Д. Звежинский, 5.12.2018

Dmitry.Zvezhinsky@sas.com

Моделирование продаж или спроса?

- На прошлом семинаре мы видели, что дефицит (остатки=0) вызывает замену продаж одного товара другим. Видимо, это значимый фактор для прогнозирования.
- Однако, на практике прогнозируются не продажи, а спрос.
 - Вопросы для обсуждения: почему не продажи?
 - Можно ли спрогнозировать дефицит?
- Важная вводная: хотя в истории и наблюдались случаи дефицита, в прогнозной модели не будет отдельного предиктора типа «флаг дефицита» или «остатки=0».
 - Могут ли быть причины, чтобы ввести такой фактор в модель?
 - Какими значениями задать этот фактор на будущее?

- Можно выдумать разные подходы к этой задаче. Обратите внимание, что:
 - Подход должен быть не слишком «тяжелым» в смысле расчетов. Требуется обработать все входные данные в самом детальном разрезе!
 - Подход должен учитывать наличие разных компонент спроса:
 - Регулярный и промо
 - Сезонность спроса (Как восстанавливать продажи арбузов зимой, или новогодних подарков летом?)
 - Влияние внешних событий: <u>праздники</u> (Новый год, 8 марта, 1 мая, день нефтяника), <u>погодные условия</u> («пошел снег», «наступило лето»), <u>глобальные и локальные маркетинговые события</u> (чемпионат мира по футболу, цены конкурентов).
 - Подход должен обрабатывать не только неестественно маленькие, но и неестественно большие продажи.
 - Их не любят регрессионные методы.
 - Помним, что при удалении положительных выбросов есть шанс «потерять» данные о влиянии «сильных» внешних событий.

- Исходные данные могут быть очень некорректными
 - Отрицательные остатки это дефицит, или особенность учёта?
 - А положительные? Могут быть «виртуальными» и не продаваться.
- При восстановлении нельзя заглядывать в будущее.
 - А почему? В статье пишут по другому.
- Не на всех временных рядах имеет смысл восстанавливать спрос
 - Ряд слишком плохой. Товара почти никогда не бывает на полках (дефицит).
 - Или товар плохо продаётся
 - Нет истории продаж, товар новый.
 - Или ряд слишком хороший, чтобы его восстанавливать. Примеры таких рядов встречаются на верхних уровнях логистической цепочки, например, на производстве или складе. Ежедневное производство шоколадок: ноль может быть только по аномалии (сломалось оборудование) или в случае забастовки. Обычно этих случаев 5-10 в течение нескольких лет и без восстановления спроса МОЖНО прогнозировать такие временные ряды (в ARIMA добавить член, учитывающий только определённый выброс Event).

- На реальных проектах подход к восстановлению спроса это вопрос консенсуса. Самые простые методы уже дадут 80%-90% результата: результат тут будет ошибка прогноза, который строится на восстановленном спросе.
- Т.к. процесс восстановления спроса самый базовый, с ним очень тяжело ставить эксперименты. По опыту, на ошибку прогноза будут сильнее влиять модели прогнозирования спроса, чем реализация методики восстановления спроса.
- Поэтому развитие методов восстановления спроса это вопрос больше бизнесовый, нежели исследовательский.

Иерархическое прогнозирование ВР

- Обзор иерархического прогнозирования (Линара К.) по мотивам https://otexts.org/fpp2/hts.html
- Как устроен этот подход?
- Особенности этого подхода.
 - Иерархия прогнозирования как её выбрать?
 - Уровень согласования как его выбрать?
 - Что может учесть этот подход?
 - Внешние факторы?
 - Каннибализацию?
 - Что он не может учесть?
 - Характеристики "соседних" рядов?
 - Его плюсы и минусы.

Визуализация случаев каннибализации

- Какие выводы мы можем сделать из визуализаций случаев каннибализации? (Илья П.)
 - Видны ли на данных типичные примеры каннибализации, т.е. когда в сумме продажи остаются примерно одинаковыми, но меняются пропорции между группами (ABC)?
 - Что происходит при вводе нового товара?
 - Какие "бизнес"-правила могут быть полезны для поиска (фильтрации) случаев каннибализации?

Домашнее задание

- Обзор статей по моделированию каннибализации.
 - https://www.evernote.com/shard/s250/sh/2c4136c1-ff1f-4d5c-9a15c3cb09d2c698/e00560308598133afe0f1dd706c38563
 - https://www.evernote.com/shard/s250/sh/9311c45f-8fd3-4b26-b6d0-166d61a9c874/05001e2c3984ca345c4e96cdf0307ecd