




Разработка скоринговых моделей

Очный тренинг (35 академических часов)



Общая информация

Параметр	Описание
Целевая группа	<ol style="list-style-type: none"> 1. Аналитики финансовых организаций и банков, занимающиеся моделированием кредитных рисков, коллекторской работой 2. Аналитики отделов маркетинга банковских организаций 3. Бизнес-аналитики подразделений продаж и маркетинга розничных и оптово-розничных дистрибьюторских компаний, работающих с клиентской базой
Длительность	4 тренинговых дня, всего 35 академических часов, не включая перерывы на обед и кофе-брейки
Формы работы	Мини-лекции, дискуссии, тематические беседы, метод кейсов, анализ практических ситуаций, мозговой штурм, работа в микро-группах, выполнение заданий с использованием компьютера и специализированного ПО
Число участников	7-14 человек
Место проведения	г. Рязань, офис BaseGroup Labs
Бизнес-тренер	Николай Паклин

Цель тренинга

Научить аналитиков разрабатывать эффективные и качественные скоринговые модели (отклика, вероятности просрочки по кредиту, вероятности взыскания долга и т.п.), объективно оценивать их качество и правильно документировать модели.

Задачи тренинга

1. Создание рабочей атмосферы на тренинге.
2. Диагностика уровня владения темой участниками.
3. Получение знаний по темам тренинга в форме активного обучения.
4. Комплексная отработка основных этапов разработки скоринговых моделей с использованием прикладного ПО: MS Excel, Deductor Studio 5.3.
5. Подведение итогов и организация обратной связи от участников.

Ожидаемые результаты

Для компании	Для участников
<ul style="list-style-type: none">▪ Компания станет более конкурентной благодаря наличию персонала с компетенциями разработки скоринговых моделей▪ Сотрудники смогут применять полученные знания и навыки на практике для повышения эффективности принятия управленческих решений	<ul style="list-style-type: none">▪ Узнают что такое скоринговая модель, скоринговая карта▪ Получат навыки подготовки и аудита данных для дальнейшего анализа и моделирования▪ Освоят стандартный цикл построения скоринговой модели с отработкой в Deductor Studio и научатся оценивать качество скоринговой карты▪ Сформируют компетенции оформления результатов в виде паспорта скоринговой модели/карты

Темы тренинга

- Введение в скоринг
- Определение события
- Элементы аудита данных
- Подготовка выборок для моделирования
- Построение модели (логистическая регрессия)
- Постобработка и валидация скоринговой модели
- Документирование скоринговых моделей/карт

Тематический план

День 1 – Понятие и сущность скоринга. Подготовка данных. Определение события. Элементы аудита данных

- Знакомство группы. Маршрут тренинга. Постановка целей.
- Определение скоринга. Сущность скоринга. Скоринговая модель, скоринговая карта. Преимущества и недостатки статистического скоринга.
- Основные этапы разработки скоринговой карты. Наиболее и наименее трудоемкие этапы. Применение скоринга
- Типовая схема разработки моделей, источники данных для разработки скоринговых моделей (кредитных, отклика, коллекшн и т.д.). Расчетные и производные показатели.
- Определение события. Дискуссия и групповые задания из различных областей: ритейл, телеком, МФО, розничный банк.
- Элементы аудита данных. Выбросы, экстремальные значения. Узел «Качество данных» в Deductor Studio.
- Задача откликов ОТП-банка. Мастер-класс по аудиту данных задачи откликов ОТП-банка.

Результат дня для участников:

- Понимание определения скоринга, его сущности, достоинств и недостатков.
- Знание основных этапов разработки скоринговой карты.
- Отработка на практике техник генерации гипотез и производных показателей (на примере скоринга откликов)
- Понимание проблемы выбора события и владение методами ее решения.
- Отработка на практике элементов аудита данных на примере реальной задачи.

День 2 – Подготовка выборок для моделирования

- Что такое сэмплинг и в каких ситуациях он необходим аналитику?
- Какие методы сэмплинга существуют? Случайный, равномерный случайный, стратифицированный сэмплинг, отбор со смещением.
- Разбиение на множества: обучающая и тестовая выборки.
- Ситуация, когда разбиение на множества задано извне. Какие «подводные» камни здесь могут быть для последующего моделирования? Поиск и автоматическая обработка редко встречающихся уникальных значений в характеристиках набора данных.
- Одномерный и двумерный анализ. Анализ взаимоотношений двух переменных. События и не-события. WoE-индексы. Анализ и интерпретация WoE-диаграмм.
- IV-индекс. Оптимальное квантование. Предсказательная сила характеристики.
- Сокращение числа уникальных значений характеристик (Fine&Coarse Classing). Какие плюсы дает Coarse Classing с точки зрения моделирования?
- Узел «Конечные классы» в Deductor Studio. Настройка конечных классов задачи откликов ОТП-банка.
- Эмпирические правила аналитика по настройке конечных классов.

Результат дня для участников:

- Четкое понимание необходимости сэмплинга и его различных видов.
- Владение на практике техниками обработки редко встречающихся уникальных значений в характеристиках набора данных.
- Знание и понимание методов оценки предсказательной силы как отдельного атрибута характеристики, так и всей характеристики целиком.
- Умение интерпретировать WoE-диаграммы характеристик будущей скоринговой модели.
- Понимание алгоритмов оптимального квантования (для непрерывной характеристики) и снижения разнообразия уникальных значений (для категориальной характеристики).
- Знание и отработка на практике настройки конечных классов в Deductor Studio.

День 3 – Построение модели (логистическая регрессия). Постобработка и валидация скоринговой модели (часть 1)

- Математическое понятия шансов и их связь с вероятностью наступления события. Отношения шансов. Примеры.
- Почему линейная регрессия плохо подходит для моделирования бинарной выходной переменной? Логит-преобразование.
- Модель парной логистической регрессии. Значимость модели и отдельного фактора. Тест Вальда. Множественная логистическая регрессия
- Узел «Логистическая регрессия» в Deductor Studio.
- В чем преимущества применения категоризации непрерывных характеристик при использовании логистической регрессии?
- Постобработка скоринговой модели. Калибровка, масштабирование, сдвиг.
- Бизнес-смысл понятий «лифт» и «лифт-диаграмма» на примере задачи отклика на рассылку. «Идеальная» лифт-кривая. Profit-диаграмма.
- Основные метрики и диаграммы для оценки качества скоринговой карты: индекс Джини, AUC, ROC-кривая, статистика KS и др.

Результат дня для участников:

- Четкое понимание связи отношений шансов, вероятности наступления события и коэффициентов логистической регрессии.
- Владение методикой последовательной оценки скоринговой модели: от оценки статистической значимости регрессионной модели до
- Понимание того, как строятся и рассчитываются главные метрики качества скоринговой карты – индекс Джини и статистика KS.
- Знание и отработка на практике построения модели логистической регрессии в Deductor Studio и получение из нее скоринговой карты.
- Владение методами калибровки скоринговой модели на истинное соотношение событий и не-событий в совокупности.
- Умение ориентироваться и отличать «плохие» от «хороших» диаграмм, оценивающих качество скоринговой карты (Lift-кривая, распределение скорингового балла, риск-сегменты и др.)

День 4 – Постобработка и валидация скоринговой модели (часть 2). Документирование скоринговых моделей

- Встроенные диаграммы и метрики для оценки качества скоринговой модели в Deductor Studio.
- Дополнительная специализированная аналитическая панель для оценки качества скоринговой модели **Scorecard KPI Panel** (готовый шаблон-сценарий для Deductor Studio) . Оценка качества построенной модели вероятности отклика для задачи клиентов ОТП-банка.
- Критический анализ качества «сторонних» скоринговых моделей (карт). На что нужно обращать внимание при анализе скоринговой модели?
- Паспорт скоринговой модели (карты). Структура, основные блоки документа.
- Профессиональные навыки скоринг-аналитика.
- Завершение тренинга.

Результат дня для участников:

- Умение быстро получить набор метрик и диаграмм для оценки качества скоринговой модели (карты) на основе готового шаблона-сценария для Deductor Studio.
- Отработанные навыки анализа качества скоринговых моделей, построенных сторонним аналитиком с критическим обсуждением результатов в мини-группах.
- Умение составить паспорт скоринговой карты.
- Понимание основных профессиональных навыков, которыми должен обладать аналитик по разработке скоринговых карт.
- Четкая картина всех основных этапов разработки скоринговых моделей.

В третий день тренинга может быть предусмотрена культурная программа по достопримечательностям г. Рязани или области.

Тренер



Бизнес-тренер, эксперт по бизнес-аналитике и Data Mining, руководитель проектов корпоративного обучения **BASEGROUP LABS**.

Окончил Ижевский государственный технический университет по специальности «Информационные системы и технологии», в 2004 г. защитил диссертацию на соискание ученой степени кандидата технических наук. С 2005 г. — сотрудник компании **BASEGROUP LABS**.

Выполнил несколько BI-проектов в качестве аналитика: компания «Сонекс», банк «Дельта-кредит», банк «Стройкредит», «МТС-банк», «Москоммерцбанк» и др.

Опыт тренерской работы 5 лет (тренинги в компаниях «Данон», «Руна», «МТС-Беларусь», «Дельта-Кредит», банк «Первомайский», банк «ИТБ», «Миг-кредит», «Связь-банк», СКБ-банк» и другие). Автор цикла очных тренингов и семинаров по кредитному скорингу и разработке скоринговых моделей. Главный разработчик конфигурации **DEDUCTOR CREDIT SCORECARD MODELER**.

Автор книги «**БИЗНЕС-АНАЛИТИКА: ОТ ДАННЫХ К ЗНАНИЯМ**» (издательство «Питер», совокупный тираж 4500 экз., 2009, 2010, 2013 годы).

Доцент РГУ имени С.А. Есенина, г. Рязань. Опыт преподавательской работы в вузах более 10 лет. Сертификат бизнес-тренера (компания «Искусство тренинга», 2015).

BaseGroup Labs

Компания BaseGroup Labs занимается разработкой аналитического программного обеспечения, консультациями и обучением в области анализа данных. Фирма основана в 1995 году. Партнерская сеть насчитывает более 40 фирм в России, ближнем и дальнем зарубежье. Аналитическая платформа Deductor включена в официальную образовательную программу 100 вузов России и стран СНГ.

Компанией BaseGroup Labs реализовано более ста проектов в различных отраслях бизнеса.



Контакты

Адрес: Россия, 390023, г. Рязань, ул. Новая 53в

Телефон: Москва - +7 (495) 222-71-17, Рязань - +7 (4912) 24-09-77, 24-06-99

Факс: +7 (4912) 24-09-77

E-mail: sale@basegroup.ru