



Силабус навчальної дисципліни

«Високопродуктивні системи оброблення та аналізу великих даних»

Спеціальність	122 Комп'ютерні науки
Освітня програма	Комп'ютерні науки
Освітній рівень	Другий (магістерський) рівень вищої освіти
Статус дисципліни	Обов'язкова
Мова викладання, навчання та оцінювання	Українська
Курс / семестр	1 курс, 1 семестр
Кількість кредитів ЄКТС	5 кредитів
Розподіл годин за формами освітнього процесу та видами навчальних занять	Лекції – 14 год.
	Практичні (семінарські) – 0 год.
	Лабораторні – 26 год.
	Самостійна робота – 110 год.
Форма семестрового контролю	Іспит
Кафедра	Кафедра інформаційних систем, ауд. 413 головного корпусу, тел. (057) 702-18-31 (дод. 4-37), сайт кафедри: https://kafis.hneu.net/
Викладач	Мінухін Сергій Володимирович, д.т.н., професор;
Контактна інформація викладача	Мінухін С.В. serhii.minukhin@hneu.net
Дні навчальних занять	Лекція: згідно діючого розкладу занять Лабораторні згідно діючого розкладу занять
Консультації	На кафедрі інформаційних систем, очні, заочні відповідно до графіку консультацій, індивідуальні, чат в ПНС

Мета навчальної дисципліни є надання здобувачам вищої освіти системи теоретичних знань і придбання практичних навичок розуміння сутності проблем, які виникають при використанні великих даних, сучасних підходів та інструментів їх оброблення та аналізу.

Структурно-логічна схема вивчення навчальної дисципліни

Пререквізити	Постреквізити
	Курсова робота: Розроблення комп'ютерних інформаційних систем
	Комплексний тренінг
	Дипломна робота

Зміст навчальної дисципліни

Змістовий модуль 1. Основні поняття, сутність та особливості великих даних. Принципи організації побудови систем для роботи з великими даними.

Тема 1. Поняття, характеристики великих даних та системи їх оброблення.

Тема 2. Сучасні системи оброблення великих даних. Склад компонентів та їх призначення.

Тема 3. Apache Hadoop: фреймворк для оброблення великих даних. Базові складові для побудови Hadoop: Google's MapReduce, Google File System.

Тема 4. Архітектура Apache Hadoop.

Змістовий модуль 2. Apache Spark: універсальна платформа для обробки та аналітики великих даних

Тема 5. Архітектура Apache Spark.

Тема 6. Режими розгортання Apache Spark.



Тема 7. Планування виконання завдань в Apache Spark.

Тема 8. Робота з базами та сховищами даних в SparkSQL. RDD, Dataframe і Dataset.

Тема 9. Розгортання та налаштування фреймворків Apache Spark та Apache Hadoop в розподіленому та віртуальному середовищах.

Матеріально-технічне (програмне) забезпечення дисципліни

Засоби дистанційного навчання: Сайт персональних навчальних систем: <https://pns.hneu.edu.ua> Бібліотека: <http://library.hneu.edu.ua> Репозитарій: <http://www.repository.hneu.edu.ua>

Аудиторії університету (м. Харків, пр. Науки 9А).

Мультимедійне обладнання: проектор, ноутбук/комп'ютер, доступ до мережі Інтернет,

ПЗ: Microsoft Windows, Microsoft Office, Vagrant, Virtualbox, ОС Ubuntu, Apache Spark

Форми та методи оцінювання

Університет використовує 100 бальну накопичувальну систему оцінювання результатів навчання здобувачів вищої освіти.

Поточний контроль здійснюється під час проведення лекційних та лабораторних занять і має на меті перевірку рівня підготовленості здобувача вищої освіти до виконання конкретної роботи і оцінюється сумою набраних балів.

Підсумковий контроль включає семестровий контроль, який проводиться у формі екзамену.

Максимально можлива кількість балів за поточний контроль упродовж семестру для дисципліни форма контролю якої екзамен – 60 та мінімально можлива кількість балів – 35.

Більш детальна інформація щодо системи оцінювання та накопичування балів з навчальної дисципліни наведена у робочому плані (технологічній карті) з навчальної дисципліни.

Політики навчальної дисципліни

Викладання навчальної дисципліни ґрунтується на засадах академічної доброчесності. Порушеннями академічної доброчесності вважаються: академічний плагіат, фабрикація, фальсифікація, списування, обман, хабарництво, необ'єктивне оцінювання. За порушення академічної доброчесності здобувачі освіти притягуються до такої академічної відповідальності: повторне проходження оцінювання відповідного виду навчальної роботи.

Більш детальну інформацію щодо компетентностей, результатів навчання, методів навчання, форм та методів оцінювання, самостійної роботи наведено у Робочій програмі навчальної дисципліни.