



Харківський національний економічний університет імені Семена Кузнеця  
**Силабус навчальної дисципліни**

*«Високопродуктивні системи обробки 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	Українська
Курс / семестр	1 курс, 1 семестр
Кількість кредитів ЄКТС	5 кредити
Розподіл за видами занять та годинами навчання	Лекції – 12 год. Лабораторні – 28 год. Самостійна робота – 110 год.
Форма підсумкового контролю	Екзамен
Кафедра	Кафедра інформаційних систем, 702-18-31, <a href="http://www.is.hneu.edu.ua/">http://www.is.hneu.edu.ua/</a>
Викладач (-і)	Мінухін Сергій Володимирович, доктор технічних наук, професор кафедри
Контактна інформація викладача (-ів)	<a href="mailto:serhii.minukhin@hneu.net">serhii.minukhin@hneu.net</a>
Дні занять	Згідно з розкладом занять
Консультації	Дистанційні: вівторок 10-15-11-50.

Метою викладання навчальної дисципліни "Високопродуктивні системи обробки та аналізу великих даних" є формування системи теоретичних знань і придбання практичних умінь і навичок з питань використання технологій розподілених систем для застосування в надпродуктивних обчисленнях на базі технологій розподілених файлових систем та баз даних.

**Передумови для навчання**

*Комп'ютерні мережі, Операційні системи, Об'єктно-орієнтоване програмування.*

**Зміст навчальної дисципліни**

**Змістовий модуль 1. Стандарти, архітектура та принципи побудови високопродуктивних систем на базі моделей сервісів і технологій розподілених систем.**

**Тема 1. Базові архітектури розподілених високопродуктивних систем.**

**Тема 2. Розподілені файлові системи. Принципи організації та функціонування.**

**Тема 3. Розподілена система Apache Hadoop.**

**Тема 4. Принципи організації та архітектура розподіленої системи зберігання даних HDFS.**

**Змістовий модуль 2. Програмна модель та архітектура роботи Apache Spark.**

**Тема 5. Програмна модель розподілених обчислень MapReduce.**

**Тема 6. Організація архітектури фреймворку Apache Spark.**

**Тема 7. Принципи функціонування фреймворку Apache Spark. Основні компоненти та їхнє призначення.**

**Тема 8. Режими розгортання фреймворку Apache Spark.**

**Тема 9. Принципи роботи з базами даних: RDD та dataframes в Apache Spark.**

**Матеріально-технічне (програмне) забезпечення дисципліни**

Наявність безкоштовного ПЗ Virtualbox та Apache Spark.

Сторінка курсу на платформі

<https://pns.hneu.edu.ua/>

Moodle(персональна навчальна система)

**Система оцінювання результатів навчання**

*ХНЕУ ім. С. Кузнеця використовує накопичувальну (100 бальну) систему оцінювання. Підсумкова кількість балів, яку може накопичити здобувач вищої освіти складається з суми балів за поточну та підсумкову успішність. Максимальна кількість балів за поточну успішність – 60. Мінімальна кількість балів – 35 балів. Максимальна кількість балів за підсумкову успішність (екзамен) – 40. Мінімальна кількість балів за підсумкову успішність (екзамен) – 25*

**Накопичування рейтингових балів з навчальної дисципліни**

Види навчальної роботи	Мах кількість балів
Лекції	5
Лабораторні заняття	10
Захист лабораторних робіт	26
Письмові контрольні роботи	19
Екзамен	40
Максимальна кількість балів	100



**Політики навчальної дисципліни**

*Політика щодо виконання завдань пізніше встановленого терміну: враховує особливості певної лабораторної роботи, оцінка може бути знижена у разі не виконання завдань у повному обсязі та внаслідок перевищення строку здачі за робочим планом; зниження оцінки за контрольну роботу можливе у разі невиконання в повній мірі завдань (питань) та перевищення строку її написання відповідно до технологічної карти з дисципліни.*

***Більш детальну інформацію щодо компетентностей, результатів навчання, методів навчання, форм оцінювання, самостійної роботи наведено у робочій програмі навчальної дисципліни***

*Силабус затверджено на засіданні кафедри « 22 » червня 2021 р. Протокол № 13.*