

Алгоритми та методи обробки великих масивів даних

Анотація. Дисципліна «Алгоритми та методи обробки великих масивів даних» належить до переліку дисциплін вільного вибору аспіранта. Вона забезпечує особистісний та професійний розвиток аспіранта і спрямована на формування висококваліфікованого фахівця вищої школи, здатного застосовувати алгоритми обробки даних при проведенні наукових досліджень, коректно інтерпретувати інформаційні потоки, аналізувати дані шляхом застосування методів математичної статистики, штучних нейронних мереж, теорії нечітких множин та еволюційних алгоритмів, що в подальшому сприятиме їх успішному використанню у професійній діяльності. Особлива увага приділяється практичному застосуванню програмних моделей обробки даних.

Кількість кредитів: 4

Викладач: Самохвалов Юрій Якович, д.т.н., професор кафедри кібербезпеки та захисту інформації факультету інформаційних технологій.

Мета навчальної дисципліни: дати сучасні знання про проведення глибоких досліджень даних з метою виявлення залежностей, аномалій, відповідності умовам аналізу, для того щоб отримати його якісну модель; поглибити знання про технології обробки даних, що сприятиме отриманню якісних наукових результатів завдяки обробці різномірної інформації.

Попередні вимоги:

Аспірант повинен знати: технології зберігання та алгоритми обробки великих масивів даних, математичні інструменти їх аналізу, моделі організації паралельних обчислень в залежності від архітектурних рішень, механізми автоматичного керування обчисленнями.

Аспірант повинен вміти: застосовувати алгоритми аналізу великих масивів даних для побудови аналітичних і статистичних моделей для виявлення прихованих закономірностей або взаємозв'язків між змінними у великих масивах необроблених даних, отримувати на основі комплексного аналізу даних якісні нові знання, робити ефективний прогноз та приймати рішення.

Змістові модулі:

- Обробка експоненціально зростаючих масивів даних
- Методи аналізу і обробки даних
- Підтримка розподілених баз даних, проміжні комунікації даних
- Моделювання та організація паралельних обчислень
- Моделі програмування і платформи для обробки великих масивів даних, процес виконання обчислень
- Розбиття та розподіл даних між процесорами, планування та балансування навантаження, обробка відмов, збір та агрегація проміжних результатів
- Розподіл задач обробки та зберігання даних
- Оцінка продуктивності обробки даних, обробка відмов та моніторинг
- Підтримка сховищ даних
- Глибинний аналіз даних, машинне навчання
- Задачі класифікації, моделювання та прогнозування

Мова викладання: українська.

Місце у структурно-логічній схемі: ДВА.3.02.15 читається на другому році навчання.

Термін вивчення: дисципліна вивчається на 2 році навчання за освітньо-науковим рівнем «доктор філософії» в обсязі 120 годин, у тому числі 24 годин аудиторних занять (18 год. – лекційні заняття, 4 год. – практичні заняття, 2 години – консультація), 96 годин самостійної роботи.