



Силабус освітнього компонента Програма навчальної дисципліни



Машинне навчання BigData

Шифр та назва спеціальності

126 – Інформаційні системи та технології

Інститут

ННІ Комп'ютерних наук та інформаційних технологій

Освітня програма

Програмне забезпечення інформаційних систем

Кафедра

Інформаційні системи та технології(329)

Рівень освіти

Магістр

Тип дисципліни

Профільна, вибіркова

Семестр

3

Мова викладання

Українська

Викладачі, розробники



Москаленко Валентина Володимирівна

Valentyna.Moskalenko@khpi.edu.ua

Д.т.н., професор, професор кафедри ICT

Кількість наукових та навчальних публікацій – більше 130.

(<https://publons.com/researcher/1588564/valentyna-moskalenko/>;

Web of Science ResearcherID R-9960-2018;

[https://scholar.google.com.ua/citations?user=eUIdJHIAAAAJ&hl](https://scholar.google.com.ua/citations?user=eUIdJHIAAAAJ&hl;);

<https://www.scopus.com/authid/detail.uri?authorId=36021571200>;

<https://orcid.org/0000-0002-9994-5404>).

Провідний лектор з дисциплін: «Теорія ймовірності та математична статистика», «Основи інформаційних систем та технологій», «Методи бізнес-аналізу для управління вимогами», «Наукові напрямки дослідження інформаційних систем та технологій», «Аналіз та управління вимогами до програмного забезпечення інформаційних систем», «Вступ до Big Data», «Машинне навчання Big Data».

Наукові напрямки: розробка інформаційних систем для стратегічного управління компанією; застосування методів та моделей обчислюваного інтелекту для розв'язання задач управління складними організаційними системами; бізнес -аналітика..

[Детальніше про викладача на сайті кафедри](#)

Загальна інформація

Анотація

Спрямування курсу на знання методів машинного навчання, які використовуються для аналізу великих даних. Вивчення дисципліни дозволить студенту аналізувати новітні тенденції у Machine Learning. Особливістю дисципліни є розгляд найбільш актуальних напрямків машинного навчання: Supervised learning, Unsupervised learning, Reinforcement learning для аналізу великих даних.

Дисципліна інтегрує теоретичні та практичні знання з машинного навчання. Отримані знання щодо розв'язання задач класифікації, кластеризації, зменшення розмірності та пошуку правил дозволять студенту ефективно займатися розробкою алгоритмів для аналізу великих даних

Мета та цілі дисципліни

Метою курсу є формування у студентів теоретичних знань та практичних навичок з основ машинного навчання для аналізу великих даних, а також придбання навичок дослідника і розробника моделей та алгоритмів машинного навчання для розробки інформаційних систем.

Формат занять

Лекції, лабораторні роботи, самостійна робота, консультації. Підсумковий контроль – залік.

Компетентності

ЗК01. Здатність до абстрактного мислення, аналізу та синтезу.

СК04. Здатність розробляти математичні, інформаційні та комп'ютерні моделі об'єктів і процесів інформатизації.

СК05. Здатність використовувати сучасні технології аналізу даних для оптимізації процесів в інформаційних системах.

Результати навчання

РН01. Відшукувати необхідну інформацію в науковій і технічній літературі, базах даних, інших джерелах, аналізувати та оцінювати цю інформацію

РН08. Розробляти моделі інформаційних процесів та систем різного класу, використовувати методи моделювання, формалізації, алгоритмізації та реалізації моделей з використанням сучасних комп'ютерних засобів.

РН09. Розробляти і використовувати сховища даних, здійснювати аналіз даних для підтримки прийняття рішень

Обсяг дисципліни

Загальний обсяг дисципліни 120 год. (4 кредити ECTS): лекції – 30 год., лабораторні роботи – 30 год., самостійна робота – 60 год.

Передумови вивчення дисципліни (пререквізити)

Для ефективного засвоєння дисципліни необхідні знання з програмування, теорії баз даних, вищої математики, теорії ймовірностей та математичної статистики, основ обчислюваного інтелекту, планування, аудит та супровід інформаційних систем

Особливості дисципліни, методи та технології навчання

Методи викладання та навчання:

інтерактивні лекції з презентаціями, дискусії, лабораторні заняття, командна робота, кейс-метод, метод зворотного зв'язку з боку студентів, проблемне навчання.

Форми оцінювання:

оцінювання знань на лабораторних заняттях (CAS), експрес-опитування (CAS)

Програма навчальної дисципліни

Теми лекційних занять

Тема 1. Методи машинного навчання з вчителем (Supervised learning) для аналізу великих даних

Машинне навчання як підрозділ штучного інтелекту. Основні положення класичного машинного навчання. Типи машинного навчання. Задачі навчання. Потреби МН

Алгоритми машинного навчання з вчителем для задачі класифікації. Регресія та інші методи класифікації. Метод k-найближчих сусідів. Алгоритм Naïve Bayes. Метод опорних векторів (support vector machine). Алгоритми побудови дерева рішень (Decision tree).

Тема 2. Методи машинного навчання без вчителя (Unsupervised learning) для аналізу великих даних

Алгоритми машинного навчання без вчителя для задач кластеризації. Методи кластерного аналізу. Ієрархічні алгоритми (Hierarchy Algorithms). Ієрархічні дивизивні (розділові) методи (Divisive ANALysis, DIANA). Неієрархічні методи кластерного аналізу: DBSCAN - просторова кластеризація, алгоритм кластеризації Mean-shift. Оцінка якості у задачі кластеризації.

Методи зменшення розмірності та пошуку правил. Методи вибору ознак (Feature selection). Методи виділення ознак (Feature extraction). методи вибору ознак: фільтри; обгортувальні методи; вбудовані методи; гібридні; ансамблеві. Пошук асоціативних правил.

Тема 3. Алгоритми машинного навчання з підкріпленням (Reinforcement learning) та глибоке навчання для аналізу великих даних.

Алгоритми машинного навчання з підкріпленням (reinforcement learning). Постановка задачі навчання з підкріпленням. Стратегії навчання, жадібні і ϵ -жадібні стратегії (greedy & ϵ -greedy). Моделі машинного навчання з підкріпленням: Марківський процес прийняття рішення; Q-навчання. Генетичні алгоритми. Використання Deep neural network для аналізу великих даних.

Теми практичних занять

Практичні заняття не передбачені навчальним планом.

Теми лабораторних робіт

Тема 1. Методи машинного навчання з вчителем для розв'язання задачі класифікації

Тема 2. Методи машинного навчання без вчителя для розв'язання задачі кластеризації

Тема 3. Методи машинного навчання з підкріпленням для розв'язання задач штучного інтелекту

Самостійна робота

Тема 1. Методи машинного навчання з вчителем (Supervised learning) для аналізу великих даних.

Використання бібліотек Python (Scikit-Learn) для реалізації алгоритмів машинного навчання, Використання Machine Learning and Deep Learning та Statistics and Machine Learning Toolbox у пакеті Matlab для розв'язання задач методами машинного навчання.

Тема 3. Алгоритми машинного навчання з підкріпленням (Reinforcement learning) та глибоке навчання для аналізу великих даних.

Інструментарій Reinforcement Learning Designer середовища MatLab для розв'язання задач за допомогою методів машинного навчання з підкріпленням. Можливості нейропакетів для моделювання нейронних систем.

Індивідуальних завдань не передбачено навчальним планом.

Студентам рекомендовані додаткові матеріали (відео, статті) для самостійного вивчення та опрацювання

Література та навчальні матеріали

Основна література

1. Mathematics Of Big Data And Machine Learning // <https://ocw.mit.edu/courses/res-ll-005-mathematics-of-big-data-and-machine-learning-january-iap-2020/pages/lecture-notes/>
2. Flach, P. (2012). Machine learning: The art and science of algorithms that make sense of data. Cambridge University Press. <https://doi.org/10.1017/CBO9780511973000>
3. Müller A. C., Guido S. (2016) Introduction to Machine Learning with Python. O'Reilly Media, Inc., 392 p.
4. Zgurovsky M.Z., Zaychenko Y.P. The Fundamentals of Computational Intelligence: System Approach. Springer International Publishing Switzerland, 2016. – 375 p.
5. Sutton R. S., Barto A. G. (2018) Reinforcement Learning: An Introduction. Second Edition. MIT Press, Cambridge, MA, 552 p. .

Додаткова література

- 1 Luge G. F., Stubblefield W. A (2013) Artificial intelligence: structures and strategies for complex problem solving. 3rd ed. Harlow, England; Reading, Mass. Addison-Wesley, 740 p.

2. Mathar R. Fundamentals of Big Data Analytics/ Lecture Notes, 2019// <https://www.ti.rwth-aachen.de/teaching/BigData/FBDA.pdf>
3. Big Data Analytics [R17a0528] Lecture Notes. 2021
//[https://mrcet.com/downloads/digital_notes/CSE/IV%20Year/\(R17A0528%20\)%20Big%20Data%20Analytics%20Digital%20notes.pdf](https://mrcet.com/downloads/digital_notes/CSE/IV%20Year/(R17A0528%20)%20Big%20Data%20Analytics%20Digital%20notes.pdf)
4. 3Zgurovsky M.Z., Zaychenko Y.P. Big Data: Conceptual Analysis and Applications. Springer, 2020. – 298p.
5. Géron A. (2019) Hands-On Machine Learning with Scikit-Learn, Keras, and Tensorflow: Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media

Web-ресурси

1. Machine Learning in MATLAB //<https://www.mathworks.com/help/stats/machine-learning-in-matlab.html>.
2. Mastering Machine Learning: A Step-by-Step Guide with MATLAB
//<https://www.mathworks.com/campaigns/offers/mastering-machine-learning-with-matlab.html>
3. Machine Learning Tutorial //<https://www.javatpoint.com/machine-learning>.

Система оцінювання

Критерії оцінювання успішності студента та розподіл балів

100% підсумкове оцінювання у вигляді заліку
100% поточне оцінювання:
3 лабораторні роботи (по 30%)
експрес-опитування (по 10%)

Шкала оцінювання

Сума балів	Національна оцінка	ECTS
90–100	Відмінно	A
82–89	Добре	B
75–81	Добре	C
64–74	Задовільно	D
60–63	Задовільно	E
35–59	Незадовільно (потрібне додаткове вивчення)	FX
1–34	Незадовільно (потрібне повторне вивчення)	F

Норми академічної етики і політика курсу

Студент повинен дотримуватися «Кодексу етики академічних взаємовідносин та доброчесності НТУ «ХПІ»: виявляти дисциплінованість, вихованість, доброзичливість, чесність, відповідальність. Конфліктні ситуації повинні відкрито обговорюватися в навчальних групах з викладачем, а при неможливості вирішення конфлікту – доводитися до відома співробітників дирекції інституту. Нормативно-правове забезпечення впровадження принципів академічної доброчесності НТУ «ХПІ» розміщено на сайті: <http://blogs.kpi.kharkov.ua/v2/nv/akademichna-dobrochesnist/>

Погодження

Силабус погоджено

30.08.2023



Завідувач кафедри
Олена НІКУЛІНА

30.08.2023



Гарант ОП
Олена НІКУЛІНА