

## ПРИМЕНЕНИЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ КЛАССИФИКАЦИИ СТОХАСТИЧЕСКИХ ОБЪЕКТОВ

---

Рассматривается проблема построения моделей прогноза при помощи нейронных сетей с применением генетического алгоритма. Решается задача оценки структуры нейронной сети с использованием генетического алгоритма, который удобен тем, что данные легко распараллеливать, разбивать поколение на несколько групп и работать с каждой из них независимо, обмениваясь время от времени несколькими хромосомами. Анализируются подходы к выбору структуры модели.

### 1. Введение

Для решения функциональных задач в информационно-управляющих системах существует достаточно большое число методов и алгоритмов: выборочный метод, метод имитационного моделирования, метод исследования алгоритмов идентификации, графические методы, эволюционное программирование, регрессионный анализ, дискриминантный анализ, методы факторного анализа, методы линейного программирования, задачи нелинейного программирования, динамическое программирование, метод распознавания образов, стохастическое программирование, робастные методы, кластерный анализ, мартигальные методы анализа случайных социально-экономических процессов, Марковские модели и др.

Выбор того или иного метода обработки данных для конкретного объекта представляет определенную трудность, связанную с анализом и оценкой его эффективности в зависимости от класса объекта и его статистических характеристик. Несмотря на то, что арсенал методов и алгоритмов велик, возникает немало препятствий их эффективному использованию. Каждый метод ориентирован, как правило, на решение конкретного класса задач и обеспечивает эффективность решения лишь при выполнении некоторых теоретических предположений.

В связи с нелинейным характером большинства методов обработки данных их аналитическое исследование представляет достаточно сложную задачу. Из-за наличия специфических особенностей объектов управления реальные данные содержат шумы различной природы, имеют нестационарные и нелинейные характеристики, меняющиеся во времени. В таких условиях только тщательная предварительная обработка данных может сэкономить время и привести к успеху. Здесь необходимо решить следующие вопросы: сбор данных; их фильтрацию и анализ; преобразование с целью сделать информацию более содержательной и удобной для используемого метода обработки. На этом этапе важно определить совокупность информативных переменных, участвующих в решении задач моделирования и принятии эффективных управляющих воздействий.

Решение задачи классификации является одним из важнейших применений нейронных сетей. Она представляет собой задачу отнесения образца к одному из нескольких попарно не пересекающихся множеств.

При решении задач классификации необходимо отнести имеющиеся *статистические образцы* (характеристики ситуации на валютном рынке) к *определенным классам*.

### 2. Использование нейронных сетей в качестве классификатора

Сети с прямой связью являются универсальным средством аппроксимации функций, что позволяет их использовать в решении задач классификации. Как правило, нейронные сети оказываются наиболее эффективным способом классификации, потому что генерируют фактически большое число регрессионных моделей (которые используются в решении задач классификации статистическими методами).

К сожалению, при применении нейронных сетей в практических задачах возникает ряд проблем. Например, заранее не известно, какой сложности (размера) может потребоваться сеть для достаточно точной реализации отображения. Эта сложность может оказаться чрезмерно высокой, что потребует сложной архитектуры сетей.

### 3. Подготовка исходных данных

Для построения классификатора необходимо определить, какие параметры влияют на принятие решения о том, к какому классу принадлежит образец. При этом могут возникнуть две проблемы. Если количество параметров мало, то может возникнуть ситуация, при которой один и тот же набор исходных данных соответствует примерам, находящимся в разных классах. Тогда невозможно обучить нейронную сеть, и система не будет корректно работать.

Временные ряды, как правило, имеют динамическую природу и представляют собой сложный стохастический процесс, который со временем меняет свою структуру и характеризуется неожиданным сдвигом в уровне. Разобьем временной ряд с данными о курсах валют на несколько типовых состояний, которые характеризуются качественными и количественными показателями. Исходные данные представлены в табл. 1. Множество исходных данных было разбито на пять групп (рис. 1).

Таблица 1

Таблица исходных данных

№	Date	Time	Close	Volume	Абс. цепн. прирост
1	10.10.2003	00:00	1.1744	27	$\Delta Y_{ц_i} = Y_i - Y_{i-1}$
2	10.10.2003	00:05	1.1745	13	1E-04
3	10.10.2003	00:10	1.1744	22	-1E-04
4	10.10.2003	00:15	1.174	28	-0.0004
5	10.10.2003	00:20	1.1743	25	0.0003
...	...	...	...	...	...
2012	20.10.2003	23:45	1.1628	18	-1E-04
2013	20.10.2003	23:50	1.1628	20	0
2014	20.10.2003	23:55	1.1628	34	0

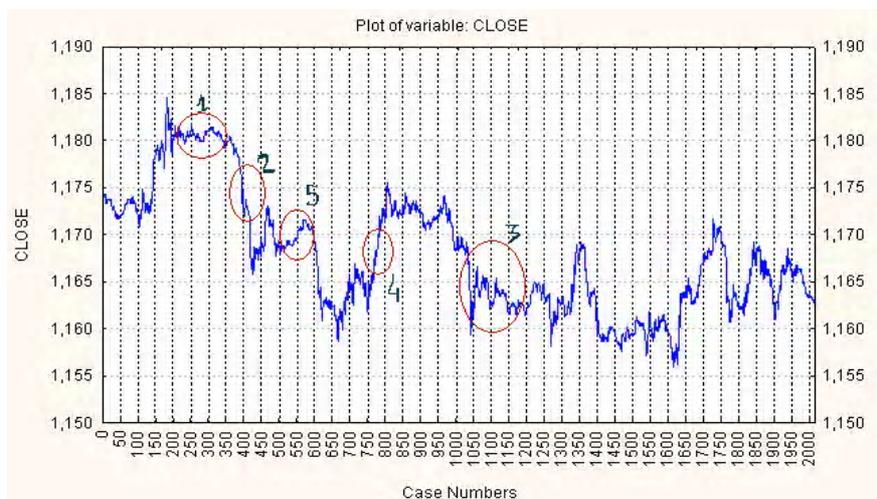


Рис. 1. Пример подмножеств данных, входящих в каждую из групп

В качестве признаков классификации определим следующие критерии:

1) среднее значение цепного темпа роста  $Y = (\Delta Y_1 + \Delta Y_2 + \dots + \Delta Y_N) / n$ , где  $\Delta Y_1$  – цепной темп роста,  $n$  – количество данных в подмножестве;

2) среднее значение размаха:  $R = (Y_{\max} - Y_{\min}) / n$ , где  $Y_{\max}$  – максимальное значение в подмножестве данных,  $Y_{\min}$  – минимальное значение в подмножестве данных;

3)  $D$  – дисперсия.

В табл. 2 представлены значения качественных и количественных показателей.

Исходные данные разбили на группы и для каждой из них рассчитали значение критериев классификации. Каждая из групп была отнесена в соответствии со значениями критериев к одному из пяти классов. Результаты приведены в табл. 3.

Таблица 2

Диапазоны допустимых значений критериев в каждой из групп

	Var1	Var2	Var3	Var4	Var5
Y	-0.000009 - 0.000009	0.0001 - 0.001	-(0.00009 - 0.00001)	-(0.001 - 0.0001)	0.00001 - 0.00009
R	0 - 0.00003	0.0001 - 0.0005	0.00004 - 0.00009	0.0001 - 0.0005	0.00004 - 0.00009
D	0 - 0.000002	0.000008 - 0.000015	0.000003 - 0.000007	0.000008 - 0.000015	0.000002 - 0.000008

Таблица 3  
Исходная обучающая выборка

	$\Delta Y$	Razmah	Dispersia	Class
1	-0.001429	0.028571	0.000732	Clas1
2	-0.001424	0.021469	0.000330	Clas1
3	-0.020859	0.026994	0.000965	Clas1
4	-0.012097	0.029032	0.001043	Clas1
5	0.010952	0.028095	0.001352	Clas1
6	-0.018056	0.079167	0.001575	Clas1
7	0.185106	0.221277	0.005230	Clas2
8	0.122835	0.118110	0.011585	Clas2
9	0.075000	0.115625	0.009020	Clas2
10	0.003448	0.081034	0.010248	Clas3
11	-0.015730	0.046067	0.003274	Clas3
12	-0.231746	0.223810	0.018347	Clas4
13	-0.115464	0.124742	0.012992	Clas4
14	-0.165000	0.165000	0.007373	Clas4
15	0.034194	0.042581	0.001949	Clas5
16	0.041667	0.087121	0.007783	Clas5

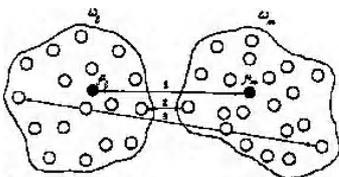


Рис. 2. Способы определения расстояния между кластерами

Существуют стандартные критерии, которым должна удовлетворять мера сходства, чтобы быть метрикой:

$$d(x, y) = d(y, x) \geq 0.$$

1) Симметрия. Даны два объекта  $x$  и  $y$ ; расстояние между ними удовлетворяет условию

$$d(x, y) = d(y, x) \geq 0.$$

2) Неравенство треугольника. Даны три объекта  $x, y, z$ ; расстояния между ними удовлетворяют условию

$$d(x, y) \leq d(x, z) + d(y, z).$$

3) Различимость нетождественных объектов. Даны два объекта  $x$  и  $y$ : если  $d(x, y) \neq 0$ , то  $x \neq y$ .

Неразличимость идентичных объектов. Для двух идентичных объектов  $x$  и  $x'$   $d(x, x') = 0$ , т.е. расстояние между этими объектами равно нулю.

На рис. 2 представлены различные способы определения расстояния между кластерами  $w_1$  и  $w_m$ : 1 — по центрам тяжести, 2 — по ближайшим объектам, 3 — по самым далеким объектам.

Расстояние ближайшего соседа есть расстояние между ближайшими объектами кластеров:

$$q_{\min}(W_1, W_m) = \min d(x_i, x_j).$$

Расстояние дальнего соседа — расстояние между самыми дальними объектами кластеров:

$$q_{\max}(W_1, W_m) = \max d(x_i, x_j).$$

Новая группа относится к тому классу, для которого классификационное значение максимально.

Проводя нейросетевой анализ данных, будем работать отдельно с каждой группой данных, используя STATISTICA Neural Networks, где реализованы основные алгоритмы обучения многослойных перцептронов: методы обратного распространения, сопряженных градиентов и Левенберга-Маркардта. Обучающие данные тут хранятся в виде набора данных (*Data Set*), содержащего некоторое количество наблюдений, для каждого из которых заданы значения нескольких входных и выходных переменных. Каждому наблюдению соответствует одна строка таблицы.

Алгоритм обратного распространения последовательно обучает сеть на данных из обучающего множества. На каждой итерации (эпохе) все наблюдения из обучающего множества по очереди подаются на вход сети. Сеть обрабатывает их и выдает выходные значения.

Эти выходные значения сравниваются с целевыми выходными значениями, которые также содержатся в наборе исходных данных, и ошибка, т.е. разность между желаемым и реальным выходом используется для корректировки весов сети так, чтобы уменьшить эту ошибку.

Алгоритм должен находить компромисс между различными наблюдениями и менять веса таким образом, чтобы уменьшить суммарную ошибку на всем обучающем множестве, поскольку алгоритм обрабатывает наблюдения по одному, общая ошибка на отдельных шагах не обязательно будет убывать.

#### **4. Генетические алгоритмы и нейронные сети**

Прежде всего, генетические алгоритмы обеспечивают глобальный просмотр пространства весов и позволяют избегать локальных минимумов. Кроме того, они могут использоваться в задачах, для которых информацию о градиентах получить очень сложно либо она оказывается слишком дорогостоящей.

В задачах нелинейной природы среди параметров могут быть взаимозависимые и избыточные. Единственный способ получить полную гарантию того, что входные данные выбраны наилучшим образом, состоит в том, чтобы перепробовать все возможные варианты входных наборов данных и архитектур сетей и выбрать из них наилучший.

Самое действенное средство решения данного вопроса – это применение генетического алгоритма, который выполняет множество экспериментов с различными комбинациями входных данных, строит для каждой из них вероятностную либо обобщенно-регрессионную сеть, оценивает ее результаты и использует их в дальнейшем поиске наилучшего варианта.

Важнейшим достоинством нейросетей в системах прогнозирования на финансовых рынках является отсутствие всяких ограничений на характер входной информации. Нейросетевые системы прогнозирования основываются в своих предсказаниях на оптимальных значениях, полученных в результате решения оптимизационных задач. Более того, эти результаты могут быть скорректированы в соответствии с изменением исходной информации. Специфика использования нейросетей для прогнозирования финансовых рынков состоит в изменчивом характере стратегии моделирования. Причиной изменчивости является хаотический характер исходных данных.

Используем генетические алгоритмы для выбора архитектуры сети. Алгоритм выполняет большое число экспериментов с различными комбинациями входных данных, строит для каждой из них вероятностную либо обобщенно-регрессионную нейронную сеть (сети этих типов выбраны потому, что для них общее время обучения/оценки очень мало, а также потому, что они очень сильно страдают от присутствия входных переменных и поэтому являются хорошим средством их обнаружения), оценивает ее результаты и использует их в дальнейшем поиске наилучшего результата. Алгоритм следит за некоторым набором таких маскирующих строк, оценивая каждую из них по контрольной ошибке. По значениям ошибки приводится отбор лучших вариантов масок, которые комбинируются друг с другом с помощью искусственных генетических операций: скрещивания и мутации.

Основной генетический алгоритм состоит из следующих шагов:

- инициализация, или выбор исходной популяции хромосом;
- оценка приспособленности хромосом в популяции;
- проверка условия остановки алгоритма;
- селекция хромосом;

- применение генетических операторов;
- формирование новой популяции;
- выбор «наилучшей» хромосомы.

Проведем эксперимент с каждой из пяти групп данных и построим нейронную сеть, которая изображена на рис. 3.

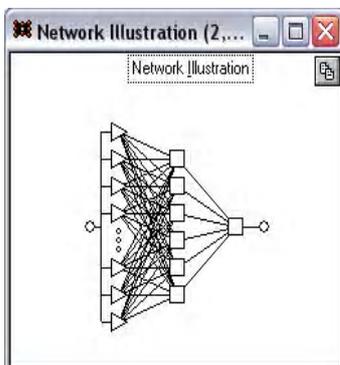


Рис.3. Нейронная сеть

Проанализировав полученные ошибки, весовое распределение и проекции временного ряда, приходим к выводу, что полученная ошибка в результате обучения методами обратного распространения, быстрого распространения и Quasi-Newton не попадает в интервал 0,025-0,35, что говорит о неэффективности их использования при обучении нейронной сети с учетом указанных выше данных. Метод Левенберга-Маркардта в результате проведенных исследований оказался более эффективным.

После проведенных исследований, на основании результатов, можно предложить математическую модель для конкретного временного ряда, что является научной новизной. Сумма умножается на

сигмоидальную функцию  $\frac{k}{1 + \exp^{-\alpha v}}$ . Сигмоидальная функция была

выбрана по таким причинам:

- простая функция для взятия производной;
- имеет перегиб на 0,5, т. е. все изменения можно выбрать от 0 до 1.

Функция для первого слоя нейронов (60):

$$y_1 = \frac{k}{1 + \exp^{-\alpha v_1}} \sum_{j=1}^5 x_j = \frac{k}{1 + \exp^{-\alpha v_1}} (x_{j1} + x_{j2} + \dots + x_{j5}) = \frac{k}{1 + \exp^{-\alpha v_1}} u_1 w_1,$$

где  $U_1 = 60$ .

Функция для второго слоя нейронов (30):

$$y_2 = \frac{k}{1 + \exp^{-\alpha v_2}} \sum_{j=1}^{60} x_j = \frac{k}{1 + \exp^{-\alpha v_2}} (x_{j1} + x_{j2} + \dots + x_{j60}) = \frac{k}{1 + \exp^{-\alpha v_2}} u_2 w_2,$$

где  $U_2 = 30$ .

Функция для третьего слоя нейронов (1):

$$y_3 = \frac{k}{1 + \exp^{-\alpha v_3}} \sum_{j=1}^{30} x_j = \frac{k}{1 + \exp^{-\alpha v_3}} (x_{j1} + x_{j2} + \dots + x_{j30}) = \frac{k}{1 + \exp^{-\alpha v_3}} u_3 w_3,$$

где  $U_3 = 1$ .

Предложенная математическая модель:

$$y = \frac{k}{1 + \exp^{-\alpha v}} \left( \sum_{j=1}^5 x_j + \sum_{j=1}^{60} x_j + \sum_{j=1}^{30} x_j \right) = \frac{k}{1 + \exp^{-\alpha v}} (u_1 w_1 + u_2 w_2 + u_3 w_3),$$

где  $k$  - коэффициент усиления = 0,1;  $\alpha$  - коэффициент сигмоиды = 0,1 или 0,2;  $v$  - сигнал, идущий от сумматора;  $j$  - количество сигналов, приходящих на сумматор;  $w$  - вес, изменяющийся от 0 до 1.

## 5. Заключение

На субъективном уровне можно сделать следующий вывод: хотя нейросеть и является механизмом, способным решить широкий класс задач, но правила управления этим механизмом пока что находятся на уровне исследований. Однако результаты проведенных исследований показали достаточную эффективность нейросетевых моделей, которые могут быть успешно использованы при принятии перспективных решений в дилинговых информационных системах.

**Список литературы:** 1. Бэстэнс Д.-Э., Ван Ден Берг В.-М., Вуд Д. Нейронные сети и финансовые рынки. Принятие решений в торговых операциях. М.: Научное издательство ТВП, 1997. 236 с. 2. Руденко О.Г., Бодянский Е.В. Основы теории искусственных нейронных сетей. Х.: ТЕЛТЕХ, 2002. 317 с. 3. Рутковская Д., Пилинский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы: Пер. с польск. И.Д.Рудинского. М.: Горячая линия–Телеком, 2004. 452 с. 4. Лиховидов В.Н., Сафин В.И. Технический анализ валютных рынков. Владивосток: Изд-во ДВГАЭУ, 1998. 200 с. 5. Bartlett P., Downs T. Training a neural networks with a genetic algorithm, Tehnical Report, Dept. of Elec. Eng., Univ. of Queensland, 1990. 264 p. 6. Shonkwiler R., Miller K.R. Genetic algorithm, neural network synergy for nonlinearly constrained optimization problems// Proceedings of International Workshop on Combinations of Genetic Algorithms and Neural Networks, COGANN-92, 1992. 348 p. 7. Лукашин Ю. П. Адаптивные методы краткосрочного прогнозирования временных рядов. М.: Финансы и статистика, 2003. 416 с. 8. Дубров А.М., Мхитарян В.С., Трошин Л.И. Многомерные статистические методы: Учебник. М.: Финансы и статистика, 2000. 352 с. 9. Марк Т. Технический анализ - новая наука. М.: Диаграмма, 1997. 10. Davis L., Handbook of Genetic Algorithms. Van Nostrand Reinhold, NY, 1991. 563 p. 11. Аналитический пакет. Deductor Professional. Руководство пользователя. BaseGroup Labs, 2002. 196 с. 12. Уолем Т. Дж., Паррамоу К. Количественные методы в финансах: Учебное пособие для ВУЗов: Пер. с англ. под ред. М. Р. Ефимовой. М: Финансы, Юнити, 1999. 527 с.

*Поступила в редколлегию 30.06.2008*

**Шамша Борис Владимирович**, канд. техн. наук, профессор кафедры ИУС ХНУРЭ. Научные интересы: разработка эффективных методов кластеризации в ИУС. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 702-14-51.

**Морозова Олеся Сергеевна**, аспирантка кафедры ИУС ХНУРЭ. Научные интересы: разработка эффективных методов кластеризации в ИУС. Адрес: Украина, 61166, Харьков, ул. Бакулина, 14, кв. 6, тел. 702-25-42.

---

УДК 007.51

*М.С. КУДРЯВЦЕВА*

## **МАТЕМАТИЧЕСКИЕ МОДЕЛИ ОЦЕНКИ И ФОРМИРОВАНИЯ ВАРИАНТОВ РЕШЕНИЙ ЗАДАЧИ ЛИКВИДАЦИИ НАРУШЕНИЙ РАБОТЫ ЭЛЕКТРОЭНЕРГЕТИЧЕСКОГО ОБОРУДОВАНИЯ**

---

Предлагается общая структурная схема описания процесса функционирования устройств релейной защиты и противоаварийной автоматики электроэнергетического комплекса, на основании которой разрабатывается математическая модель описания процесса функционирования устройств релейной защиты и противоаварийной автоматики. В рамках данной структурной схемы разрабатываются алгоритм функционирования и предварительной оценки работы устройств релейной защиты и автоматики, алгоритм оценки работы устройств РЗА. На основании данных алгоритмов с использованием модифицированного аппарата регулярных схем алгоритмов разрабатываются модель формирования вариантов решений задачи ликвидации нарушений работы оборудования, связанных с неправильными срабатываниями и отказами устройств релейной защиты и автоматики, модель оценки работы устройств РЗА.

### **1. Введение**

Электроэнергетическое оборудование является основой построения любых энергетических систем. Его уровень определяет надежность и устойчивость всей системы, качество энергоснабжения потребителей.

Повышение надежности поставки электроэнергии в Украине, и в первую очередь снижение и быстрое устранение возникающих аварий, накладывает жесткие требования к электроэнергетическому и электротехническому оборудованию и, главным образом, к системам контроля и управления – устройствам автоматического управления электроэнергетическим комплексом (устройствам релейной защиты и автоматики (РЗА), устройствам противоаварийной автоматики (ПА), автоматизированным системам управления, системам контроля и диагностики электрооборудования и энергетического хозяйства в целом [1]).

За последние двадцать лет в ведущих зарубежных странах произошла революция в области аппаратных средств и функциональных возможностей таких систем контроля и