedman J., Tibshirani R. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. New York: Springer, 2017. URL: <https://web.stanford.edu/~hastie/ElemStatLearn>.
3. Інтелектуальний аналіз даних : практикум / М. Т. Фісун, І. О. Кравець, П. П. Казмірчук, С. Г. Ніколенко. Львів: "Новий Світ-2000", 2019. 162 с.
4. Ng A. Machine Learning Yearning (draft). Deeplearning.ai, 2018.
5. Bishop C. Pattern recognition and machine learning. Springer, 2006. URL: <https://www.microsoft.com/en-us/research/people/cmbishop/prml-book/>.
6. Mitchell T. Machine learning. Singapore: McGraw-Hill, 1997. URL: <http://www.cs.cmu.edu/afs/cs.cmu.edu/user/mitchell/ftp/mlbook.html>.

Допоміжна

1. MacKay D. Information theory, inference, and learning algorithms. Cambridge: Cambridge University Press, 2019. URL: <http://www.inference.org.uk/mackay/itila>.
2. Burkov A. The hundred-page machine learning book.
3. Daume H. A Course in Machine Learning (v0.9), 2017. URL: <http://ciml.info>.
4. Wasserman L. All of Statistics. New York, NY: Springer, 2013. URL: <http://www.stat.cmu.edu/~larry/all-of-statistics/index.html>.
5. Коельо Л.П., Ричарт В. Побудова систем машинного навчання на мові Python. 2016. 302 с.
6. Ng A. Machine Learning course notes. Coursera, 2011.
7. Domingos, P. The Master Algorithm: How machine learning is reshaping how we live. 2015.

Корисні інтернет-ресурси

1. Конспект лекцій курсу «Machine Learning» (проф. Ендрю Ін, Coursera), автор – Алекс Холхаус. URL: <http://www.holehouse.org/mlclass/>.
2. Спеціалізація «Machine Learning Specialization» (Coursera), автор – проф. Ендрю Ін. URL: <https://www.coursera.org/specializations/machine-learning-introduction>.

3. Курс «Machine Learning» (Udacity) розробник – Технологічний інститут Джорджії. URL: <https://www.udacity.com/course/machine-learning--ud262>.
4. Курс CS229 – «Machine Learning» (Stanford), автор – проф. Ендрю Ін. URL: <https://see.stanford.edu/Course/CS229>.
5. Курс «Machine Learning» (YouTube), автор – mathematicalmonk. URL: <https://bit.ly/2YmuvYN>.
6. Курс «Machine Learning with Python» (YouTube), автор – mathematicalmonk. URL: <https://bit.ly/3h2Srae>.
7. Курс «Scikit-learn Machine Learning with Python and SKlearn» (YouTube), автор – sentdex. URL: <https://bit.ly/38I5qts>.
8. Дбірка корисних матеріалів з машинного навчання (Medium), автор – Роббі Аллен. URL: <https://bit.ly/3kQKzK8>.