

Предиктивная аналитика (Predictive analytics) на базе регрессионных моделей. Уровень 1

1 день. Основы моделирования и прогнозирования

- Типизация исходных данных для анализа
- Big Data и традиционные выборки
- Классификатор инструментов предиктивной аналитики по типу решаемой задачи
- Активный и пассивный эксперименты
- Основные понятия моделирования: модель, прогноз, отклонения (остатки), гипотеза, фактор, отклик, статистическая значимость
- Подходы к валидации модели (деление набора данных на обучающий и тестовый и другие способы)
- Проблемы моделирования: переобучение, нерепрезентативность выборки и другие
- Подходы к оценке точности модели, понятие обучения модели
- Базовые статистические показатели и графики: среднее и медиана, размах, стандартное отклонение, дисперсия, стандартная ошибка, доля, доверительные интервалы, гистограмма, ящичная диаграмма, вероятностный график
- Различные виды распределений (нормальное, экспоненциальное, Вейбулла, Пуассона, биномиальное и т.д.). Проверка соответствия виду распределения и практические выводы
- Предварительный графический анализ влияния факторов на отклик
- Оценка тесной связи количественных признаков с помощью коэффициентов корреляции; корреляция и причинность
- Практическая часть – отработка примеров в Minitab

2 день. Моделирование и прогнозирование на основе многофакторной регрессии

- Многофакторный регрессионный анализ – метод моделирования и прогнозирования количественного показателя:
 - требования к виду и количеству исходных данных
 - необходимые и достаточные теоретические основы, идеи и принципы, плюсы, минусы и особенности подхода
 - требования к факторам, которые планируется учитывать в анализе
 - пошаговый алгоритм регрессионного анализа
 - способ проверки адекватности модели (анализ отклонений «факт-модель»)
 - интерпретация результатов: как влияют факторы на отклик, ранжирование по силе влияния
 - трансформации исходных данных при нарушении требований линейного регрессионного анализа
 - оценка эффекта взаимодействия факторов
 - оптимизация на основе построенной модели – поиск оптимальных значений факторов для достижения требуемых значений отклика на основе пассивного эксперимента

- Практическая часть – отработка примеров в Minitab

3 день. Моделирование и прогнозирование на основе других видов регрессионного анализа

- Бинарная логистическая регрессия – метод моделирования и прогнозирования одного из двух исходов:
 - требования к виду и количеству исходных данных для каждой модели
 - необходимые и достаточные теоретические основы, идеи и принципы, плюсы и минусы метода
 - валидация и интерпретация полученных результатов
 - оценка эффекта взаимодействия факторов
 - построение прогнозов с использованием Рос-анализа
- Прогнозирование отказов оборудования:
 - Прогнозирование вероятности отказов на основе анализа интенсивности отказов за период
 - Анализ влияния различных факторов на интенсивность отказов и затраты на их устранение с использованием регрессии:
 - требования к виду и количеству исходных данных
 - необходимые и достаточные теоретические основы, идеи и принципы, плюсы и минусы подхода
 - требования к факторам, которые планируется учитывать в анализе
 - пошаговый алгоритм анализа
 - способ проверки адекватности модели
 - интерпретация результатов: как влияют факторы на исследуемые показатели
- Практическая часть – отработка примеров в Minitab

С уважением,
Команда Novo BI
customer@4analytics.ru
www.4analytics.ru
www.novoforecast.com
www.4promotion.ru