

АННОТАЦИЯ РАБОЧЕЙ ПРОГРАММЫ ДИСЦИПЛИНЫ Б1В.ДВ.04.01 «Визуализация данных»

Автор: Щербинин Павел Алексеевич, технический Директор проектов Mail group

Код и наименование направления подготовки, профиля: 38.04.02 Менеджмент
«Прорывные финансовые и цифровые технологии»

Квалификация (степень) выпускника: Магистр

Форма обучения: очная

Цель освоения дисциплины:

Формирование у студентов необходимой теоретической базы и практических навыков, которые позволят всесторонне и системно понимать современные проблемы прикладной математики и информатики, проблемы обработки и анализа информации, а также разрабатывать и анализировать концептуальные и теоретические модели при решении научных и прикладных задач в области информационных технологий. Одна из главных проблем современной обработки и анализа данных - рост объемов данных, поэтому вопросам обработки большого объема данных посвящена данная дисциплина. Главная задача курса - сформировать целостное представление о современных проблемах анализа и обработки больших данных, помочь овладеть опытом разработки и анализа концептуальных и теоретических моделей прикладных задач анализа больших данных с применением моделей Data Mining.

План курса:

Тема 1. Технологии анализа данных

1. Большие данные (Big Data): современные подходы к обработке и хранению. Проблема множественного сравнения данных.
2. Процесс анализа. Общая схема анализа. Извлечение и визуализация данных. Этапы моделирования. Процесс построения моделей. Формы представления данных, типы и виды данных. Представления наборов данных.
3. 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.
4. Программное обеспечение в области анализа данных. Аналитические платформы: классификация и особенности применения. Языки визуального моделирования.
5. Начало работы. Понятие сценария и узла обработки. Консолидация данных. Трансформация данных. Визуализация данных.

6. Проработка необходимых вопросов для подготовки к лекциям и практическим занятиям.

Тема 2. Интеллектуальный анализ данных

1. Ассоциативные правила. Аффинитивный анализ, предметный набор. Поддержка и достоверность ассоциативного правила. Значимость ассоциативных правил, лифт и левередж. Поиск ассоциативных правил. Частые предметные наборы и их обнаружение. Алгоритм генерации ассоциативных правил. Иерархические ассоциативные правила. Методы поиска иерархических ассоциативных правил.
2. 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.
3. Применение классификации и регрессии. Обзор методов классификации и регрессии. Статистические методы. Методы, основанные на обучении, разнообразие подходов.
4. Основные понятия теории нейронных сетей. Основные парадигмы нейронных сетей. Многослойный персептрон: класс решаемых задач, архитектура.
5. Определение дерева решений. Причины популярности и условия применимости. Структура дерева решений. Выбор атрибута разбиения в узле. Алгоритм ID3, критерий выбора атрибута разбиения ID3, пример работы алгоритма. Проблема переобучения, Неизвестные значения атрибутов, алгоритм C4.5.
6. Ассоциативные правила. Поиск ассоциативных правил.
7. Кластеризация. Алгоритм кластеризации k-means.
8. Прогнозирование с помощью линейной регрессии.
9. Классификация с помощью нейросети.
10. Классификация с помощью деревьев решений.
11. Проработка необходимых вопросов для подготовки к лекциям и практическим занятиям.

В результате освоения дисциплины у студентов должны быть сформированы:

Код компетенции	Наименование компетенции	Код этапа освоения компетенции	Наименование этапа освоения компетенции
ПК-4	способность управлять организациями, подразделениями, группами (командами) сотрудников, проектными сетями	ПК-4.1	способен управлять организациями, подразделениями, группами (командами) сотрудников, проектными сетями
ПК-7	способность представлять результаты проведенного исследования в виде научного отчета, статьи или доклада	ПК-7.1	способен представлять результаты проведенного исследования в виде научного отчета, статьи или доклада
ПК ОС-13	способность принимать	ПК ОС-13.1	способен принимать

	управленческие решения на основе анализа структурированных и неструктурированных данных		управленческие решения на основе анализа структурированных и неструктурированных данных
ОК-1	способность к абстрактному мышлению, анализу и синтезу	ОК-1.1	обладает абстрактным мышлением, анализом и синтезом

Формы и методы текущего контроля и промежуточной аттестации:

В ходе реализации дисциплины используются следующие методы текущего контроля и успеваемости обучающихся:

Опрос, тест, диспут, итоговый зачет с оценкой

Основная литература:

1. Браэм Г., Психология цвета, АСТ/Астрель, 2009. 2. Уилсон Р.А., Квантовая психология. Управление сознанием, София, 2014. 3. Джуан С., Странности нашего мозга, Рипол Классик, 2011.
2. Боровков, А.А. Математическая статистика: учебник, СПб.: Лань, 2010, URL <https://e.lanbook.com/book/3810>
3. Жуковский, О.И. Информационные технологии и анализ данных : учебное пособие, Томск : Эль Контент, 2014, URL <http://biblioclub.ru/index.php?page=book&id=480500>
4. Мхитарян В.С. - Отв. ред. АНАЛИЗ ДАННЫХ. Учебник для академического бакалавриата: Гриф УМО ВО, URL <https://biblio-online.ru/book/CC38E97A-CCE5-4470-90F1-3B6D35ACC0B4>
5. Edward R. Tufte “The Visual Display of Quantitative Information”
6. Cole Nussbaumer Knaflie “Storytelling With Data: A Data Visualization Guide for Business Professionals”
7. Kieran Healy “Data Visualization – A Practical Introduction”
8. Scott Berinato “Good Charts: The HBR Guide to Making Smarter, More Persuasive Data Visualizations”
9. Steven Heller and Rick Landers “Infographics Designers’ Sketchbooks”
10. Stephen Few “Information Dashboard Design: Displaying Data for At-a-glance Monitoring”
11. David McCandless “Knowledge is Beautiful”
12. Manuel Lima “The Book of Circles: Visualizing Spheres of Knowledge”
13. Julie Steele, Noah Iliinsky “Beautiful Visualization, Looking at Data Through the Eyes of Experts”