

# Силлабус навчальної дисципліни

Назва дисципліни (інформація з ЦІСТ)

## Аналіз великих даних на Scala та Spark

Факультет

КОМП'ЮТЕРНИХ НАУК ▾

Рівень вищої освіти

Перший (бакалавр) ▾

Код і назва спеціальності

F3 Комп'ютерні науки

Тип і назва освітньої програми

ОПП «Штучний інтелект»

Кількість кредитів ЄКТС

5

Структура дисципліни (розподіл за видами та годинами навчання)

Лекції - 30 годин, практичні заняття - 30 годин, консультації - 10 годин, самостійна робота - 80 години.

Семестровий контроль - залік.

Графік та терміни вивчення дисципліни

4-й рік, 8-й семестр

Передумови для навчання за дисципліною

Раніше мають бути вивчені дисципліни: Функціональний дизайн програм на Scala; Паралельне програмування на Scala.

## Анотація та зміст дисципліни

Метою курсу є ознайомлення студентів з можливостями використання мови Scala для обробки великих даних за допомогою фреймворку Spark та як парадигма паралельних даних може бути розширена до розподіленого випадку з використанням Spark.

Курс складається з:

Тема 1. Введення до предмету

Введення, логістика, те, що ви дізнаєтеся. Від паралельних даних до розподілених паралельних даних. Латентність паралельних даних.

Тема 2. Основи роботи з RDD

Why We Need Apache Spark. RDD та розподілені колекції Spark. Трансформації та дії. Обчислення в Spark: відмінність від Scala колекцій.

Тема 3. Операції на парних RDD

Операції зведення. Розподілені пари ключ-значення (парні RDD). Перетворення та дії на парних RDD. Зрощення

Тема 4. Додаткові можливості оптимізації

Перемішування. Перерозподіл. Оптимізація з розподілювачами. Широкі та вузькі залежності.

Тема 5. Інструменти вирішення складних задач

Spark SQL. DataFrames частина. Datasets.

## Компетентності, знання, вміння, розуміння, які набуде здобувач вищої освіти в процесі навчання

ЗК 1 - Здатність до абстрактного мислення, аналізу та синтезу.

ЗК 2 - Здатність застосовувати знання у практичних ситуаціях.

ЗК 3 - Знання та розуміння предметної області та розуміння професійної діяльності.

ЗК 5 - Здатність спілкуватися іноземною мовою.

ЗК 6 - Здатність вчитися та оволодівати сучасними знаннями.

ЗК 7 - Здатність до пошуку, оброблення та аналізу інформації з різних джерел.

ЗК 9 - Здатність працювати в команді.

ЗК 10 - Здатність бути критичним і самокритичним.

ЗК 11 - Здатність приймати обґрунтовані рішення.

ЗК 12 - Здатність оцінювати та забезпечувати якість виконуваних робіт.

ЗК 13 - Здатність діяти на основі етичних міркувань.

ФК 3 - 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.

ФК 8 - 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.

ФК 9 - Здатність реалізувати багаторівневу обчислювальну модель на основі архітектури клієнт-сервер, включаючи бази даних, знань сховища даних, виконувати розподілену обробку великих наборів даних на кластерах стандартних серверів для забезпечення обчислювальних потреб користувачів, у тому числі на хмарних сервісах.

ФК 16 - 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 здобувача вищої освіти

ПРН 17 - 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.

#### Система оцінювання відповідно до кожного завдання для складання заліку/екзамену

Оцінювання складається з відвідування занять - 25 балів, 3 завдань до практичних робіт по 20 балів - 60 балів, 3 есе по результатах виконаних робіт по 5 бали - 15 балів.

Підсумкова рейтингова оцінка за семестр обчислюється у 100-бальній системі.

#### Якість освітнього процесу

Здійснюється шляхом застосування пункту 5 "Система забезпечення якості освітнього процесу" [Положення про організацію освітнього процесу в ХНУРЕ](#).

Серед засобів оцінки використовуються анкети опитування студентів про зміст та якість навчання по завершенню вивчення курсу.

## Методичне забезпечення

Комплекс навчально-методичного забезпечення навчальної дисципліни «Аналіз великих даних на Scala та Spark» для студентів усіх форм навчання за спеціальністю 122 Комп'ютерні науки, ОПП «Штучний інтелект» [Електронний ресурс] / ХНУРЕ; розроб. Гребенюк В.О. - Х. ХНУРЕ, 2018.- 22 с.

## Розробник силлабусу (посада, ПІБ, ел. пошта)

Старший викладач кафедри Штучного Інтелекту,  
В'ячеслав ГРЕБЕНЮК, vyacheslav.grebenyuk@nure.ua