**Прикладная статистика в машинном обучении**

1. Введение. Основные задачи и методы теории статистических выводов (1 лекция, 1 семинар)

Параметрические и непараметрические модели. Основные задачи: точечное оценивание, доверительные множества, тестирование гипотез, исследование зависимостей. Эмпирическая функция распределения. Статистические функционалы.

1. Бутстреп (1 лекция, 1 семинар)

Моделирование Монте-Карло, бутстреп. Оценка дисперсии на основе бутстрепа. Оценка доверительных интервалов на основе бутстрепа. Метод складного ножа.

1. Параметрическое оценивание (2 лекции, 2 семинара)

Метод моментов. Метод максимального правдоподобия и его свойства. Дельта-метод. Случай векторного параметра. Параметрический бутстреп. Доверительное оценивание. Достаточная статистика. Экспоненциальное семейство распределений. EM-алгоритм.

1. Проверка гипотез (2 лекции, 2 семинара)

Основные понятия теории проверки гипотез. Критерий Вальда. P-значение. Распределение хи-квадрат и критерий Пирсона. Критерий перестановок. Критерий на основе отношения правдоподобия. Множественные тесты. Критерий согласия. Критерий Неймана-Пирсона для случая двух простых гипотез, t-критерий.

1. Линейная и логистическая регрессия (2 лекции, 2 семинара)

Стандартная линейная регрессия. Метод оценивания на основе минимизации невязок/максимизации правдоподобия. Свойства оценок метода наименьших квадратов. Прогнозирование. Множественная регрессия. Выбор модели. AIC, BIC. Логистическая регрессия.

1. Непараметрическое оценивание сигналов (1 лекция, 1 семинар)

Выбор оптимального соотношения между смещением и дисперсией. Гистограммы. Ядерная оценка плотности. Непараметрическая регрессия.

1. Нелинейные методы построения регрессионных зависимостей (1 лекция, 1 семинар)

Аддитивные модели. Аппроксимационные модели на основе параметрических словарей (сигмоидальные функции, гауссоподобные функции, и т.д.). Построение аппроксимационных моделей на основе параметрических словарей: методы инициализации структуры; методы оценки параметров, учитывающие структуру модели; методы адаптивной регуляризации в процесс оценки параметров модели.

1. Регрессионные модели на основе гауссовских процессов (1 лекция, 1 семинар)

Гауссовские процессы. Прогноз и оценка точности прогноза на основе гауссовских процессов. Оценка параметров ковариационной функции. Регуляризация логарифма правдоподобия. Модель нестационарной ковариационной функции. Регрессия на основе гауссовских процессов как ядерная гребневая регрессия. Глобальная оптимизация на основе гауссовских процессов. Планирование экспериментов на основе гауссовских процессов.

1. Оценка чувствительности модели (1 лекция, 1 семинар)

Введение, начальные условия, области применения. Постановка задачи оценки глобальной чувствительности модели. Обзор методов оценки чувствительности модели в случае выборки данных. Обзор методов оценки чувствительности модели в случае черного ящика. Выделение параметров, которые сильнее всего влияют на уровень радиоактивных изотопов в воде (Level E test case). Обзор теоретических основ метода оценки чувствительности модели на основе индексов Соболя.

1. Методы планирования экспериментов (Design of Experiments) (1 лекция, 1 семинар)

Space-filling DoE. Планирование экспериментов для линейных регрессионных моделей.

1. Снижение размерности многомерных данных (1 лекция, 1 семинар)

Снижение размерности в суррогатном моделировании. Постановка задачи снижения размерности. Обзор линейных методов снижения размерности (метод главных компонент, целенаправленное проектирование и т.п.). Обзор локальных и нелинейных методов снижения размерности (метод нелинейных главных компонент, метод локального линейного вложения и т.п.). Новые постановки задач снижения размерности. Аппроксимация многомерных зависимостей как решение задачи снижения размерности.