

ПРОГНОЗИРОВАНИЕ ОБЪЁМА ПРОДАЖ НЕФТЕПРОДУКТОВ НА АЗС

Волобой В. С., студентка 4 курса кафедры информационно-аналитических систем СПбГУ, voloboylera@mail.ru

Аннотация

Грамотное прогнозирование объёма продаж является одной из самых важных концепций успеха любой компании. Для того, чтобы наиболее эффективно вести бизнес, необходимо правильно рассчитать прогноз, что позволит компании оптимизировать и контролировать расходы. Когда же речь идёт о продукции, то качественно проведённое прогнозирование объёмов продаж, позволит так же сформировать оптимальный запас продукции, что не повлечёт за собой нехватки или недостачи продукта.

В данной работе рассматривается возможность применения алгоритмов прогнозирования для создания ансамбля методов, который позволит строить ежедневный прогноз на различные промежутки времени объёмов продаж нефтепродуктов на АЗС компании ПАО "Газпром нефть".

Введение

Одна из важных составляющих успеха компании – качественное прогнозирование продаж. Правильно рассчитанный прогноз позволяет более эффективно вести бизнес, прежде всего, контролировать и оптимизировать расходы. Кроме того, если речь идет о продукции, это позволяет сформировать оптимальные (а не завышенные или заниженные) запасы продукции на складе.

По данным, предоставленным компанией ПАО "Газпром нефть" для создания прогноза объёма продаж нефтепродуктов на следующий день используется одна из встроенных функции Excel и тратится 4.5 человека/дня, что является очень дорогостоящим показателем. Поэтому можно выделить ряд прикладных задач, позволяющих упростить и модифицировать этот процесс. Одна их таких задач – это задача разработки ансамбля из разных моделей алгоритмов прогнозирования, который позволит выбрать лучший из возможных методов для каждого прогнозируемого объекта, сможет накапливать статистику о параметрах выбранной модели для

дальнейшего анализа, а так же позволит сделать процесс создания прогноза более автоматизированным.

Обзор

На данный момент в компании ПАО «Газпром нефть» реализован кейс по прогнозированию объемов продаж сопутствующих товаров и нефтепродуктов на АЗС на один месяц вперед. Создание подневного прогноза на более мелкие промежутки времени позволит получить более детальную картину для составления бизнес-стратегии и планирования.

Алгоритмы прогнозирования применялись к области составления прогноза объема продаж и ранее. В этих исследовательских работах рассматривались различные методы, такие как: ARIMA – модели, методы экспоненциального сглаживания, метод Хольта-Винтерса [1], [2], [3], [4], [5]. Однако многие из существующих исследовательских работ посвящены одному из методов прогнозирования и не рассматривают возможности создания ансамбля нескольких методов, из которых будет выбираться наилучшая модель. Ансамблирование методов позволит делать более точный прогноз, так как для различных рассматриваемых объектов, лучший результат могут показывать различные модели.

Также стоит отметить, что данные работы не подразумевают возможность накопления статистики параметров лучшей модели для рассматриваемых объектов, что существенно может облегчить задачу создания прогноза в случае потери или отсутствия входных данных.

В качестве вспомогательных пакетов для исследования использовались следующие пакеты языка R: `neuralnet` [6] и `forecast` [7].

Ход работы

В качестве набора данных используется детальная история продаж нефтепродуктов на АЗС компании ПАО «Газпром нефть» с 01.01.2015 по 31.03.2016.

В качестве исследуемых объектов были выделены следующие нефтепродукты:

1. Бензин Аи-92
2. Бензин Аи-95

Для подготовки данных были предприняты следующие шаги:

1. Агрегация: в данных присутствуют записи, которые необходимо агрегировать
2. Очистка: не рассматриваются временные ряды с пропущенными зна-

чениями

Прогноз на строился на следующие промежутки времени:

1. на 1 день вперёд
2. на 7 дней вперёд
3. на 14 дней вперёд
4. на 30 дней вперёд

К подготовленным данным применялись следующие алгоритмы:

1. Holt-Winters
2. ETS
3. ARIMA
4. Нейронные сети (прогноз строился на 1 день и на 7 дней вперёд)

Каждый из рассмотренных алгоритмов содержал в себе ещё несколько моделей, которые были обусловлены следующими показателями:

1. длиной моделируемого промежутка
2. длиной сезонного цикла
 - 7 – недельный
 - 14 – двухнедельный
 - 30 – месячный
 - 365 - годовой

Для обучения и тестирования нейронных сетей было выбрано несколько моделей, принимающих в качестве входного сигнала различные варианты данных:

1. по 3-м предыдущим дням
2. по 7-ми предыдущим дням
3. по 14-ти предыдущим дням
4. по 7-ми предыдущим дням через 7

Для накопления и анализа статистики в дальнейшем, для каждой АЗС сохраняются данные о следующих параметрах:

1. лучший алгоритм
2. лучшая длина моделируемого промежутка
3. лучшее количество скрытых нейронов
4. лучшая длина сезонного цикла

Понятие лучшего параметра основано на оценке абсолютной средней процентной ошибки MAPE.

$$MAPE = \frac{1}{N} * \sum_{i=1}^N \frac{|X_i - \hat{X}_i|}{X_i}, \text{ где } \hat{X}_i - \text{спрогнозированное значение,}$$

X_i - фактическое значение, N – длина прогноза.

Результаты

Рассмотренные алгоритмы показали следующие результаты MAPE (показанные здесь результаты усреднены по продукту, длине прогнозируемого промежутка и отделению):

1. Holt-Winters : 6.8%
2. ETS: 8.1%
3. ARIMA: 8.6 %
4. Нейронные сети: 10.4%

На основании полученных результатов были сделаны следующие выводы:

1. Точность прогноза Бензина АИ-92 в 100 % случаях выше точности прогноза Бензина АИ-95
2. Прогноз АЗС отделения Санкт-Петербург лучше прогноза АЗС отделения Москва в 95% случаях
3. Точность прогнозирования нейронных сетей на 1 день значительно хуже точности других алгоритмов

Заключение

В данной статье была рассмотрена задача прогнозирования объёмов продаж основных нефтепродуктов на АЗС компании ПАО "Газпром нефть". Были рассмотрены некоторые возможности подготовки входных данных, а так же применены различные модели алгоритмов прогнозирования, объединённые в единый ансамбль, позволяющий накапливать статистику о моделях, показавших лучшие результаты.

Для улучшения качества прогноза в дальнейшем можно учитывать факт изменения формата АЗС, который зависит от площади, занимаемой автозаправочной станцией; в данных не учитываются возможные простые АЗС, учёт которых позволил бы не исключать из рассмотрения временные

ряды с пропущенными значениями, а так же возможно введение меток для входных данных, таких как: “праздничный день”, “пятница” и др., что позволит сделать дополнительный акцент на характере временного ряда.

Литература

- [1] Chatfield, C, (2004). The analysis of time series: An introduction, 6th edition, Boca Raton: Chapman & Hall/CRC Press.
- [2] Holt, C.C. (2004a). Forecasting seasonals and trends by exponentially weighted moving averages, International Journal of Forecasting, 20, 5-10.
- [3] Nau Robert, (2015), Statistical forecasting: notes on regression and time series analysis
- [4] Prajakta S. Kalekar Time series Forecasting using Holt-Winters Exponential Smoothing // Kanwal Rekhi School of Information Technology, 2006.
- [5] Rob J. Hyndman, Anne B. Koehler, J. Keith Ord, Ralph Snyder, (2008). Forecasting with Exponential Smoothing, The State Space Approach, Springer
- [6] Package ‘neuralnet’ // cran.r-project.org URL: <https://cran.r-project.org/web/packages/neuralnet/neuralnet.pdf> (дата обращения: 10.02.2016).
- [7] Package ‘forecast’ // cran.r-project.org URL: <https://cran.r-project.org/web/packages/forecast/forecast.pdf> (дата обращения: 02.02).