ЛР\_5

**Прогнозирование данных на основе временного ряда**

Прогнозирование результата на определенное время вперед, основываясь на данных за прошедшее время – задача, встречающаяся довольно часто. Например, перед большинством торговых фирм стоит задача оптимизации складских запасов, для решения которой требуется знать, чего и сколько должно быть продано через неделю и т.п.; задача предсказания стоимости акций какого-нибудь предприятия через день и т.д. *Deductor Studio* предлагает для этого инструмент «Прогнозирование».

Прогнозирование появляется в списке мастера обработки только после построения какой-либо модели прогноза: нейросети, линейной регрессии и т.д. Прогнозировать на несколько шагов вперед имеет смысл только временной ряд (к примеру, если есть данные по недельным суммам продаж за определенный период, можно спрогнозировать сумму продаж на две недели вперед). Поскольку при построении модели прогноза необходимо учитывать много факторов (зависимость результата от данных день, два, три, четыре назад), то методика имеет свои особенности.

1. Моделирование и ввод исходных данных

Вначале производится моделирование в Excel, затем копируются 2 выделенных серым цветом колонки в текстовый редактор.

x\_t=0,2t+0,2tsin(t)+N(0.0,5)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| t | 0,2\*t | sin t | 0,2\*t\*sin(t) | N(0. 0,5) | Row |  | Арг | Ряд |
| 1 | 0,2 | 0,841471 | 0,1682942 | 1,084045 | 1,5 |  | 1 | 1 |
| 2 | 0,4 | 0,909297 | 0,36371897 | 0,393287 | 1,2 |  | 2 | 1 |
| 3 | 0,6 | 0,14112 | 0,084672 | -0,12923 | 0,6 |  | 3 | 1 |
| 4 | 0,8 | -0,7568 | -0,605442 | 0,825453 | 1,0 |  | 4 | 1 |
| 5 | 1 | -0,95892 | -0,95892427 | -0,2291 | -0,2 |  | 5 | 0 |
| 6 | 1,2 | -0,27942 | -0,3352986 | 0,989876 | 1,9 |  | 6 | 2 |
| 7 | 1,4 | 0,656987 | 0,91978124 | 0,147677 | 2,5 |  | 7 | 2 |
| 8 | 1,6 | 0,989358 | 1,58297319 | 0,21563 | 3,4 |  | 8 | 3 |
| 9 | 1,8 | 0,412118 | 0,74181327 | 0,433672 | 3,0 |  | 9 | 3 |
| 10 | 2 | -0,54402 | -1,08804222 | -0,13272 | 0,8 |  | 10 | 1 |
| 11 | 2,2 | -0,99999 | -2,19997845 | -0,06924 | -0,1 |  | 11 | 0 |
| 12 | 2,4 | -0,53657 | -1,287775 | 0,476022 | 1,6 |  | 12 | 2 |
| 13 | 2,6 | 0,420167 | 1,0924343 | 0,189922 | 3,9 |  | 13 | 4 |
| 14 | 2,8 | 0,990607 | 2,7737006 | -0,77458 | 4,8 |  | 14 | 5 |
| 15 | 3 | 0,650288 | 1,95086352 | -0,17439 | 4,8 |  | 15 | 5 |
| 16 | 3,2 | -0,2879 | -0,92129061 | 0,623152 | 2,9 |  | 16 | 3 |
| 17 | 3,4 | -0,9614 | -3,26875147 | -0,26085 | -0,1 |  | 17 | 0 |
| 18 | 3,6 | -0,75099 | -2,70355409 | 0,185898 | 1,1 |  | 18 | 1 |
| 19 | 3,8 | 0,149877 | 0,5695334 | 0,323199 | 4,7 |  | 19 | 5 |
| 20 | 4 | 0,912945 | 3,651781 | -0,66525 | 7,0 |  | 20 | 7 |
| 21 | 4,2 | 0,836656 | 3,51395368 | 0,068314 | 7,8 |  | 21 | 8 |
| 22 | 4,4 | -0,00885 | -0,03894576 | 0,47446 | 4,8 |  | 22 | 5 |
| 23 | 4,6 | -0,84622 | -3,89261386 | 0,638842 | 1,3 |  | 23 | 1 |
| 24 | 4,8 | -0,90558 | -4,34677614 | 0,552413 | 1,0 |  | 24 | 1 |
| 25 | 5 | -0,13235 | -0,66175875 | -0,18381 | 4,2 |  | 25 | 4 |
| 26 | 5,2 | 0,762558 | 3,96530394 | -0,06963 | 9,1 |  | 26 | 9 |
| 27 | 5,4 | 0,956376 | 5,16443001 | 0,430671 | 11,0 |  | 27 | 11 |
| 28 | 5,6 | 0,270906 | 1,51707241 | -0,10349 | 7,0 |  | 28 | 7 |
| 29 | 5,8 | -0,66363 | -3,84907653 | -0,28297 | 1,7 |  | 29 | 2 |
| 30 | 6 | -0,98803 | -5,92818974 | 0,254903 | 0,3 |  | 30 | 0 |
| 31 | 6,2 | -0,40404 | -2,5050334 | 0,05675 | 3,8 |  | 31 | 4 |
| 32 | 6,4 | 0,551427 | 3,52913076 | 0,691755 | 10,6 |  | 32 | 11 |
| 33 | 6,6 | 0,999912 | 6,59941828 | -0,74754 | 12,5 |  | 33 | 12 |
| 34 | 6,8 | 0,529083 | 3,59776227 | -0,38413 | 10,0 |  | 34 | 10 |
| 35 | 7 | -0,42818 | -2,99727869 | -0,35838 | 3,6 |  | 35 | 4 |
| 36 | 7,2 | -0,99178 | -7,14080774 | -0,15758 | -0,1 |  | 36 | 0 |
| 37 | 7,4 | -0,64354 | -4,76218219 | -0,83012 | 1,8 |  | 37 | 2 |
| 38 | 7,6 | 0,296369 | 2,2524012 | -0,24779 | 9,6 |  | 38 | 10 |
| 39 | 7,8 | 0,963795 | 7,51760401 | 0,050748 | 15,4 |  | 39 | 15 |
| 40 | 8 | 0,745113 | 5,96090528 | 0,346997 | 14,3 |  | 40 | 14 |
| 41 | 8,2 | -0,15862 | -1,30070588 | -0,93376 | 6,0 |  | 41 | 6 |
| 42 | 8,4 | -0,91652 | -7,698781 | -1,09429 | -0,4 |  | 42 | 0 |
| 43 | 8,6 | -0,83177 | -7,15326279 | -0,10142 | 1,3 |  | 43 | 1 |
| 44 | 8,8 | 0,017702 | 0,15577694 | -0,26093 | 8,7 |  | 44 | 9 |
| 45 | 9 | 0,850904 | 7,65813172 | -0,15934 | 16,5 |  | 45 | 16 |
| 46 | 9,2 | 0,901788 | 8,2964528 | -0,03616 | 17,5 |  | 46 | 17 |
| 47 | 9,4 | 0,123573 | 1,16158735 | -0,24688 | 10,3 |  | 47 | 10 |
| 48 | 9,6 | -0,76825 | -7,37524475 | 0,202273 | 2,4 |  | 48 | 2 |
| 49 | 9,8 | -0,95375 | -9,346776 | 0,376567 | 0,8 |  | 49 | 1 |
| 50 | 10 | -0,26237 | -2,62374854 | -0,42181 | 7,0 |  | 50 | 7 |
| 51 | 10,2 | 0,670229 | 6,83633759 | -0,18328 | 16,9 |  | 51 | 17 |
| 52 | 10,4 | 0,986628 | 10,260927 | 0,413788 | 21,1 |  | 52 | 21 |
| 53 | 10,6 | 0,395925 | 4,19680659 | 0,142929 | 14,9 |  | 53 | 15 |
| 54 | 10,8 | -0,55879 | -6,03492173 | 0,201194 | 5,0 |  | 54 | 5 |
| 55 | 11 | -0,99976 | -10,9973069 | 0,547233 | 0,5 |  | 55 | 1 |
| 56 | 11,2 | -0,52155 | -5,84137122 | -0,29961 | 5,1 |  | 56 | 5 |
| 57 | 11,4 | 0,436165 | 4,97227821 | -0,25447 | 16,1 |  | 57 | 16 |
| 58 | 11,6 | 0,992873 | 11,5173227 | 0,487091 | 23,6 |  | 58 | 24 |
| 59 | 11,8 | 0,636738 | 7,51350848 | 0,162164 | 19,5 |  | 59 | 19 |
| 60 | 12 | -0,30481 | -3,65772745 | 0,55178 | 8,9 |  | 60 | 9 |
| 61 | 12,2 | -0,96612 | -11,7866368 | 0,045176 | 0,5 |  | 61 | 0 |
| 62 | 12,4 | -0,73918 | -9,16584064 | -0,44823 | 2,8 |  | 62 | 3 |
| 63 | 12,6 | 0,167356 | 2,10868182 | -0,25477 | 14,5 |  | 63 | 14 |
| 64 | 12,8 | 0,920026 | 11,7763333 | -0,68212 | 23,9 |  | 64 | 24 |
| 65 | 13 | 0,826829 | 10,7487728 | 0,400468 | 24,1 |  | 65 | 24 |
| 66 | 13,2 | -0,02655 | -0,35047523 | 0,17317 | 13,0 |  | 66 | 13 |
| 67 | 13,4 | -0,85552 | -11,4639677 | -0,04295 | 1,9 |  | 67 | 2 |
| 68 | 13,6 | -0,89793 | -12,2118165 | 0,212402 | 1,6 |  | 68 | 2 |
| 69 | 13,8 | -0,11478 | -1,58403043 | -0,68115 | 11,5 |  | 69 | 12 |
| 70 | 14 | 0,773891 | 10,8344695 | -1,34942 | 23,5 |  | 70 | 23 |
| 71 | 14,2 | 0,951055 | 13,5049761 | -0,01268 | 27,7 |  | 71 | 28 |
| 72 | 14,4 | 0,253823 | 3,65505642 | -0,05102 | 18,0 |  | 72 | 18 |
| 73 | 14,6 | -0,67677 | -9,88087057 | -0,19991 | 4,5 |  | 73 | 5 |
| 74 | 14,8 | -0,98515 | -14,5801647 | 0,086587 | 0,3 |  | 74 | 0 |
| 75 | 15 | -0,38778 | -5,81672453 | 0,962614 | 10,1 |  | 75 | 10 |
| 76 | 15,2 | 0,566108 | 8,60483608 | 0,593566 | 24,4 |  | 76 | 24 |
| 77 | 15,4 | 0,99952 | 15,3926104 | -0,36165 | 30,4 |  | 77 | 30 |
| 78 | 15,6 | 0,513978 | 8,01806391 | -0,65627 | 23,0 |  | 78 | 23 |
| 79 | 15,8 | -0,44411 | -7,01698017 | 0,394277 | 9,2 |  | 79 | 9 |
| 80 | 16 | -0,99389 | -15,9022185 | -0,20398 | -0,1 |  | 80 | 0 |
| 81 | 16,2 | -0,62989 | -10,2041855 | 1,011999 | 7,0 |  | 81 | 7 |
| 82 | 16,4 | 0,313229 | 5,13695203 | -0,70587 | 20,8 |  | 82 | 21 |
| 83 | 16,6 | 0,968364 | 16,0748501 | -0,06928 | 32,6 |  | 83 | 33 |
| 84 | 16,8 | 0,73319 | 12,3175974 | 0,325986 | 29,4 |  | 84 | 29 |
| 85 | 17 | -0,17608 | -2,99328554 | 0,931791 | 14,9 |  | 85 | 15 |
| 86 | 17,2 | -0,92346 | -15,8834853 | 0,016776 | 1,3 |  | 86 | 1 |
| 87 | 17,4 | -0,82182 | -14,2996304 | 0,919276 | 4,0 |  | 87 | 4 |
| 88 | 17,6 | 0,035398 | 0,62301013 | 0,098416 | 18,3 |  | 88 | 18 |
| 89 | 17,8 | 0,860069 | 15,3092354 | 0,163172 | 33,3 |  | 89 | 33 |
| 90 | 18 | 0,893997 | 16,0919399 | 0,754517 | 34,8 |  | 90 | 35 |
| 91 | 18,2 | 0,105988 | 1,92897271 | 0,480554 | 20,6 |  | 91 | 21 |
| 92 | 18,4 | -0,77947 | -14,3421757 | 0,171627 | 4,2 |  | 92 | 4 |
| 93 | 18,6 | -0,94828 | -17,6380478 | -0,11642 | 0,8 |  | 93 | 1 |
| 94 | 18,8 | -0,24525 | -4,61073733 | -0,84015 | 13,3 |  | 94 | 13 |
| 95 | 19 | 0,683262 | 12,9819726 | -0,34968 | 31,6 |  | 95 | 32 |
| 96 | 19,2 | 0,983588 | 18,8848847 | 0,522097 | 38,6 |  | 96 | 39 |
| 97 | 19,4 | 0,379608 | 7,36439014 | -0,20572 | 26,6 |  | 97 | 27 |
| 98 | 19,6 | -0,57338 | -11,2382847 | 0,26102 | 8,6 |  | 98 | 9 |
| 99 | 19,8 | -0,99921 | -19,7842953 | -0,2357 | -0,2 |  | 99 | 0 |
| 100 | 20 | -0,50637 | -10,1273128 | -0,1737 | 9,7 |  | 100 | 10 |

1. Мастер импорта

Показаны некоторые шаги работы.

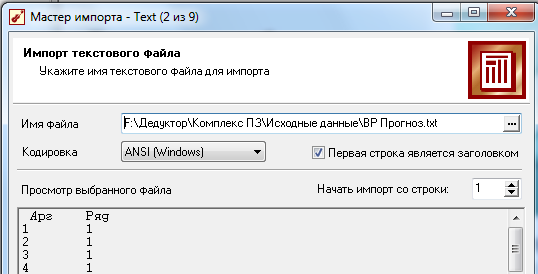


Рис.1 Импорт файла

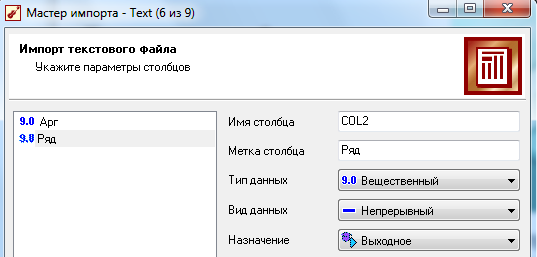


Рис.2 Указание параметров столбцов

При отображении на 8 шаге указать в качестве визуализаторов Таблица и Диаграмма. На 9 шаге отметить Ряд (рис.3).

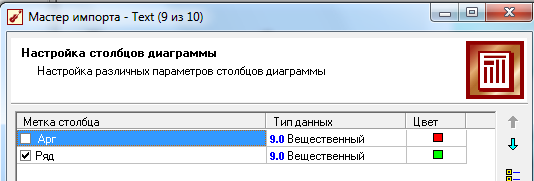


Рис.3 Настройка столбцов диаграммы

На 10 шаге вставить следующую фразу: «Импорт примера для демонстрации предобработки данных».

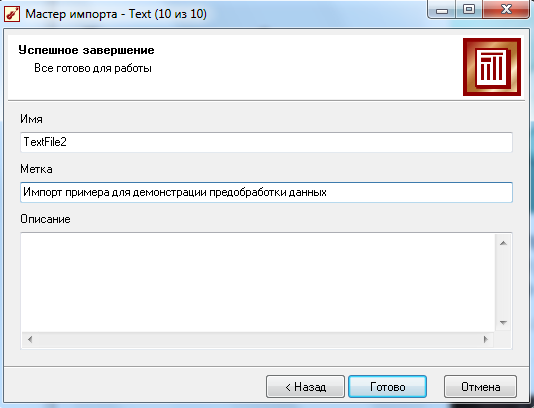


Рис.4 Последний шаг работы мастера импорта

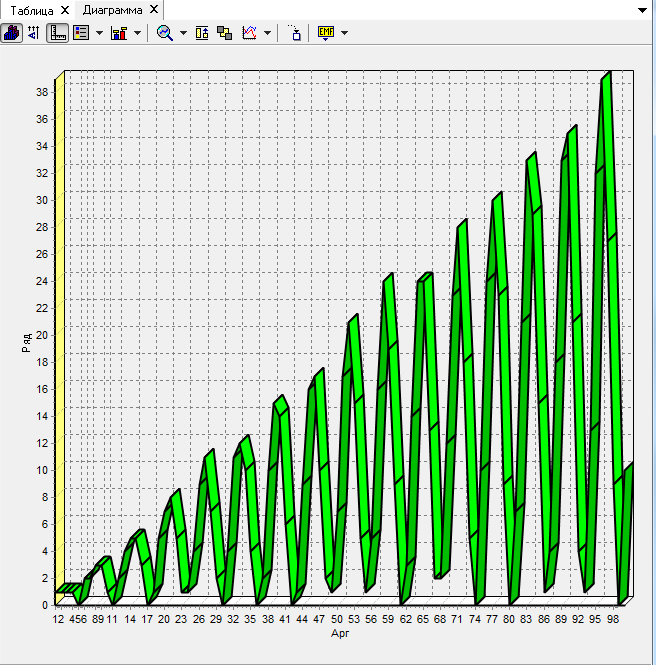


Рис.5 Исходный ряд

1. Мастер обработки. Устранение аномалий

Перед прогнозированием необходимо удалить аномалии и сгладить данные. Сделать это можно при помощи спектральной обработки. Запустим мастер обработки (рис.6), выберем в качестве обработки данных спектральную обработку и перейдем на следующий шаг мастера. Параметр Ряд выделим как Используемый.

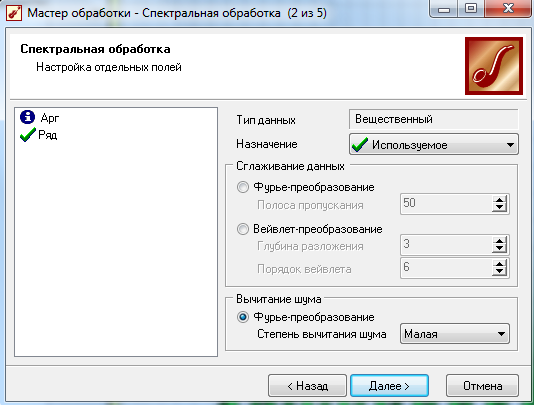


Рис.6 Переход к спектральной обработке

Следующий шаг отвечает за удаление аномалий из исходного набора. Выберем поле для обработки «Вычитание шума» (степень вычитания – малая).

В качестве визуализаторов выберем Таблица и Диаграмма. На шаге 5 отметим Ряд и получим на рис.7 результат спектральной обработки.



Рис.7 Результат спектральной обработки

Видно, что данные сгладились, аномалии и шумы исчезли. Также видна тенденция. Теперь перед аналитиком встает вопрос, а как, собственно, прогнозировать временной ряд. Имеется только один столбец данных.

Строить прогноз на будущее необходимо, основываясь на данных прошлых периодов. Предполагается, что количество продаж на следующий месяц зависит от количества продаж за предыдущие месяцы. Входными факторами для модели могут быть продажи за текущий месяц, продажи за месяц ранее и т.д., а результатом должны быть продажи за следующий месяц. Здесь явно необходимо трансформировать данные к *скользящему окну*.

Запустим мастер обработки, выберем в качестве обработчика скользящее окно (рис.8) и перейдем на следующий шаг.

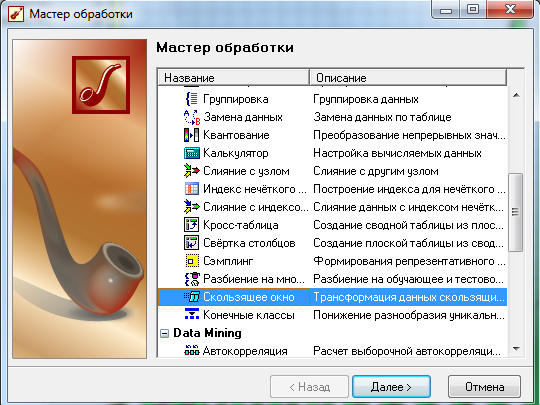


Рис.8 Выбор скользящего окна

Построим прогноз на неделю вперед, основываясь на данных за 12, 11 месяцев назад, два месяца назад и месяц назад. Поэтому необходимо, назначив поле «Ряд» используемым, выбрать глубину погружения 12. Тогда данные трансформируются к скользящему окну (рис.9).

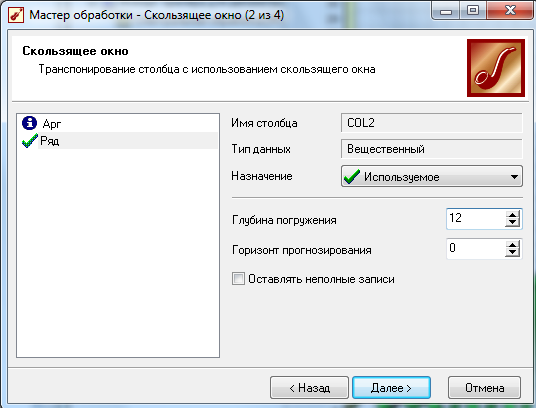


Рис.9 Переход к скользящему окну

Просмотреть полученные данные можно в виде таблицы (рис.10). Как видно, теперь в качестве входных факторов можно использовать «Ряд- 12», «Ряд - 11» - данные по количеству 12 и 11 месяцев назад (относительно прогнозируемого месяца) и остальные необходимые факторы. В качестве результата прогноза буден указан столбец «Ряд».

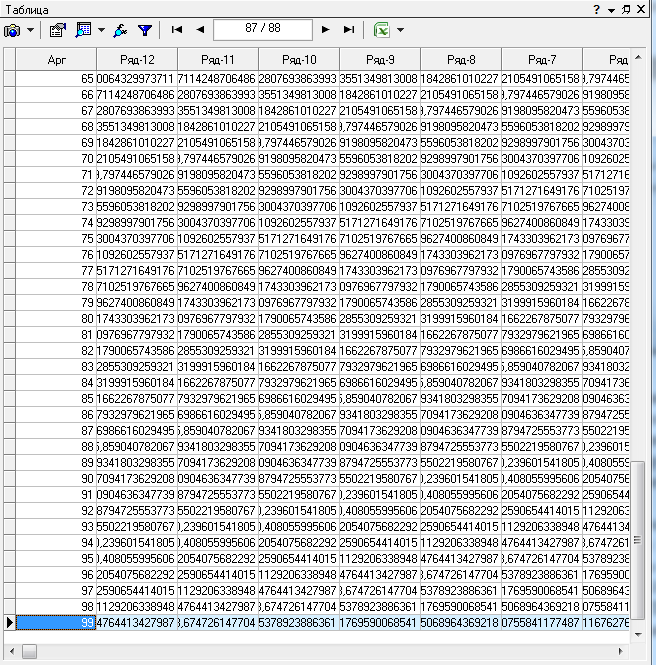


Рис.10 Таблица скользящего окна

1. Мастер обработки. Нейронная сеть

Перейдем непосредственно к самому построению модели прогноза. Откроем мастер обработки и выберем в нем нейронную сеть. На втором шаге мастера, согласно с принятым ранее решениям, установим в качестве входных поля «Ряд-12», «Ряд -11», «Ряд -2» и «Ряд -1», а в качестве выходного - «Ряд». Остальные поля сделаем информационными (рис.11).

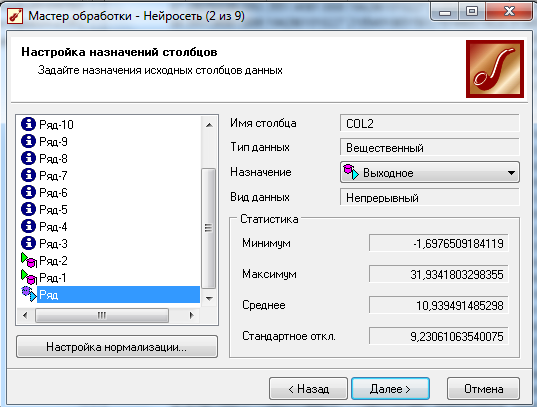


Рис.11 Настройка столбцов

Оставив остальные параметры построения модели по умолчанию, только количество нейронов скрытого слоя поставим равным 5 (рис.12).

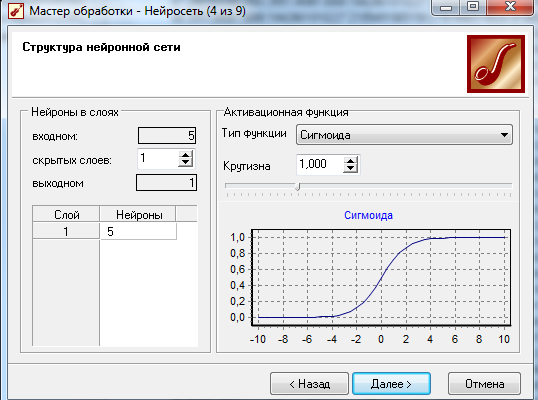


Рис.12 Структура нейронной сети

Далее обучим нейросеть (рис.13, 14).

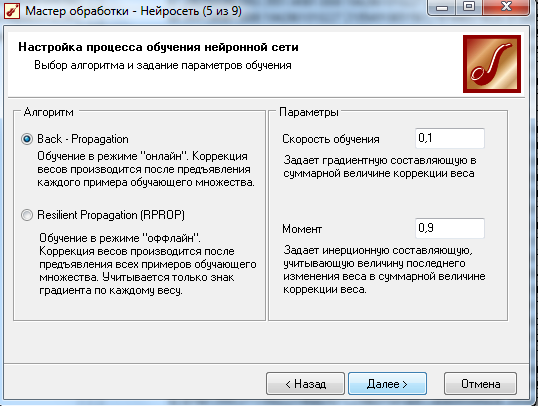


Рис.13 Выбор метода обучения

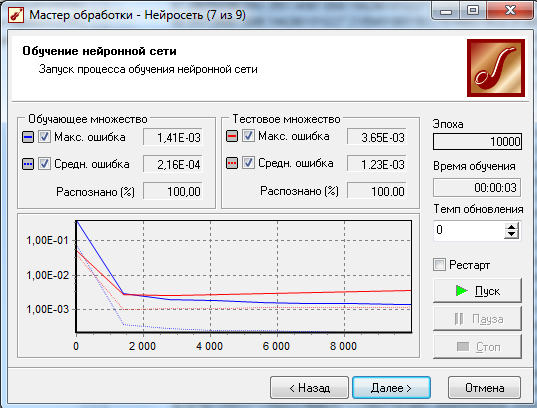


Рис.14 Обучение нейросети

После построения модели для просмотра качества обучения представим полученные данные в виде диаграммы, диаграммы рассеяния и графа нейросети. В мастере настройки диаграммы выберем для отображения поля «Ряд» и «Ряд\_OUT» - реальное и спрогнозированное значение (рис.15).

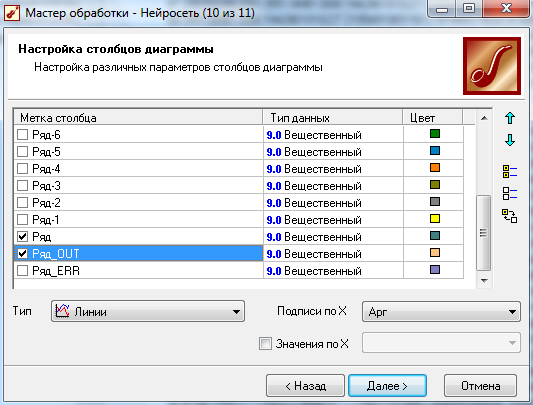


Рис.15 Выбор полей отображения

Результат работы - графики, показанные на рис.16-18.

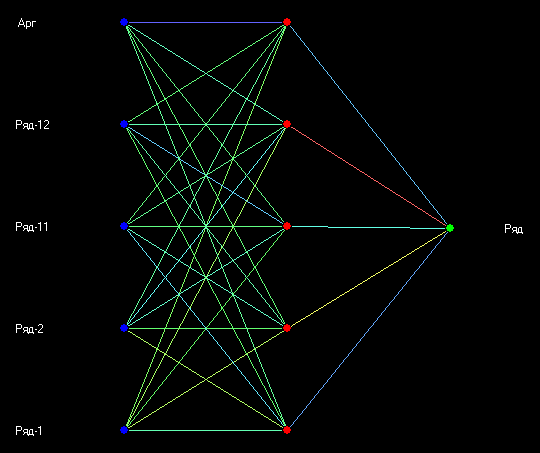


Рис.16 Граф нейросети

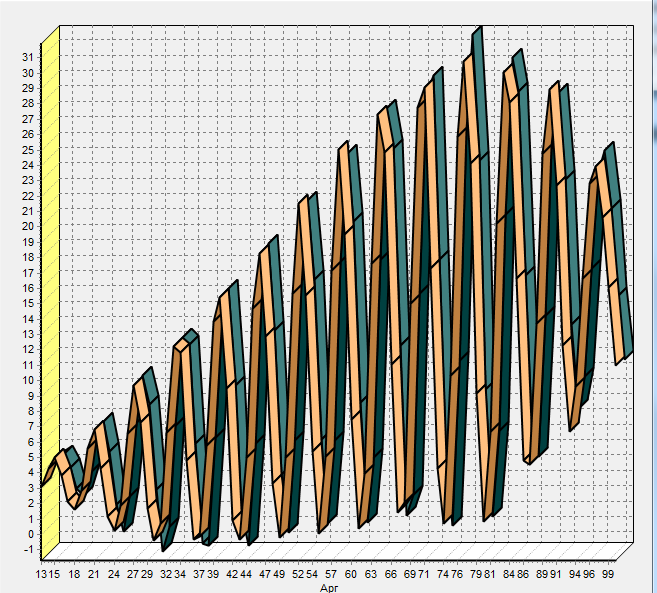


Рис.17 График ряда

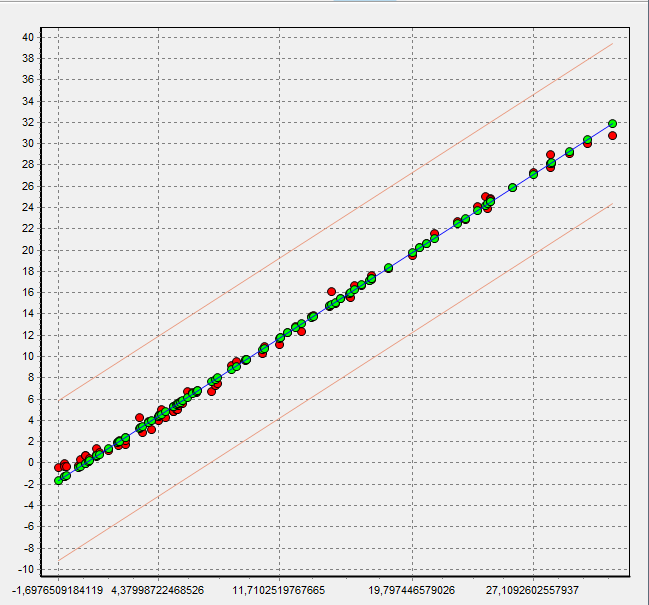


Рис.18 Диаграмма рассеяния

Диаграмма рассеяния более наглядно показывает качество обучения.

1. Мастер обработки. Прогнозирование

Нейросеть обучена, теперь осталось самое главное – построить требуемый прогноз. Для этого открываем мастер обработки (рис. 19) и выбираем появившийся теперь обработчик «Прогнозирование».

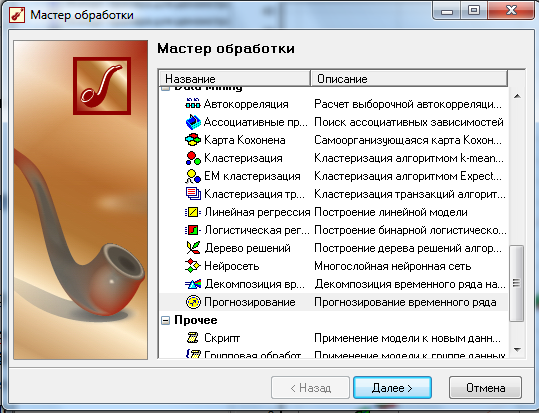


Рис.19 Выбор обработчика «Прогнозирование»

На втором шаге мастера предлагается настроить связи столбцов для прогнозирования временного ряда – откуда брать данные для столбца при очередном шаге прогноза (рис.20).

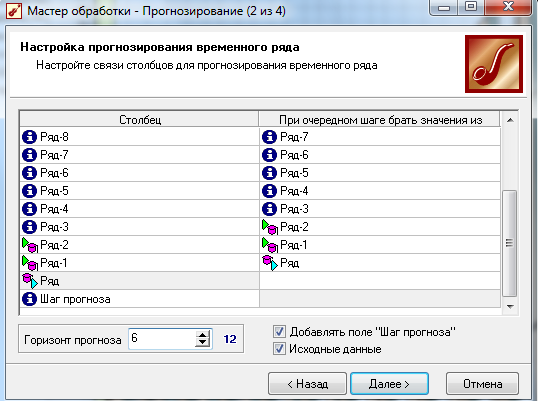


Рис.20 Настройка связей столбца

В итоге сделаем прогноз на 6 временных интервалов вперед (рис.21).

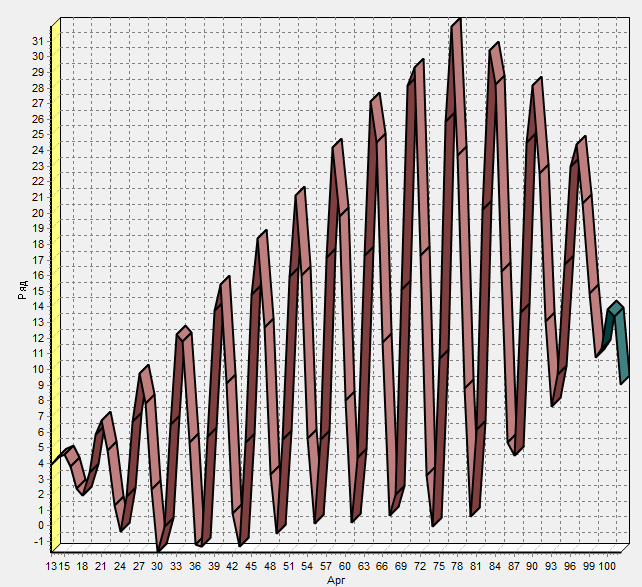


Рис.21 Прогноз на 6 точек вперед

На рис.22 приведен отрезок спрогнозированного ряда с указанными значениями.

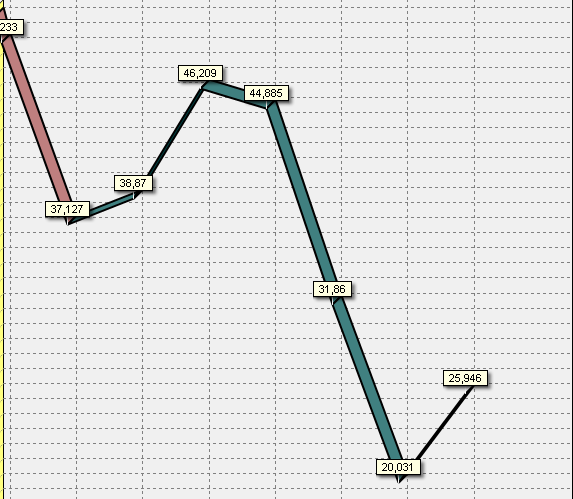


Рис.23 Отрезок спрогнозированного ряда

После завершения анализа данные можно экспортировать. Данные можно выгрузить либо в текстовый файл, либо в собственный проект программы с расширением «\*.*ded*». Мастер экспорта имеет точно такие же настройки, как и мастер импорта. Более того, если экспорт данных совершить в текстовый фал, то далее данные можно скопировать в файл табличного процессора *Excel* и достаточно комфортно с ними работать.

**Варианты заданий**

1. ВР разыгрывается по следующей формуле

xt = 0,2\*t+0,2 \*t\* sin(t)+N(0. 1)

Здесь шум имеет среднее значение, равное 0, и СКО, равное 1.

2. ВР разыгрывается по следующей формуле

xt = 0,5\*t+0,2 \* t \* cos(t) + N(0. 0,5)

3. ВР разыгрывается по следующей формуле

xt = 0,5 \*t + 0,5 \*t\* sin2(t)+ N(0. 1)

4. ВР разыгрывается по следующей формуле

xt = 0,5 \*t + 0,5 \*t\* cos2 (t)+ N(0. 1)

Варианты расчета выбираются в соответствии со списком группы в ЛК.