

АННОТАЦИЯ РАБОЧЕЙ ПРОГРАММЫ ДИСЦИПЛИНЫ

АНАЛИЗ ДАННЫХ

(НАИМЕНОВАНИЕ ДИСЦИПЛИНЫ)

38.03.05 «Бизнес-информатика»

1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ: формирование у студентов теоретических знаний, практических умений и навыков по применению современных методов аналитической обработки электронных массивов данных в различных сферах человеческой деятельности.

Задачи дисциплины:

- изучение существующих технологий подготовки данных к анализу,
- изучение основных методов поиска в данных внутренних закономерностей, взаимосвязей, тенденций,
- овладение практическими умениями и навыками реализации технологий аналитической обработки данных, формирования и проверки гипотез о их природе и структуре, варьирования применяемыми моделями.
- формирование умений и навыков применения универсальных программных пакетов и аналитических платформ для анализа данных.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП ВО:

Дисциплина «Анализ данных» относится к базовой части учебного плана ОПОП.

3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Процесс изучения дисциплины направлен на формирование следующей компетенции:

ПК-18: способность использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования

Знать:

- проблемы и направления научных изысканий в области аналитической обработки данных;
- тенденции разработки универсальных программных средств и аналитических платформ, предназначенных для аналитической обработки данных, построения прогнозов и аналитических сценариев,
- основные методы консолидации, трансформации, визуализации, оценки качества, очистки и переработки данных,
- принципы построения и структурную организацию хранилища данных,
- алгоритмы поиска ассоциативных правил и кластерного анализа,
- статистические и машинные методы классификации и регрессии,
- методики анализа и прогнозирования временных рядов,
- технологию построения ансамблей и сравнения моделей,
- возможности отечественных и зарубежных универсальных программных средств и аналитических платформ, применяемых для анализа данных,
- проблемные вопросы внедрения аналитических программных продуктов и технологий в профессиональную деятельность организаций и учреждений.

Уметь: практически применять методы консолидации, трансформации, визуализации, оценки качества, очистки и переработки данных для качественной подготовке данных к анализу,

- применять технологии анализа электронных массивов данных для решения конкретных практических проблем,

- использовать возможности отечественных и зарубежных универсальных программных средств и аналитических платформ для аналитической обработки данных, построения прогнозов и аналитических сценариев,

- свободно ориентироваться на современном динамичном рынке аналитических программных продуктов.

Владеть: технологиями аналитической обработки электронных массивов данных в целях решения практических проблем предметной области.

- навыками выбора и применения отечественных и зарубежных аналитических платформ, используемых для анализа табулированных массивов электронных данных.

4. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Разделы:

1. Аффинитивный анализ. Поиск последовательных шаблонов. Введение в аффинитивный анализ. Алгоритм априори. Иерархические ассоциативные правила.

2. Кластерный анализ. Введение в кластеризацию. Классификация методов кластеризации. Алгоритм кластеризации k-means. Сети Кохонена. Карты Кохонена. Проблемы алгоритмов кластеризации.

3. Классификация и регрессия. Статистические методы. Введение в классификацию и регрессию. Простая линейная регрессия. Оценка соответствия простой линейной регрессии реальным данным. Простая регрессивная модель. Множественная линейная регрессия. Модель множественной линейной регрессии. Регрессия с категориальными входными переменными. Методы отбора переменных в регрессионные модели. Ограничения применимости регрессионных моделей. Основы логической регрессии. Интерпретации модели логической регрессии.

4. Классификация и регрессия. Машинное обучение. Введение в деревья решений. Алгоритмы построения деревьев решений. Алгоритмы ID3, C4.5., CART. Упрощение деревьев решений. Введение в нейронные сети. Искусственный нейрон. Принципы построения нейронных сетей. Алгоритмы обучения нейронных сетей. Алгоритм обратного распространения ошибки.

5. Анализ и прогнозирование временных рядов. Временной ряд и его компоненты. Модели прогнозирования. Прогнозирование в торговле и логистике.

6. Ансамбли моделей. Введение в ансамбли моделей. Бэггинг. Бустинг. Альтернативные методы построения ансамблей.

Составитель

зав. кафедрой БИЭ Тесленко И.Б.

должность, ФИО, подпись

Заведующий кафедрой:

Бизнес-информатика и экономика

название кафедры

Тесленко И.Б.

ФИО, подпись

Директор Института экономики и менеджмента

Захаров

П.Н.

ФИО, подпись

Дата:

Печать института

