

Система прогнозирования потребительского спроса Goods4Cast

*О. В. Барина, А. С. Вальков, К. В. Воронцов,
С. А. Громов, А. Н. Ефимов, Ю. В. Чехович*
(Москва)

Задача прогнозирования потребительского спроса для розничной сети магазинов состоит в ежедневном построении прогнозов спроса на каждый товар из ассортимента каждого магазина сети. Прогнозы спроса используются при планировании закупок, планировании промо-акций (рекламных кампаний), управлении ценами и ассортиментом.

Данная прикладная задача имеет ряд существенных особенностей, накладывающих определённые ограничения на методы решения.

1. Количество прогнозируемых временных рядов имеет порядок 10^7 – 10^8 , при этом расчет прогнозов должен занимать не более нескольких часов. Проблема не может быть решена простым наращиванием вычислительных мощностей, требуется высокая скорость вычисления прогнозов и настройки алгоритмов прогнозирования.

2. Статистические характеристики временных рядов продаж существенно различны для разных товаров. Штучным товарам соответствуют дискретные ряды, весовым товарам — вещественные. Существуют товары, устойчиво продающиеся каждый день, однако подавляющее большинство товаров продаются не чаще, чем раз в неделю.

3. Прогнозируемые ряды не являются стационарными. Они могут претерпевать значительные разовые всплески (многократные изменения продаж во время промо-акций или праздников), либо, наоборот, падения до нуля (временное прекращение поставок). Они могут претерпевать как резкие, так и плавные изменения статистических характеристик, связанные с вытеснением одних товаров другими, изменением режима поставок, и другими факторами.

4. Завышение и занижение прогнозов приводят к различной величине потерь. Поэтому строить критерий качества прогнозирования на основе стандартной квадратичной функции потерь нецелесообразно.

5. Требуется прогнозировать временной ряд спроса, хотя наблюдается временной ряд продаж. Отсутствие продаж может быть вызвано не только отсутствием спроса, но и другими причинами: отсутствием товара на складе, недоступностью товара на полках, и т. д.

Перечисленными особенностями задачи продиктованы следующие особенности решения.

1. Для прогнозирования используются либо достаточно простые алгоритмы, время настройки которых не более чем линейно по длине ряда, либо динамически обучаемые (on-line) алгоритмы, параметры которых подстраиваются по мере поступления данных.

2. Поскольку такие алгоритмы прогнозирования являются, по сути дела, эвристическими и некорректными, используется техника построения алгоритмических композиций, основанная на идеях алгебраического подхода [3]. Для каждого временного ряда строится несколько альтернативных моделей алгоритмов. В простейшем случае из них выбирается алгоритм, дающий лучшее качество прогнозов по критерию скользящего контроля на последнем отрезке ряда. В более общем случае строится динамически адаптируемая композиция базовых алгоритмов [2]. Для каждого ряда в каждый момент времени наибольший вес получает наиболее адекватная модель. Это позволяет полностью автоматизировать построение прогнозов для большого количества нестационарных временных рядов, имеющих существенно различное поведение.

3. Перед настройкой алгоритмов производится фильтрация исторических данных. Из обучения исключаются интервалы аномально высоких и аномально низких продаж, связанные скорее с разовыми событиями, чем с регулярным спросом.

4. В качестве критерия настройки алгоритмов прогнозирования применяется неквадратичный несимметричный функционал потерь, оценивающий не абстрактную «точность прогноза», а величину потерь в рублях от завышенного или заниженного прогноза [1].

Рассматриваемые в докладе методы и подходы реализованы в автоматизированной системе прогнозирования потребительского спроса Goods4Cast. Данная система строит прогнозы для 140 магазинов и 250000 товаров. Время расчета прогнозов для одного магазина около 5 минут на рабочей станции 2×Xeon 3.2 ГГц, 3 Гб ОЗУ, 2×250 Гб raid.

Работа выполнена при поддержке РФФИ, проект № 04-01-08063-офи.

Литература

1. Барина О. В. Об одном методе прогнозирования временных рядов с несимметричным функционалом потерь // В настоящем сборнике. 2005.
2. Воронцов К. В., Егорова. Е. В. Динамически адаптируемые композиции алгоритмов прогнозирования // В настоящем сборнике. 2005.
3. Журавлев Ю. И. Корректные алгебры над множествами некорректных (эвристических) алгоритмов. I–III // Кибернетика, 1977. №4. С. 5–17, 1977. №6. С. 21–27, 1978. №2. С. 35–43.