



## Силабус освітнього компонента

Програма навчальної дисципліни

# Основи розподілених та паралельних обчислень



**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**  
Системний аналіз та інформаційно-аналітичні технології (322)

**Рівень освіти**  
Бакалавр

**Тип дисципліни**  
Спеціальна (фахова), Обов'язкова

**Семестр**  
8

**Мова викладання**  
Українська

## Викладачі, розробники



**Любченко Наталія Юріївна**

[Nataliia.Liubchenko@khp.edu.ua](mailto:Nataliia.Liubchenko@khp.edu.ua)

Кандидат технічних наук, доцент, доцент кафедри системного аналізу та інформаційно-аналітичних технологій

Досвід роботи – 20 років. Автор понад 60 наукових та навчально-методичних праць. Провідний лектор з дисциплін: «Основи програмування», «Технології та засоби великих даних», «Семантичний WEB», «Об'єктно-орієнтоване програмування», «Основи розподілених та паралельних обчислень».

[Детальніше про викладача на сайті кафедри](#)

## Загальна інформація

### Анотація

Дисципліна забезпечує теоретичну і практичну підготовку в області паралельних та розподілених обчислень, оволодіння концепціями сучасного програмування в рамках парадигм паралельного та розподіленого програмування. Основу вивчення складають підходи до програмування в багатопотокових системах, розподілених системах, системах паралельних розрахунків, розглядаються проблеми сумісної роботи процесів паралельної програми та їх синхронізації.

### Мета та цілі дисципліни

Вивчення теоретичних основ та практичних аспектів використання паралельних обчислювальних систем для вирішення складних прикладних задач з великим об'ємом обчислень.

### Формат занять

Лекції, лабораторні роботи, самостійна робота, консультації. Індивідуальне завдання - курсова робота. Підсумковий контроль – залік.

## Компетентності

ЗК1. Здатність до абстрактного мислення, аналізу та синтезу.

ЗК2. Здатність застосовувати знання у практичних ситуаціях.

ЗК3. Знання та розуміння предметної області та розуміння професійної діяльності.

ЗК6. Здатність вчитися й оволодівати сучасними знаннями.

ЗК7. Здатність до пошуку, оброблення та аналізу інформації з різних джерел.

ЗК8. Здатність генерувати нові ідеї (креативність).

ЗК11. Здатність приймати обґрунтовані рішення.

ЗК12. Здатність оцінювати та забезпечувати якість виконуваних робіт.

ЗК13. Здатність діяти на основі етичних міркувань.

СК1. Здатність до математичного формулювання та досліджування неперервних та дискретних математичних моделей, обґрунтування вибору методів і підходів для розв'язування теоретичних і прикладних задач у галузі комп'ютерних наук, аналізу та інтерпретування.

СК2. Здатність до виявлення статистичних закономірностей недетермінованих явищ, застосування методів обчислювального інтелекту, зокрема статистичної, нейромережевої та нечіткої обробки даних, методів машинного навчання та генетичного програмування тощо.

СК3. Здатність до логічного мислення, побудови логічних висновків, використання формальних мов і моделей алгоритмічних обчислень, проектування, розроблення й аналізу алгоритмів, оцінювання їх ефективності та складності, розв'язності та нерозв'язності алгоритмічних проблем для адекватного моделювання предметних областей і створення програмних та інформаційних систем.

СК4. Здатність використовувати сучасні методи математичного моделювання об'єктів, процесів і явищ, розробляти моделі й алгоритми чисельного розв'язування задач математичного моделювання, враховувати похибки наближеного чисельного розв'язування професійних задач.

СК5. Здатність здійснювати формалізований опис задач дослідження операцій в організаційно-технічних і соціально-економічних системах різного призначення, визначати їх оптимальні розв'язки, будувати моделі оптимального управління з урахуванням змін економічної ситуації, оптимізувати процеси управління в системах різного призначення та рівня ієрархії.

СК6. Здатність до системного мислення, застосування методології системного аналізу для дослідження складних проблем різної природи, методів формалізації та розв'язування системних задач, що мають суперечливі цілі, невизначеності та ризики.

СК7. Здатність застосовувати теоретичні та практичні основи методології та технології моделювання для дослідження характеристик і поведінки складних об'єктів і систем, проводити обчислювальні експерименти з обробкою й аналізом результатів.

СК8. Здатність проектувати та розробляти програмне забезпечення із застосуванням різних парадигм програмування: узагальненого, об'єктно-орієнтованого, функціонального, логічного, з відповідними моделями, методами й алгоритмами обчислень, структурами даних і механізмами управління.

СК9. Здатність реалізувати багаторівневу обчислювальну модель на основі архітектури клієнт-сервер, включаючи бази даних, знань і сховища даних, виконувати розподілену обробку великих наборів даних на кластерах стандартних серверів для забезпечення обчислювальних потреб користувачів, у тому числі на хмарних сервісах.

СК11. Здатність до інтелектуального аналізу даних на основі методів обчислювального інтелекту включно з великими та погано структурованими даними, їхньої оперативної обробки та візуалізації результатів аналізу в процесі розв'язування прикладних задач.

СК16. Здатність реалізовувати високопродуктивні обчислення на основі хмарних сервісів і технологій, паралельних і розподілених обчислень при розробці й експлуатації розподілених систем паралельної обробки інформації.

## Результати навчання

РН1. Застосовувати знання основних форм і законів абстрактно-логічного мислення, основ методології наукового пізнання, форм і методів вилучення, аналізу, обробки та синтезу інформації в предметній області комп'ютерних наук.

РН4. Використовувати методи обчислювального інтелекту, машинного навчання, нейромережевої та нечіткої обробки даних, генетичного та еволюційного програмування для розв'язання задач розпізнавання, прогнозування, класифікації, ідентифікації об'єктів керування тощо.

PH9. Розробляти програмні моделі предметних середовищ, вибирати парадигму програмування з позицій зручності та якості застосування для реалізації методів та алгоритмів розв'язання задач в галузі комп'ютерних наук.

PH12. Застосовувати методи та алгоритми обчислювального інтелекту та інтелектуального аналізу даних в задачах класифікації, прогнозування, кластерного аналізу, пошуку асоціативних правил з використанням програмних інструментів підтримки багатовимірного аналізу даних на основі технологій DataMining, TextMining, WebMining.

PH16. Виконувати паралельні та розподілені обчислення, застосовувати чисельні методи та алгоритми для паралельних структур, мови паралельного програмування при розробці та експлуатації паралельного та розподіленого програмного забезпечення.

PH17. Розробляти системи штучного інтелекту на основі використання моделей, методів та засобів інженерії даних та знань.

## Обсяг дисципліни

Загальний обсяг дисципліни 120 год. (4 кредитів ECTS): лекції – 20 год., лабораторні роботи – 20 год., самостійна робота – 80 год.

## Передумови вивчення дисципліни (пререквізити)

Для успішного проходження курсу необхідно мати знання та практичні навички з наступних дисциплін: «Бази даних в інформаційних системах», «Штучні нейронні мережі», «Основи штучного інтелекту», «Математичні методи дослідження операцій та прийняття рішень».

## Особливості дисципліни, методи та технології навчання

На лекційних заняттях викладання матеріалу здійснюється в усній формі із записом основних положень лекції у конспект. Для демонстрації презентацій застосовується медіа проектор та комп'ютер. На лабораторних заняттях студенти виконують індивідуальні завдання на комп'ютерах у середовищі IntelliJ IDEA, Eclipse для роботи с проектами Scala, за допомогою інструментів Oracle JDK, Scala Build Tool (SBT), Scala IDE, Apache Spark, Apache Hadoop.

## Програма навчальної дисципліни

### Теми лекційних занять

#### Тема 1. Введення в розподілені та паралельні обчислення.

Поняття про паралельні та розподілені обчислення. Послідовні обчислення. Паралельні обчислення. Засоби для здійснення паралельних обчислень.

#### Тема 2. Введення в розподілені розрахунки. Модель, спеціалізації розподілених розрахунків.

Поняття розподілених обчислень та розподіленої системи. Цілі побудови розподілених систем. Вимоги до розподілених систем.

#### Тема 3. Розуміння великих даних.

Концепції та термінологія. Характеристики великих даних. Проблеми використання великих даних. Визначення еталонної архітектури. Обґрунтування використання архітектурних патернів. Огляд інструментів для роботи з великими даними та їх аналізу.

#### Тема 4. Концепції зберігання та обробки великих даних.

Кластери. Файлові системи та розподілені файлові системи. База даних NoSQL. Шардінг (Sharding). Реплікація. Поєднання шардінга і реплікації. Теорема CAP. Принцип проектування бази даних ACID. Принцип BASE. Паралельна обробка даних. Розподілена обробка даних. Обробка робочих завдань. Кластер. Обробка в пакетному режимі. Обробка в режимі реального часу.

#### Тема 5. Розробка додатків обробки даних, що виконуються в розподіленому обчислювальному середовищі.

Apache Hadoop. Екосистема Hadoop. Архітектурна концепція Hadoop. Демони Hadoop. Методи розгортання Hadoop. Розподілена файлова система Hadoop. Архітектура HDFS. Реплікації даних. Файлові операції HDFS (операція запису, операція читання, операція видалення).

#### Тема 6. Розподілена пакетна і потокова обробки неструктурованих і слабоструктурованих даних.

Фреймворк Apache Spark. Особливості, обмеження, архітектура і принцип роботи компонентів Apache Spark. Робота з даними в Spark – RDD, DataFrame, DataSet. Операції Spark RDD. Spark SQL DataFrame.

**Тема 7. Управління ресурсами кластерів і планування завдань.**

YARN – планувальник ресурсів Hadoop. Архітектура YARN. Ключові компоненти Hadoop YARN.

**Тема 8. Модель програмування для паралельної і розподіленої обробки великих наборів даних.**

Apache Hadoop MapReduce. Стадії MapReduce. Завдання Map и Reduce. Приклади MapReduce

## Теми практичних занять

Практичні роботи в рамках дисципліни не передбачені.

## Теми лабораторних робіт

**Лабораторна робота №1.**

Знайомство з основними сервісами Apache Hadoop.

**Лабораторна робота №2.**

Рішення задачі запису даних в HDFS. Особливості управління ресурсами в режимі Hadoop кластера з YARN.

**Лабораторна робота №3.**

Знайомство з Hadoop MapReduce.

**Лабораторна робота №4.**

MapReduce, додаткові функції. Ланцюжки MapReduce завдань.

**Лабораторна робота №5.**

Мова програмування Scala. Ознайомлення з синтаксисом і основними можливостями мови Scala.

**Лабораторна робота №6.**

Знайомство з Apache Spark. Архітектура Spark. Варіанти запуску Spark- додатків.

**Лабораторна робота №7.**

Конфігурація і системні утиліти Apache Spark, взаємодія з класичними та розподіленими файловими системами.

**Лабораторна робота №8.**

Використання різних форматів зберігання даних і використання Spark RDD, dataframes і datasets.

**Лабораторна робота №9.**

Machine Learning в Apache Spark.

**Лабораторна робота №10.**

Streaming в Apache Spark.

## Самостійна робота

Опрацювання лекційного матеріалу. Підготовка до лабораторних занять та модульних контролів. Самостійне вивчення тем та питань, які не викладаються на лекційних заняттях. Виконання, оформлення звітних матеріалів та захист курсової роботи.

## Література та навчальні матеріали

### Основна література

1. Минайленко Р.М. Паралельні та розподілені обчислення : навч. посіб. — Кропивницький : Видавець Лисенко В. Ф., 2021. — 153 с.

<https://dspace.kntu.kr.ua/server/api/core/bitstreams/396e02d2-725b-47b5-a1c0-ae07a9bec326/content>

2. Олещенко Л.М. Технології оброблення великих даних: навч. посіб. / Л.М. Олещенко; КПІ ім. Ігоря Сікорського. –Київ: КПІ ім. Ігоря Сікорського, 2021. – 227 с.

[https://ela.kpi.ua/bitstream/123456789/42206/1/%D0%9AonspLekts Tekhnolohii-obroblennia-velykykh-danykh\\_%D0%9Eleshchenko.pdf](https://ela.kpi.ua/bitstream/123456789/42206/1/%D0%9AonspLekts%20Tekhnolohii-obroblennia-velykykh-danykh_%D0%9Eleshchenko.pdf)

3. Jules S. Damji. Learning Spark. Lightning-Fast Data Analytics. Jules S. Damji, Brooke Wenig, Tathagata Das, Denny Lee / Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472. - 2020. - 399 p.

<https://pages.databricks.com/rs/094-YMS-629/images/LearningSpark2.0.pdf>

4. Spark by Robert D. Schneider and Jeff Karmiol / Published by John Wiley & Sons, Inc. 111 River St. Hoboken, NJ 07030-5774. – 2019. – 80 p. <https://www.ibm.com/downloads/cas/WEB4XBOR>

### Додаткова література

5. Любченко Н. Ю. Основи мови програмування SCALA [Електронний ресурс] : навч.-метод. посібник / Н. Ю. Любченко, А. О. Подорожняк, О. П. Черних ; Нац. техн. ун-т "Харків. політехн. ін-т". – Електрон. текст. дані. – Харків, 2023. – 148 с. <https://repository.kpi.kharkov.ua/items/3f5ee53b-258b-472a-ba35-3fdc412bfbed>

6. Проект Apache Hadoop [Електронний ресурс]. – Режим доступу: <https://hadoop.apache.org/>

7. Unified engine for large-scale data analytics [Електронний ресурс]. – Режим доступу: <https://spark.apache.org/>

8. Корочкін О.В. Паралельні та розподілені обчислення. Вибрані розділи: Навч. посібник. / О.В. Корочкін, Русанова О.В. – Київ : КПІ ім. Ігоря Сікорського, 2020. – 123 с <https://ela.kpi.ua/bitstream/123456789/48224/1/Paralelni%20.pdf>

9. Big Data [Електронний ресурс]. Режим доступу: <https://www.it.ua/knowledge-base/technology-innovation/big-data-bolshie-dannye>

## Система оцінювання

### Критерії оцінювання успішності студента та розподіл балів

Оцінювання проводиться за 100-бальною шкалою. Бали нараховуються за наступним співвідношенням:  
модульні тести - 20 балів; лабораторні роботи - 40 балів, курсова робота – 20 балів; залік - 20 балів

### 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

29.08.2023

Завідувач кафедри  
Юрій ДОРОФЄЄВ

29.08.2023

Гарант ОП  
Марина ГРИНЧЕНКО