Силабус освітнього компонента
Програма навчальної дисципліни

Основи Big Data

Шифр та назва спеціальності
122 – Комп’ютерні науки

Інститут
ННІ Комп'ютерних наук та інформаційних технологій

Освітня програма
Комп'ютерні науки. Штучний інтелект та управління проєктами

Кафедра
Інформаційні системи та технології(329)

Рівень освіти
Бакалавра

Тип дисципліни
Профільна, вибіркова

Семестр
5

Мова викладання
Українська

Викладачі, розробники

Москаленко Валентина Володимирівна
Valentyna.Moskalenko@khpi.edu.ua
Д.т.н., професор, професор кафедри ІСТ

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://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

Дисципліна присвячена вивченню базових понять великих даних, технологій аналізу даних, вона висвітлює питання аналізу великих даних і пов'язаних з ними технічні, концептуальні та етичні проблеми. Передбачається вивчення різних методів розв’язання задач, які потребують обробку та аналіз великих даних і надання практичних навичок розробки програмних рішень з аналізу великих даних для конкретних предметних областей

Мета та цілі дисципліни

Надання знань із застосування основних сучасних технологій обробки великих даних для розв'язання задач у різних галузях, а також надання практичних навичок, достатніх для успішного виконання професійних обов'язків в області видобутку нових знань з масивів великих даних в інформаційних системах.

Формат занять

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

Компетентності

КЗ 1. Здатність до абстрактного мислення, аналізу та синтезу.
КЗ 2. Здатність застосовувати знання у практичних ситуаціях
КЗ 6. Здатність до пошуку, оброблення та узагальнення інформації з різних джерел.
КС 6. Здатність використовувати сучасні інформаційні системи та технології (виробничі, підтримки прийняття рішень, інтелектуального аналізу даних та інші), методики й техніки кібербезпеки під час виконання функціональних завдань та обов'язків.
КС 11. Здатність до аналізу, синтезу і оптимізації інформаційних систем та технологій з використанням математичних моделей і методів.

Результати навчання

ПР 2. Застосовувати знання фундаментальних і природничих наук, системного аналізу та технологій моделювання, стандартних алгоритмів та дискретного аналізу при розв'язанні задач проектування і використання інформаційних систем та технологій.
ПР 3. Використовувати базові знання інформатики й сучасних інформаційних систем та технологій, навички програмування, технології безпечної роботи в комп'ютерних мережах, методи створення баз даних та інтернет-ресурсів, технології розроблення алгоритмів і комп'ютерних програм мовами високого рівня із застосуванням об'єктно-орієнтованого програмування для розв'язання задач проектування і використання інформаційних систем та технологій.

Обсяг дисципліни

Загальний обсяг дисципліни 150 год. (5 кредити ECTS): лекції – 32 год., лабораторні роботи – 32 год., самостійна робота – 86 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 та планування.

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), підсумковий/семестровий контроль у формі заліку, відповідно до графіку навчального процесу (FAS)

Програма навчальної дисципліни

Теми лекційних занять

Тема 1. Поняття та характеристики аналізу великих даних
Тема 2. Життєвий цикл аналізу великих даних
Тема 3. Попередня обробка і аналіз даних.

Тема 4. Парадигма MapReduce.
Тема 5. Інструменти для великих даних. Hadoop та Apache Spark.
Тема 6. Методи машинного навчання для обробки великий даних

Теми практичних занять

Практичні заняття не передбачені навчальним планом.

Теми лабораторних робіт

Тема 1. Робота з великими файлами та великими даними у MatLab (або з використанням бібліотек Python).
Тема 2. Робота з великими даними у вигляді розкладу.
Тема 3. Аналіз великих даних за допомогою MapReduce.

Самостійна робота

Тема РЗ. Візуалізація великих даних.
Студентам рекомендуються додаткові матеріали (відео, статті) для самостійного вивчення та опрацювання..

Література та навчальні матеріали

Основна література

1. Mathar R. Fundamentals of Big Data Analytics/ Lecture Notes, 2019// https://www.ti.rwth-aachen.de/teaching/BigData/FBDA.pdf
2. Big Data Analytics [R17a0528] Lecture Notes. 2021 //https://mrcet.com/downloads/digital_notes/CSE/IV%20Year/(R17A0528%20)%20Big%20Data%20Analytics%20Digital%20notes.pdf
3. Zgurovsky M.Z., Zaychenko Y.P. Big Data: Conceptual Analysis and Applications. Springer, 2020. – 298p
4. Олійник А. О. Інтелектуальний аналіз даних : Навчальний посібник / А. О. Олійник, О. О. Олійник, С. О. Субботін. – Запоріжжя : ЗНТУ, 2011. – 278 с.
5. Черняк О.І. Інтелектуальний аналіз даних: підручник. - К: Знання, 2014. -599с.

Додаткова література

1 BIG DATA. Зброя математичного знищення. Як великі дані збільшують нерівність. Book Chef, 2019 – 336 с.
2 Ghavami Peter. Big Data Governance: Modern Data Management Principles for Hadoop, NoSQL & Big Data Analytics. CreateSpace Independent Publishing Platform, 2016. — 204 p
3. Martin Hilbert. Big Data for Development: A Review of Promises and Challenges https://doi.org/10.1111/dpr.12142
4. EECS E6893: Big Data Analytics. EECS E6895: Advanced Big Data and AI // https://www.ee.columbia.edu/~cylin/course/bigdata/index.html
5. Setareh Majidian, Iman Raeesi Vanani Literature Review on Big Data Analytics Methods, 2019. DOI:10.5772/intechopen.86843

Система оцінювання

Критерії оцінювання успішності студента та розподіл балів

100% підсумкове оцінювання у вигляді заліку (10%) та поточного оцінювання (90%).
10% залік, відповідно до графіку навчального процесу;
90% поточне оцінювання:
3 лабораторні роботи (по 30%)

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	30.08.2024		Завідувач кафедри Олена НІКУЛІНА
	30.08.2024		Гарант ОП Марина ГРИНЧЕНКО