

НЕЙРОННЫЕ СЕТИ

Нейронные сети, в отличие от статистических методов многомерного классификационного анализа, базируются на параллельной обработке информации и обладают способностью к самообучению, то есть получать обоснованный результат на основании данных, которые не встречались в процессе обучения.

Эти свойства позволяют нейронным сетям решать сложные (масштабные) задачи, которые на сегодняшний день считаются трудноразрешимыми. Основными достоинствами нейронных сетей перед традиционными вычислительными методами являются:

1. Процесс создания нейронной сети больше относится к процессу обучения, нежели к программированию
2. Нейрокомпьютеры особенно эффективны там, где нужно подобие человеческой интуиции, в частности к таким задачам относятся принятие решений в процессе оценки финансового состояния некоторого экономического объекта.
3. Гибкость структуры нейронных сетей позволяет различными способами комбинировать простые составляющие нейрокомпьютеров – нейроны и связи между ними. За счет этого один нейрокомпьютер можно применять для решения различных задач, зачастую, никак не связанных между собой.
4. Нейронные сети позволяют создать эффективное программное обеспечение для высокопараллельных компьютеров. Создавая математическое обеспечение на базе нейронных сетей, можно для широкого класса задач решить проблему эффективности одновременного решения параллельных задач. Кроме того параллельная обработка информации обеспечивает высокую скорость решения задач.
5. Решение задач в условиях неопределенности – способность обучения нейронной сети позволяет решать задачи с неизвестными закономерностями и зависимостями между входными и выходными данными,

что позволяет работать с неполными данными. Кроме того, взаимоотношения между величинами заранее не устанавливаются, поскольку метод предусматривает изучение существующих взаимосвязей на готовых моделях.

6. Устойчивость к шумам во входных параметрах – нейронная сеть может самостоятельно определять неинформативные для анализа параметры и производить их отсев, в связи с чем, пропадает необходимость дополнительного анализа информационного вклада входных данных.

7. Адаптирование к изменениям окружающей среды – нейронные сети могут быть переучены в новых условиях окружающей среды, описываемых незначительными колебаниями параметров этой среды. То есть можно производить переобучение нейронных сетей на основе незначительных колебаний параметров среды. Если задача решается в условиях нестационарной среды (где статистика изменяется с течением времени), то могут быть созданы нейронные сети, переучивающиеся в реальном времени. Чем выше адаптивные способности системы, тем более устойчивой будет ее работа в нестационарной среде.

8. Потенциальная отказоустойчивость нейронных сетей обоснована незначительным снижением их производительности при неблагоприятных условиях. Эта особенность объясняется распределенным характером хранения информации в нейронной сети, благодаря чему можно утверждать, что только серьезные повреждения структуры нейронной сети существенно повлияют на ее работоспособность.

Имеется определенное сходство в использовании нейронных сетей и методов многомерного статистического анализа при оценке финансово-экономического состояния предприятия. Преимущество использования нейронных сетей, как инструмента оценки финансово-экономического состояния предприятия состоит в том, что взаимоотношения между величинами заранее не устанавливаются, поскольку метод предусматривает изучение существующих взаимосвязей на готовых моделях. Для нейронных сетей также не требуется никаких предположений относительно основного распределения

совокупности, а также, в отличие от многих традиционных статистических методов, они могут работать с неполными данными.

Нейронные сети представляют собой новую и весьма перспективную вычислительную технологию, дающую новые подходы к исследованию динамических задач в финансовой области. Способность к моделированию нелинейных процессов, работе с зашумленными данными и адаптивность дают возможность применять нейронные сети для решения широкого класса финансовых задач [3].

Актуальность исследования финансово-экономического состояния предприятия, включая оценку его кредитоспособности, связана с ростом влияния малого бизнеса на развитие мировой хозяйственной системы и признание его особой роли в решении ряда важнейших социально-экономических задач рыночной экономики.

Нейронная сеть – параллельно распределенная структура обработки информации, состоящая из нейронов, которые соединены между собой связями [7].

Нейронные сети - мощный метод моделирования, позволяющий воспроизводить сложные нелинейные зависимости. Способность к моделированию нелинейных процессов, работе с зашумленными данными и адаптивность дают возможность применять нейронные сети для решения широкого класса экономических задач. В последние несколько лет на основе нейронных сетей было разработано много программных систем для применения в таких вопросах, как прогнозирование и оценка экономических параметров, например, состояния предприятия, кредитоспособности заёмщика, вероятности банкротства предприятия, контроль за инвестициями, размещение займов и т.д.

Нейронные сети обучают на примерах. Разработчик нейронной сети подбирает представительные данные, а затем запускает алгоритм обучения, который автоматически воспринимает структуру данных. При этом от разработчика требуется набор эвристических знаний о том, как следует

отбирать и подготавливать данные, выбирать нужную архитектуру сети и интерпретировать результаты [9].

Для эффективного использования нейронных сетей необходимо наличие достаточного объема обучающей выборки, используя которую нейронную сеть можно обучить.

СТРУКТУРА СЕТИ

Нейрон представляет собой единицу обработки информации в нейронной сети. Общая модель нейрона состоит в следующем: нейрон имеет несколько каналов ввода информации – *дендриды* и канал вывода информации – *аксон*.

Аксон нейрона соединен с дендридами других нейронов с помощью *синапсов*. При возбуждении нейрон посылает сигнал по своему аксону. Через синапсы сигнал передается другим нейронам, которые в свою очередь могут возбуждаться или переходить в состояния тормажения. Нейрон возбуждается если суммарный уровень сигналов, прешедший в него превышает некоторый уровень – *порог* возбуждения или активации.

На рисунке 1 представленна модель нейрона, лежащего в основе искусственных нейронных сетей [12].

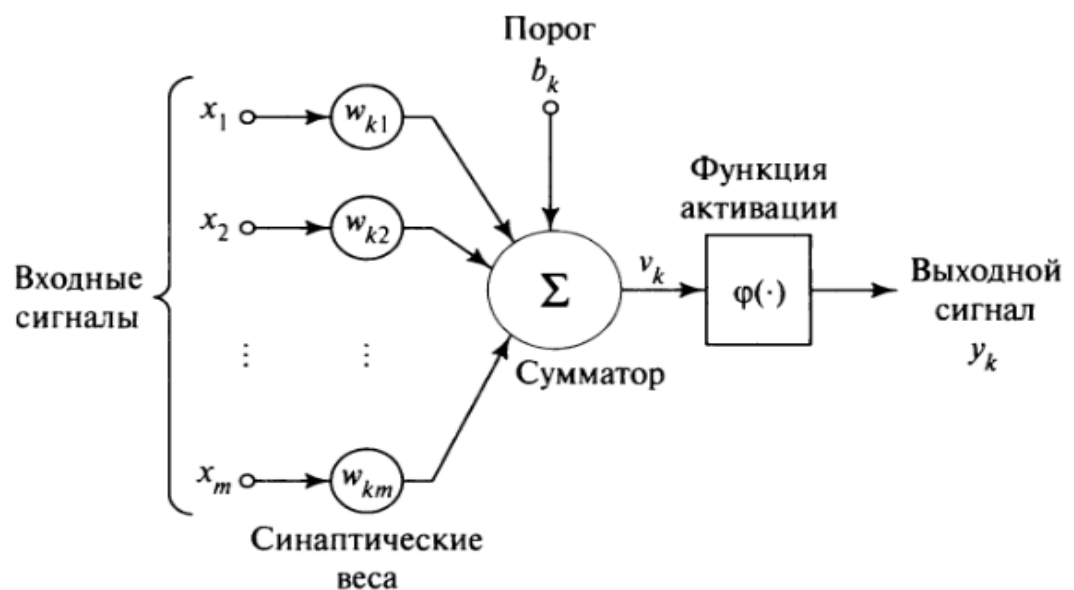


Рисунок 1 – Модель нейрона

Из рисунка видно, что на вход нейронна поступает сигналы x_i через входные каналы, каждый из которых проходит через соединение, имеющее определенную силу и вес w_{ki} . Коэффициенты w_{ki} называются *весами* синаптических связей, положительное значение которых соответствует возбуждающим синапсам, отрицательное значение – тормозящим синапсам. Если $w_{ki}=0$, то говорят что связь между нейроном k и нейроном i отсутствует. Далее производится преобразование сигналов в сумматоре и добавляется порог b_k , а результат преобразуется с помощью функции φ и подается на выход нейрона. В математическом представлении функционирование нейронна k можно описать следующей парой уравнений:

$$v_k = \sum_{i=1}^m w_{ki} x_i \quad (1)$$

где w_{ki} – синаптические веса нейрона k ;

x_i – входные сигналы;

$$y_k = \varphi(v_k + b_k) \quad (2)$$

где φ – функция активации;

v_k – линейная комбинация входных воздействий;

b_k – порог активации.

Полученный нейронами сигнал преобразуется с помощью нелинейной функции активации φ в выходной сигнал $y_k = \varphi(v_k)$.

Функция активации φ – это некоторая нелинейная функция, моделирующая процесс передачи возбуждения.

Используют несколько функций активации: функция единичного скачка, сигмоидальная функция, кусочно-линейная функция, функция знак. Одной наиболее распространенных является нелинейная функция активации с насыщением, так называемая *логистическая* функция или *сигмоидальная*.

$$\varphi(v) = \frac{1}{1 + \exp(-av)} \quad (2)$$

где a – параметр наклона сигмоидальной функции;

v – линейная комбинация входных воздействий.

Изменяя параметр a , можно построить функции с различной крутизны. Сигмоидальная функция может принимать бесконечное множество значений в диапазоне от 0 до 1, как показано на рисунке 2 [12].

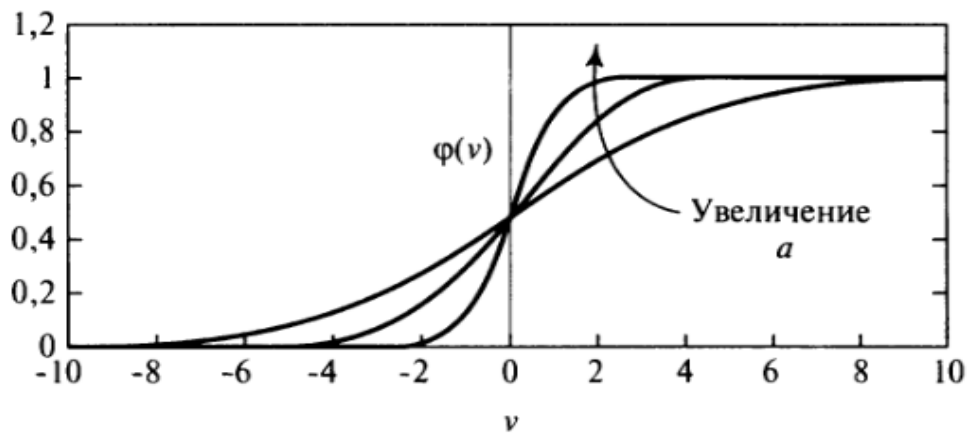


Рисунок 2 – график сигмоидальной функции.

Искусственная нейронная сеть построена из нейронов, связанных друг с другом. Организация нейронов и их связей в определенную структуру (архитектуру) оказывает значительное влияние на вычислительные возможности нейронной сети. Можно выделить три основных типа нейронных сетей: полносвязные сети (рис. 3-а), многослойные сети (рис. 3-б), слабосвязные сети (нейронные сети с локальными связями) (рис. 3-в).

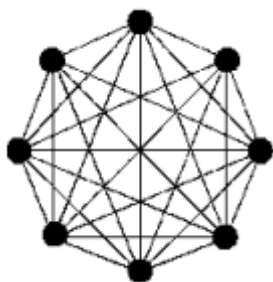


Рис.3-а

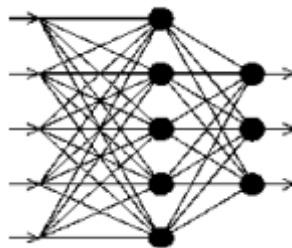


Рис. 3-б

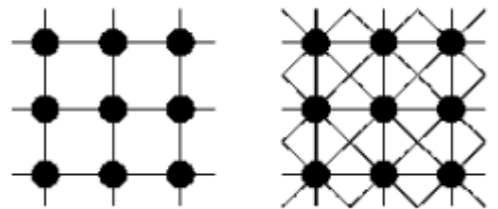


Рис. 3-в

ТОПОЛОГИИ НЕЙРОННЫХ СЕТЕЙ

Топологии сети можно разделить на две основные группы: сети прямого распространения и рекуррентные сети.

1. Сети прямого распространения.

В многослойной нейронной сети, нейроны располагаются по слоям. Нейроны первого слоя получают сигналы, преобразуют их и передают нейронам второго слоя, далее срабатывает второй слой, и т.д. до k -ого слоя, который выдает выходные сигналы. Такая сеть называется *многослойной сетью прямого распространения*. Промежуточные слои между внешним входным сигналом и выходным слоем называется *скрытым*, рисунок 4. Если все узлы каждого слоя соединены со всеми узлами смежных слоев, то такая сеть называется *полносвязной*. Если некоторые из синаптических связей отсутствуют, то сеть называется *неполносвязной*.

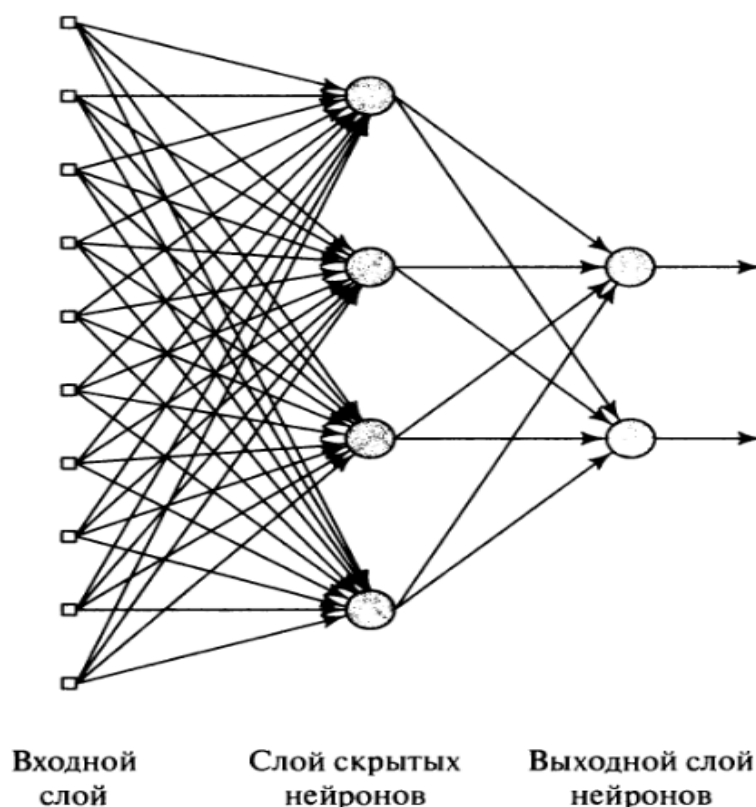


Рисунок 4 – Полносвязная многослойная сеть прямого распространения с одним скрытым слоем.

Сети с прямой связью могут служить универсальным средством для аппроксимации, но нет никакого правила, позволяющего найти оптимальную топологию сети для данной задачи.

2. Рекуррентные сети.

Рекуррентные сети отличаются от сетей прямого распространения наличием хотя бы одной обратной связи. Наличие обратных связей в сетях оказывает непосредственное влияние на их способность к обучению и на их производительность [12]. Такие сети представляют большой интерес для исследователей в области нейронных сетей. Однако при решении практических задач наиболее полезными оказались сети прямого распространения.

Выбор структуры (топологии) нейронной сети осуществляется в соответствии с особенностями и сложностью задачи. На рисунке 5 представлены наиболее распространенные архитектуры нейронных сетей [6].

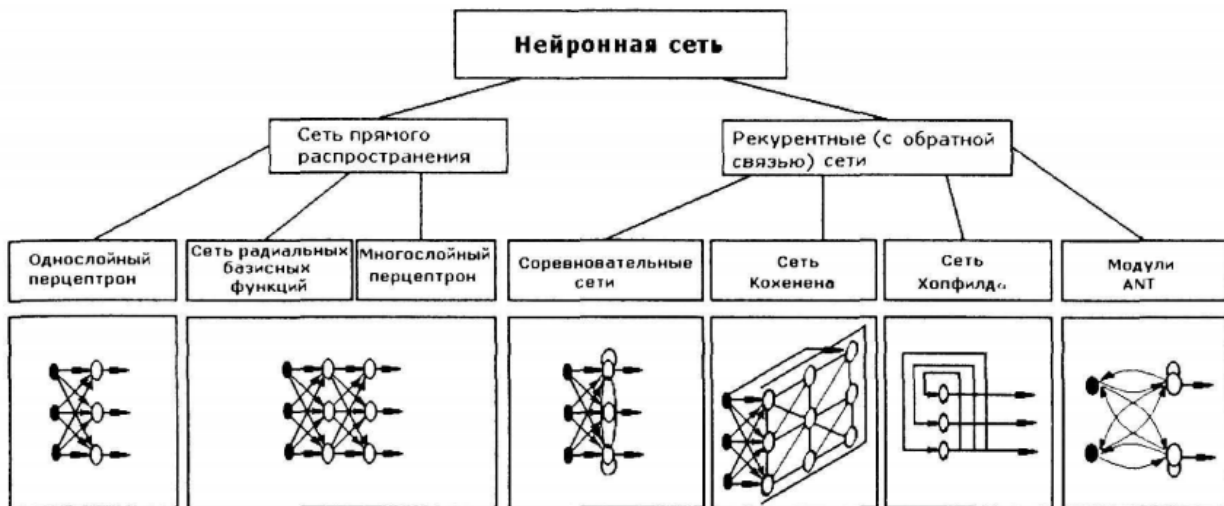


Рисунок 5 – Архитектуры нейронных сетей

ОБУЧЕНИЕ НЕЙРОННОЙ СЕТИ

Обучение сети – это подгонка ее внутренних параметров с целью адаптации к внешним воздействиям для решения конкретной задачи.

Процесс обучения нейронной сети является итеративным, а его шаги называют *эпохами* или *циклами*. На каждом цикле обучения происходит вычисление синаптических коэффициентов w_{ki} на основе примеров сгруппированных в обучающие множества (обучающую выборку). Такое множество состоит из данных, с указанным для каждого из них значением, выходного параметра, которое желательно получить. Действие, которое при этом происходит, можно назвать контролируемым обучением: учитель подает на вход сети вектор исходных данных, а на выходной узел сообщает желаемое значение результата. Контролируемое обучение нейронной сети можно рассмотреть как решение оптимизационной задачи. Разность между правильным решением и откликом представляет собой ошибку, которую следует уменьшить с помощью настройки свободных параметров. Корректируя сеть с целью минимизировать ошибку, мы не можем быть уверены что можно добиться еще меньшей ошибки[3].

Поверхность ошибки имеет сложное строение и обладает рядом неприятных свойств, в частности может иметь локальные минимумы, плоские участки, седловые точки и длинные узкие овраги. Аналитическими свойствами невозможно определить положение глобального минимума на поверхности ошибок, поэтому обучение нейронной сети, по сути заключается в исследовании этой поверхности. Достижение глобального минимума называется *сходимостью процесса обучения*. Разработано более сотни обучающих алгоритмов, отличающихся друг от друга стратегией оптимизации и критерием ошибок. Обычно, в качестве меры ошибки берутся среднеквадратичная ошибка E (3), которая определяется как сумма квадратов между желаемой величиной выхода и реально полученными на сети значениями y_k для каждого примера [3]

$$E = \frac{1}{P} \sum_{k=1}^p (d_k - y_k)^2 \quad (3)$$

где P – число примеров в обучающем множестве;

d_k – желаемая величина выхода;

y_k – значения полученные на сети.

Наиболее распространенный алгоритм обучения нейронных сетей с прямой связью – алгоритм *обратного распространения ошибки*. Этот метод основан на вычислении вектора градиента поверхности ошибок, который указывает направление кратчайшего спуска по поверхности из данной точки. Последовательность шагов приводит после ряда итераций к минимуму поверхности ошибок. Очевидную трудность здесь представляет выбор длины шага. При большой длине сходимость более быстрая, но есть опасность «перескочить» через решение, особенно в случаях, когда поверхность отклика имеет форму узкого оврага. При малом шаге направление продвижения выбирается правильное, но требуется много итераций для достижения минимума. На практике величина шага принимается пропорциональной крутизне склона с некоторой постоянной, называемой скоростью обучения.

Алгоритм обратного распространения ошибки является алгоритмом градиентного спуска, минимизирующим квадратичную ошибку E (4):

$$E = \frac{1}{2} \sum_{k=1}^p \sum_i (d_k^i - y_k^i)^2 \quad (4)$$

Индекс i пробегает все выходы многослойной сети.

На каждой эпохе на вход сети поочередно подается все обучающие наблюдения, выходные значения сети сравниваются с целевыми значениями, и вычисляется ошибка. Значение ошибки, а так же градиента поверхности ошибок используются для корректировки весов, после чего все действия повторяются. Процесс обучения прекращается либо когда пройдено определенное количество эпох, либо ошибка достигла некоторого уровня или вовсе перестает уменьшаться.

Опасность метода кроется в том, что таким образом мы минимизируем ошибку на обучающем множестве, и это не то же самое что минимизировать

«настоящую» ошибку на поверхности ошибок. Сильнее всего это различие сказывается в проблеме переобучения, или излишней подгонки.

ПРИМЕНЕНИЕ НЕЙРОННЫХ СЕТЕЙ В ЗАДАЧАХ КЛАССИФИКАЦИИ

Рассмотрим как нейронные сети с прямой связью используются в задачах классификации. В любой задачи классификации требуется отнести имеющиеся образы, например характеристики финансового положения, к определенным классам. Классификация и регрессия – основные области практического приложения нейронных сетей. Применительно к задачам менеджмента классификациям могут подвергаться фирмы, предприятия, поставщики и т.д. В структурном распознавании образов, образцы описываются тем, как они составлены из своих компонент, т.е. структурой. При статистическом подходе к распознаванию, образец представляется вектором x принадлежащим N -мерному пространству R^N , компоненты которого представляют собой различные характеристики образца. Классификатор относит объект x^k к тому или иному классу C в соответствии с определенным разбиением N -мерного пространства, которое называется *пространством входов*.

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).

НЕЙРОННЫЕ СЕТИ В ПАКЕТЕ STATISTICA 6.

В задачи классификации сеть должна отнести каждое наблюдение к одному из нескольких классов. В пакете ST Neural Networks американской фирмы Statsoft для классификации используется номинальная выходная переменная – различные ее значения соответствуют различным классам. Данный пакет относится к современным универсальным нейросетевым

продуктам и поэтому более совершенен по сравнению с ранее выпущенными. Он имеет мощные алгоритмы обучения сети (включая методы сопряженных градиентов и Левенберга-Маркара), возможность создания сложных комбинаций из сетей различных архитектур. Для этого пакета характерны простота в использовании и аналитические мощности, например Automatic Network Desingner (автоматический конструктор сети) определит наилучшую архитектуру для конкретной задачи, осуществит отбор переменных.

Однако в применении к практическим задачам имеется ряд проблем. Во-первых, заранее неизвестно какой сложности (размера) сеть потребуется для достаточно точной реализации отображения. Эта сложность может оказаться чрезмерно большой. Архитектура сети, т. е. число слоев, и число элементов в каждом слое, должны быть определены до начала обучения. Эта архитектура порождает сложные нелинейные разделяющие поверхности в пространстве входов. В сети с одним скрытым слоем, векторы образцов сначала преобразуются (нелинейным образом) в новое пространство представлений (пространство скрытого слоя), а затем гиперплоскости, соответствующие выходным узлам, располагаются так, чтобы разделить классы уже в этом новом пространстве. Тем самым, сеть распознает уже другие характеристики – «характеристики характеристик», полученные в скрытом слое. Все это подчеркивает важность этапа предварительной обработки данных.

При построении классификатора на основе нейронной сети можно выделить следующие этапы:

1 Данные:

а) Составить базу данных из примеров, характерных для поставленной задачи.

б) Разбить всю совокупность данных на три множества: обучающее, контрольное и тестовое.

2 Предварительная обработка:

а) Выбрать систему признаков, характерных для данной задачи, и преобразовать данные соответствующим образом для подачи на вход сети

(нормировка, стандартизация и т.д.). В результате желательно получить линейно отделимое представление множества образцов.

б) Выбрать систему кодирования выходного значения или значений

3 Конструирование, обучение и оценка качества сети:

а) Выбрать топологию сети: число элементов и структуру связей (входы, слои, выходы).

б) Выбрать функцию активации, которая будет использоваться.

в) Выбрать подходящий алгоритм обучения сети.

г) Оценить качество работы сети (по подтверждающему множеству или по информационному критерию) в зависимости от ее сложности с целью оптимизировать архитектуру (уменьшение весов, прореживание и т.д.).

д) Остановиться на варианте сети с наилучшей способностью к «обобщению» и оценить качество работы по тестовому множеству.

4. Использование и диагностика:

а) Выяснить степень влияния различных факторов на принимаемое решение (эвристики).

б) Убедиться, что доля случаев неправильной классификации достаточно мала.

в) При необходимости вернуться к шагу 2, изменив способ представления образцов или предварительно «почистив» базу данных.

г) Приступить к практическому использованию сети.

В рамках работы строится нейронная сеть, классифицирующая предприятия по 21 финансовому показателю его экономической деятельности с использованием пакета прикладных программ Statistica.

Для того чтобы сконструировать хороший классификатор, очень важно иметь в своем распоряжении высококачественные данные. Никакой метод построения классификаторов, будь то в области распознавания образов, машинного обучения или многомерной статистики, никогда не выдаст классификатор нужного качества, если имеющийся набор примеров не будет

достаточно богатым и представительным для той выборки, с которой придется работать в данной модели.

Обобщенная схема моделирования системы методами искусственных нейронных сетей представлена на рисунке 6 [1].

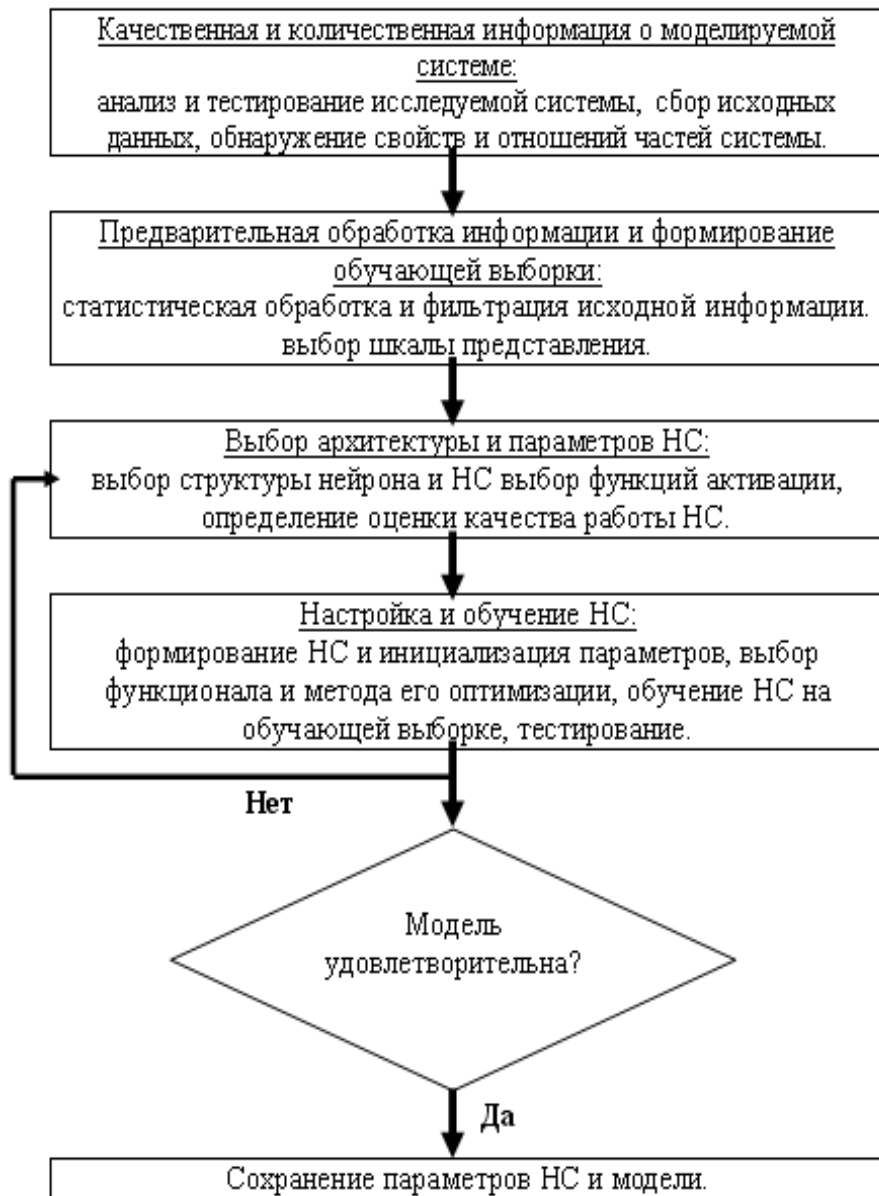


Рисунок 6– Обобщенная схема моделирования методами нейронных сетей

ПОСТРОЕНИЕ НЕЙРОСЕТЕВОЙ МОДЕЛИ МОНИТОРИНГА ПРЕДПРИЯТИЯ

При построении нейронной сети для оценки финансово-экономического состояния предприятия в работе используется следующий разработанный автором алгоритм:

Первым этапом при разработке нейронной сети определялся тип анализа, который необходимо провести. Целью создания нейронной сети является классификация предприятий с очень высоким, высоким, средним, низким и очень низким уровнями риска банкротства.

Вторым этапом является определение состава исходных данных и сбор этих данных. Обучающий набор данных представляет собой ряд наблюдений, для которых указаны значения входных и выходных переменных. Данные были получены из исследования [14], в котором сгенерирована виртуальная модельная база данных предприятий, состоящая из 6000 наблюдений, для которых известны значения 21 входного параметра выбранной системы финансовых показателей. На основе исследования все предприятия были разбиты на пять групп:

1 группа – предприятия очень высокого уровня риска банкротства, что соответствует номинальному значению «отказ» (1517 предприятий);

2 группа – предприятия, находящиеся в зоне высокого риска банкротства, что соответствует номинальному значению «высокий» (572 предприятий);

3 группа – предприятия, находящиеся в зоне среднего риска банкротства, что соответствует номинальному значению «средний» (1687 предприятий);

4 группа – предприятия, находящиеся в зоне низкого риска банкротства, что соответствует номинальному значению «низкий» (1537 предприятий);

5 группа – кредитоспособные предприятия, которым можно выдать кредит, что соответствует номинальному значению «кредит» (687 предприятий).

В таблице 1 приведены примеры 10 наблюдений по два из каждого класса.

Таблица 1– наблюдения

	1 R3	2 R4	3 R5	4 L1	5 L2	6 P1	7 A2	8 A4	9 1/A5	10 A6	11 F1	12 F2	13 F3	14 F4	15 R6	16 L3	17 1/A1	18 1/A3	19 1/F8	20 F11	21 P2	22 Класс
1	-0.00694	-0.00092	0.09183	0.39524	0.18569	0.53405	0.49529	0.31418	0.23320	1.66186	1.19827	0.07208	0.13616	0.08479	-0.58589	0.16522	0.03280	0.73688	0.09503	0.13369	0.05034	ВЫСОКИЙ
2	-0.08258	-0.03248	0.09086	0.44960	0.16754	0.94607	0.57086	0.77651	0.26511	1.35202	1.12044	0.23395	0.19959	0.06148	-0.01530	0.14728	0.03062	0.74217	0.11137	0.20347	0.06992	ВЫСОКИЙ
3	0.61608	0.17160	0.64619	1.32549	0.80329	1.84599	1.92202	2.08267	0.00000	5.62034	0.32024	0.78473	0.82480	0.30507	2.89910	0.81597	0.00740	0.21136	0.65704	1.19460	0.95331	НИЗКИЙ
4	0.91176	0.19073	0.76829	1.02518	0.56681	1.60651	2.25945	3.96261	0.00000	6.98772	0.30990	0.79202	0.94991	0.40102	1.59351	0.75325	0.00847	0.23068	0.65171	1.38642	0.65768	НИЗКИЙ
5	-0.67949	-0.17843	-0.45054	0.00000	0.06703	0.00000	0.00000	-2.83541	2.20408	0.00000	-1.77486	-0.58513	-1.15624	-0.14857	-3.90088	0.02611	0.04381	-0.13025	-0.32048	-0.65816	-0.65203	ОЧЕНЬ ВЫСОКИЙ
6	-0.10082	-0.16753	-0.74218	0.00000	0.00000	0.30747	0.00000	-2.82965	1.34353	0.29164	2.61788	-0.13309	-0.26190	-0.28449	-3.91033	0.08949	0.04339	-0.17560	-0.01636	-0.45504	-0.07352	ОЧЕНЬ ВЫСОКИЙ
7	0.42651	0.01892	0.15144	0.93878	0.21502	1.31476	1.58670	4.55395	0.04459	3.24411	0.59574	0.40926	0.49738	0.25544	0.04143	0.43129	0.02858	0.58104	0.57394	0.71893	0.34119	СРЕДНИЙ
8	0.18374	0.09899	0.25435	0.96933	0.42556	1.05109	0.95603	7.38152	0.03220	2.31966	0.68967	0.43020	0.32249	0.20534	0.68460	0.52499	0.01460	0.55728	0.20408	0.53027	0.12736	СРЕДНИЙ
9	1.14564	0.16127	1.07780	1.14026	0.88782	3.14031	5.11855	1.15435	0.00000	10.00853	0.00000	1.00000	1.57473	0.75231	5.84793	1.26116	0.00000	0.01189	1.69721	1.84663	1.01747	ОЧЕНЬ НИЗКИЙ
10	1.53158	0.13404	1.84675	1.46049	0.52838	3.33755	4.32331	1.52229	0.00000	9.35722	0.00000	1.00000	3.27565	0.54961	3.01624	1.23373	0.00000	0.10618	1.05763	2.15816	1.25398	ОЧЕНЬ НИЗКИЙ

Построенную модельную базу предприятий разобьем на две выборки – обучающую и контрольную. В качестве тестового множества будем рассматривать данные 400 реально существующих предприятий, прообраз которых лег в основу генерации модельной базы.

Третьим этапом является построение и обучение нейронной сети. Мы строили сеть, обучение которой есть минимизация квадратичной ошибки на обучающем множестве с использованием градиента. Градиент оценки вычисляли методом двойственности, а именно, методом обратного распространения ошибки, который является алгоритмом градиентного спуска. Функцией активации выбрали логистическую функцию, а функцией ошибки – среднеквадратичную. Для выбора минимального числа нейронов и определения структуры сети использовали процедуру контрастирования.

Если ошибка на обучающем множестве не удовлетворительная то необходимо вернуться и изменить обучающую выборку или метод обучения сети.

Четвертый этап - тестирование и использование нейронной сети. Для тестирования применяется тестовая выборка и анализируются значения классификации. Если сеть обученна не верно и результаты классификации неудовлетворительны необходимо изменить архитектуру сети.

Для запуска нейронных сетей в пакете Statistica Neural Networks необходимо зайти в меню Анализ стартовой панели инструментов и выбрать нейронные сети.

Откроется окно *Нейронные сети* как показано на рисунке 7, которое содержит три вкладки: быстрый, дополнительно и сети/ ансамбли. На вкладке быстрый необходимо указать тип решаемой задачи. В нашем случае тип решаемой задачи - классификация. Так же возможен выбор инструмента – мастер решений, для анализа данных и создания сетей, требующий минимального вмешательства, и проводящий все необходимые фазы анализа.

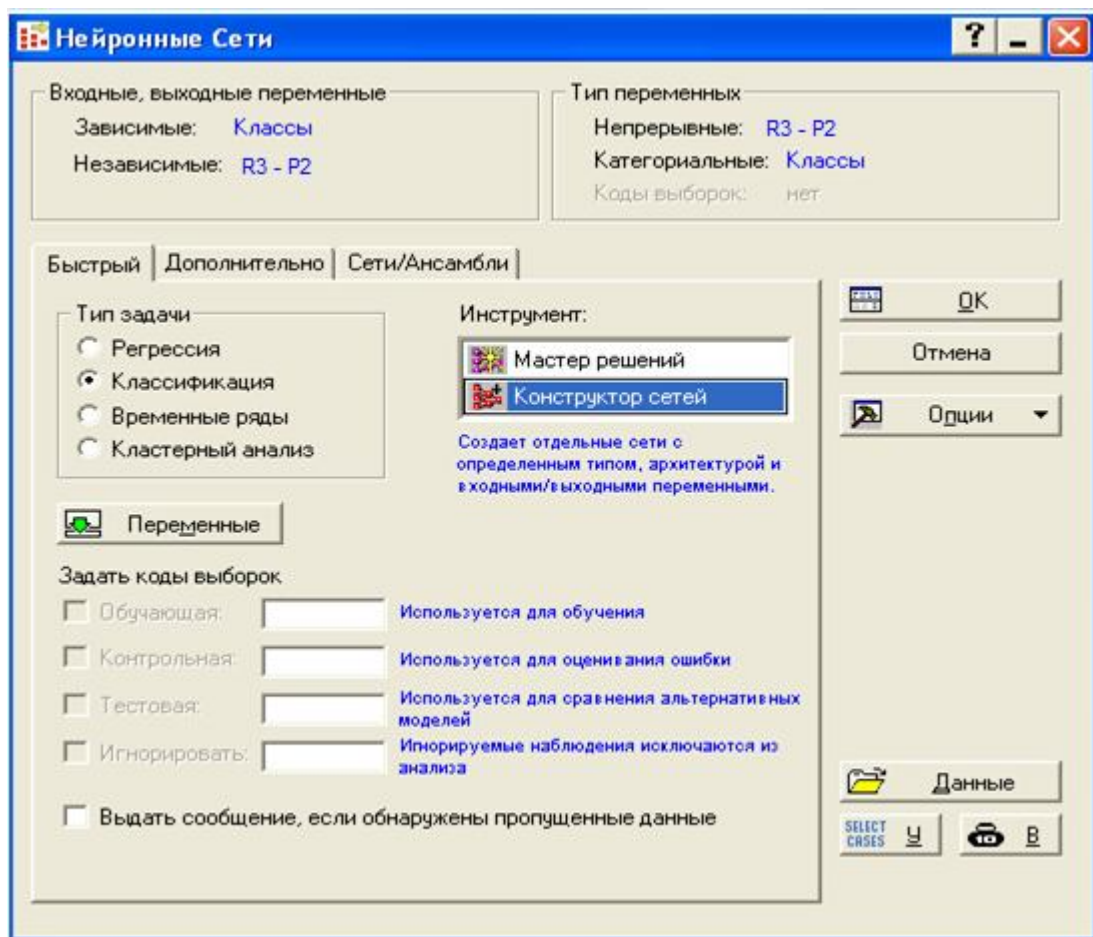


Рисунок 7 – окно нейронные сети

Нажимаем кнопку *Переменные* и выберем зависимые и независимые переменные (рис. 8).

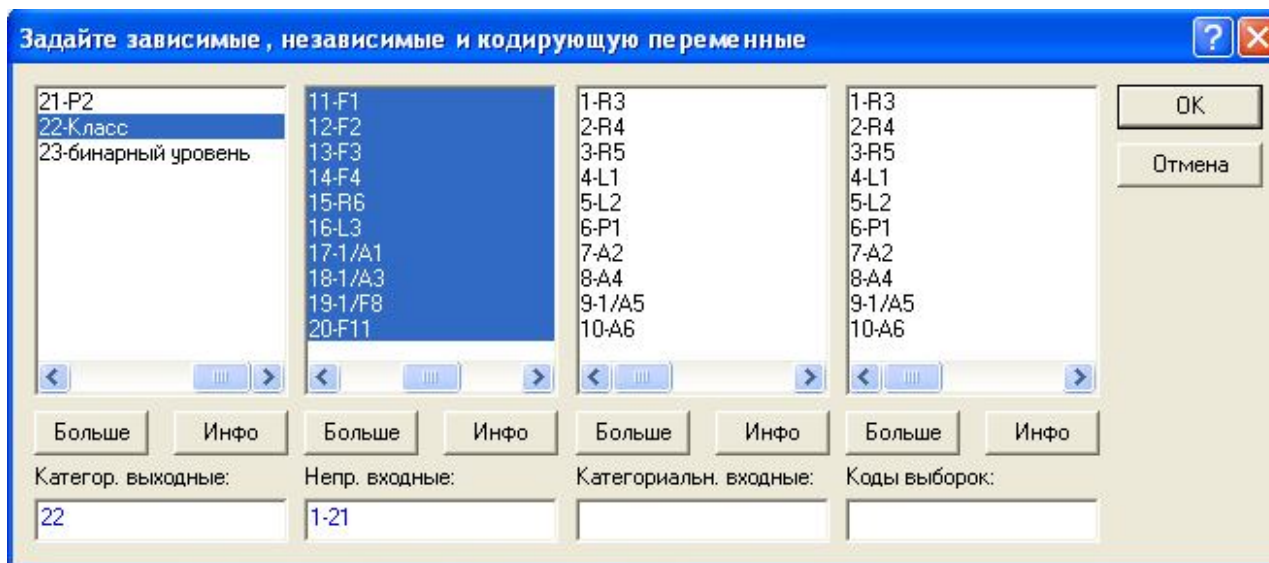


Рисунок 8 – Окно выбора переменных

В качестве категориальной выходной переменной выбираем *Класс*, а 21 финансовый показатель выбираем как непрерывные входные значения. Коды выборок будут активны, если задать переменную, с кодами выборок. Нажатие кнопки *OK* возвращает нас к окну *Нейронные сети*.

Для выбора архитектуры сети в окне нейронные сети выбираем *конструктор сети* – это общий инструмент, позволяющий выбрать индивидуальные особенности архитектуры сетей и обучающие алгоритмы, в соответствии с решаемой задачей. Вкладка *Сети/ансамбли* позволяет выбрать сеть, если она уже построена и сохранена в отдельном файле. К данной вкладке мы вернемся позже.

Нажатие кнопки *OK* переводит нас в окно *Конструктор сети* (рис. 9), в котором, на вкладке *Быстрый*, есть возможность выбора типа используемой сети (топологии сети). Задача классификации может решаться с помощью многослойного персептрона (MPL), радиальной базисной функции (RBF), вероятностной нейронной сети и сетей Кохонена. Сети Кохонена обычно применяются когда располагают только входными значениями.

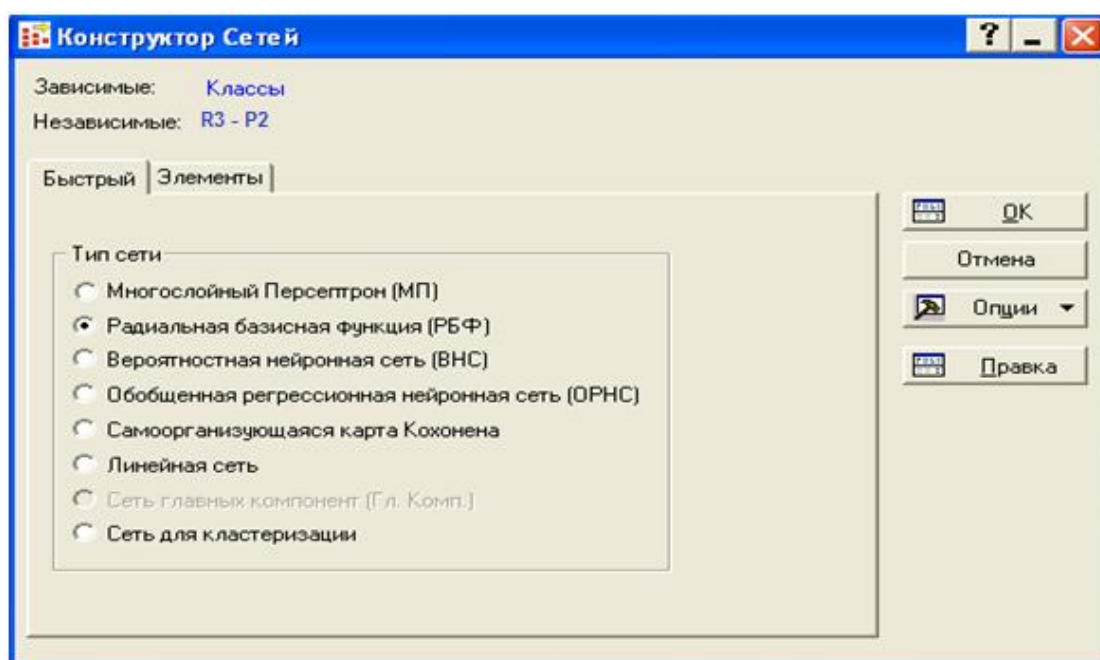


Рисунок 9 – Окно конструктора сети, вкладка быстрый

Хотя байесова сеть обычно находит примерный путь решения задачи, она, вообще говоря, не выдает результат такого качества, как сети других архитектур.

Сеть на основе радиальной базисной функции и многослойный персептрон являются примерами нелинейной многослойной сети прямого распространения. Всегда существует сеть RBF, способная имитировать многослойный персептрон, но эти два типа всегда отличаются по некоторым важным аспектам:

- Сети RBF (в своей основной форме) имеют один скрытый слой, в то время как MLP может иметь большее количество слоев.
- Обычно вычислительные узлы MLP в скрытых и выходном слоях используют одну и ту же модель нейрона. Вычислительные узлы RBF могут в корне отличаться от узлов выходного слоя и служить разным целям.
- Скрытый слой в RBF является нелинейным, в то время, как выходной – линейным. В сетях MLP, построенных для задач классификации, и скрытые, и выходной слои являются нелинейными.
- Аргумент функции активации каждого скрытого узла сети RBF представляет собой Евклидову норму (расстояние) между входным вектором и центром радиальной функции. Аргумент функции активации каждого скрытого узла сети MLP является скалярным произведением входного вектора и вектора синаптических весов данного нейрона.
- MLP обеспечивает глобальную аппроксимацию нелинейного отображения. RBF с помощью уменьшающихся локализованных нелинейностей (т.е. функций Гаусса) обеспечивает локальную аппроксимацию нелинейного отображения.

Согласно следствию теоремы Колмогорова-Арнольда-Хехт-Нильсена [монография Muller B., Reinhart J. Neural Networks: an introduction, Springer-Verlag, Berlin Heidelberg, 1990.] показано, что для любого множества пар $(X_k; Y_k)$ (где X_k и Y_k - скаляры) существует однородная нейронная сеть первого порядка с одинаковыми функциями активации, с одним промежуточным слоем, с последовательными связями и с конечным числом нейронов, которая выполняет отображение $X \rightarrow Y$, выдавая на каждый входной сигнал X_k правильный выходной сигнал Y_k . Нейроны в такой нейронной сети должны иметь сигмоидальные передаточные функции.

Количество скрытых слоев в RBF соответствует 1, как и в трехслойном персептроне (входной, скрытый, выходной). Поэтому при тестировании в качестве типа нейронной сети мы выбирали трехслойный персептрон и радиальную базисную функцию.

Во всех случаях нужно помнить, что на результат обучения сети могут сильно влиять параметры алгоритма обучения. Окончательный выбор зависит от того какая требуется скорость выполнения. Как правило, чем меньше сеть тем дольше она обучается и тем быстрее она работает [12]. Самое важное – выбрать число элементов в промежуточном слое, что можно сделать на вкладке *Элементы окна Конструктор сети* (рис. 10а-б).

Заметим, что в зависимости от выбора топологии сети (на вкладке *Быстрый*) вкладка с заданием параметров сети так же изменяется. Если мы в качестве типа сети выбрали многослойный персептрон, то на вкладке *Элементы* необходимо задать количество промежуточных слоев и число элементов в них, а так же функцию классификации (рис 10 а). Если же мы строим сеть на основе радиальной базисной функции, то количество скрытых слоев соответствует одному по определению, поэтому вкладка *Элементы* допускает возможность выбора лишь числа нейронов в промежуточном слое и вида функции классификации (рис 10 б)

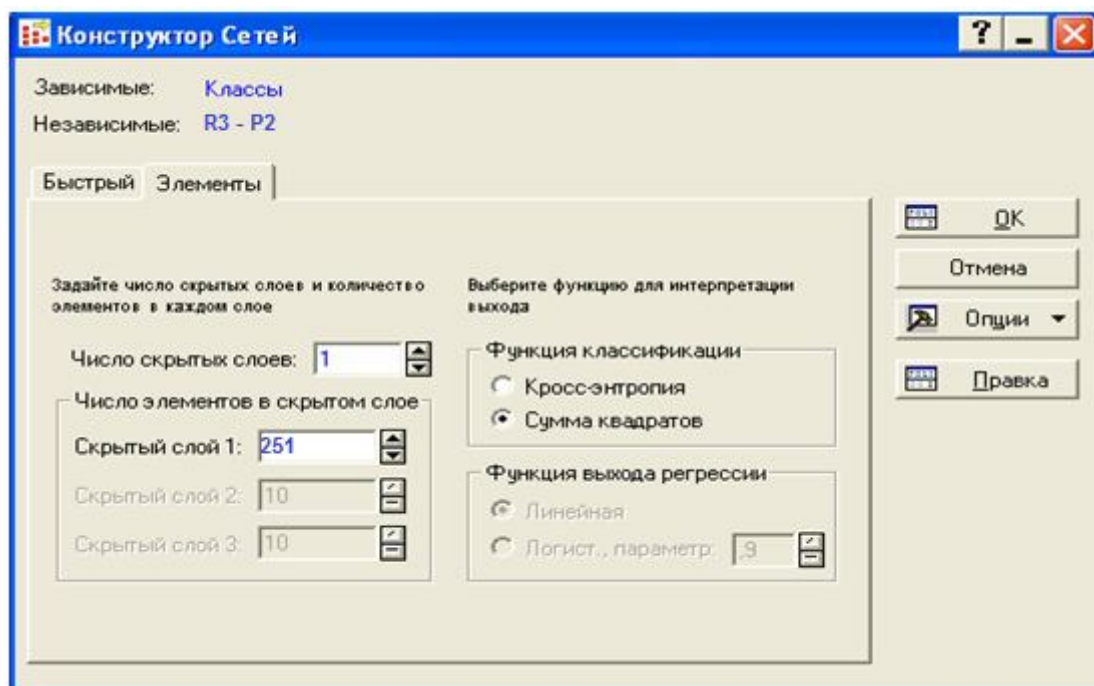


Рисунок 10а – Конструктор сетей, вкладка элементы для сети MLP

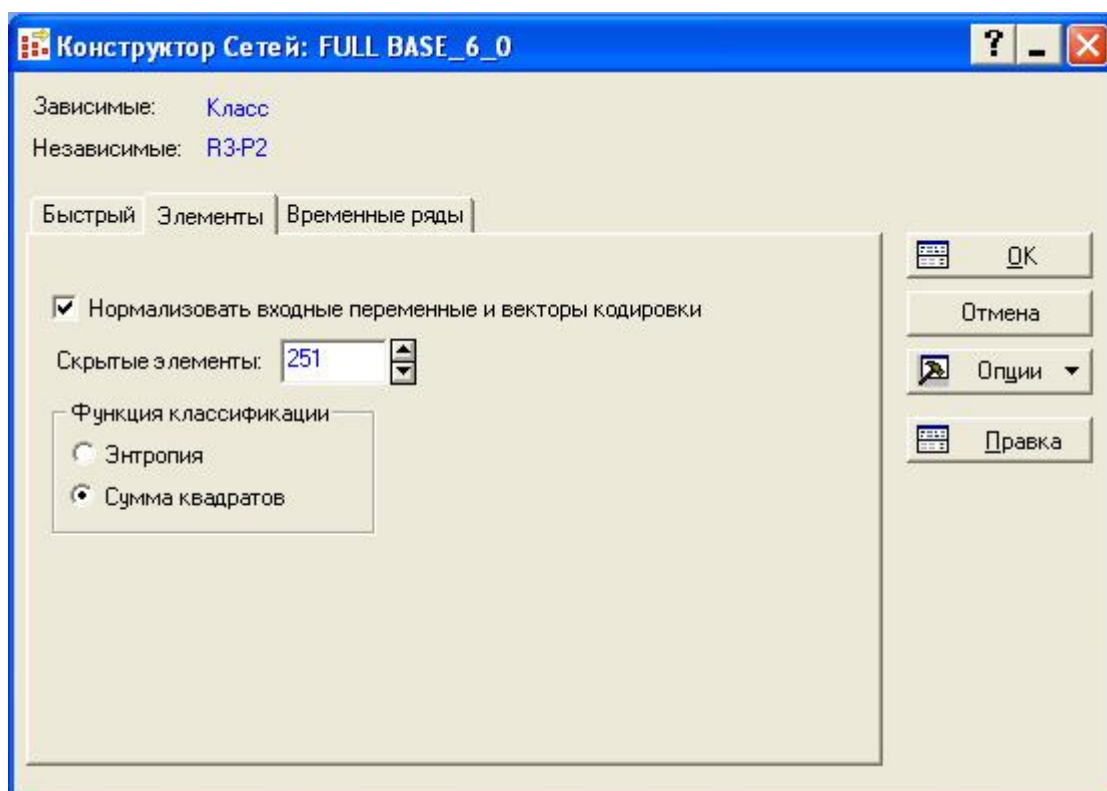


Рисунок 10б – Конструктор сетей, вкладка элементы для сети RBF

По условию задачи размерность входного вектора равна $N_x=21$, по этому входной слой содержит 21 нейрон; число нейронов в выходном слое соответствует числу классов ($N_y=5$), на которые предполагается разбить выборку данных. Неизвестным является число нейронов в промежуточном слое N . Для оценки числа нейронов с скрытом слое однородной нейронной сети мы воспользовались формулой для оценки необходимого числа синаптических весов N_w в многослойной сети с сигмоидальными передаточными функциями [Widrow B., Lehr M.A. 30 years of adaptive neural networks: perceptron, madaline, and backpropagation // Proceedings of the IEEE, vol. 78, No. 9, September, 1990, p. 1415-1442.]:

$$\frac{N_y N_p}{1 + \log_2(N_p)} \leq N_w \leq N_y \left(\frac{N_p}{N_x} + 1 \right) (N_x + N_y + 1) + N_y,$$

где N_y - размерность выходного сигнала, N_p - число элементов обучающей выборки, N_x - размерность входного сигнала.

$$\frac{5 \cdot 6000}{1 + 12,6} \leq N_w \leq 5 \cdot \left(\frac{6000}{21} + 1 \right) (21 + 5 + 1) + 5$$

Округляя до целых, получим

$$2222 \leq N_w \leq 7746$$

Оценив необходимое число весов, можно рассчитать число нейронов в скрытых слоях. Так, число нейронов в нейронной сети с одним скрытым слоем составит:

$$N = \frac{N_w}{N_x + N_y}$$

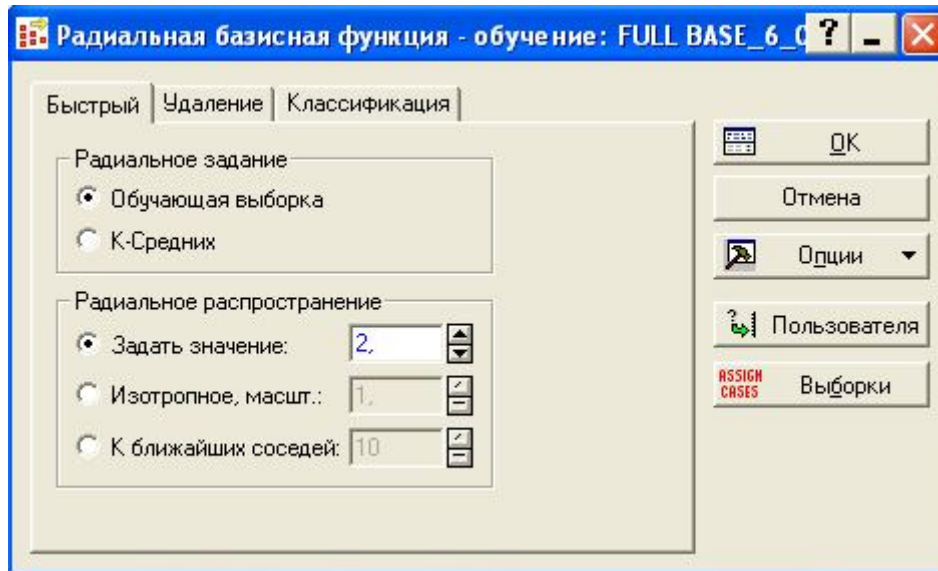
$$N_{\min} = \frac{2222}{26} \approx 85 \text{ и } N_{\max} = \frac{7746}{26} \approx 297$$

То есть $85 \leq N \leq 297$.

Экспериментируя с количеством элементов в скрытом слое из полученного диапазона значений N и типом сети (RBF или MLP) приходим к

выводу, что минимальное значение квадратичной ошибки на обучающем множестве соответствует сети RBF с 251 элементом в скрытом слое.

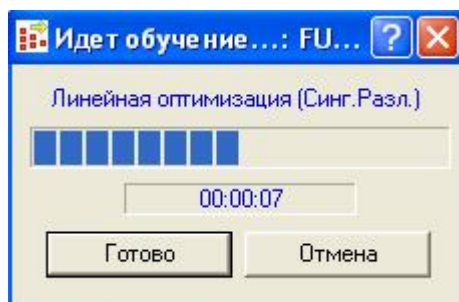
В случае выбора построения сети на основе радиальной базисной функции после нажатия кнопки ОК открывается дополнительное окно ввода параметров обучения сети (рис.11а-б). На вкладке Быстрый (рис. 11а)



На вкладке Классификация (рис.11 б)

На вкладке удаление предоставляется возможность удалить скрытые элементы с малыми выходными весами, задав порог и входные элементы с низкой чувствительностью после обучения сети. Параметрами данной вкладки мы не будем пользоваться, так как построение сети ориентируем на максимально точный результат.

Выбрав опции, нажав *ОК* в окне задания параметров обучения, сеть обучается (рис. 12)



Когда сеть обучена перед пользователем появляется окно вывода результатов (рис.13). В нем присутствуют вкладки *Быстрый*, *Дополнительно*, *Предсказанные*, *Чувствительность* и *Описательные*. Рассмотрим основные возможности на каждой из вкладок.

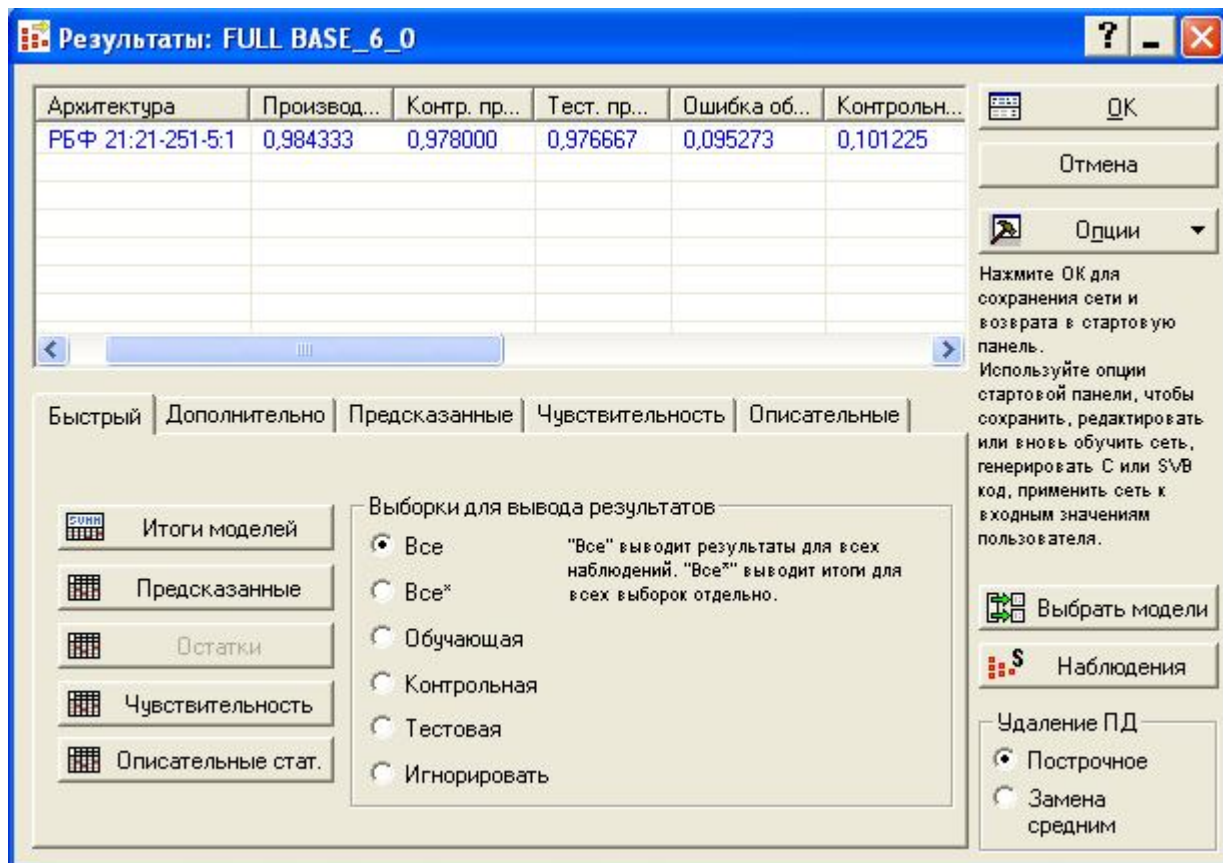


Рисунок 13 – Окно вывода результатов

В верхней части окна вывода результатов представлена информация о параметрах построенной модели нейронной сети. В более удобном (табличном) варианте эту информацию можно получить, если воспользоваться кнопкой *Итоги моделей* вкладки *Быстрый* (таблица 2)

Таблица 2 – Итоги моделей

Архитект. Сети	Производ. обучения	Контрольн. Производит	Тест. производ.	Ошибка обуч.	Контрольн. ошибка	Тестовая ошибка	Входы	Скрытый слой
РБФ 21:21- 251-5:1	0,984333	0,978000	0,976667	0,095273	0,101225	0,104674	21	251

Воспользовавшись опцией *Предсказанные* вкладки *Быстрый* мы получаем таблицу данных со значениями входных данных (финансовых коэффициентов), исходным классом принадлежности, предсказанным классом, значениями уровней доверия по классам, значениями элементов вектора кодировки по числу входных параметров и выигравший нейрон. Вывод тех или иных перечисленных параметров таблицы можно задать на вкладке *Предсказанные* окна вывода результатов (рис. 14)

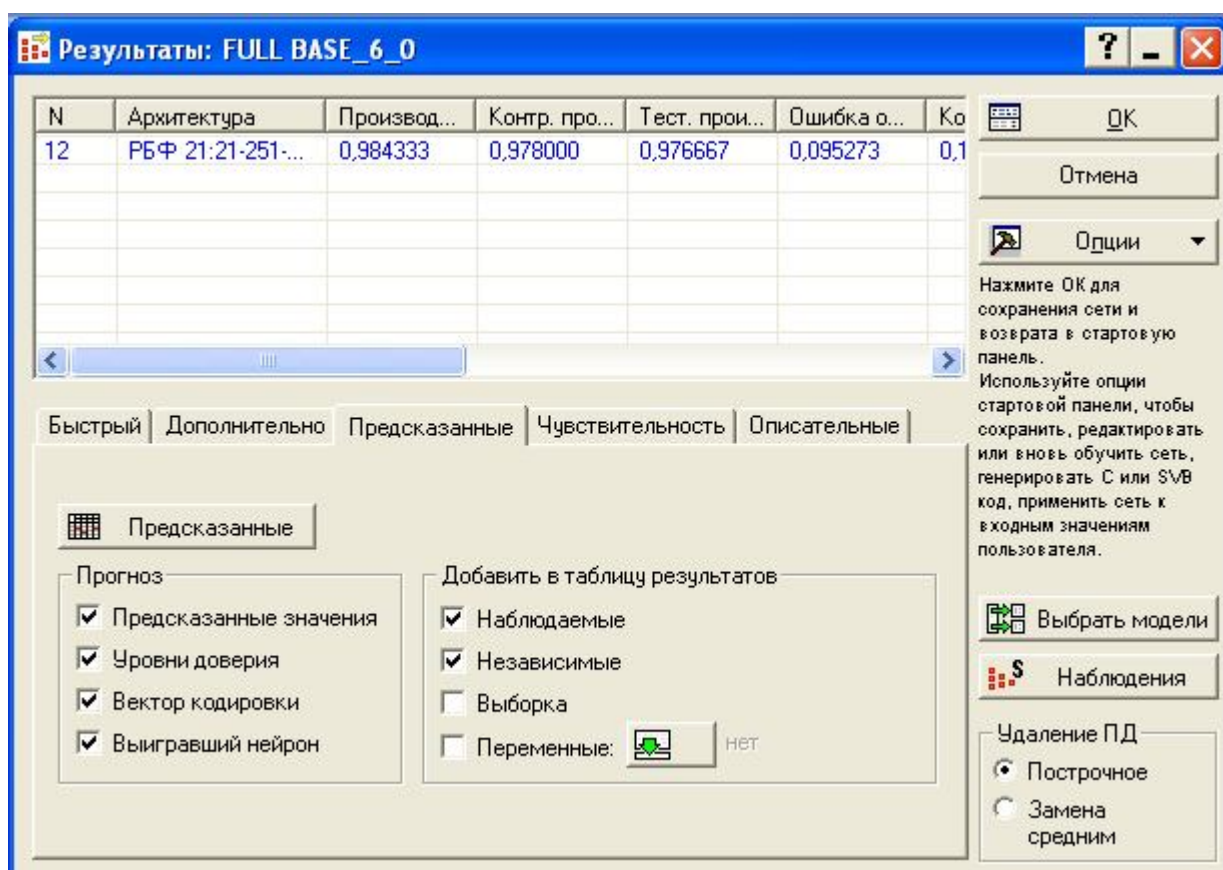


Рисунок 14

На вкладке *Быстрый* опция *Описательные статистики* позволяет вывести матрицу ошибок классификации (таблица 3) и таблицу результатов классификации (таблица 5). В матрице ошибок классификации отображена вся информация о том, сколько объектов каждого из наблюдаемых классов зависимой переменной отнесено по результатам классификации к тому или иному классу. По диагонали показано количество правильно классифицированных переменных. Также таблице указан объем обучающей

выборки. Исходные классы соответствуют столбцам матрицы, предсказанные классы – строкам.

	Очень низкий	Низкий	Средний	Высокий	Очень высокий
Очень низкий	683	20	0	0	0
Низкий	4	1510	29	0	0
Средний	0	7	1652	42	0
Высокий	0	0	4	523	0
Очень высокий	0	0	2	7	1517

Данные результатов классификации представлены в таблице 4. Статистика выдаётся отдельно для обучающей и тестовой выборок. В верхней части таблицы можно увидеть суммарные статистики (общее число предприятий в каждом классе, число классифицированных правильно, ошибочно и неизвестно), а в нижней части – кросс-результаты классификации в % (какой процент предприятий из данного столбца был отнесён к данной строке).

Таблица 4

	Очень низкий	Низкий	Средний	Высокий	Очень высокий
Всего	687,0000	1537,000	1687,000	1537,000	572,0000
Правильно	683,0000	1510,000	1652,000	1510,000	523,0000
Ошибочно	4,0000	27,000	35,000	27,000	49,0000
Неизвестно	0,0000	0,000	0,000	0,000	0,0000
% правильных	99,4178	98,243	97,925	98,243	91,4336
% ошибочных	0,5822	1,757	2,075	1,757	8,5664
% неизвестных	0,0000	0,000	0,000	0,000	0,0000

Из таблицы видно, что процент верно классифицированных предприятий для каждого из классов высок: 99,42%, 98,24%, 97,93%, 98,24% и 91,43% для

классов очень низкий, низкий, средний, высокий и очень высокий уровни риска классификации. Общий процент неверно классифицированных составил всего 2,367% (142 предприятия из 6000). Так же заметим, что неклассифицированных предприятий нет.

Возможность вывода матрицы ошибок классификации и результатов классификации задается на вкладке *Описательные* (рис.15)

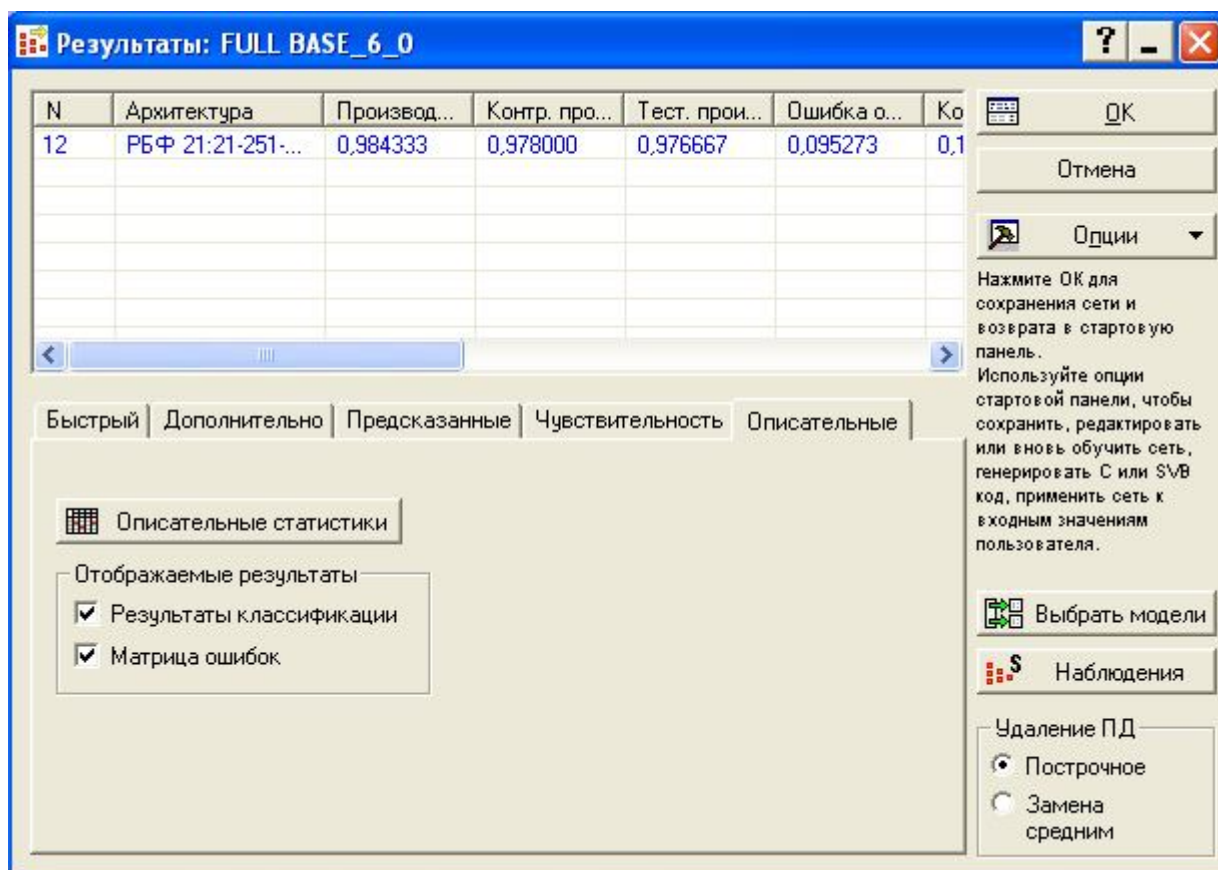


Рис.15

На вкладке *Чувствительность* пользователю по кнопке представляются результаты анализа чувствительности в виде таблицы (рис 16).

Анализ чувствительности - 12 (FULL BASE_6_0)																			
	R3	R4	R5	L1	L2	P1	A2	A4	1/A5	A6	F1	F2	F3	F4	R6	L3	1/A1	1/A3	1/F8
Отношение:	1,09	1,05	1,05	1,13	1,28	1,23	1,36	1,18	1,05	1,47	1,01	1,10	1,03	1,07	1,08	1,20	1,09	1,49	1,05
Ранг.	12,00	17,00	19,00	9,00	4,00	5,00	3,00	7,00	16,00	2,00	21,00	10,00	20,00	14,00	13,00	6,00	11,00	1,00	18,00
																			8,00
																			15,00

Рис.16

Анализ чувствительности позволяет сделать вывод о важности входных переменных для конкретной нейронной сети. При необходимости можно удалить входную переменную с низким показателем чувствительности. Он

позволяет выделить ключевые переменные, без которых анализ невозможен, и идентифицировать те, которые можно без ущерба исключить из рассмотрения. Вообще говоря анализ чувствительности не позволяет надежно определить полезность переменных, по этому выводы о важности переменных нужно делать аккуратно. Ранги переменным присваиваются по порядку, начиная с переменных с самым высоким значением отношения шансов. Максимальное значение отношения шансов соответствует переменной $1/A3$, поэтому именно ей присвоен ранг со значением 1.

На вкладке Дополнительно (рис 17) опция *Архитектура сети* выводит окно *Параметры сети*, в котором на вкладке *Наблюдения пользователя* (рис.18) предоставляется возможность просмотра классификации на основе значений финансовых показателей каждого из 6000 виртуальных предприятий модельной базы.

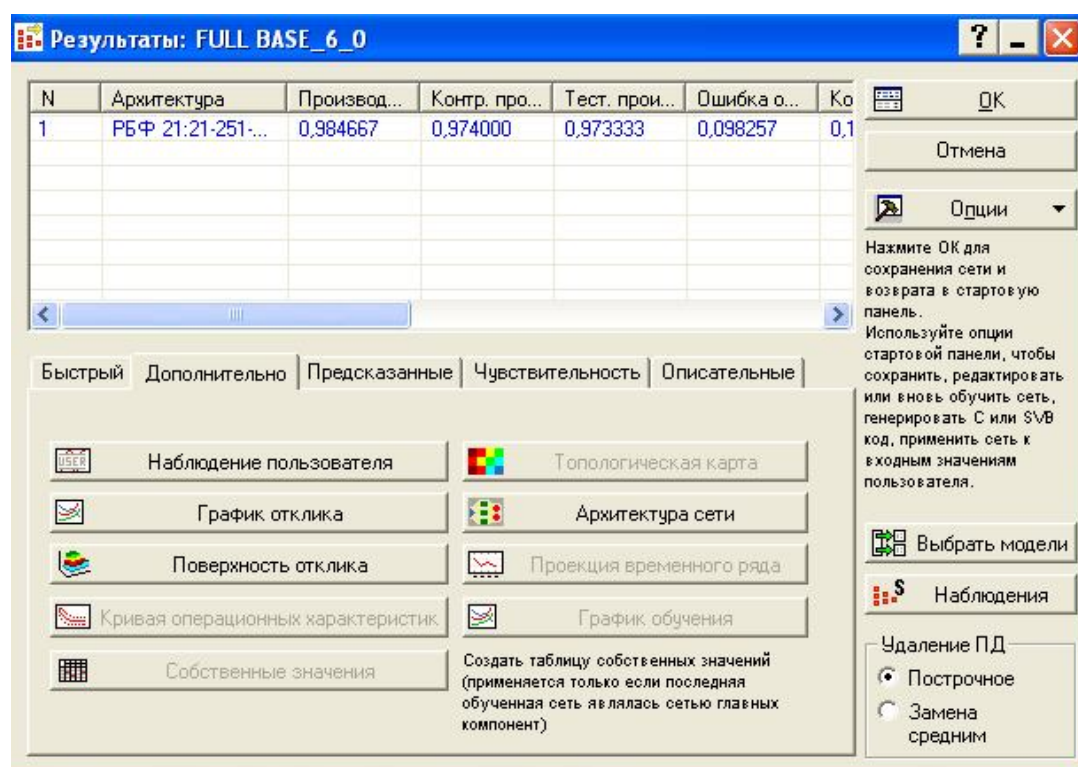


Рис.17

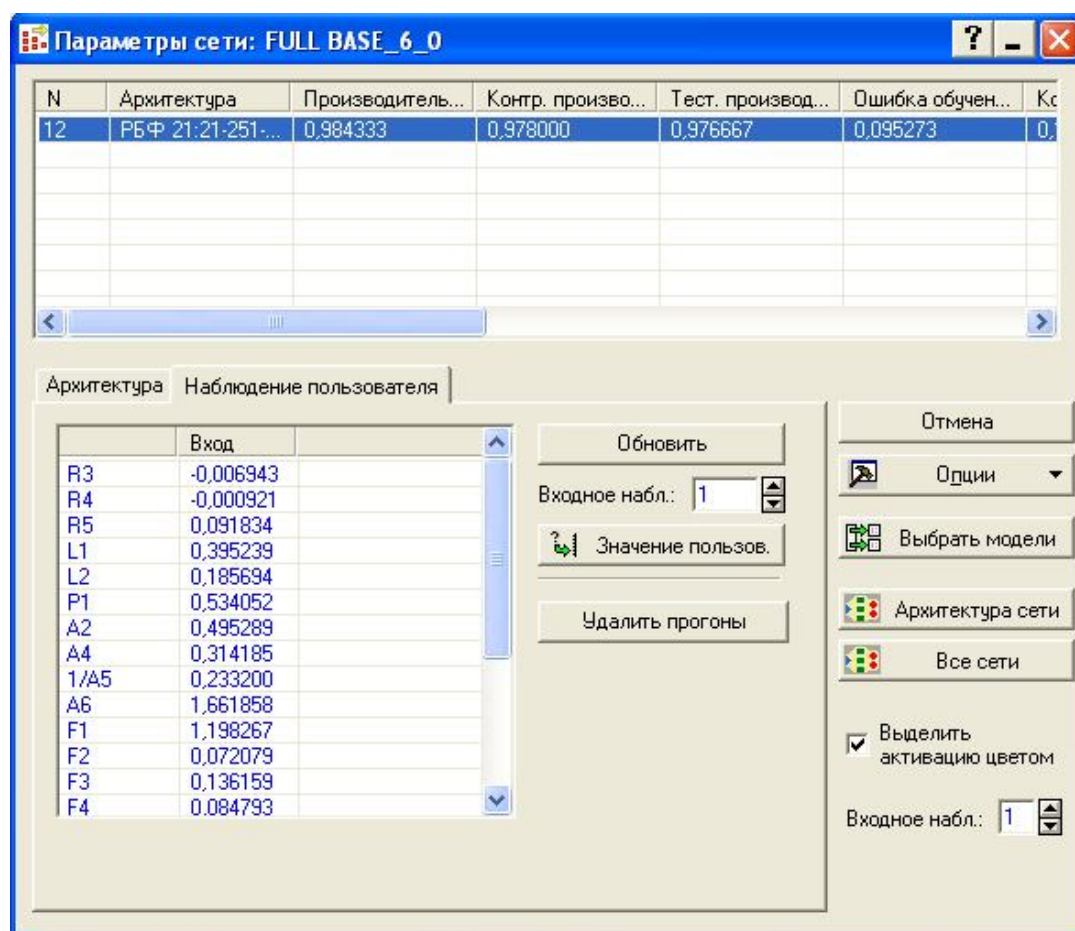


Рис.18

На данной вкладке есть возможность ввода номера наблюдения или данных финансовых показателей абсолютно нового предприятия, согласно которым будет произведена классификация на основе построенной нейронной сети. На вкладке *Архитектура* выводится архитектура построенной сети (рис.19) с указанием активации каждого нейрона для того наблюдения, которое было выбрано или задано на вкладке *Дополнительно*. Интенсивность окраса нейронов соответствует их активациями, показывая визуальную индикацию активности сети. Видно, что на выходном слое положительная активация только у одного нейрона (темный окрас), который и определяет уровень очень низкий риск.

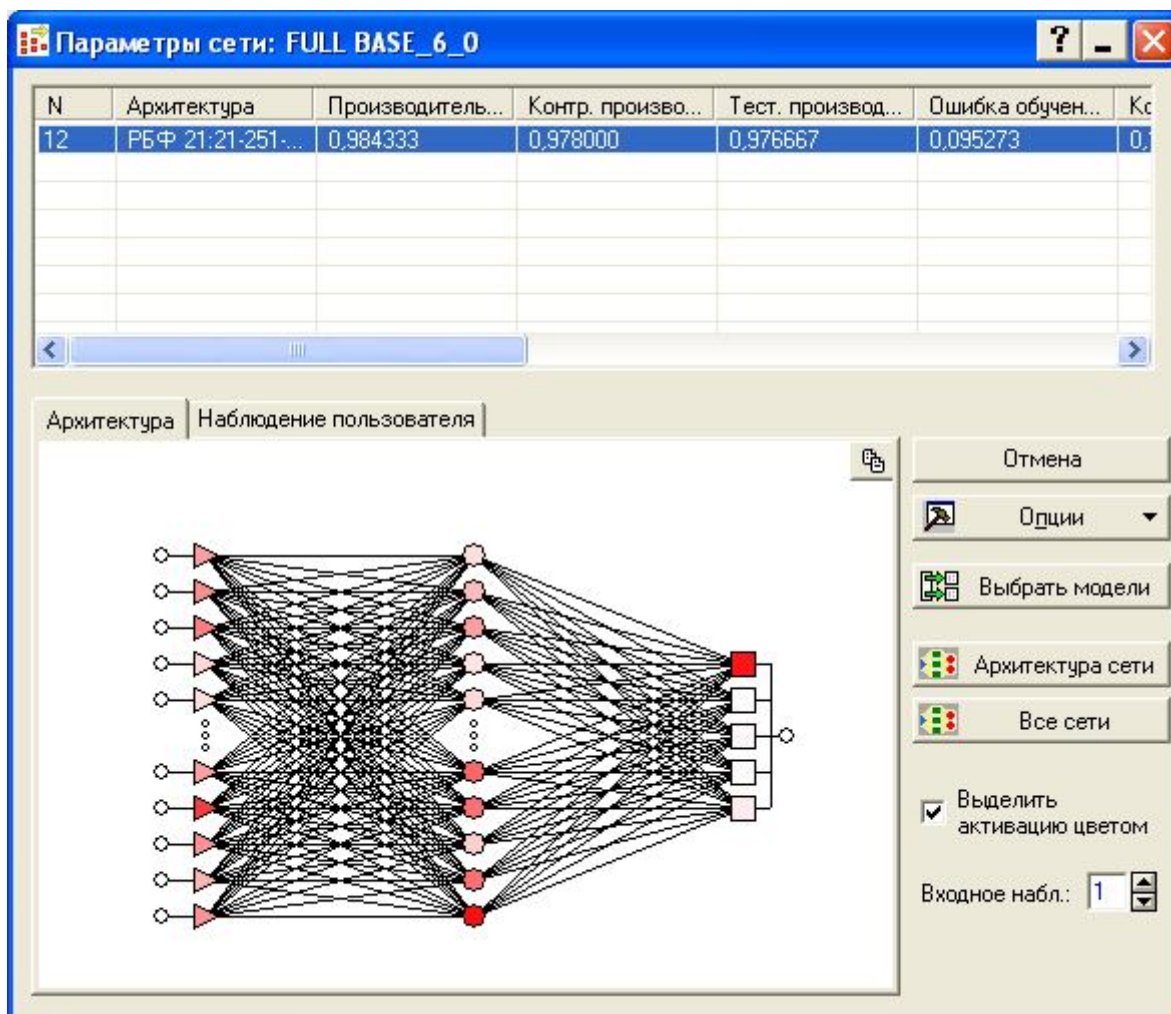


Рис.19

Нажатие кнопки *Отмена* вернет нас в окно результатов на вкладку *Дополнительно*. Теперь рассмотрим опцию *Наблюдения пользователя* данной вкладки, при выборе которой выводится окно *Прогноз для наблюдения пользователя* (рис. 20), содержащее две вкладки: *Быстрый* и *Дополнительно*.

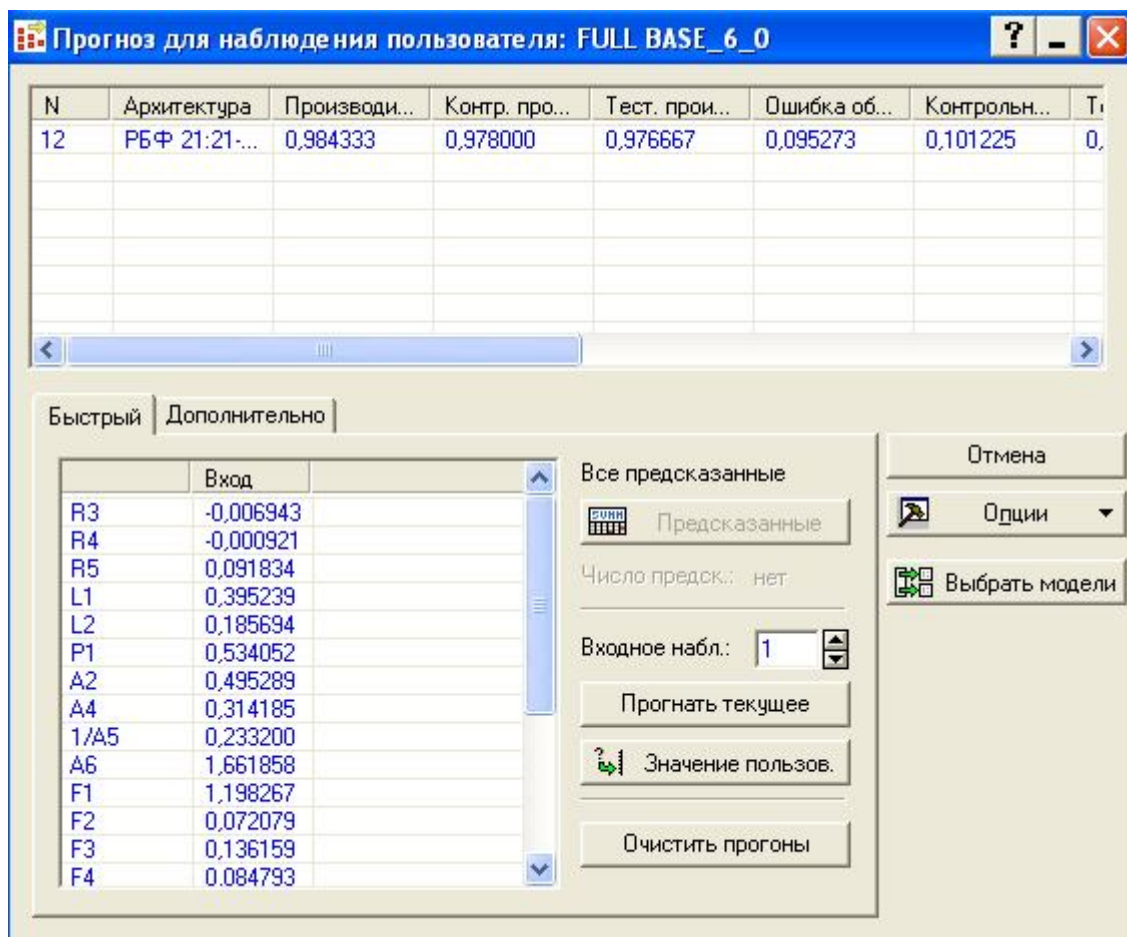


Рис.20

Здесь можно либо задать пользовательские значения финансовых показателей (кнопка *Значения пользователя*), либо использовать имеющиеся, указав номер наблюдения в текстовом поле *Входное наблюдение*. Опция *Прогнать текущее* позволяет классифицировать заданное или выбранное наблюдение согласно построенной модели нейронной сети. После «прогона» становится активной кнопка *Предсказанные*, при нажатии которой выдается таблица с результатами классификации. Состав введенных в таблице можно корректировать путем выбора соответствующих элементов на вкладке *Дополнительно* (рис.21): предсказанный уровень риска банкротства, уровни доверия по классам, вектор кодировки входных параметров, выигравший нейрон, входные значения параметров. Заметим, что доверительные значения нельзя интерпретировать как оценки вероятностей. Если они выходят за пределы пороговых значений (класс не предсказан), то уровни активации отражают не вероятность принадлежности к классу, а меру ошибки

классификации. Для всех предприятий тестовой выборки доверительные значения не больше нижнего предела или не меньше верхнего. Это значит, что уровни доверия отражают принадлежность предприятия к уровням риска неплатёжеспособности.

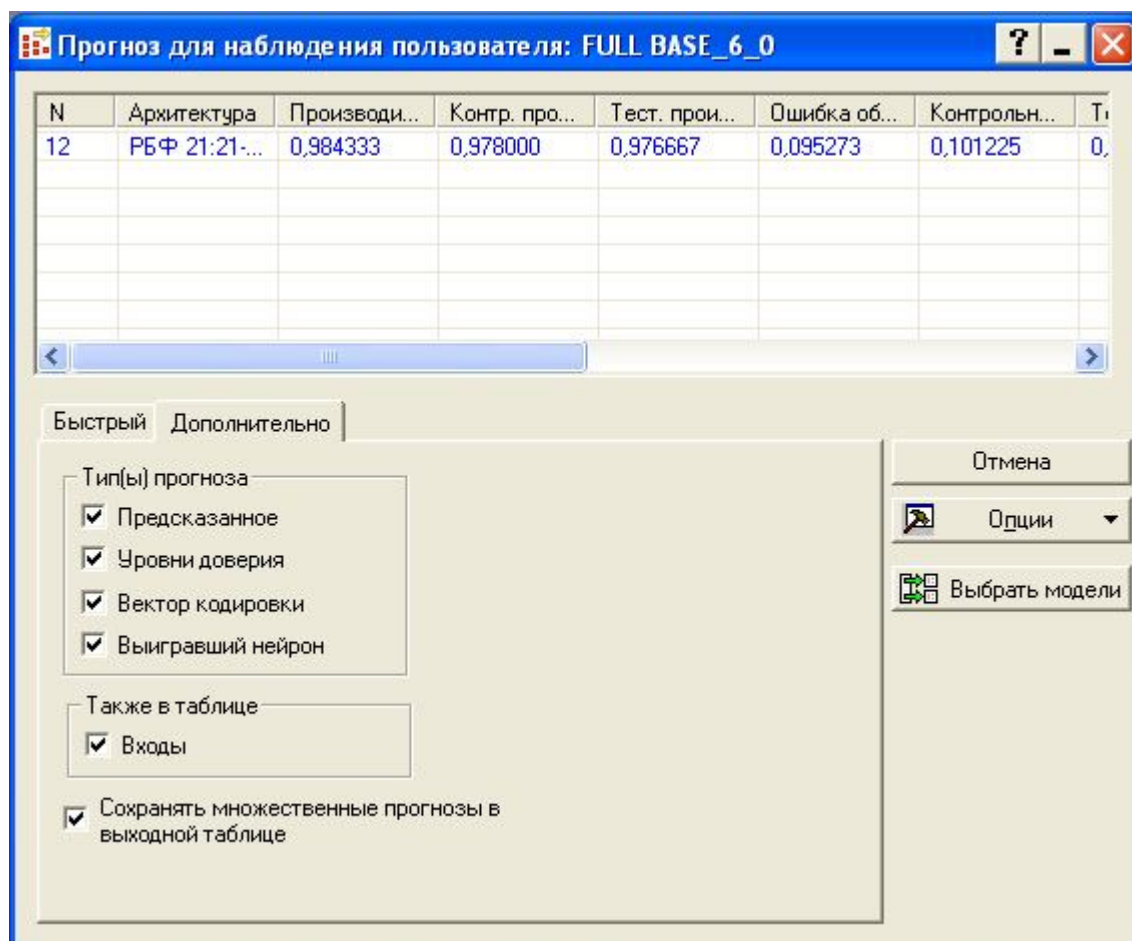


Рис.21

Так же имеется возможность накапливать «прогоны» различных наблюдений и возвращаться к их результатам по кнопке *Предсказанные*. В случае необходимости все результаты «прогонов» могут быть удалены при использовании опции *Очистить прогоны*.

Мы рассмотрели все вкладки окна вывода результатов нейронной сети (см. рис.14), в которое мы можем вернуться, нажав Отмена в окне прогноза наблюдений пользователя.

Для повторного обращения к построенной нейронной сети необходимо сначала сохранить файл сети на вкладке *Сети/Ансамбли* окна *Нейронные сети* (рис.22), а затем выбрать этот созданный файл при помощи опции *Открыть файл сети*.

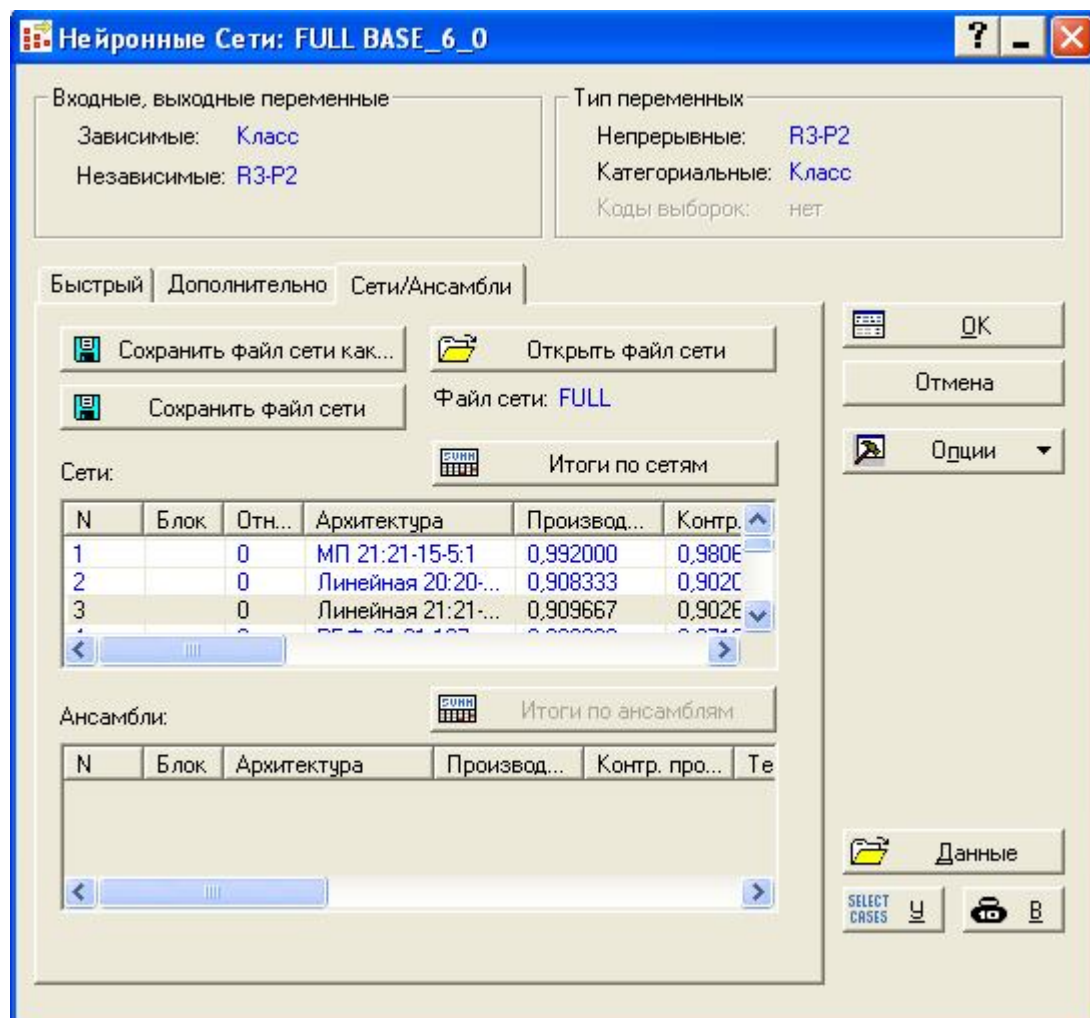


Рис.22

Построенную сеть можно использовать для создания программного модуля на любом языке программирования. Для этого необходимо использовать данные весов нейронов каждого слоя. Данные могут быть переписаны вручную в каждую переменную созданных массивов, а могут быть занесены в файл при помощи функции генерации кода вкладки дополнительно (см. рис.7). Нами был создан программный продукт с использованием построенной нейронной сети. Его работу можно протестировать на сайте лаборатории статистических исследований statlab.kubsu.ru.

Список использованных источников

1. Артёмкин, Д.Е. Разработка математического и программного обеспечения автоматизированного прогнозирования временных рядов на основе нейрокомпьютерных технологий: дис. ...канд. тех. наук/ Д.Е. Артёмкин. – Рязань, 2003. – 140 с.
2. Боровиков, В.П. Нейронные сети. Statistica Neural Networks. Методология и технологии современного анализа данных / В.П. Боровиков. – 2-е изд. – М.: Горячая линия – Телеком, 2008. – 392 с.
3. Бэстенс, Д.-Э. Нейронные сети и финансовые рынки. Принятие решений в торговых операциях/ Д.-Э. Бэстенс, В.-М. Ван Ден Берг, Д. Вуд . – М.: ТВП, 1997. – 236с.
4. Ендовицкий, Д.А. Анализ и оценка кредитоспособности заёмщика: уч.-практ. пос./ Д.А. Ендовицкий, И.В. Бочарова. - М.: Кнорус, 2005. -272 с.
5. Илларионов, А. В. Разработка математических моделей и алгоритмов принятия решения по кредитованию предприятий малого (среднего) бизнеса на основе аппарата теории нечётких множеств: дис. ... канд. экон. наук / А. В. Илларионов. - Владимир, 2006. – 231с.
6. Комашинский , В.И. Нейронные сети и их применение в системах управления и связи./ В.И. Комашинский , Д.А Смирнов.– М:Горячая линия Телеком 2003. – 94с.
7. Кричевский ,М.Л. Интеллектуальный анализ данных в менеджменте:Учеб. пособие/ М.Л. Кричевский. – СПбГУАП. СПб., 2005. – 208с.
8. Лаврушин, О.И. Банковское дело: Современная система кредитования: Учеб. пособие для вузов / О.И. Лаврушин, О.Н. Афанасьева, С.Л. Корниенко; под ред. Лаврушина О.И. – 3-е изд., перераб. и доп. – М.: КНОРУС, 2005. – 256 с.
9. Лобанов А.А. Энциклопедия финансового риск-менеджмента/Под ред. А.А. Лобанова и А.В. Чугунова – М: Алытина Паблшер, 2003 – 786с.

10. Современные математические методы анализа финансово-экономического состояния предприятия: монография. / Т.П. Барановская, В.Н.Кармазин, А.В Коваленко, М.Х. Уртенев – Краснодар: КубГАУ, 2009. – 250 с.
11. Халафян, А.А. STATISTICA 6. Статистический анализ данных. 3-е изд. учеб. / А.А. Халафян. – М.: Бином-Пресс, 2007. – 512с.
12. Хайкин, С. Нейронные сети:полный курс, 2-е изд., испр.:Пер. с англ. /С. Хайкин.– М.:ООО И.Д. Вильямс, 2006. – 1104с.
13. Нейроинформатика / А.Н.Горбань, В.Л.Дунин-Барковский, А.Н.Кирдин и др. - Новосибирск: Наука. Сибирское предприятие РАН, 1998. - 296с.
14. Шевченко И.В., Халафян А.А, Пелипенко Е.Ю. Создание виртуальной клиентской базы для анализа кредитоспособности российских предприятий.// Финансы и кредит. №1, 2010. С. 13 – 18.