

Современные методы анализа данных – геостатистика, нейронные сети Нужный А.С.

Курс посвящен вопросам моделирования и прогнозирования на основе данных измерений. Изучаются методы статистического пространственного анализа (геостатистика) и методы искусственного интеллекта (нейронные сети). Методы рассматриваются применительно к задачам регрессии и классификации, а также для задач кластеризации. Рассматриваются подходы для многопеременного анализа и получения вероятностных оценок.

№	Тема	Содержание
1	Анализ и моделирование пространственных корреляций пространственно-распределенных данных	<ol style="list-style-type: none"> 1. Вариограмма как мера, описывающая пространственную корреляцию. 2. Вычисление вариограммы по исходным данным (экспериментальная вариограмма). 3. Математические модели вариограмм и их особенности.
2	Геостатистические методы пространственных интерполяций	<ol style="list-style-type: none"> 1. Основные ограничения классической геостатистики (стационарность 2 порядка и внутренняя гипотеза). 2. Простой кригинг – предположения и вывод, свойства. 3. Обычный кригинг и его свойства. 4. Кригинговая вариация и ее интерпретация.
3	Стохастическое моделирование	<ol style="list-style-type: none"> 1. Пространственная неопределенность – пространственная совместная функция распределения. 2. Последовательный принцип стохастического моделирования и алгоритм его практической реализации. 3. Геостатистические методы стохастического моделирования, основанные на последовательном принципе. 4. Моделирование отжига.
4	Вероятностное картирование	<ol style="list-style-type: none"> 1. Задача оценки локальной функции распределения, как общая задача анализа данных. 2. Индикаторный подход. 3. Стохастическое моделирование.
5	Искусственные нейронные сети (ИНС)	<ol style="list-style-type: none"> 1. Принципы искусственной нейронной сети. 2. Элементы искусственной нейронной сети. 3. Структуры искусственных нейронных сетей. 4. Подходы к обучению искусственных нейронных сетей. 5. Обучение ИНС «без учителя».
6	Многослойный персептрон (МСП) как ИНС прямого распространения	<ol style="list-style-type: none"> 1. Элементы и структура МСП. 2. Инициализация и обучение МСП (обучение «с учителем»). 3. Метод обратного распространения ошибки.
7	Применение нейронных сетей для анализа и моделирования данных	<ol style="list-style-type: none"> 1. Моделирование нелинейного крупномасштабного тренда. 2. Нейронная сеть обобщенной регрессии.
8	Обзор и сопоставление существующих геоинформационных систем	
9	Классификация и форматы	<ol style="list-style-type: none"> 1. Классификация топографических карт. 2. Форматы хранения картографической информации.

Вопросы по курсу

1. Персептрон. Метод обратного распространения ошибки.
2. Персептрон как некорректная задача. Валидация (метод раннего останова).
3. Байесовская регуляризация обучения персептрона.
4. Байесовская регуляризация в задаче аппроксимации функции по точкам с помощью базисных функций.
5. Аппроксимации функции по точкам с помощью ортогонализованного базиса.
6. Метод главных линейных компонент.
7. Сеть с узким горлом — нелинейные главные компоненты.
8. Энтропийный анализ, взаимная информация.
9. Метод независимых компонент.
10. Кластерный анализ. K-means.
11. Гауссова смесь.
12. Карта Кохонена.
13. Меры пространственной корреляции. Вариограммы. Моделирование вариограмм.
14. Простой кригинг.
15. Обычный кригинг.
16. Деревья решений. Критерии GAIN и GINI.
17. Голосование независимых классификаторов. Случайный лес.
18. Стохастическое моделирование. Последовательное гауссово моделирование.

Литература

1. В. Демьянов, Е. Савельева, Геостатистика. Теория и практика. Издательство "Наука", Москва, 2010
2. Саймон Хайкин, Нейронные сети. Издательство "Вильямс", Москва 2006
3. С. А. Шумский, Байесова регуляризация обучения. Материалы школы-семинара "Современные проблемы нейроинформатики", Москва 2002
4. А. С. Нужный, Байесовская регуляризация в задаче аппроксимации функции по точкам с помощью ортогонализованного базиса, математическое моделирование, 23:9 (2011), 33-42