



АННОТАЦИЯ К РАБОЧЕЙ ПРОГРАММЕ ДИСЦИПЛИНЫ

09.03.01 Информатика и вычислительная техника

Профиль: Автоматизированные системы обработки информации и управления

Б1.В.ДВ.04.01 «Методы анализа данных»

| Индекс | Наименование | Семестр 5 | | | | | | | | | | Итого за курс | | | | | | | | | |
|---------------|-----------------------|-----------|---------------------|-----------|-----|-----|----|-----|----|----------|------|---------------|---------------------|-----------|-----|-----|----|-----|----|----------|------|
| | | Контроль | Академических часов | | | | | | | | з.е. | Контроль | Академических часов | | | | | | | | з.е. |
| | | | Всего | Кон такт. | Лек | Лаб | Пр | КРП | СР | Контроль | | | Всего | Кон такт. | Лек | Лаб | Пр | КРП | СР | Контроль | |
| Б1.В.ДВ.04.01 | Методы анализа данных | Экз РГР | 180 | 66 | 34 | 16 | 16 | | 78 | 36 | 5 | Экз РГР | 180 | 66 | 34 | 16 | 16 | | 78 | 36 | 5 |

Формируемые компетенции: УК-1, УК-2, ПК-1, ПК-2.

Содержание дисциплины

Лекционные занятия 17 шт. по 2 часа:

1.1. Работа с данными: Понятие данных. Измерения. Типы шкал измерений. Дискретные и непрерывные данные. Этапы решения задачи анализа данных и их взаимосвязи. Особенности анализа данных на качественном и количественном уровне. Использование математических методов обработки, анализа и синтеза результатов в профессиональных исследованиях. Обзор математического аппарата и инструментальных средств, используемых для решения практических задач анализа данных.

1.2 Распределение вероятностей: Вероятность. Случайные величины. Понятие, свойства и виды распределения вероятностей. Числовые характеристики дискретных распределений вероятностей. Законы дискретных и непрерывных распределений вероятностей.

1.3 Статистический анализ данных: Выборочный метод. Понятие случайных переменных, выборки, случайной выборки данных. Проверка статистических гипотез. Анализ одной выборки (критерий согласия (Пирсона)). Анализ двух выборок: параметрические критерии (t-критерий Стьюдента, критерий Фишера), непараметрические критерии.

1.4 Методы статистического анализа взаимосвязи признаков, их отличие от методов описательной статистики. Задачи и функции статистических методов. Измерение связи и статистической значимости. Парная и множественная корреляция. Коэффициенты связи. Проверка на статистическую значимость. Ковариация и корреляция Пирсона, их свойства и интерпретация. Ранговые корреляции: Спирмена и Кендалла.

1.5 Основы регрессионного анализа. Парная линейная регрессия: Понятие «регрессия». Простая линейная взаимосвязь. Построение моде-

ли парной линейной регрессии, проверка качества модели регрессии, интерпретация параметров регрессии.

1.6 Основы регрессионного анализа. Множественная регрессия: Уравнение множественной регрессии: построение модели множественной регрессии, проверка качества модели регрессии. Нелинейная регрессия.

1.7 Оценка адекватности модели регрессии: Проверка предпосылок МНК. Использование фиктивных переменных в уравнении множественной регрессии.

1.8 Моделирование временных рядов. Обобщенная модель динамического ряда: Понятие обобщенной модели динамического ряда. Классификация моделей динамического ряда. Порядок построения обобщенной модели динамического ряда. Трендовая компонента. Выявление наличия трендовой компоненты. Построение трендовой компоненты.

1.9 Моделирование временных рядов. Обобщенная модель динамического ряда: Определение типа модели динамического ряда. Понятие сезонной (циклической) компоненты. Построение сезонной компоненты. Аналитическое выравнивание сезонных колебаний с помощью ряда Фурье. Оценка параметров модели и ее адекватности. Анализ модели.

1.10 Моделирование временных рядов. Обобщенная модель динамического ряда: Сезонная корректировка временного ряда.

1.11 Прогнозирование временных рядов: Методы и алгоритмы прогнозирования временных рядов. Оценка точности прогноза временного ряда. Статистические оценки взаимосвязи двух временных рядов.

1.12 Основы дисперсионного анализа: Основные понятия дисперсионного анализа. Однофакторный дисперсионный анализ. Многофакторный дисперсионный анализ.

1.13 Сокращение размерности данных: Метод главных компонент. Факторный анализ.

1.14 Сокращение размерности данных: Кластерный анализ.

1.15 Технологии анализа данных: Оперативный анализ данных (OLAP). Технологии анализа данных. Причины развития и составляющие. Оперативный анализ данных – OLAP. Реляционный оперативный анализ данных. ROLAP. Работа с реляционным хранилищем. Гибридный оперативный анализ данных (HOLAP). Многомерные таблицы для агрегатов. WOLAP, Web-based OLAP – OLAP ориентированный на Web.

1.16 Data Mining. Visual Mining. Text Mining. Internet: Основные понятия, особенности, тенденции, перспективы.

1.17 Программные средства анализа данных: Современные программные средства для оперативного анализа данных. SAP OLAP Server. SAP BW. IBM Cognos PowerPlay. Microstrategy Intelligence Server. Mondrian.

Лабораторные работы 4 шт. по 4 часа

2.1 Предварительная обработка данных. Построение дискретного вариационного ряда. Расчет и анализ его характеристик. Построение интервального вариационного ряда. Расчет и анализ его характеристик.

2.2 Исследование корреляционных зависимостей между признаками. Отсев грубых погрешностей. Проверка требований к отбору исходных факторов для изучения корреляционной зависимости между ними. Поле корреляции. Коэффициенты парной линейной корреляции и их анализ. Построение корреляционной матрицы и ее анализ. Множественная корреляция.

2.3 Исследование модели линейной регрессии. Парная линейная регрессия. Оценка качества модели. Множественная линейная регрессия. Оценка качества модели.

2.4 Исследование модели регрессии. Использование фиктивных переменных в уравнении множественной линейной регрессии. Проверка предпосылок МНК.

Практические занятия 8 шт. по 2 часа:

- 3.1 Дискретные и непрерывные случайные величины, их основные характеристики.
- 3.2 Статистический анализ данных. Проверка статистических гипотез.
- 3.3 Методы статистического анализа взаимосвязи признаков.
- 3.4 Построение и анализ модели линейной регрессии.
- 3.5 Построение обобщенной модели динамического ряда. Аналитическое выравнивание сезонных колебаний с помощью ряда Фурье.
- 3.6 Построение обобщенной модели динамического ряда. Сезонная корректировка временного ряда.
- 3.7 Построение авторегрессионных моделей.
- 3.8 Изучение методов сокращения размерности данных. Применение кластерного анализа

Расчетно-графическая работа

- 3.1 Расчетно-графическая работа «Разработка модели объекта, системы, процесса с использованием методов анализа данных»

Год начала подготовки (по учебному плану) 2020

Образовательный стандарт (ФГОС) №929 от 19.09.2017