

| Название кейса                                    | Содержание   | Вопросы по кейсу  | Критерии оценки   |
|---|--|---|---|
| Оценка финансового результата от внедрения модели | <p>В процессе рассмотрения заявок на кредиты Вы используете модель классификации, которая присваивает каждому заемщику некоторый скоринговый балл. Чем выше значение балла, тем ниже вероятность дефолта.</p> <p>У вас есть историческая выборка с параметрами выданных кредитов (сумма, срок, ставка), факторами заемщиков, значениями баллов и известным фактом наступления (или ненаступления) события дефолта.</p> | <p>Как оценить финансовый результат использования данной модели?</p> <p>Какие еще параметры в исторической выборке Вам могут потребоваться и как бы Вы их использовали?</p>   | <ol style="list-style-type: none"> <li>1. Использует понятие ошибки первого и второго рода</li> <li>2. Предлагает оценить оптимальный уровень отсечения</li> <li>3. Рассчитывает прибыль при оптимальном уровне отсечения</li> <li>4. Указывает на проблему ненаблюдаемой ошибки второго рода</li> <li>5. Дополнительные параметры - оценить вероятность досрочного погашения и недозаработанный доход, транзакционные данные, данные внешних источников</li> </ol>   |
| Рекомендательная система для up-sell              | <p>Перед Вами стоит задача построения рекомендательной системы продуктов банка. На данный момент таких продуктов несколько десятков (кредиты, кредитные карты, дебетовые карты, депозиты и т.д.), и нужно научиться понимать для каждого клиента, какой из этих продуктов наиболее интересен. При этом численность клиентских менеджеров позволяет совершить всего с 1000 клиентами в день.</p>                        | <p>Какие данные могут быть полезны для данной задачи? Как Вы будете собирать разметку для обучения модели?</p> <p>Допустим, Вы построили модель, и она для каждого клиента выдаёт наиболее интересный ему продукт. Как понять, приобретёт ли клиент этот продукт самостоятельно, либо же обязательно нужна коммуникация от банка?</p> <p>Как оценивать качество модели, какую метрику Вы можете предложить? Достаточно ли для этого использовать исторические данные?</p> | <ol style="list-style-type: none"> <li>1. Порассуждал о полезных данных, сказал про транзакции, данные по счетам, историю обращений и подключения/отключения услуг, предыдущие коммуникации с клиентом.</li> <li>2. Предложил использовать предыдущие коммуникации для формирования разметки; порассуждал о том, как понять, по какой причине подключена услуга — благодаря коммуникации или же самоходом.</li> <li>3. Предложил сделать в модели управляющий параметр, который говорит о том, будет ли сделана коммуникация.</li> <li>4. Относительно метрики качества заметил, что для модели имеет значение, что происходит в топ-1000 клиентов, и придумал метрику на основе этого (Lift, precision@1000).</li> <li>5. Для обогащения исторических данных предложил дизайн пилота для небольшого потока клиентов в целях сбора данных об откликах клиентов</li> </ol> |