

МЕТОДЫ И МОДЕЛИ ПРОГНОЗИРОВАНИЯ ПОКАЗАТЕЛЕЙ ХОЗЯЙСТВЕННОЙ ДЕЯТЕЛЬНОСТИ

УДК 519.24

ЛЕПА Евгений Владимирович

к.т.н., доцент кафедры информационных технологий Херсонского национального технического университета.

Научные интересы: системы принятия и поддержки решений, технологии интеллектуального анализа данных.

e-mail: e.lepa@mail.ru

ВВЕДЕНИЕ

В организациях осуществляется постоянная фиксация показателей, отражающих их деятельность в определенные, периодические моменты времени, например, ежемесячно, ежесуточно и т.д. К таким показателям могут относиться объемы производства, продаж и предоставления услуг, как в количественном, так и денежном выражении. Такие данные называются временным рядом и отображают результат деятельности организации за прошедший период [1].

На основании временных рядов можно прогнозировать значение показателя на будущие периоды времени. Для обеспечения хороших результатов, исходные данные необходимо очистить, выполнив предварительно парциальную обработку.

Цели и задачи исследования. На основании временного ряда, отражающего объем продаж за прошедший период необходимо выполнить прогноз продаж на будущие периоды. Для прогнозирования использованы методы, реализованные в аналитической платформе Deductor [2]. Сравнение результатов, полученных различными методами, позволяет получить достаточно обоснованный прогноз.

ИССЛЕДОВАНИЕ И АНАЛИЗ РЕЗУЛЬТАТОВ

При прогнозировании на основе временных рядов любым методом, нужно подавать на вход модели значения нескольких сопредельных значений из исходного набора данных. Такой метод отбора данных называется скользящим окном [3]. Окно выделяет только некоторый непрерывный участок данных и «скользит» (перемещается) по всем набору данных.

При этом эффективность реализации заметно повышается, если не выбирать данные каждый раз из нескольких последовательных записей, а последовательно расположить данные, относящиеся к конкретной позиции окна, в одной записи.

Значения в одном из полей записи будут относиться к текущему отсчету, а в других – смещены от текущего отсчета «в будущее» или «в прошлое».

Таким образом, преобразование скользящего окна имеет два параметра:

- глубина погружения – количество «прошлых» отсчетов, попадающих в окно,
- горизонт прогнозирования – количество «будущих» отсчетов.

Для правильного формирования скользящего окна данные должны быть соответствующим образом упорядочены, например, для временных рядов они должны быть в порядке возрастания даты.

Если задать глубину погружения 2 и горизонт прогнозирования 1, то будет получена таблица с неполными записями (табл. 1).

Выделенная часть таблицы представляет собой таблицу с полными записями.

Таким образом, имея достаточную длину временного ряда, можно с помощью скользящего окна получить необходимые данные для прогноза, где входными факторами для модели могут быть продажи за текущий месяц, продажи за месяц ранее и т.д., а результатом (выходным или прогнозируемым фактором) должны быть продажи на следующий месяц.

Таблиця 1 –

Применение скользящего окна

Первый день месяца	Объем продаж 2 месяца назад	Объем продаж 1 месяц назад	Объем продаж в текущем месяце	Объем продаж на следующий месяц
				1000
январь			1000	1160
февраль		1000	1160	1210
март	1000	1160	1210	1130
апрель	1160	1210	1130	1250
май	1210	1130	1250	1300
июнь	1130	1250	1300	
	1250	1300		
	1300			

Часто временные ряды имеют периодичность (повторяемость), что выражается в падении или росте показателя. Для торговли, например, это связано с сезонностью продаж какого-либо товара (летом возрастает объем продаж прохладительных напитков).

Наличие сезонности во временных рядах определяется с помощью обработчика Автокорреляция. Если, например, обнаружена сезонность продаж с периодом

12 месяцев, то для прогнозирования количества продаж на месяц вперед можно в качестве входного фактора указать не только значение количества продаж за предыдущий месяц, но и за 12 месяцев назад. Наличие годовой сезонности позволяет строить прогноз на месяц вперед, основываясь на данных за 1, 2, 11 и 12 месяцев назад. Поэтому требуется выбрать глубину погружения 12.

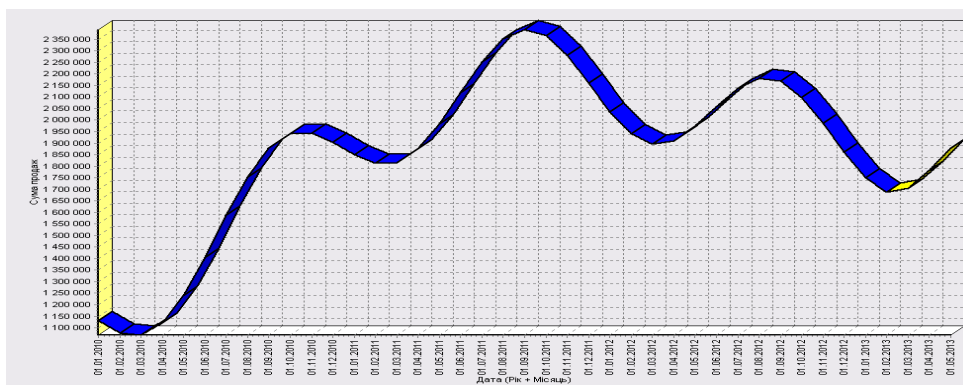


Рисунок 1 – Диаграмма прогнозирования

Архитектура нейронной сети для прогнозирования состоит из трех слоев – входного, выходного и одного скрытого (внутреннего). На входном слое будет четыре нейрона по количеству входных полей данных. На выходном слое один нейрон, соответствующий одному выходному параметру Сумма продаж. На скрытом слое два нейрона.

Качество моделей в используемой аналитической платформе можно оценить двумя способами.

1. Сравнить диаграммы исходных данных и данных, полученных с помощью модели сети за те же пе-

риоды времени. При хорошем качестве они должны практически совпадать.

2. Построить диаграмму рассеивания, которая служит для наглядной оценки качества обучения модели с помощью результатов сравнения непрерывных значений выходного поля и непрерывных значений того же поля, но рассчитанных по модели.

Применение этих способов показали, что для конкретных исходных данных качество нейронной сети достаточно высокое, и она может быть использована для прогнозирования. В результате получена диаграм-

ма прогнозирования для нейронной сети конкретной архитектуры (рис. 1).

На диаграмме изображены значения сумм продаж за прошедшие периоды времени (месяцы) и на три периода вперед (горизонт прогнозирования). Для указанных периодов определены числовые результаты прогнозирования (табл. 2).

Таблица 2 –

**Результаты прогнозирования
(нейронная сеть)**

Период	1 период	2 период	3 период
Сумма продаж	1704168	1780701	1879621

Таблица 3 –

**Результаты прогнозирования (уравнение
регрессии)**

Период	1 период	2 период	3 период
Сумма продаж	1668264	1695525	1767912

Для сравнения и оценки качества прогнозирования была построена модель в виде уравнения регрессии, предполагая, что прогнозируемый процесс близок к линейному. При построении линейной модели использованы те же исходные данные, полученные приведением к скользящему окну.

Результат построения линейной модели можно представить в виде уравнения регрессии

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 \quad (1)$$

где a_i – коэффициенты регрессии;

x_i – входные факторы;

y – выходной параметр (результат).

В окончательном виде уравнение регрессии имеет вид.

$$y = 138720 - 0,459x_1 + 0,46403x_2 - 0,60155x_3 + 1,5207x_4$$

На основании уравнения регрессии построена сравнительная диаграмма исходных данных и данных, полученных на основании линейной модели, а также

диаграмма рассеивания. Все точки (исходные и рассчитанные) находятся практически на линии идеальных значений и не выходят за границы доверительного интервала. Диаграмма рассеивания для модели в виде нейронной сети имеет большие отклонения от линии идеальных значений. Это дает основание предполагать, что модель линейной регрессии позволяет получить более достоверный прогноз, на который и следует ориентироваться.

Получены результат прогнозирования на основе уравнения регрессии (табл. 3).

Чтобы повысить достоверность прогноза, можно получить и другие подходящие модели. В конкретном случае, автокорреляционная функция достаточно гладкая, что позволяет сделать предположение о наличии тренда в данных. Таким образом, можно построить еще и трендовую модель прогнозирования.

Какой прогноз выбрать полностью зависит от лица принимающего решение, который должен хорошо знать весь процесс деятельности организации, а также характеризующие ее показатели.

ВЫВОДЫ

Для прогнозирования показателей хозяйственной деятельности организации использованы две модели в виде нейронной сети и линейного уравнения регрессии, реализованных в аналитической платформе Deductor. Выполнен сравнительный анализ качества полученных моделей сравнением исходных данных и данных, полученных на основании моделей, а также диаграмм рассеивания. Для конкретных исходных данных было определено, что модель в виде уравнения регрессии лучше, чем модель в виде нейронной сети.

Получены количественные характеристики прогнозируемого показателя на три периода.

Использование рассмотренных методов прогнозирования улучшить планирование деятельности организации и повышение эффективности ее работы.

ЛИТЕРАТУРА:

1. Dyuk V., Samoilenko P. Data Mining. Educational course. – SPb.: Piter, 2002. – 368 s.
2. Guidance of analyst. Version of 5.2. – M.: Basegroup Labs, 2010. – 122 s.
3. Palkin N.B., Oreshkov V.I. Business of analyst: from information to knowledges. – SPb.: Piter, 2010. – 352 s.

Рецензент: д.т.н., проф. Соколова Н.А.,
Херсонский национальный технический университет.