Электронный научный журнал «ИССЛЕДОВАНО В РОССИИ» 1270 http://zhurnal.ape.relarn.ru/articles/2006/136.pdf

Применение нейронных сетей для решения задач прогнозирования

Солдатова О.П., Семенов В.В. (vlad-eraser@mail.ru)

Самарский государственный аэрокосмический университет имени академика С.П. Королева (СГАУ)

Нейронные сети – это раздел искусственного интеллекта, в котором для обработки сигналов используются явления, аналогичные происходящим в нейронах живых существ.

Важнейшая особенность сети, свидетельствующая о ее широких возможностях и огромном потенциале, состоит в параллельной обработке информации всеми звеньями, что позволяет значительно ускорить процесс обработки информации. Кроме того, при большом числе межнейронных соединений сеть приобретает устойчивость к ошибкам, возникающим на некоторых линиях.

Другое не менее важное свойство – способность к обучению и обобщению накопленных знаний. Нейронная сеть обладает чертами искусственного интеллекта. Натренированная на ограниченном множестве данных сеть способна обобщать полученную информацию и показывать хорошие результаты на данных, не использовавшихся при ее обучении [1].

В настоящее время нейронные сети используются для решения целого ряда задач, одной из которых является задача прогнозирования.

Прогнозирование – это предсказание будущих событий. Пусть заданы n дискретных отсчетов $\{y(t_1),\ y(t_2)...,\ y(t_n)\}$ в последовательные моменты времени $t_1,\ t_2,...,\ t_n$. Тогда задача прогнозирования состоит в предсказании значения $y(t_{n+1})$ в некоторый будущий момент времени t_{n+1}.

Целью прогнозирования является уменьшение риска при принятии решений. Прогноз обычно получается ошибочным, но ошибка зависит от используемой прогнозирующей системы. Предоставляя прогнозу больше ресурсов, можно увеличить точность прогноза и уменьшить убытки, связанные с неопределенностью при принятии решений. Типичными приложениями техники прогноза являются предсказание цен на фондовой бирже, прогноз погоды, прогноз потребления электроэнергии, прогноз отказов технических систем и пр.

В настоящей работе рассматривается применение нейронной сети для решения задачи прогнозирования временных рядов. Пользователь выбирает произвольный временной ряд, содержащий N отсчетов, и разбивает его на три множества: обучающую, тестирующую и контрольную выборки, которые затем подаются на вход сети. Результатом прогнозирования является значение временного ряда в требуемый момент времени.

Для повышения качества прогноза необходимо произвести предварительную (препроцессорную) обработку информации. Поскольку временной ряд представляет собой последовательность числовых отсчетов, препроцессорная обработка, как правило, сводится к масштабированию значений отсчетов с целью их приведения в единый диапазон.

Каждая выборка представляет собой дискретную функцию, заданную в точках на интервале [0,N] с шагом 1, где N – максимальное значение аргумента этой функции.

При решении задач прогнозирования роль нейронной сети состоит в предсказании будущей реакции системы по ее предшествующему поведению. Обладая информацией о значениях переменной x в моменты, предшествующие прогнозированию $x(k-1)$, $x(k-2)$, ..., $x(k-N)$, сеть вырабатывает решение, каким будет наиболее вероятное значение последовательности $\bar{x}(k)$ в текущий момент k. Для адаптации весовых коэффициентов сети используются фактическая погрешность прогнозирования $\varepsilon = x(k) - \bar{x}(k)$ и значения этой погрешности в предшествующие моменты времени [2].

При выборе архитектуры сети обычно опробуется несколько конфигураций с различным количеством элементов. Исходя из того, что задача прогнозирования является частным случаем задачи регрессии, следует, что она может быть решена следующими типами нейронных сетей: многослойным персептроном (MLP), радиально-базисной сетью (RBF), обобщенно-регрессионной сетью (GRNN), сетью Вольтерри и сетью Эльмана.

При решении задачи прогнозирования временных рядов в качестве нейронной сети была выбрана обобщенно-регрессионная сеть, реализующая методы ядерной аппроксимации. В задачах регрессии выход сети может рассматриваться как ожидаемое значение модели в данной точке пространства входов. Это ожидаемое значение связано с плотностью вероятности совместного распределения входных и выходных данных. В точку расположения каждого обучающего наблюдения помещается гауссова ядерная функция. Считается, что каждое наблюдение свидетельствует о некоторой уверенности в том, что поверхность отклика в данной точке имеет определенную высоту, и эта уверенность убывает при отходе в сторону от точки. GRNN-сеть копирует внутрь себя все обучающие наблюдения и использует их для оценки отклика в произвольной точке. Окончательная выходная оценка сети получается как взвешенное среднее выходов по всем обучающим наблюдениям, где величины весов отражают расстояние от этих наблюдений до той точки, в которой производится оценивание (и, таким образом, более близкие точки вносят больший вклад в оценку).

Структура нейронной сети GRNN представлена на рисунке 1.

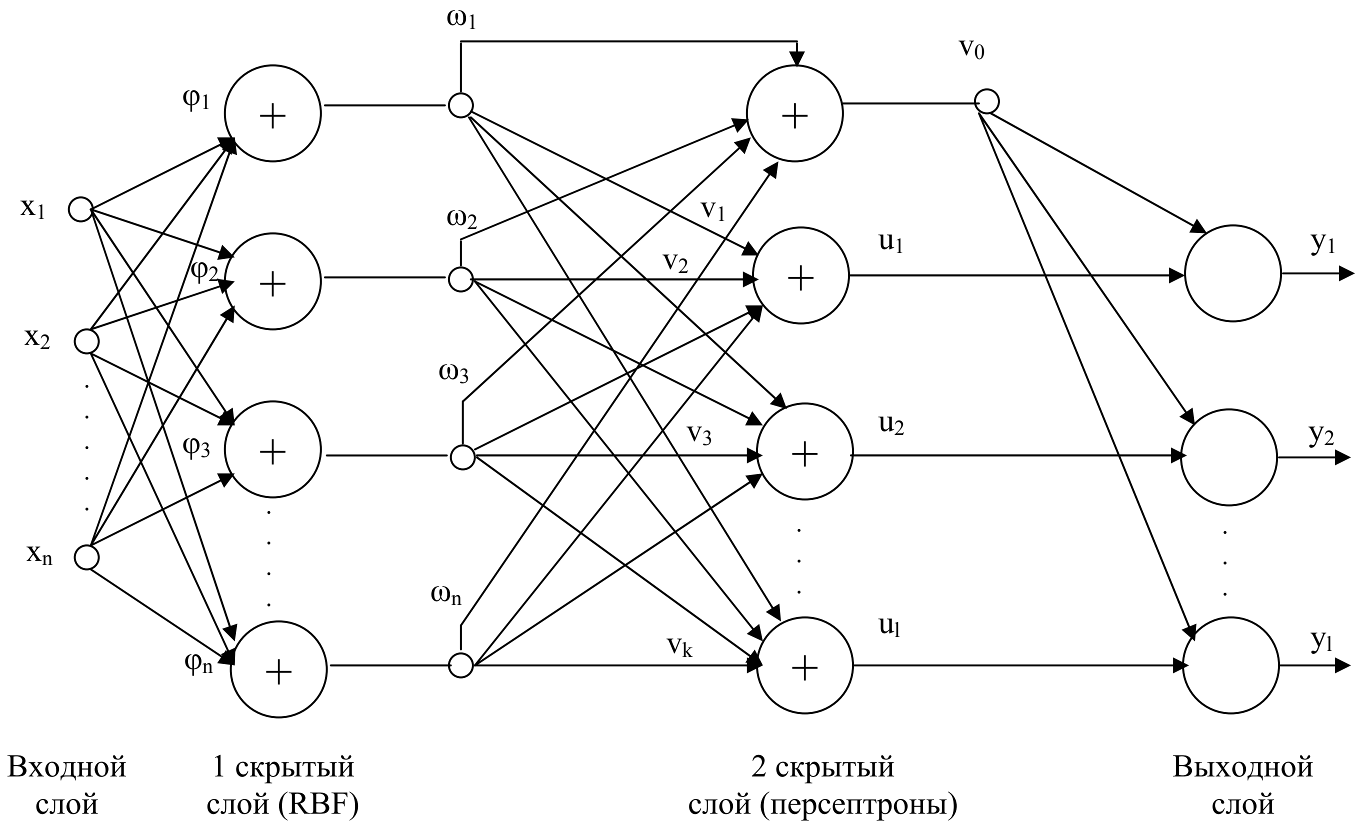


Рисунок 1 – Обобщенная структура сети GRNN

GRNN-сеть имеет два скрытых слоя: слой радиальных элементов и слой элементов, которые формируют взвешенную сумму для соответствующего элемента выходного слоя. В выходном слое определяется взвешенное среднее путем деления взвешенной суммы на сумму весов. В качестве радиальной функции применяется функция Гаусса.

Входной слой передает сигналы на первый промежуточный слой нейронов, являющихся радиально симметричными. Они несут в себе информацию о данных обучающих случаев или же их кластерах и передают ее во второй промежуточный слой. В нем формируются взвешенные суммы для всех элементов выходного слоя и сумма весов, вычисляемая специальным элементом. Если обозначить выход i-го нейрона RBF-слоя как v_i, то выходной сигнал l-го нейрона второго промежуточного слоя вычисляется по формуле

$$u_l = \sum_{i=1}^{k} v_i \,, \qquad (1)$$

где k – число нейронов в RBF - слое.

Обозначив теперь весовой коэффициент i-го нейрона RBF-слоя как ω_i, получим формулу для суммы весов:

$$v_0 = \sum_{i=1}^{k} \omega_i \,, \qquad (2)$$

Наконец, выходной слой делит взвешенные суммы на сумму весов и выдает окончательный прогноз. Обозначив его за y_l, получим:

$$y_l = \frac{u_l}{v_0} \tag{3}$$

Рассмотрим теперь принципы функционирования первого промежуточного слоя, структура которого представлена на рисунке 2.

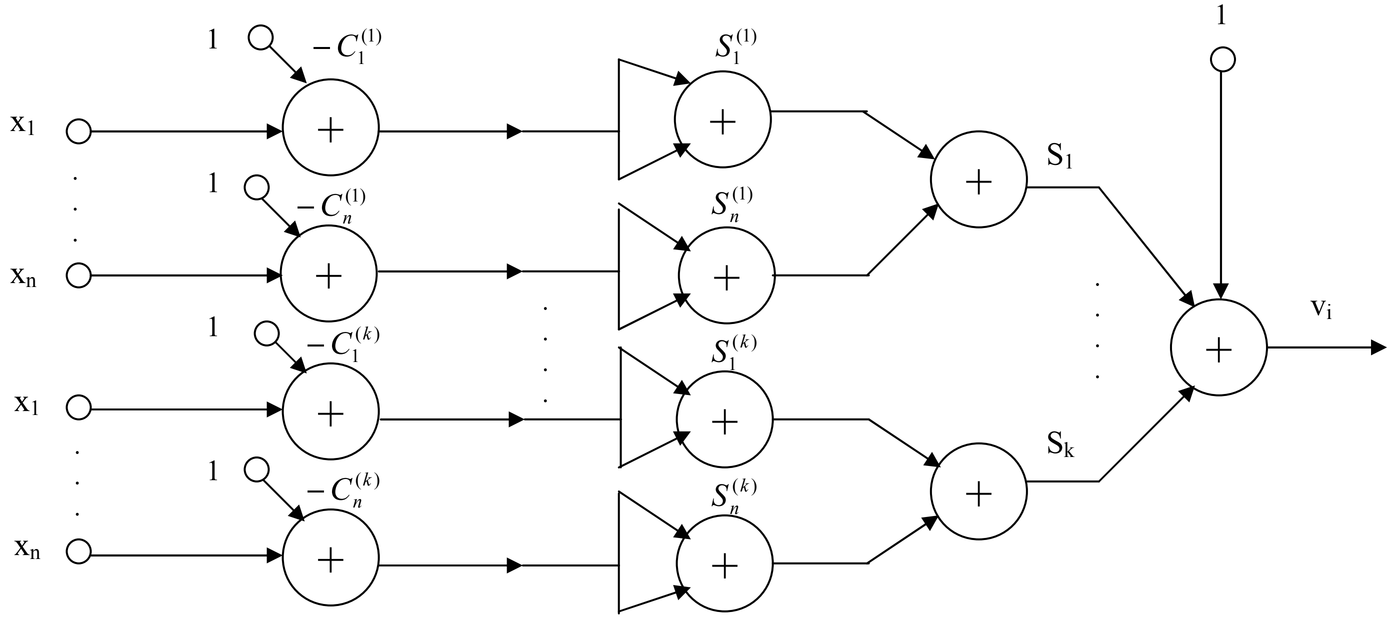


Рисунок 2 – Структура RBF-слоя сети GRNN

На вход радиальных элементов из входного слоя подается вектор *x*. Базисные функции RBF-слоя задаются матрицей Q, но в практическом плане более удобно использовать для описания элементов матрицу корреляции C, которая получается из матрицы Q следующим образом:

$$C = Q^T Q \tag{4}$$

Центр i-го нейрона радиального слоя обозначим как c_i.

Окончательный результат обработки входных сигналов S_j вычисляется по следующим формулам:

$$S_j^{(t)} = -\frac{1}{2}\sum_{t=1}^{n}(x_t - c_i^{(t)})^2, \tag{5}$$

$$S_t = \sum_{j=1}^{n} S_j^{(t)}, \tag{6}$$

$$v_i = \sum_{t=1}^{k} \exp\left(-\frac{1}{2}\frac{S_t}{\sigma_t^2}\right)\omega_t \tag{7}$$

Затем вектор выходных сигналов *v* передается на вход второго промежуточного слоя сети.

Достоинством сети GRNN можно считать определенность структуры: сеть фактически вмещает в себя все обучающие данные. С другой стороны, такая структура нейросети и является ее самым большим недостатком: при большом объеме обучающих данных скорость работы сети падает, иногда очень существенно, по причине заметного увеличения сложности архитектуры.

Выходное значение сети имеет вероятностный смысл, поэтому его легче интерпретировать. При небольшом объеме входных данных сеть очень быстро обучается.

Обучение сети необходимо выполнять отдельно для каждого временного ряда, так как попытка прогнозирования ряда, на котором сеть не была обучена, приведет к ошибочному результату [3].

В качестве алгоритма обучения использовался модифицированный алгоритм обратного распространения ошибки с автоматической коррекцией длины шага обучения (ParTan).

В дальнейшем были проведены исследования зависимости качества прогноза от параметров алгоритма обучения и структуры нейронной сети. Результаты исследований показали, что качество прогнозирования зависит, прежде всего, от разбиения отсчетов ряда на три множества – обучающее, тестирующее и контрольное. Наилучшее качество прогноза достигается при соотношении объемов выборок 60:20:20. Очевидно, что точность прогноза будет падать по мере увеличения дальности. Оптимальными значениями параметров алгоритма являются: коэффициент скорости обучения $\eta = 0.7$, коэффициент момента обучения $\mu = 0.9$, количество итераций до запоминания $N = 20$, величина изменения коэффициента скорости обучения $\alpha = 0.1$. Число нейронов в скрытых слоях сети определяется для каждого временного ряда индивидуально.

При проведении исследований использовались следующие временные ряды: прогноз погоды в Москве за 1998 год и отказы АТС за 2002 год.

Результаты прогнозирования представлены в таблицах 1 - 2 и приведены в графическом виде на рисунках 3 - 5.

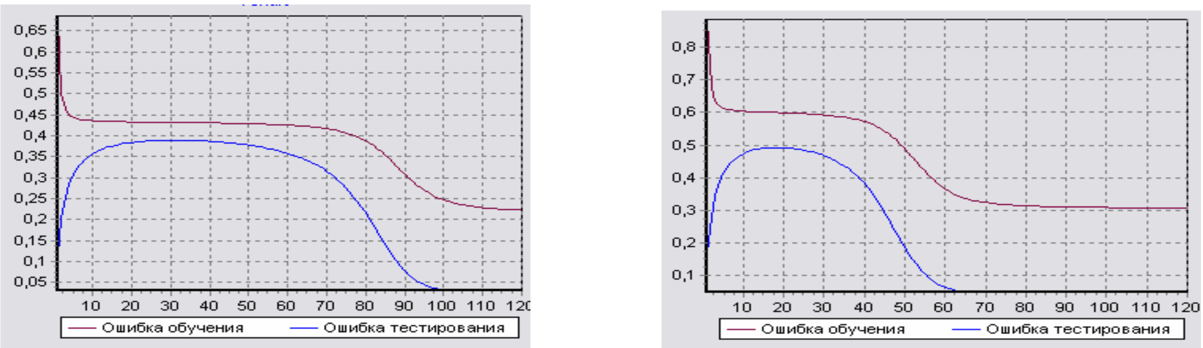
Таблица 1 – Статистика обучения сети GRNN для ряда «Прогноз погоды в Москве за 1998 год»

Итоговые статистики	**Параметры алгоритма**		
	$\eta = 0.25$ $\mu = 0.5$	**$\eta = 0.5$ $\mu = 0.9$**	**$\eta = 0.7$ $\mu = 0.9$**
Математическое ожидание ошибки	0.12474	0.05218	0.01153
Дисперсия ошибки	0.002285	0.00119	0.00087
Среднеквадратическое отклонение ошибки	0.01995	0.03184	0.01318

Таблица 2 – Статистика обучения сети GRNN для ряда «Отказы АТС за 2002 год»

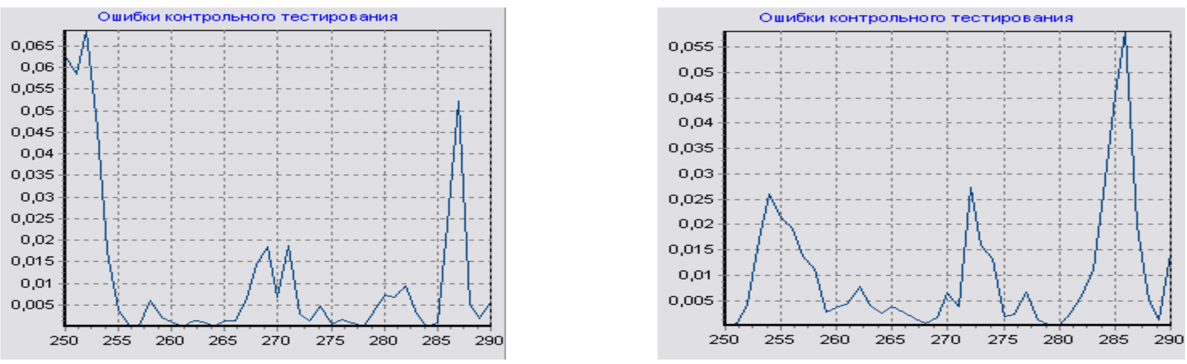
Итоговые статистики	**Параметры алгоритма**		
	η **= 0.25** μ **= 0.5**	η **= 0.5** μ **= 0.9**	η **= 0.7** μ **= 0.9**
Математическое ожидание ошибки	0.15340	0.05821	0.01286
Дисперсия ошибки	0.002651	0.00139	0.00125
Среднеквадратическое отклонение ошибки	0.02655	0.03731	0.01394

Исследование проводилось на нейронной сети следующей структуры: количество нейронов в 1 скрытом слою = 8, количество нейронов во 2 скрытом слою = 6, дальность прогноза = 5, алгоритм обучения – ParTan, параметры алгоритма – оптимальные, разбиение отсчетов ряда на множества - оптимальное.



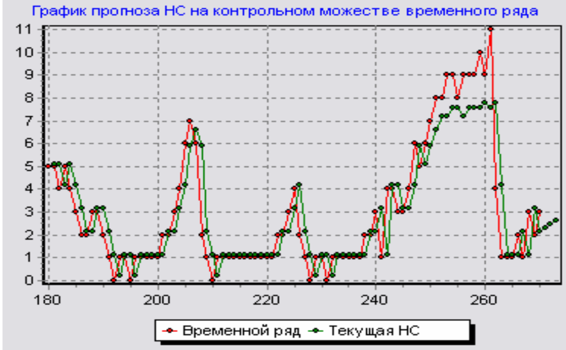
Прогноз погоды в Москве Прогноз отказов на АТС

Рисунок 3 – Ошибки обучения и тестирования сети



Прогноз погоды в Москве Прогноз отказов на АТС

Рисунок 4 – Ошибки контрольного тестирования сети



Прогноз погоды в Москве Прогноз отказов на АТС

Рисунок 5 – Результаты прогнозирования

Следует особо отметить тот факт, что эффективное решение задачи прогнозирования возможно только в том случае, если нейронная сеть обучается на большом объеме данных. В случае малоразмерной или некачественной обучающей выборки даже самый лучший алгоритм не даст удовлетворительного результата, поскольку без полноценного набора данных нейросеть принципиально не способна обучиться.

Литература

1. Осовский С. Нейронные сети для обработки информации / Пер. с польского И.Д. Рудинского. – М.: Финансы и статистика, 2002. – 344 с.: ил.
2. Обобщенная аппроксимационная теорема и вычислительные возможности нейронных сетей. А.Н. Горбань. Сибирский журнал вычислительной математики, 1998. Т.1, №1, с.12-24.
3. Уоссермен Ф. Нейрокомпьютерная техника: теория и практика / Пер. с английского Ю.А. Зуев. – М.: Мир, 1992.