

АННОТАЦИЯ РАБОЧЕЙ ПРОГРАММЫ ДИСЦИПЛИНЫ МАШИННЫЙ АНАЛИЗ ЭКОНОМИЧЕСКИХ ДАННЫХ

Автор: Стефановский Дмитрий Владимирович

Код и наименование направления подготовки, профиля: 38.04.01 Экономика («Системы больших данных в экономике»)

Квалификация (степень) выпускника: Магистр

Форма обучения: очная

Цель освоения дисциплины:

Сформировать компетенции в сфере аналитической работы в области экономики и финансов, обработки и анализа данных, а также научно-исследовательской работы

План курса:

Тема 1. Введение в операционную аналитику. Основные определения. Предварительное изучение квазиупорядоченности.

Определение операционной аналитики. Место аналитики в бизнес процессах. Направление развития в аналитике. Множества и многозначные отображения. Граф. Пути и контуры. Цепи и циклы. Квазипорядок, определяемый графом. Индуктивный граф и базы. Центры графа. Диаметр сильно связного графа. Матрица смежности графа. Основные числа теории графов. Ядра графа. Игры на графе. Общие соображения относительно бесконечных графов. Порядковая функция. Функции Гранди. Операции над графами.

Тема 2. Большие данные в аналитике.

Качество данных. Внедрение систем больших данных. Системы больших данных: контекст и антология. Применение операционной аналитики. Очистка данных. Препроцессинг. Обогащение данных. Моделирование и анализ бизнес процессов, анализ рисков, обнаружение типичных и нетипичных паттернов.

Тема 3. Принципы создания аналитической платформы.

Планирование создания, проектирование создания. Основы управления на базе операционной аналитики. Создание операционно-аналитических процессов, сравнение аналитических подходов. Управление и конфиденциальность.

Тема 4. Аналитическая команда и культура.

Подбор аналитиков. Сертификация аналитиков. Организация команды. Аналитическая культура. Оценка эффективности методов. Задача о наибольшем паросочетании. Дефицит простого графа. Венгерский алгоритм. Обобщение на бесконечный случай. Приложение к теории матриц. Цикломатическое число. Хроматическое число. Число внутренней устойчивости. Число внешней устойчивости. Теоремы существования и единственности. Приложение к функциям Гранди. Игра Ним. Общее определение игры (с полной информацией). Стратегии. Центры и радиус графа. Общие свойства сильно связных графов без петель. Диаметр. Применение обычных матричных операций. Задачи на подсчет. Задача о лидере. Применение булевых операций.

Тема 5. Исследование взаимосвязей между переменными. Проверка гипотез.

Таблица сопряженности; формулировка гипотез; этапы проверки гипотез; уровень значимости и ошибка первого рода; тест Хи-квадрат; построение диаграммы рассеяния; парные коэффициенты корреляции (Пирсона, Кендалла, Спирмана). Частные корреляции; сравнение средних (t-тест для независимых и зависимых выборок, однофакторный дисперсионный анализ).

Тема 6. Регрессионный анализ.

Простая линейная регрессия. Множественная регрессия. Оценка качества модели. Анализ остатков. Бинарная логистическая регрессия. Мультиномиальная логистическая регрессия.

Использование фиктивных переменных для моделирования зависимостей от качественных признаков. Виды моделей, интерпретация коэффициентов при фиктивных переменных, фиктивные переменные сдвига и наклона.

Использование фиктивных переменных для моделирования сезонных колебаний.

Использование нескольких наборов фиктивных переменных

Компьютерный анализ временных рядов в эконометрике с использованием прикладных программных пакетов для определения: особенности моделирования временных рядов; модели тренда; сезонных колебания; проблемы автокорреляции, теста Дарбина-Уотсона.

Тема 7. Факторный анализ.

Метод главных компонент. Определение первой и m -ой главных компонент

Построение главных компонент.

Способы отбора главных компонент.

Интерпретация значений факторов. Методы факторного анализа: цепные подстановки, индексы, интегральный способ. Необходимость снижения размерности признакового пространства

Снижения размерности признакового пространства.

Матрица нагрузок.

Сфера применения методов сжатия данных.

Метод главных факторов.

Разложение дисперсии исходного признака.

Алгоритм метода главных факторов.

Метод максимального правдоподобия.

Качественный анализ результатов факторного анализа.

Тема 8. Кластерный анализ

Иерархический кластерный анализ; кластерный анализ методом k -средних; содержательная характеристика кластеров.

Пакеты прикладных программ анализа и моделирования в экономических данных.

Исследование операций и компьютерный анализ задач микроэкономики.

Системный анализ и машинные методы анализа в задачах региональной и мировой экономике.

Аудиторные часы: 108

Формы текущего контроля и промежуточной аттестации: опросы, домашнее задание, контрольная работа, зачет

Основная литература:

1. Белов В.С. Информационно-аналитические системы. Основы проектирования и применения: учебное пособие, руководство, практикум / Московский государственный университет экономики, статистики и информатики. — М., 2005. — 111 с.
2. Барсегян, А. А. Анализ данных и процессов: учеб. пособие / А. А. Барсегян, М. С. Куприянов, И. И. Холод, М. Д. Тесс, С. И. Елизаров. — 3-е изд., перераб. и доп. — СПб.: БХВ-Петербург, 2009. — 512 с.: ил. + CD-ROM — (Учебная литература для вузов) ISBN 978-5-9775-0368-6.