

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Петренко Вячеслав Иванович

Должность: и.о. директора Института цифрового развития

Дата подписания: 06.06.2024 09:42:20

Уникальный программный ключ:

b47f96eea65a18e179f94dbf7fb0b10000e7ad4d

Аннотация к рабочей программе дисциплины

Наименование дисциплины	Анализ больших данных
Содержание	<i>Методы многомерного статистического анализа и анализа нечисловой информации.</i> Факторный анализ. Дискриминантный анализ. Кластерный анализ. Кластерный анализ. Многомерное шкалирование. Методы контроля качества. <i>Технологии анализа данных.</i> Большие данные (Big Data): современные подходы к обработке и хранению. Процесс анализа. Общая схема анализа. Извлечение и визуализация данных. Этапы моделирования. Процесс построения моделей. Формы представления данных, типы и виды данных. Представления наборов данных. Технологии KDD и Data Mining. Подготовка данных к анализу. Методика извлечения знаний. Data Mining. Мультидисциплинарный характер Data Mining. Причины распространения KDD и Data Mining. Актуальность технологий Data Mining как средств обработки больших объемов информации. Программное обеспечение в области анализа данных. Аналитические платформы: классификация и особенности применения. Языки визуального моделирования. <i>Интеллектуальный анализ данных.</i> Ассоциативные правила. Аффинитивный анализ, предметный набор. Поддержка и достоверность ассоциативного правила. Значимость ассоциативных правил, лифт и леввередж. Поиск ассоциативных правил. Частые предметные наборы и их обнаружение. Алгоритм генерации ассоциативных правил. Иерархические ассоциативные правила. Методы поиска иерархических ассоциативных правил. Определение кластеризации. Постановка задачи кластеризации. Цели кластеризации в Data Mining. Примеры кластеризации в различных областях. Виды метрик. Шаги алгоритма. Меры расстояний. Пример работы алгоритма k-means. Проблемы алгоритмов кластеризации. <i>Программирование обработки и загрузки больших данных.</i> Системы управления потоками данных. Системы хранения Больших данных. Платформы Больших данных. Обработка данных в реальном времени. Системы управления. Большими данными. Аналитические платформы <i>Аналитика в больших данных.</i> Аналитика Big Data — реалии и перспективы в России и мире. Технологии и методы анализа, которые используются для анализа Big Data. Big data: применение и возможности. Решения на основе Big data. Рынок Big data в России. Big data в банках. Big data в бизнесе. Big data в маркетинге.
Результаты освоения дисциплины (модуля)	Имеются знания теоретических основ функционирования интеллектуальных моделей и особенности их обучения и оптимизации Применяет знания функционирования интеллектуальных моделей и особенности их обучения и оптимизации Имеются знания об интеллектуальных методах Анализирует и применяет интеллектуальные методы согласно решаемой задаче Демонстрирует знания инструментальных средства визуализации данных, в том числе больших данных Сформированы умения и навыки применения инструментальных средств визуализации данных, в том числе больших данных Имеются знания, настраиваемых параметров и гиперпараметров интеллектуальных моделей для достижения высокой эффективности обучения и функционирования Демонстрирует настройку параметров и гиперпараметров интеллектуальных моделей для достижения высокой эффективности обучения и функционирования

	Имеются знания об инструментальных средствах интеллектуальной обработки данных различной природы (текстовые, изображения, видео, аудио) Применяет инструментальные средства интеллектуальной обработки данных различной природы (текстовые, изображения, видео, аудио) Имеются знания в области архитектуры систем управления базами данных, типологии баз данных, методов оптимизации баз данных Сформированы умения и навыки в области архитектуры систем управления базами данных, типологии баз данных, методов оптимизации баз данных Имеются знания об инструментальных средствах для проектирования баз данных различного типа в том числе распределенных Использует инструментальных средств для проектирования баз данных различного типа, в том числе распределенных Имеются знания о методиках оптимизации систем хранения данных на основе интерпретируемых критериев эффективности Сформированы умения и навыки владения методикой оптимизации систем хранения данных на основе интерпретируемых критериев эффективности Демонстрирует знания методов анализа предметной области в рамках решения профессиональных задач Имеются знания о методах, технологиях и средствах проектирования архитектуры систем обработки данных на основании типов данных, конфигурации потоков данных Применяет методы анализа предметной области в рамках решения профессиональных задач Сформированы умения и навыки проектирования архитектуры систем обработки данных на основании типов данных, конфигурации потоков данных
Трудоемкость, з. е.	3 з. е.
Формы отчетности (в т.ч. по семестрам)	Зачет с оценкой 6 семестр
Перечень основной и дополнительной литературы, необходимой для освоения дисциплины	
Основная литература	1. Келлехер, Д. Наука о данных: базовый курс : учебное пособие : [16+] / Д. Келлехер, Б. Тирни ; науч. ред. З. Мамедьяров ; пер. с англ. М. Белоголовского. – Москва : Альпина Паблишер, 2020. – 224 с. 2. Анализ данных : учебник для вузов / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2022. — 490 с. ЭБС
Дополнительная литература	1. Аврунев, О. Е. Модели баз данных : учебное пособие / О. Е. Аврунев, В. М. Стасышин. – Новосибирск : Новосибирский государственный технический университет, 2018. – 124 с. ЭБС 2. Бутаков, Н. А. Обработка больших данных с Apache Spark : учебно-методическое пособие : [16+] / Н. А. Бутаков, М. В. Петров, Д. Насонов. – Санкт-Петербург : Университет ИТМО, 2019. – 52 с. ЭБС