

АНОТАЦІЯ ДИСЦИПЛІНИ

Назва дисципліни	Основи Big Data
	3 кредити / 90 годин
Вид підсумкового контролю	Залік
Викладач	Батишкіна Юлія Валеріївна
Профайл викладача (ів) на сайті кафедри	http://www.iktmvi.rv.ua/pro-kafedru/teachers/teacher/shlihta-ganna-oleksandrivna.html
Е-mail викладача:	kiktmvi@rshu.edu.ua
Посилання на освітній контент дисципліни в CMS Moodle (за наявності) або на іншому ресурсі	http://do.iktmvi.rv.ua/course/view.php?id=160
Мова викладання	українська
Консультації	<i>Очні консультації:</i> щопонеділка, з 12.45 до 14.05 (2 академічні години) <i>Он лайн- консультації:</i> щовівторка, з 14.00 до 16.00; щосереди з 14.00 до 16.00

Мета курсу – оволодіння компетентностями здатність забезпечити організацію обчислювальних процесів витягнення та інтерпретації прихованих знань з використанням пакетів прикладних програм різних типів (tensor toolbox, Hadoop, Map reduce та ін.), орієнтованих на середовище MatLab; здатність до розробки інтерфейсного програмного забезпечення для поєднання джерела великих даних (ВД) з відповідним програмним середовищем; здатність реалізовувати обчислювальний процес на основі хмарних сервісів і технологій, паралельних і розподілених обчислень. набуття знань, практичних навичок та умінь використання сучасних методів, алгоритмів обробки високо вимірних, велико об'ємних, швидко змінюваних різноформатних даних, в т. ч. спотворених зовнішніми неконтрольованими завадами з урахуванням умов неповної визначеності, для витягнення прихованих знань, необхідних для підтримки прийняття рішень;

Очікувані результати навчання

знати: характеристики нового об'єкта інформаційних технологій, відомого як ВД, розуміти його роль та місце в системах автоматизованого проектування в легкій промисловості, основні методи, моделі, програмно-технічні засоби, що дозволяють реалізувати обробку багато

вимірних, швидко-змінюваних та високо-об'ємних масивів даних;

вміти: застосувати набуті знання в професійній діяльності під час розробки, налагодження та експлуатації інформаційних систем з використанням штучного інтелекту та за нестандартних умов проведення дослідження;

володіти: навичками застосування математичних моделей представлення ВД, використання сучасних програмних засобів обробки ВД, інтерпретації отриманих результатів аналізу, пошуку прихованих знань, розумінням шляхів розвитку інтелектуальних технологій при дослідженні складних об'єктів, систем, обладнання та технологічних процесів в промисловості;

здатен продемонструвати: методи обробки ВД на прикладах представлення високо-вимірних (3D) об'єктів моделями низької (1D-2D) вимірності, реальну підтримку прийняття рішень при наявності високо об'ємних масивів даних.

Передумови

Необхідні навчальні компоненти (пререквізити, кореквізити і постреквізити): вища математика, алгоритмізація і програмування, фундаментальні принципи розробки програмного забезпечення, основні системні принципи та інтелектуальні системи .

Програма курсу

Зміст дисципліни: Тема 1. Аналітика Великих Даних. Вступ в аналітику Великих Даних (ВД). Основні поняття і визначення. Історія питання. Джерела ВД. Технології формування і обробки ВД. Застосування ВД: в економіці, бізнесі, охороні здоров'я, промисловості, приклади використання. Великі дані в наукових областях, Особливості застосування. Big Data, Data Mining. Datafication. МатЛаб як потужний і головний засіб обробки ВД і прийняття рішень за умов невизначеності в середовищі ВД.

Тема 2. Управління даними, життєвий цикл ВД. Фази життєвого циклу даних, умовність циклу, принципи формування. Управління даними, 7 головних міркувань у створенні гібридної стратегії інтеграції.

Тема 3. Проблеми уявлення і моделювання знань і їхній зв'язок з проблемою ВД. Комп'ютерні знання – база даних + база правил. Концептуальні та емпіричні моделі знання. Data Mining - дослідження і виявлення «машиною» в сирих даних прихованих знань. «Штучний інтелект» як синтез "data mining", "knowledge discovery" і "hidden knowledge". Методи вилучення нових знань з баз фактів. Матричні декомпозиції та їхня роль в аналізі даних.

Тема 4. Багатовимірні масиви, тензорні моделі. Традиційна DataScience - Big Data - загальні риси та відмінності. Тенденція сучасності – від 2D матриць до 3D тензорів. Доцільність і необхідність машинного навчання при опрацюванні результатів обробки ВД. База головного понятійного апарат для аналізу ВД: методи чисельного аналізу на підставі тензорних представлень даних; новітні методи і моделі обробки великих даних, заснованих на тензорних мережах та тензорних декомпозиціях. Тензор – головний об'єкт аналізу багатовимірних великих даних. Матриці, матричні розкладення.

Тема 5. Тензорна алгебра, застосування тензорів. Коротка історія сучасного тензорного аналізу. Тензори, основа, нотація, мотивація. Волокна та слайси – головні форми представлення багатовимірних тензорів. Узагальнення математичних операцій в тензорній та матричній алгебра. Ранг матриці і ранг тензора. Тензорні декомпозиції, високо порядкові декомпозиції. Оператори векторизації та реорганізації. Перетворення тензорів у вектори. Розгортання тензорів засобами МатЛаб. Суми rank-1 матриць та rank-1 тензорів. Розгортання 3D тензорів по одному з вимірів.

Тема 6. Інструментальні та програмні засоби роботи з великими даними. Основні принципи роботи з даними; інструменти; приклади розв'язання практичних завдань. Машинне

навчання. Принципи роботи з великими даними, парадигма MapReduce, Hadoop і корпоративні системи, Hadoop і СУРБД. Сучасні головні засоби MatLab, орієнтовані на роботу з ВД: Parallel Computing Toolbox, Database Toolbox, MATLAB Distributed Computing Server. Пакети прикладних програм (Tensor Matlab Toolbox) для роботи з 3D тензорами..

Політика дисципліни

При організації освітнього процесу здобувачі вищої освіти, викладачі, методисти та адміністрація діють відповідно до: Положення про організацію освітнього процесу у РДГУ, Положення про академічну доброчесність, Положення про оцінювання знань і умінь здобувачів вищої освіти, Положення про практики, Положення про внутрішнє забезпечення якості освіти.

Здобувачам вищої освіти необхідно зареєструватись в системі CMS MOODLE, отримавши кодове слово, де розміщені опорні конспекти лекцій, завдання та методичні вказівки до виконання лабораторних робіт, завдання для самостійної роботи та тести.

Присутність на заняттях не обов'язкова для студентів, які навчаються за дуальною формою навчання, офіційно працевлаштовані і мають дозвіл від деканату на вільне відвідування занять.

Дозволяється користування мобільним телефоном, планшетом чи іншими мобільними пристроями під час занять, окрім тих, де проводиться тестування.

Несвоєчасне виконання завдань самостійної роботи та ІНДЗ не вітається: виставляються нульові бали без права перездачі.

Політика доброчесності

Здобувач вищої освіти виконуючи самостійну роботу повинен дотримуватись політики доброчесності, робити посилання на джерела, звідки взято матеріал. У разі наявності плагіату в будь-яких видах робіт здобувача вищої освіти він отримує незадовільну оцінку і повинен повторно виконати завдання.