

Построение прогностических моделей в IBM SPSS Modeler

01 Подготовка данных к моделированию

Чистка данных

Балансировка данных

Преобразования количественных данных

Категоризация данных

Разбиение данных

Обнаружение аномальных данных

Отбор предикторов для моделей

02 Сжатие данных. Метод главных компонент

Применение главных компонент для моделей прогноза и кластерного анализа

На что следует обратить внимание при выполнении факторного анализа или анализа главных компонент

Принципы анализа

Сравнение факторного анализа и анализа главных компонент

Число компонент

Вращения

Значения главных компонент

Объем выборки

Методы

Общие рекомендации

Пример: регрессия на основе главных компонент

03 Деревья решений / Индукция правил

Сравнение моделей деревьев решений

Узел C5.0

Просмотр модели

Построение набора правил и их просмотр

Изучение правил и определение точности

Важнейшие предикторы в прогнозировании

Подробнее про C5.0

Моделирование категориального целевого поля алгоритмами CHAID, C&RT и Quest

Категориальное целевое поле в CHAID

Категориальное целевое поле в C&R Tree

Категориальное целевое поле в QUEST

Интерактивные деревья

Количественное целевое поле

04 Нейронные сети

Узел Neural Net

Первичный разбор построенной модели

Проверка состава предикторов

Интерпретация нейронной сети

Логика прогнозов модели

Применение узла Analysis

Заключение по модели

Методы построения нейронных сетей

Методы обучения

Многослойный перцептрон

Нейросеть на основе радиальных базисных функций

Какой метод и когда?

Предупреждение переобучения

Пропущенные значения в нейронных сетях

Другие варианты построения нейронной сети





05 Линейная регрессия

Прицип регрессионного анализа

Пример: обнаружение ошибок или случаев мошенничества в заявленных суммах страховок

06 Анализ временных рядов

Определение временного ряда

Файл данных временного ряда

Тренд, сезонная и циклическая компоненты

Понятие модели временного ряда

Интервенции

Экспоненциальное сглаживание

ARIMA

Требования к данным

Автоматическое построение прогноза в производственном режиме

Прогнозирование использования каналов широкополосной связи на нескольких рынках

Применение моделей для нескольких рядов

07 Логистическая регрессия

Мультиномиальная логистическая регрессия: прогнозирование кредитного риска

Интерпретация коэффициентов

08 Получение максимальной отдачи от моделей

Комбинирование моделей с помощью узла Ensemble

Применение скоринговых баллов

Мета-моделирование

Моделирование ошибок

09 Методы кластеризации

На что следует обращать внимание при проведении кластерного анализа?

Кластеризация методом K-средних

Узел K-MEANS

Изучение профилей кластеров

Кластеризация при помощи сети Кохонена

Узел KOHONEN

Двухэтапный кластерный анализ

10 Введение в правила ассоциаций

Узел APRIORI

Применение ассоциаций

11 Применение правил ассоциаций

Правила ассоциаций

Алгоритм APRIORI

Алгоритм CARMA

Дополнительные настройки узла APRIORI

Дополнительные настройки узла CARMA

Выбор метода и дополнительные настройки

Обработка пропущенных значений в правилах ассоциаций

12 Машина опорных векторов

Структура моделей SVM

Модель SVM при прогнозировании оттока

Анализ модели

Модель с различными функциями ядра

Настройка RBF модели





13 Анализ данных о выживаемости на основе регрессии Кокса

Что такое анализ выживаемости?

Регрессия Кокса

Регрессия Кокса для предсказания оттока

Проверка допущения о пропорциональности рисков

Прогнозы модели Кокса

14 Дискриминантный анализ

Принцип работы дискриминантного анализа

Дискриминантная модель

Классификация наблюдений

Допущения дискриминантного анализа

Рекомендации по проведению анализа

Сравнение дискриминантного анализа и логистической регрессии

Пример дискриминантного анализа

15 Байесовские сети

Основы байесовских сетей

Типы байесовских сетей в IBM SPSS Modeler

Создание модели на основе байесовской сети

Изменение настроек модели байесовской сети

