

Домашнее задание №2: CausallImpact

дедлайн - 30 октября

Домашнее задание присылайте в [anytask](#). Вопросы по этой теме стоит отправлять в anytask или на causalimpact@yandex.ru. Задания можно присылать в любой адекватной форме. Если присылаете код, то он должен быть таким, чтобы его можно было запустить и получить ваши результаты. Я буду читать присланное вами раз в неделю, и давать фидбек на ваши решения, соответственно, если вы будете быстро делать задание, то у вас будет 2 возможности дослать исправления, но можно и впервые отправить задание в конце третьей недели, тогда уже без возможности прислать доработки.

1. (3 балла) Изучите модель, используемую в качестве дефолтной в CausallImpact. Объясните «на пальцах» о том, как работает дефолтная модель, и сделайте выводы о границах её применимости.
2. (3 балла + бонусные баллы за редкие и хорошие примеры) У модели, используемой в CausallImpact, есть множество различных дополнительных параметров (параметры сезонности, динамические параметры модели). Проведите эксперименты с различными данными и на их основе создайте список эмпирических правил, когда использование дополнительных параметров позволит улучшить результат, а когда, несмотря на их кажущуюся «логичность», их использование портит результат. Постарайтесь в этих правилах рассмотреть максимальное число ситуаций "с подвохом", когда человек, не проделавший данное упражнение, скорее попытался бы воспользоваться неудачными параметрами. Все правила проиллюстрируйте примерами, и желательно чтобы данные в примерах были "правдоподобными".
3. (3 балла) Допустим несколько месяцев назад Яндекс значительно улучшил качество поиска картинок. Вы хотите узнать увеличилось ли это изменение активность пользователей на основном поиске Яндекса. Какими рядом-таргетом и рядом-предиктором вы бы воспользовались? Обоснуйте свой ответ, а также укажите возможные проблемы вашего решения и как бы вы их обошли.
4. (10 баллов) Представьте, что у Яндекса есть ещё один сервис – Яндекс.Котики. Вам нужно оценить эффективности рекламных кампаний, шедших в период с августа по ноябрь (включительно) 2015 года (для каждой отдельная оценка). У этой задачи может быть много способов её решения, ваша цель достигнуть максимально точной и корректной итоговой оценки. В [первом файле](#) вы можете найти информацию об аудитории данного сервиса (target), различные оценки

аудитории, интересующейся котиками (market), а так же показатели различных срезов аудитории сервиса, которые могли быть или не быть подвержены воздействию рекламных кампаний. Во [втором файле](#) вы найдете информацию о датах рекламных кампаний и оценку (возможно неверную) их размеров. В этом же файле есть информация и о рекламных кампаниях некоторых других сервисов, которые потенциально могли воздействовать на Яндекс.Котиков (например, реклама с посылом "весь Яндекс - это круто" могла помочь нашему сервису, а могла и не повлиять на его популярность).

Распространенные ошибки по задачам среди уже сданных:

- Рассказ не о том. В задании интересует то, что было "черным ящиком" во время задания: рассказ о самой дефолтной модели в CausalImpact, предсказывающей по предиктору таргет, а не рассказ о самой идее метода в духе того, что рассказывалось на занятии.
- Сыпать формулами, и не пояснять суть. Представьте, что вам нужно за несколько минут объяснить вашему нерадивому одногруппнику как работает эта модель, чтобы он успел все понять и запомнить, и сумел ответить на зачете на вопрос по этой теме.
- Нет выводов о границах применимости модели. Возможно, что ваше объяснение вышло не достаточно простым и понятным, чтобы из него напрашивались подобные выводы.
- Правила стоит доводить до вида "Когда данные соответствуют <выявленные вами признаки>, (не)стоит использовать <параметр, для которого вы нашли это правило>".
- Исследуются не параметры, с которыми вызывается функция CausalImpact, а что-то еще (что в целом для вас полезно, но не отвечает на вопрос задания)
- Старайтесь максимально полно рассмотреть имеющиеся возможности
- Не забывайте, что с результаты поиска по картинкам пользователь может увидеть и не заходя на соответствующую вертикаль.
- Не забывайте обосновывать все принятые вами решения: выбор таргета и предиктора, выбор периодов, выбор модели, и т.д.