

Федеральное государственное автономное образовательное учреждение
высшего образования
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ТОМСКИЙ
ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

Федеральное государственное бюджетное образовательное учреждение
высшего образования
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ
УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ

На правах рукописи

Нгуен Ань Ту

**АЛГОРИТМИЧЕСКОЕ ОБЕСПЕЧЕНИЕ НЕЙРОННОЙ СЕТИ С
ПОЛИНОМИАЛЬНЫМИ КУСОЧНО-НЕПРЕРЫВНЫМИ ФУНКЦИЯМИ
АКТИВАЦИИ ДЛЯ ОБНАРУЖЕНИЯ ЗАКОНОМЕРНОСТЕЙ В ДАННЫХ**

Специальность: 05.13.01 – Системный анализ, управление и обработка информации
(научные исследования)

Диссертация на соискание ученой степени кандидата технических наук

Научный руководитель
доктор технических наук,
профессор Корилов А.М.

Томск – 2019

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	4
1 ИСКУССТВЕННЫЕ НЕЙРОННЫЕ СЕТИ: ОСНОВНЫЕ ДОСТИЖЕНИЯ И ПРОБЛЕМЫ, НЕОБХОДИМОСТЬ СОЗДАНИЯ НОВЫХ ФУНКЦИЙ АКТИВАЦИИ	13
1.1 Искусственные нейронные сети	13
1.2 Алгоритмы обучения нейронных сетей	23
1.3 Новые функции активации, основанные на функциях принадлежности нечетких чисел.....	29
1.4 Выводы по первой главе	38
2 МНОГОСЛОЙНЫЕ ОДНОНАПРАВЛЕННЫЕ НЕЙРОННЫЕ СЕТИ С НОВЫМИ ФУНКЦИЯМИ АКТИВАЦИИ ДЛЯ ПРОГНОЗА И ОПРЕДЕЛЕНИЯ ПАРАМЕТРОВ СИГНАЛОВ.....	40
2.1 Использование новых функций активации в многослойной однонаправленной нейронной сети для прогноза значений временных рядов	40
2.1.1 Временный ряд.....	40
2.1.2 Создание нейронной сети с использованием метода скользящего окна для прогноза временных рядов.....	41
2.2 Определение амплитуды детерминированного сигнала на фоне белого шума с помощью ННС второго типа	50
2.3 Выводы по второй главе	55
3 НЕЧЕТКАЯ НЕЙРОСЕТЕВАЯ КЛАССИФИКАЦИЯ ИНТЕНСИВНОСТИ МАРКОВСКОГО СЛУЧАЙНОГО ПОТОКА СОБЫТИЙ	57
3.1 Марковские случайные процессы	57

3.2 Настройка функций активации	59
3.3 Обучение ННС	62
3.4 Эксперименты по классификации интенсивности потоков событий	66
3.5 Выводы по третьей главе	75
4 НЕЙРО-НЕЧЕТКАЯ КЛАССИФИКАЦИЯ ОБЪЕКТОВ И ИХ СОСТОЯНИЙ	76
4.1 Нейро-нечеткий классификатор	78
4.2 Обучение нейро-нечеткого классификатора	81
4.3 Тестирование нейро-нечеткого классификатора	86
4.4 Выводы по четвертой главе	89
ЗАКЛЮЧЕНИЕ	91
СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ	94
СПИСОК ЛИТЕРАТУРЫ	96
Приложение А. Акт внедрения в учебный процесс	110
Приложение Б. Акт внедрения в компании Midas Core Pte Ltd	111
Приложение В. Акт внедрения в компании Center of Education and Technology Transfer VIET4C	112

ВВЕДЕНИЕ

Актуальность темы исследования. Методы и технологии искусственного интеллекта используются для решения многих практических задач: автоматизации и роботизации технологических процессов и производств, проектирования информационно-вычислительных систем, телекоммуникационных и компьютерных сетей и т.д. Искусственные нейронные сети и технологии, базирующиеся на них, занимают лидирующие позиции среди методов и технологий искусственного интеллекта. В естественных и технических науках задачи делятся на формализуемые и неформализуемые. Постановка любой задачи заключается в том, чтобы перевести ее словесное, вербальное описание в формальное. В случае относительно простых задач такой переход осуществляется в сознании (естественном интеллекте) человека. Если полученная формальная модель (математическая зависимость между величинами в виде формулы, уравнения, системы уравнений) опирается на фундаментальный закон или подтверждается экспериментом, то этим доказывается ее адекватность отображаемой ситуации, и модель рекомендуется для решения задач соответствующего класса. Неформализуемые задачи невозможно представить, используя только естественный интеллект человека, в виде каких-то формальных моделей. С развитием науки и техники количество таких неформализуемых задач постоянно растет. К числу неформализуемых задач относятся и задачи обнаружения скрытых закономерностей в наборах данных. Для их решения развиваются специальные приемы и методы, обобщенные в теории систем и системном анализе, создаются экспертные системы – интеллектуальные системы, основанные на знаниях. Роль методов искусственного интеллекта, теории нейронных сетей и нейросетевых технологий в реализации методик системного анализа и разработке интеллектуальных систем трудно переоценить. Решение неформализуемых задач осложняется различными видами НЕ-факторов знаний: неопределенность, нечеткость, неточность, недоопределенность и неполнота знаний эксперта о

свойствах проблемной области.

Таким образом, задача обнаружения скрытых закономерностей в наборах данных на основе нейросетевых технологий является актуальной в научном аспекте и практически значимой.

Степень научной разработанности проблемы. Для повышения эффективности систем обработки информации и снижения вычислительных затрат, в теории нейронных сетей и нейросетевых технологий разрабатываются новые функции активации. Большая работа в этом направлении проведена такими авторами, как Билски Дж., Вонг К.В., Леунг К.С., Нага К., Накаямма К., и др. Билски Дж. предложил новую логарифмическую функцию активации. Подробное сравнение между этой логарифмической функцией активации и различными простыми функциями активации, основанными на экспонентах сделано Азизом С.М., Камруззаманом Дж. Авторы исследуют скорость обучения и скорость сходимости по проблеме распознавания символов и классической проблеме классификации XOR, предлагая использовать обратную функцию тангенса как функцию активации. Они так же сравнивают эффективность обучения в терминах эпох, необходимых для изучения задачи, с предложенной обратной функцией тангенса, с классическими сигмоидальными и гиперболическими функциями тангенса и новой логарифмической функции активации, найденной Билски Дж, обнаруживая значительный прирост производительности. Пиккневский Ф., Рыбицкий Л. предложили две функции активации, одна из которых основана на интеграции треугольной функции, а другая - на разнице между двумя сигмоидами (лог-экспоненциальная), и сравнивается с помощью метода барицентрического построения, который проецирует возможности отображения сети в гиперразмерный куб. Исследование показало, что лог-экспоненциальная функция медленно ускорялась, но она была эффективной в многослойном перцептроне с обучением методом обратного распространения. Гуарнери С., Камполуччи П., Каппарелли Ф., Пьяцца Ф., Унчини А. используют в качестве функции активации кусочную

интерполяцию на основе кубических сплайнов, обеспечивающую характеристики, сравнимые с сигмоидальной функцией, с уменьшенными вычислительными затратами. Ма Л., Хорасани К. предложили функцию активации, которая представляется ортонормированными полиномами Эрмита. Каждый добавленный нейрон в скрытом слое характеризуется функцией активации, которая является более сложной по сравнению с имеющимися нейронами. С помощью обширного моделирования авторы доказывают, что такая сеть демонстрирует отличные характеристики по сравнению с аналогичными нейронными сетями с идентичными сигмовидными функциями активации. Функция активации также исследована Вонгом К. В., Леунгом С. С., Чанегом С. Авторы исследовали обучение нейронных сетей с помощью алгоритма расширенного фильтра Калмана. Сеть протестирована на классические сигмоидальные и синусоидальные функции активации, распознавание рукописного ввода, прогнозирование временных рядов, создание четности и отображение XOR. Они доказывают, что предложенная периодическая функция превосходит обе классические функции активации с точки зрения конвергенции обучения. Накаямма К., Хара К., предлагают комбинацию сигмоидальной и синусоидальной и гауссовской функции активации, чтобы использовать их независимые свойства деления пространства. Авторы сравнивают гибридную структуру в задаче классификации многочастотных сигналов, делая вывод, что комбинация трех функций активации работает эффективнее, чем сигмоидальная (по скорости сходимости) и гауссовская (по шумоподавлению) и синусоидальная функция активации, даже если каждая из них обеспечивает достижение лучших результатов по отмеченным выше критериям.

Объединение нейронных сетей с функциями принадлежности из теории нечетких систем нашло применение для решения задач прогнозирования и классификации состояний объектов. Гомес-Чова Л., Кэмпс-Вальс Г., Кальпе-Маравилья Ж., Мартин-Герреро Ж.Д., Сория-Оливас Э., Серрано-Лопес А. определяют передачу гиперболического тангенса с помощью трех различных

функций принадлежности, фактически определяя классическую функцию активации с помощью нечеткой логики. Основным преимуществом на этапе обучения является низкая вычислительная стоимость, обеспечиваемая тем, что обновление веса не всегда необходимо. Дополнительное применение нейронных сетей с нечеткими функциями активации можно найти в работе Субаси А. Автор использует нейронные сети для обнаружения эпилептического припадка, обработки и классификации электроэнцефалографии сигналов. Вон И., Хунь Х.Т. предлагают метод аналитического обучения, отмеченный как экстремальное машинное обучение, для нейронных сетей с нечеткими функциями активации. Авторы тестируют методику по проблемам медицинской диагностики, классификации изображений и анализа спутниковых изображений. Ярушкиной Н.Г. представлена генетическая нейро-нечеткая система, которая построена на основе многослойной архитектуры с использованием И-, ИЛИ-нейронов и принципом «конкуренции и кооперации». В работе Ярушкиной Н.Г. также отметим совместное применение генетических алгоритмов и нейронных сетей – это настройка функций принадлежности нечетких и нейро-нечетких систем.

Таким образом, существует обширное научное поле, в котором рассматриваются проблемы разработки функций активации в теории нейронных сетей и функции принадлежности, изучаемые в теории нечетких систем, на основе которых можно создать новые нейронные сети с нечеткими функциями активации для решения задач обнаружения закономерностей в наборах данных.

Объектом исследования являются нейронные сети нового типа с функциями активации, базирующимися на функциях принадлежности теории нечетких систем.

Предметом исследования является моделирование и применение новых нейронных сетей для решения типовых задач обнаружения закономерностей в наборах данных: прогнозирования временных рядов, амплитуды детерминированного сигнала на фоне белого шума, классификации интенсивности марковского случайного потока событий и классификации объектов и их состояний.

Целью работы является разработка и исследование нейронных сетей нового типа с полиномиальными кусочно-непрерывными функциями активации для обнаружения закономерностей в наборах данных, алгоритмов обучения и программных средств создания нечетких нейронных сетей для решения типовых задач обнаружения закономерностей в наборах данных.

Задачи исследования. Для достижения намеченной цели были решены следующие задачи:

1. Анализ существующих полиномиальных кусочно-непрерывных функций принадлежности в теории нечетких систем с целью их использования в качестве функции активации нечетких нейронных сетей.
2. Разработка тестовых наборов данных для типовых задач обнаружения закономерностей в данных.
3. Построение моделей нечетких нейронных сетей для тестовых задач обнаружения закономерностей в данных:
 - прогнозирование временных рядов со стохастическим трендом;
 - определение параметров детерминированного сигнала на фоне белого шума;
 - классификация интенсивности марковского случайного потока событий;
 - классификация объектов и их состояний.
4. Разработка алгоритмов обучения нечетких нейронных сетей для выше перечисленных тестовых задач обнаружения закономерностей в данных.
5. Математическое моделирование нечетких нейронных сетей для выше перечисленных тестовых задач обнаружения закономерностей в данных.

Научная новизна работы заключается в следующем:

1. Разработаны новые функции активации для нейронных сетей, которые основаны на функциях принадлежности теории нечетких систем и теории надежности, и на их базе разработаны новые нейронные сети.
2. Модифицированы алгоритмы обучения нейронных сетей для типовых задач обнаружения закономерностей в наборах данных. Усовершенствованы метод Левенберга – Марквардта, алгоритм кластеризации К - средних, вычисление весов выходного слоя нечетких нейронных сетей с использованием псевдообратного метода, алгоритм масштабированного сопряженного градиента.
3. Созданы модели новых нейронных сетей для типовых задач обнаружения закономерностей в наборах данных: прогнозирование временных рядов, определение амплитуды детерминированного сигнала на фоне белого шума, классификация интенсивности марковского случайного потока событий и классификация объектов и их состояний.

Теоретическая значимость и практическая значимость работы. Предложенные в диссертации модели и алгоритмы могут быть применены для создания программного обеспечения нечеткого нейросетевого классификатора для повышения эффективности решения следующих практических задач:

1. Прогнозирование временных рядов со стохастическим трендом.
2. Определение параметров сигнала на фоне помех (шум) при диагностике технических систем.
3. Определение трафика в телекоммуникационных и компьютерных сетях как интенсивности марковского случайного потока событий.
4. Классификация сложных объектов и определение их состояний.

Результаты диссертации использованы в учебном процессе кафедры «Автоматизированные системы управления» в федеральном государственном бюджетном образовательном учреждении высшего образования «Томский

государственный университет систем управления и радиоэлектроники».

Методы исследования основаны на теории нечетких систем, теории принятия решений, нейроинформатике, методах обучения нейронных сетей, цифровой обработки сигналов.

Научные положения, выносимые на защиту:

1. Новые функция активации нейронных сетей, основанные на функциях принадлежности LR-типа в теории нечетких систем и теории надежности, на базе этих функций разработаны новые нейронные сети.

Соответствует пункту 1 паспорта специальности 05.13.01: Теоретические основы и методы системного анализа, оптимизации, управления, принятия решений и обработки информации.

2. Модификация алгоритмов обучения нейронных сетей для типовых задач обнаружения закономерностей в наборах данных для повышения эффективности систем обработки информации.

Соответствует пункту 7 паспорта специальности 05.13.01: Методы и алгоритмы структурно-параметрического синтеза и идентификации сложных систем.

3. Модели новых нейронных сетей для типовых задач обнаружения закономерностей в наборах данных: прогнозирование временных рядов, определение амплитуды детерминированного сигнала на фоне белого шума, классификация интенсивности марковского случайного потока событий и классификация объектов и их состояний.

Соответствует пункту 11 паспорта специальности 05.13.01: Методы и алгоритмы прогнозирования и оценки эффективности, качества и надежности сложных систем.

Апробация результатов исследования. Основные положения и результаты диссертации докладывались и обсуждались на:

Международной научно-практической конференции «Электронные средства и системы управления», г. Томск, 2016 г.

Международной научно-практической конференции студентов, аспирантов и молодых учёных «Молодёжь и современные информационные технологии», г. Томск, 2016 г.

Международной научно-практической конференции «Электронные средства и системы управления», г. Томск, 2017 г.

Международной научно-практической конференции студентов, аспирантов и молодых учёных «Молодёжь и современные информационные технологии», г. Томск, 2017 г.

Международной научной конференции «Информационные технологии в науке, управлении, социальной сфере и медицине », г. Томск, 2017 г.;

The 22nd National Scientific Conference for Fluid Mechanics, Hai Phong, Viet Nam, 2019.

The International Symposium on Applied Science (ISAS), Ho Chi Minh city, Viet Nam, 2019.

Научных семинарах кафедры «Автоматизированные системы управления» в федеральном государственном бюджетном образовательном учреждении высшего образования «Томский государственный университет систем управления и радиоэлектроники» в 2015 – 2019 гг.

Основные результаты диссертационного исследования апробированы в научных исследованиях двух учреждений (Midas Core Pte LTD, Singapore; Center of Education and Technology Transfer VIET4C, Viet Nam) и в учебный процесс в 1 образовательном учреждении (Кафедра АСУ ТУСУР).

Публикации по теме работы. По материалам диссертационных исследований опубликовано 12 работ, в том числе 5 публикаций в изданиях, включенных в перечень ведущих рецензируемых научных журналов и изданий, рекомендованных ВАК Российской Федерации; 1 публикация в WoS, 2 публикации в Scopus.

Структура и объём работы. Диссертация состоит из введения, 4 глав, заключения, списка литературы. Объем диссертации составляет 112 страниц, включая 39 рисунков, 5 таблиц и список литературы из 118 наименований.

1 ИСКУССТВЕННЫЕ НЕЙРОННЫЕ СЕТИ: ОСНОВНЫЕ ДОСТИЖЕНИЯ И ПРОБЛЕМЫ, НЕОБХОДИМОСТЬ СОЗДАНИЯ НОВЫХ ФУНКЦИЙ АКТИВАЦИИ

1.1 Искусственные нейронные сети

Искусственные нейронные сети (ИНС) можно рассматривать как реализацию систем искусственного интеллекта (ИИ), как упрощенные математические модели мозговых систем естественного интеллекта (ЕИ), а при обработке информации они функционируют как параллельные распределенные вычислительные сети. Однако, в отличие от обычных компьютеров, запрограммированных для выполнения конкретных задач, большинство ИНС должны быть обучены для решения конкретных задач. Тогда они могут создавать новые ассоциации, изучать новые функциональные зависимости и генерировать новые шаблоны. Хотя современные компьютеры превосходят как биологические, так и искусственные нейронные системы при решении задач, требующих точных и быстрых арифметических операций, искусственные нейронные системы представляют собой перспективное новое поколение сетей, реализующих параллельную обработку информации. В нашей работе исследуется применение ИНС для обнаружения (поиска) закономерностей в больших наборах данных.

Изучение математических моделей ИНС началось более 75 лет тому назад в работах Маккаллоха и Питта (1943) [1] и чуть позже в знаменитой работе, посвященной исследованию поведения организации Хебба (1949) [2]. В этой ранней работе в области ИИ появился конфликт между теми, кто считал, что интеллектуальные системы лучше всего можно построить на компьютерах, как моделях мозга, и таких, как Мински и Паперт (1969) [3], которые считали, что ИИ, в основном, предназначен для символьной обработки информации. По многим причинам подход к обработке информации как символов стал доминирующей темой в ИИ в 1970-х годах.

Однако в 1980-е годы возникло возрождение интереса к нейронным вычислениям:

В 1982 г. Хопфилд [4] разработал математическую основу для понимания динамики важного класса сетей.

В 1984 г. Кохонен [5] разработал неконтролируемые обучающие сети для отображения признаков в регулярные массивы нейронов.

В 1986 г. Румельхарт и Макклелланд [6] представили алгоритм обучения обратного распространения для сложных многослойных сетей.

Начиная с 1986-87 гг. было инициировано множество исследовательских программ нейронных сетей. Список приложений, которые могут быть решены нейронными сетями, расширился с небольших примеров тестовых размеров до крупных практических задач. Сформированы крупномасштабные интегрированные чипы нейронных сетей.

В долгосрочной перспективе можно ожидать, что искусственные нейронные системы будут использоваться в приложениях, связанных со зрением, речью, принятием решений и рассуждениями, а также в качестве сигнальных процессоров, таких как фильтры, детекторы и системы контроля качества.

ИНС являются физическими сетевыми системами, которые могут приобретать, хранить и использовать эмпирические знания. Они состоят из совокупности простых процессоров, которые обмениваются данными, передавая сигналы друг другу по большому количеству взвешенных соединений. Знания находятся в форме стабильных состояний или отображений, встроенных в сети, которые можно вызвать в ответ на поступление сигналов.

Основные элементы обработки информации в ИНС называются искусственными нейронами, просто нейронами или узлами. Каждый нейрон выполняет относительно простое задание: получает входной сигнал от соседей или внешних источников и использует его для вычисления выходного сигнала, который распространяется на другие узлы. Помимо этой обработки, второй задачей является

настройка весов. Система по своей сути параллельна, так как многие узлы могут одновременно выполнять свои вычисления. Каждый из выходов этих нейронов математически представляется реальными числами. Таким образом, каждое соединение имеет вес синаптической связи, который определяет влияние входа на уровень активации устройства. Веса могут быть положительными (возбуждающими) или отрицательными (тормозящими).

Каждый нейрон характеризуется уровнем активности, выходным значением, набором входных соединений, значением смещения и набором выходных соединений.

В нейронных системах полезно выделить три типа нейронов: входные, которые принимают данные за пределами нейронной сети, выходные нейроны, которые отправляют данные из нейронной сети; и скрытые нейроны, входные и выходные сигналы которых остаются в нейронной сети [7]. Однослойная ИНС представлена на рисунке 1.1, а нейрон как элемент обработки с одним выходом – на рисунке 1.2.

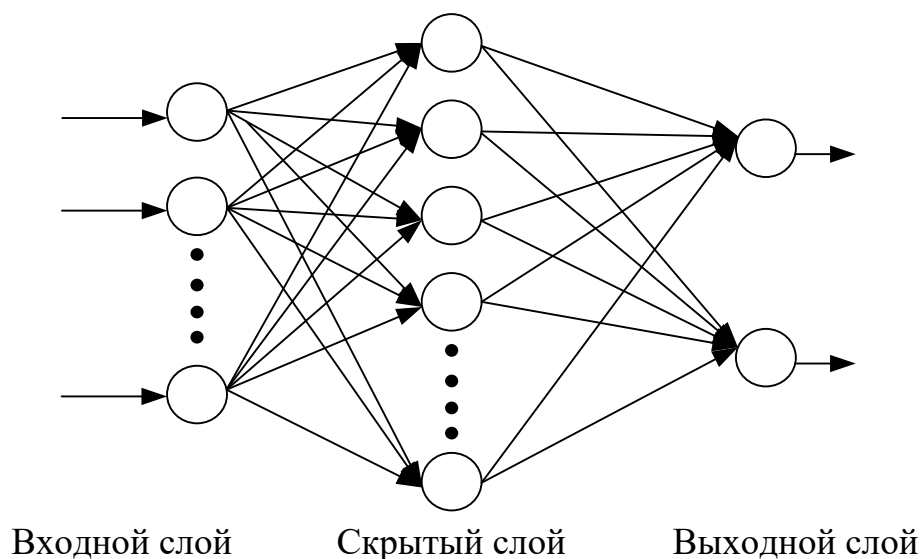


Рисунок 1.1. Однослойная нейронная система

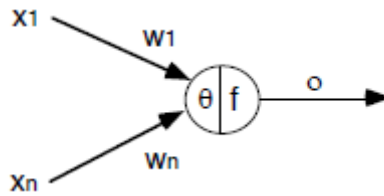


Рисунок 1.2. Элемент обработки с одним выходом

Сигнальный поток от входов нейронов x_j ($j=1,2,...,n$) считается однонаправленным, обозначенным стрелками, как поток входных сигналов нейрона. Выходной сигнал нейрона определяется следующим соотношением:

$$o = f(\langle w, x \rangle) = f(w^T x),$$

где $w = (w_1, ..., w_n)^T \in R^n$ - вектор входных весов нейрона. Функция $f(w^T x)$ часто называется как функция активации (ФА) (или функция передачи) нейрона.

ФА можно использовать в многослойной однонаправленной нейронной сети, для того, чтобы решать задачи прогнозирования или определения атрибута объектов. Многослойные однонаправленные сети называются также сетями прямого распространения, или многослойными персептронами. Сети этого типа состоят из нескольких слоев нейронов: входного слоя, выходного и нескольких «скрытых слоев». Функционирование многослойной сети осуществляется следующим образом: входной сигнал, подаваемый на сеть, поступает на нейроны входного слоя, проходит по очереди через все слои и снимается с выходов нейронов выходного слоя. По мере распространения сигнала по сети он претерпевает ряд преобразований, которые зависят от начальных условий, от преобразующих функций и от величины весов связей.

На рисунке 1.3 представлена многослойная однонаправленная ИНС. Через W_{ij}^k обозначен вес сигнала, проходящего от i -го нейрона $(k-1)$ -го слоя к j -му нейрону k -го слоя

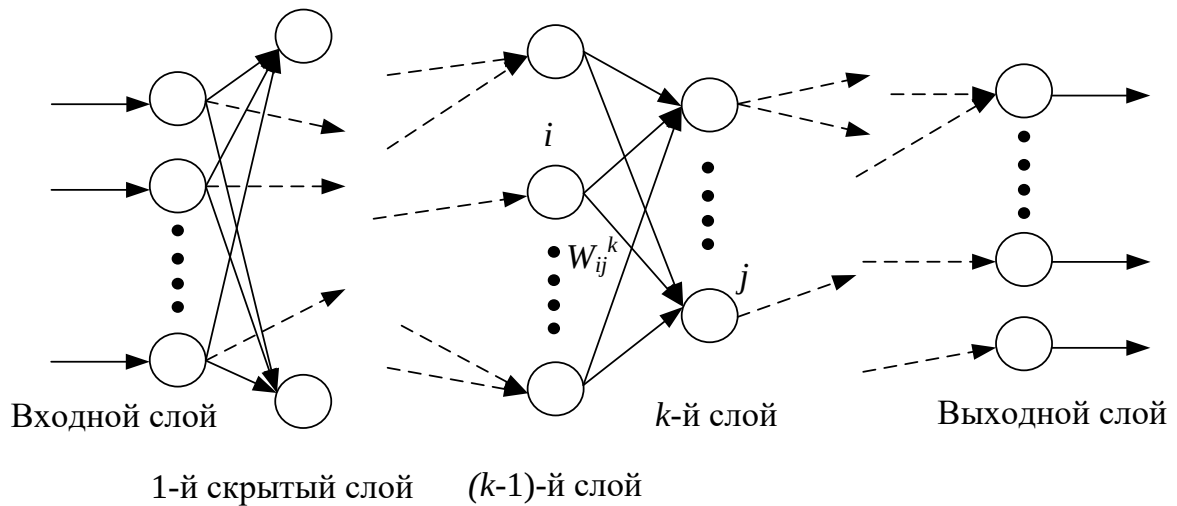


Рисунок 1.3. Многослойная однонаправленная нейронная сеть

Самым популярным классом многослойной однонаправленной НС является многослойный персептрон (МП). МП - это ИНС с прямой связью, которая генерирует набор выходов из набора входов. МП характеризуется несколькими слоями входных узлов, соединенных в виде ориентированного графа между входным и выходным слоями.

Каждый нейрон в МП принимает взвешенную сумму своих входных значений. То есть каждое входное значение умножается на коэффициент, и все результаты суммируются. Один нейрон МП представляет собой простой линейный классификатор, но сложные нелинейные классификаторы могут быть построены путем объединения этих нейронов в сеть.

Широко применяется и другой тип нейронной сети, который также находится в классе многослойных однонаправленных НС и который называется радиально-базисной функциональной нейронной сетью (РБФНС). Она была разработана Брумхедом и Лоу [8] в 1988 году. Поскольку радиальные базисные функции (РБФ) имеют только один скрытый слой, сходимость к цели оптимизации происходит намного быстрее, и, несмотря на наличие одного скрытого слоя, РБФНС доказали свою универсальность во многих приложениях.

РБФНС - это искусственная нейронная сеть с входным слоем, скрытым слоем и выходным слоем. Скрытый слой РБФНС состоит из скрытых нейронов, и функция активации этих нейронов часто является гауссовой функцией. Скрытый слой генерирует сигнал, соответствующий входному вектору во входном слое, и сеть генерирует ответ соответствующий этому сигналу. На рисунке 1.4 показана схема радиально-базисной сети.

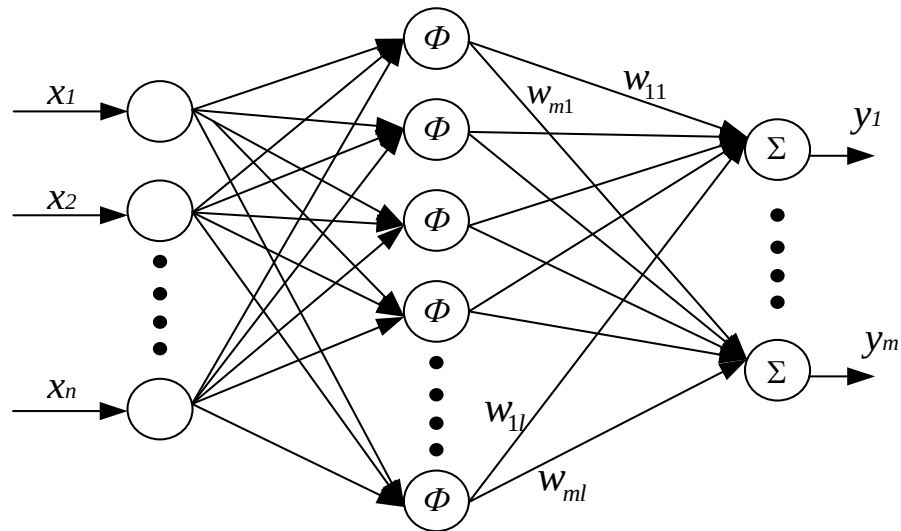


Рисунок 1.4. Радиально-базисная нейронная сеть

РБФНС на рисунке 1.4 имеет n - входов и m -выходов, выполняет нелинейное преобразование вида:

$$y_j = w_{j0} + \sum_{i=1}^l w_{ji} \varphi_i(x),$$

где $\varphi_i(x)$ – радиально-базисные функции, определяющие процесс преобразования из n - мерного пространства входов в m - мерное пространство выходов, l – количество нейронов в скрытом слое, w_{ij} – весовой коэффициент, характеризующий силу синаптической связи между нейронами i и j .

РБФНС имеют ряд преимуществ перед рассмотренными многослойными сетями прямого распространения. Во-первых, они моделируют произвольную

нелинейную функцию с помощью всего одного промежуточного слоя, тем самым, избавляя разработчика от необходимости решать вопрос о числе слоев. Во-вторых, параметры линейной комбинации в выходном слое можно полностью оптимизировать с помощью хорошо известных методов линейной оптимизации, которые работают быстро и не испытывают трудностей с локальными минимумами, так мешающими при обучении с использованием алгоритма обратного распространения ошибки. Поэтому сеть РБФ обучается очень быстро - на порядок быстрее, чем с использованием алгоритма обратного распространения [9].

Недостатки РБФНС: данные сети обладают плохими экстраполирующими свойствами и получаются весьма громоздкими при большой размерности вектора входов [9].

Для наших исследований представляет также интерес нелинейная авторегрессионная нейронная сеть (НАНС) [10, 11], которая может моделировать динамические временные ряды. НАНС использует прошлые значения временного ряда для прогнозирования его будущих значений, что можно записать в виде следующего уравнения [12]:

$$y(t) = h(y(t-1), y(t-2), \dots, y(t-p)) + \varepsilon(t).$$

Эта формула описывает, как НАНС используется для прогнозирования значения ряда данных $y(t)$ в момент времени t , используя p предыдущих значений серии. Функция $h(\cdot)$ неизвестна, и обучение нейронной сети нацелено на аппроксимацию функции с помощью оптимизации весов сети и смещения нейронов. Слагаемое $\varepsilon(t)$ обозначает ошибку аппроксимации ряда y в момент времени t .

Структура НАНС показана на рисунке 1.5, значения $y(t-1)$, $y(t-2)$, ..., $y(t-p)$, называются задержками обратной связи.

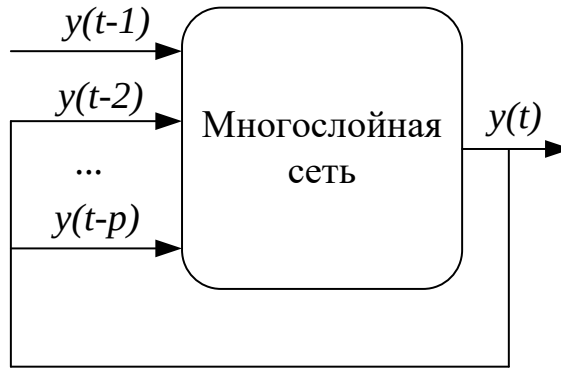


Рисунок 1.5. Авторегрессионная нейронная сеть

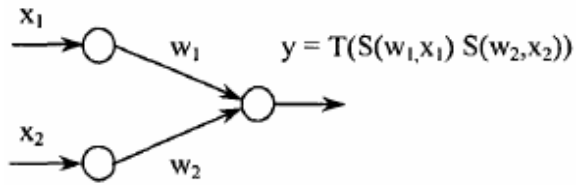
Количество скрытых слоев и нейронов в слое изменяется и может быть оптимизировано методом проб и ошибок для получения топологии сети, которая может обеспечить лучшую производительность. Тем не менее, важно иметь в виду, что увеличение числа нейронов делает систему более сложной, в то время как небольшое количество нейронов может ограничивать возможности обобщения и вычислительную мощность сети.

Нечеткой нейронной сетью (ННС) в [13] называется четкая нейронная сеть прямого распространения сигнала, которая построена на основе многослойной архитектуры с использованием И-, ИЛИ-нейронов [13].

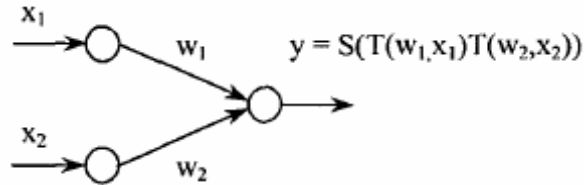
И-нейроном называется нейрон, в котором умножение веса w на вход x моделируется конормой $S(w, x)$, а сложение - нормой $T(w, x)$.

ИЛИ-нейроном называется нейрон, в котором умножение веса w на вход x моделируется нормой $T(w, x)$, а сложение весов - конормой $S(w, x)$.

Понятия T -норм и S -конорм обсуждаются в [13]. T -норм и S -конорм являются функциями со специальными свойствами. Это – функции двух действительных переменных, которые определены на интервалах $[0, 1] \times [0, 1]$, и значения этих функций находятся в интервале $[0, 1]$. Модели нечетких И-нейронов и ИЛИ-нейронов представлены на рисунке 1.6.



а)



б)

Рис. 1.6. Модели нечетких нейронов: а) И-нейрон; б) ИЛИ-нейрон

Самой популярной ННС этого класса является адаптивная НС, основанная на системе нечеткого вывода (Adaptive Network Based Fuzzy Inference System - ANFIS), которая может служить основой для построения набора нечетких правил если - то с соответствующими функциями принадлежности для генерации оговоренных пар ввода-вывода.

На рисунке 1.7 приведена структура ANFIS для двух правил:

ЕСЛИ $x_1 = A_1$ И $x_2 = B_1$ ТО $y_1(x_1, x_2)$,

ЕСЛИ $x_1 = A_2$ И $x_2 = B_2$ ТО $y_2(x_1, x_2)$.

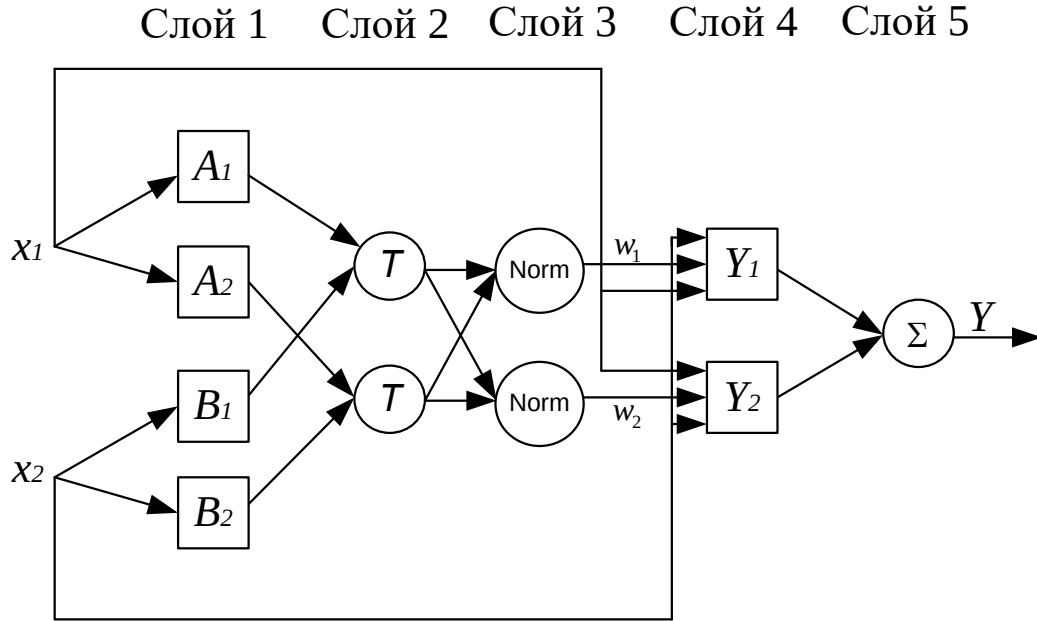


Рисунок 1.7. Структура ANFIS

Выход ННС формируется по формуле:

$$y = \frac{w_1 \cdot y_1 + w_2 \cdot y_2}{w_1 + w_2}.$$

ННС ANFIS состоит из пяти слоев. На вход подаются значения данных и вычисляются значения функций принадлежности входных переменных; второй слой выполняет агрегирование значений предпосылок для каждого правила в соответствии с выбранной Т-нормой; в третьем слое выполняется нормализация значений, поступивших на вход; четвертый слой формирует значение выхода; последний слой выполняет дефаззификацию [14].

Итак, по типу сигналов и характеру их взаимодействия в сети НС следует разделить на детерминированные НС (ДНС), вероятностные НС (ВНС) и нечеткие НС (ННС). Заметим, что в некоторых источниках [15, 16] не делается различия между ВНС и ННС, то есть ВНС называется также ННС. Этот факт мы обсуждаем далее в подразделе 1.3.

1.2 Алгоритмы обучения нейронных сетей

После того, как сеть была создана для конкретного приложения, эта сеть готова к обучению. Для запуска этого процесса исходные веса выбираются случайным образом. Затем начинается обучение.

Обучение нейронной сети - это процесс, в котором параметры нейронной сети настраиваются посредством моделирования среды, в которую эта сеть встроена [17]. Тип обучения определяется способом подстройки параметров. Определение числа промежуточных слоев и числа элементов в них является важным вопросом при конструировании нейронных сетей [18]. На рисунке 1.8 показан процесс обучения НС.

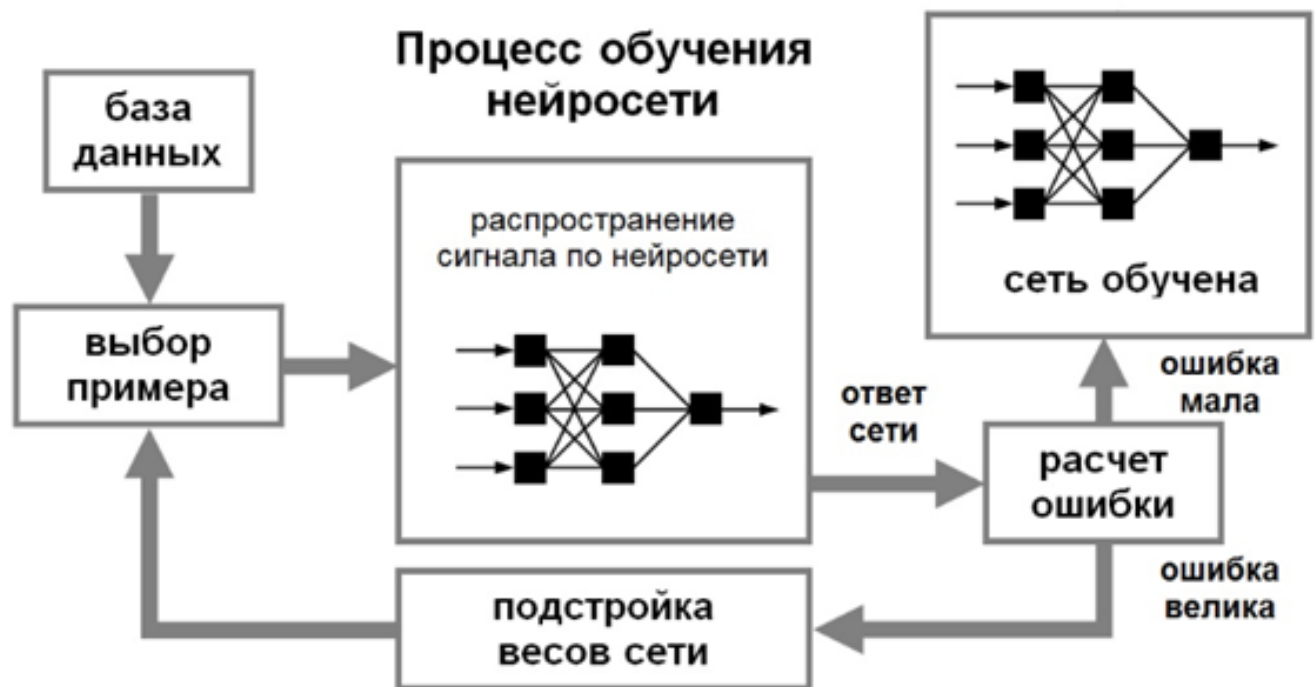


Рисунок 1.8. Иллюстрация процесса обучения НС

Существует два подхода к обучению: с учителем и без учителя. Обучение с учителем указывает на присутствие руководителя в качестве учителя. Процесс обучения с учителем представляет собой предъявление сети выборки обучающих

примеров.

Каждый образец подается на входы сети, затем проходит обработку внутри структуры НС, вычисляется выходной сигнал сети, который сравнивается с соответствующим значением целевого вектора, представляющего собой требуемый выход сети. Затем по определенному правилу вычисляется ошибка, и происходит изменение весовых коэффициентов связей внутри сети в соответствии с выбранным алгоритмом обучения. Векторы обучающего множества предъявляются последовательно, вычисляются ошибки и веса подстраиваются для каждого вектора до тех пор, пока ошибка по всему обучающему массиву не достигнет приемлемо низкого уровня. Обучение без учителя - это обучение НС использованию информации, которая предварительно не классифицирована и не структурирована. Задача НС состоит в том, чтобы группировать неструктурированную информацию в соответствии со сходствами, шаблонами и различиями без предварительной подготовки данных.

Большая часть сетей использует обучение с учителем. Отметим следующие методы обучения:

- Метод обратного распространения ошибки.
- Метод сопряженных градиентов.
- Метод Левенберга – Марквардта.

Наиболее известным и чаще всего применяемым для нейронных сетей алгоритмом обучения является метод обратного распространения ошибки [19]. Обучение на основе алгоритма обратного распространения ошибки предполагает два прохода по всем слоям сети: прямого и обратного. При прямом проходе входной вектор подается на входной слой нейронной сети, после чего распространяется по сети от слоя к слою. В результате генерируется набор выходных сигналов, который и является фактической реакцией сети на данный входной образ. Во время прямого прохода все синаптические веса сети фиксированы. Во время обратного прохода все

синаптические веса настраиваются в соответствии с правилом коррекции ошибок, а именно: фактический выход сети вычитается из желаемого, в результате чего формируется сигнал ошибки. Этот сигнал впоследствии распространяется по сети в направлении, обратном направлению синаптических связей. Отсюда и название – алгоритм обратного распространения ошибки. Синаптические веса настраиваются с целью максимального приближения выходного сигнала сети к желаемому [20].

Кроме метода обратного распространения ошибки популярен, также другой метод обучения нейронной сети - метод сопряженных градиентов [21]. Метод сопряжённых градиентов – метод нахождения локального минимума функции на основе информации об её значениях и её градиенте. В случае квадратичной функции в $\mathbb{R}^n$ минимум находится за n шагов. Метод сопряжённых градиентов является методом первого порядка, в то же время скорость его сходимости квадратична. Этим он выгодно отличается от обычных градиентных методов. Например, метод наискорейшего спуска и метод координатного спуска для квадратичной функции сходятся лишь в пределе, в то время как метод сопряженных градиентов оптимизирует квадратичную функцию за конечное число итераций. При оптимизации функций общего вида, метод сопряжённых направлений сходится в 4-5 раз быстрее метода наискорейшего спуска [22]. При этом, в отличие от методов второго порядка, не требуется трудоёмких вычислений вторых частных производных.

Рассмотрим также метод Левенберга - Марквардта. Данный метод считается самым быстрым и надёжным алгоритмом обучения. Однако его применение связано с определёнными ограничениями:

- Сети с одним выходом. Метод Левенберга - Марквардта можно применять только для сетей с одним выходным элементом. Метод Левенберга–Марквардта, согласно [23] невозможно использовать для обучения сети, предназначенной для классификации нескольких типов облачности [24 Сорокоходов, А.В. Сравнительный].

– Небольшие сети. Метод Левенберга - Марквардта требует памяти, пропорциональной квадрату числа весов в сети. Фактически это ограничение не позволяет использовать метод для сетей большого размера (порядка тысячи и более весов).

– Среднеквадратичная функция ошибок. Метод Левенберга - Марквардта применим только для среднеквадратичной функции ошибок. Если указан для сети другой вид функции ошибок, то этот выбор будет проигнорирован при обучении методом Левенберга - Марквардта. Поэтому этот метод обычно подходит только для регрессионных сетей.

Метод Левенберга - Марквардта предполагает, что функция, моделируемая нейронной сетью, является линейной. В таком предположении минимум определяется за один шаг вычислений. Затем найденный минимум проверяется, и если ошибка уменьшилась, весам присваиваются новые значения. Вся процедура последовательно повторяется. Поскольку предположение о линейности, вообще говоря, не оправдано, может получиться так, что придется проверять точки, лежащие далеко от текущей точки. В методе Левенберга - Марквардта местоположение новой точки есть результат компромисса между продвижением в направлении наискорейшего спуска и описанного выше скачка. Успешные шаги принимаются, и баланс смещается в сторону предположения линейности (которое можем оказаться неверным в окрестности точки минимума). Неудачные шаги отвергаются и согласно алгоритму выполняются более осторожные (малые) шаги вниз по склону. Таким образом, алгоритм Левенберга - Маркара все время меняет схему действия и может работать очень быстро. Алгоритм Левенберга - Марквардта специально разработан так, чтобы минимизировать среднеквадратичную функцию ошибок с помощью формулы, которая (частично) предполагает, что функция, которую моделирует сеть, является линейной. Вблизи точки минимума это предположение выполняется с большой точностью, так что алгоритм может продвигаться очень быстро. Вдали от

минимума это предположение может быть неправильным. Поэтому метод Левенберга - Марквардта находит компромисс между линейной моделью и градиентным спуском. Шаг делается только в том случае, если он уменьшает ошибку, и там, где это необходимо, для обеспечения продвижения используется градиентный спуск с достаточно малым шагом [25].

В таблице 1.1 представлены базовые алгоритм обучения и задачи, которые необходимо решать эвристическим алгоритмом у рассмотренных видов нейронной сети.

Таблица 1.1 Анализ алгоритмов обучения НС различных типов

Вид нейронной сети	Базовый алгоритм обучения	Задачи, которые необходимо решить эвристическим алгоритмом
Многослойный персептрон	Алгоритм обратного распространения ошибки	1. Выход из локальных минимумов. 2. Расширение пространства поиска оптимальной конфигурации сети. 3. Возможность блокировки отдельных синаптических связей или всей сети.
Авторегрессионная нейронная сеть	Алгоритм Левенберга-Марквардта	Минимизация среднеквадратичной функции ошибок: Предполагает, что функция, моделируемая нейронной сетью, является линейной. Определение минимум за один шаг вычислений. Проверка найденного минимума, и если ошибка уменьшилась, весам присваиваются новые значения.
Радиально-базисная нейронная сеть	Алгоритм кластеризации и метод сопряженных градиентов.	1. Уточнение параметров радиально-базисных элементов в сети. 2. Уточнение параметров весовых коэффициентов скрытого слоя. 3. Настройка полей влияния РБФ функций
ANFIS	Метод наименьших квадратов и обратное распространение ошибки	1. Прямой ход: используется метод наименьших квадратов; определяются параметры слоя 4. 2. Обратной ход: используется обратное распространение ошибки; уточняются параметры первого слоя.

Кроме вышеупомянутых алгоритмов, для обучения НС используются и другие алгоритмы, такие как стохастические алгоритмы обучения [26], байесовские

алгоритмы [27] и т.д. Чтобы выбрать какой алгоритм обучения необходимо обратить внимание на следующие факторы: размер, качество и характер данных, доступное вычислительное время, сроки решения, интерпретация и т.д.

Существует множество и других алгоритмов обучения, но приведённые выше — наиболее популярные.

1.3 Новые функции активации, основанные на функциях принадлежности нечетких чисел

Исследования биологических нейронных сетей стимулировали развитие искусственных нейронных сетей. Тем не менее, ИНС не дают даже приближенное представление о том, как работает мозг. Поэтому полезно оценить значимость функции активации в биологической НС и, соответственно, значимость её использования в искусственной нейронной сети.

Типичный нейрон имеет физическую структуру, состоящую из тела клетки [28], аксона, который отправляет сообщения другим нейронам, и дендритов, которые принимают сигналы или информацию от других нейронов.

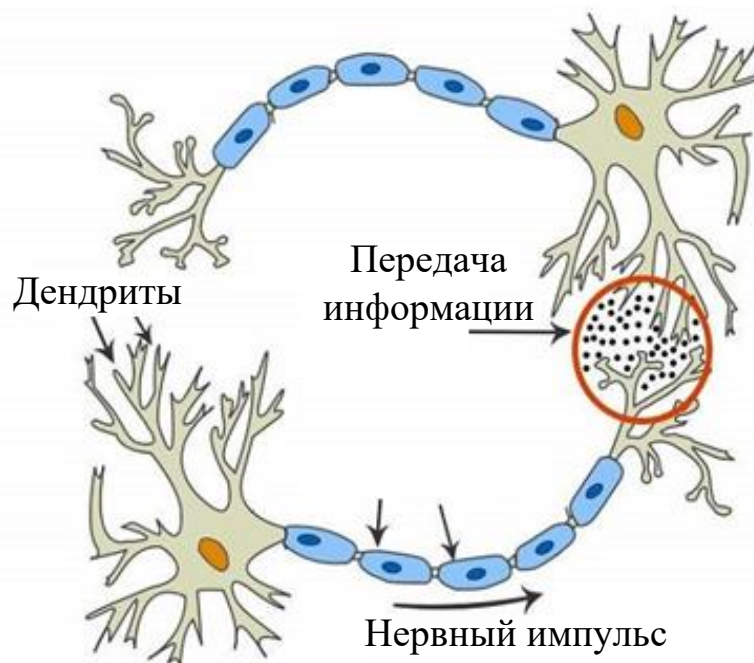


Рисунок 1.9. Биологический нейрон

На рисунке 1.9 круг указывает область, в которой взаимодействуют два нейрона. Нейрон получает сигналы от других нейронов через дендриты [29]. Вес, связанный с дендритом, называется синаптическим весом, умножается на поступающий сигнал [30]. Сигналы от дендритов накапливаются в теле клетки, и если сила результирующего сигнала выше определенного порога, нейрон передает сообщение аксону. В противном случае сигнал не усиливается нейроном и не распространяется дальше.

Функция активации (ФА) выполняет чрезвычайно важную роль в искусственных нейронных сетях. На основе её значений принимается решение о передаче или задержке сигнала.

Если в НС нет ФА, то веса и смещения будут выполнять линейные преобразования. Линейные уравнения легко решаются, но ограничены их возможности по решению сложных задач. НС без ФА по существу является только моделью линейной регрессии. ФА выполняет нелинейное преобразование на входе нейрона, что позволяет ему обучать и выполнять более сложные задачи. Мы хотели бы, чтобы наши нейронные сети могли решать сложные задачи, обнаружения закономерностей в данных такие как прогнозирование временных рядов или классификация параметров в наборах данных. Линейные преобразования никогда не смогут решать подобные сложные задачи.

Существуют разные типы функций активации. Самая простая ФА – пороговая функция. На рисунке 1.10 показана форма пороговой функции.

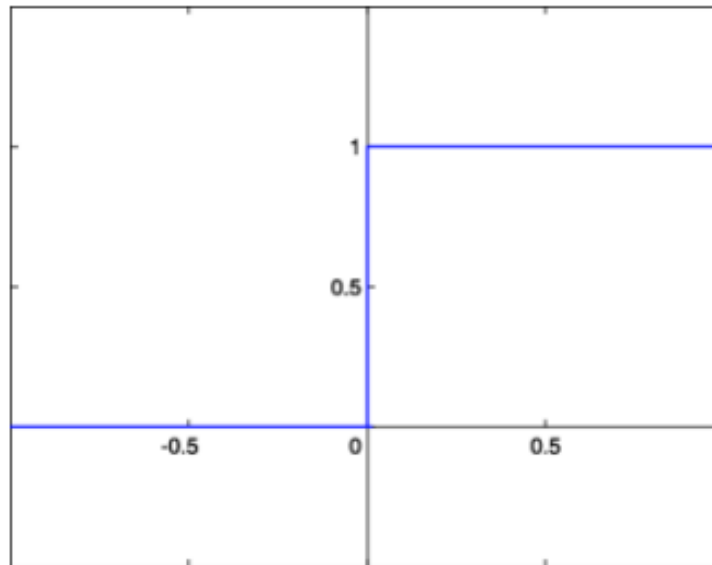


Рисунок 1.10. Пороговая функция

Если значение функции активации превышает заданное пороговое значение, тогда нейрон активируется, иначе он деактивирован. Пороговая функция чрезвычайно проста. Она может использоваться при создании бинарного классификатора. Когда нам просто нужно сказать «да» или «нет» для одного класса, пороговая функция будет лучшим выбором, поскольку она либо активирует нейрон, либо оставит его в прежнем состоянии. Такая функция чаще используется в теории, чем на практике, поскольку в большинстве случаев мы будем классифицировать данные на несколько классов, а не только на один класс. Это пороговая функция не сможет этого сделать. Более того, градиент пороговой функции равен нулю. Поэтому пороговая функция не столь полезна, поскольку во время обучения, когда градиенты функций активации отправляются для расчета ошибок для улучшения и оптимизации результатов, градиент ступенчатой функции сводит все к нулю, тогда совершенствование моделей на самом деле не происходит. Её нельзя использовать в сетях, которые обучаются по алгоритмам, требующим дифференцируемости функции активации [31]. Для преодоления этого, мы можем попытаться использовать линейную функцию активации, которая имеет уравнение, аналогичное линейной. В течение обучения, градиент ступенчатой функции не сводится к нулю,

поэтому градиент может обновиться. Однако, производная от линейной функции постоянна, т.е. не зависит от входного значения. И это большая проблема, мы не снижаем ошибку, так как градиент почти такой же. Теперь, если каждый слой имеет линейное преобразование, независимо от того, сколько слоев у нас есть окончательный результат, это не что иное, как линейное преобразование входного значения. Следовательно, линейная функция активации может быть идеальной для простых задач, где интерпретируемость очень нужна, и она используется только в одном выходном слое.

Наряду с линейной функцией активации, нелинейные функции активации являются наиболее часто используемыми функциями активации. Функция сигмоида является одной из них. На рисунке 1.11 показана форма функций сигмоида.

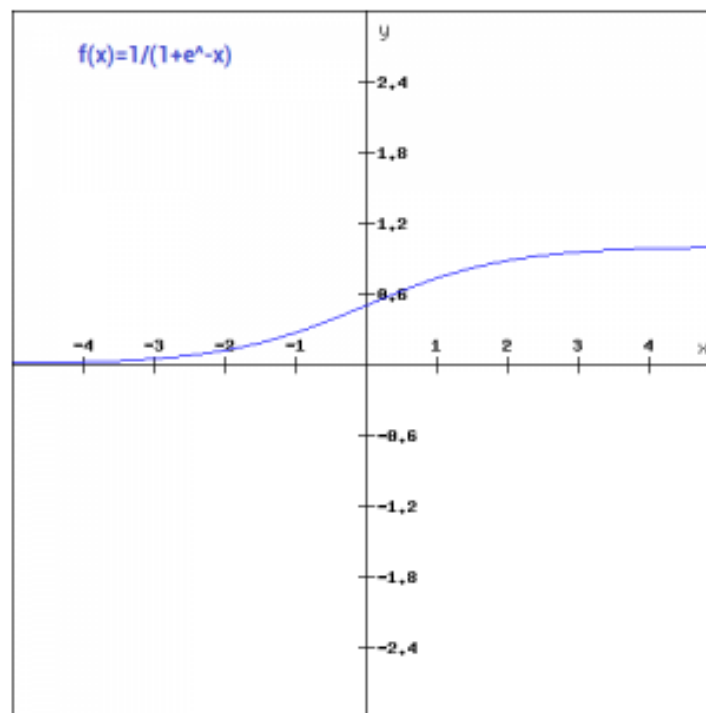


Рисунок 1.11. Сигмоидная функция

Основная причина, по которой используются сигмоидную функцию, состоит в том, что она существует между 0 и 1. Поэтому она особенно часто используется для моделей, где мы должны прогнозировать вероятность в качестве выхода. Поскольку вероятность чего-либо существует только между диапазоном от 0 до 1, функция

сигмоида является правильным выбором. График градиента сигмоидной функции, который представлен на рисунке 1.12, гладкий и зависит от входных значений. Это означает, что во время обучения веса могут быть соответственно обновлены.

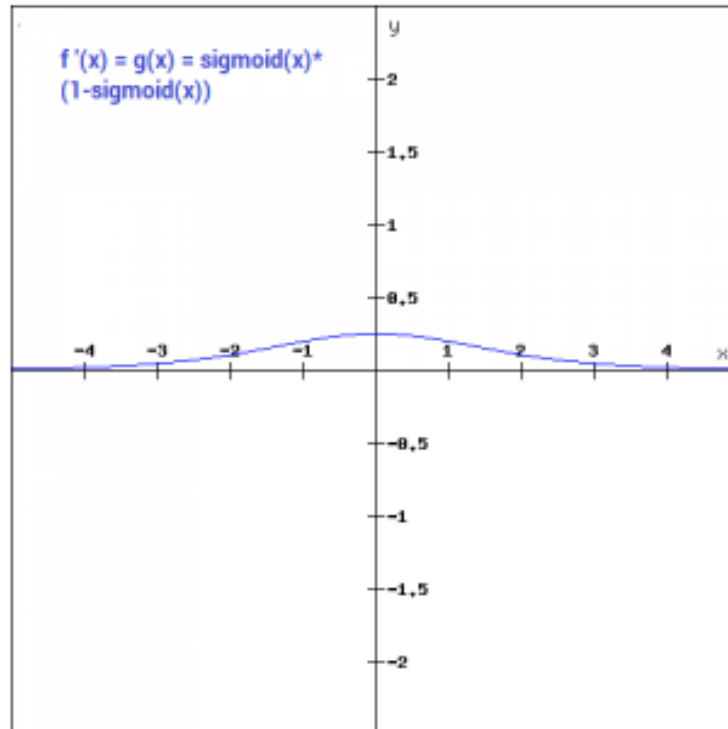


Рисунок 1.12. Градиент градиента сигмоидной функции

Однако, главный недостаток сигмоидной функцией, это то, что её градиент довольно плоский за пределами области $(-3, +3)$. Это означает, что, как только функция попадает в эту область, градиенты становятся очень малыми. Это означает, что градиент приближается к нулю, и сеть не обучается.

В некоторых приложениях используется гиперболический тангенс, дающий выходные значения в диапазоне $[-1, +1]$. Она имеет характеристики, подобные сигмоиду. Эта ФА в основном используется для классификации между двумя классами. Подобно сигмоиду, у гиперболического тангенса также есть проблема с градиентом. Известны также примеры применения импульсных функций активации [32], а также функций активации в форме гауссоиды, при этом возможно существенное упрощение структуры НС. Объясняется этот феномен усложнением

функций активации. В этой связи применение в качестве функций активации ННС функций принадлежности (ФП), разрабатываемых в теории нечетких чисел, является актуальной задачей [33, 34]. В силу введения понятий ФП приведем некоторые сведения из теории нечетких множеств.

Нечеткое множество [35] (fuzzy set) включает в себя элементы произвольной природы, относительно которых нельзя точно утверждать – обладают ли эти элементы некоторым характеристическим свойством, которое используется для задания нечеткого множества.

Пусть E есть множество, x – его элемент. Тогда нечетким подмножеством (или просто нечетким множеством) A множества E называется множество упорядоченных пар [36]:

$$\{(x | \mu_A(x))\}, \forall x \in E,$$

где μ_A – некоторое отображение элементов множества E в элементы множества значений (принадлежностей) $M = [0, 1]$. Функция μ_A называется функцией принадлежности (ФП).

Пусть $M = \{0, 1\}$. Тогда нечеткое множество A можно рассматривать как «не нечеткое» или просто обычное множество. Таким образом, обычные множества являются частным случаем нечетких.

Нечетким числом $\tilde{A} = \{(x, \mu_A(x)) | x \in \mathbb{R}\}$ называется нечеткое подмножество универсального множества действительных чисел, функция принадлежности которого $\mu_A(x): \mathbb{R} \rightarrow [0; 1]$ удовлетворяет следующим условиям [37]:

- непрерывность;

- нормальность: $\sup_{x \in \mathbb{R}} (\mu_A(x)) = 1$;

- выпуклость: $\mu_A(y) \geq \min \{ \mu_A(x), \mu_A(z) \}, x \leq y \leq z$.

Пример функции принадлежности нечеткого числа представлен на рисунке 1.13.

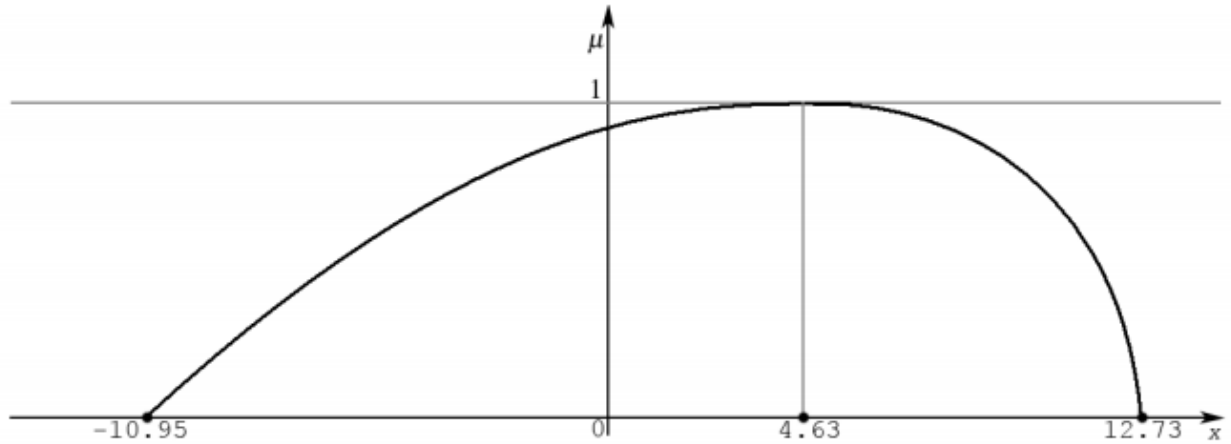


Рисунок 1.13. Пример функции принадлежности нечеткого числа

Основанием (носителем) нечеткого числа $\tilde{A}$ называется такое подмножество множества действительных чисел, для которого $\mu_A(x) > 0$, т.е.

$$\text{supp}(\tilde{A}) = [S_L, S_R] : \mu_A(x) > 0, \forall x \in [S_L, S_R]. \quad (1.1)$$

Аналогично *ядром* нечеткого числа $\tilde{A}$ называется такое подмножество множества действительных чисел, для которого $\mu_A(x) = 1$, т.е.

$$\text{ker}(\tilde{A}) = [K_L, K_R] : \mu_A(x) = 1, \forall x \in [K_L, K_R]. \quad (1.2)$$

Таким образом, нечеткое число $\tilde{A}$ определяется четверкой характерных точек:

$$\tilde{A} = \langle S_L, K_L, K_R, S_R \rangle \quad (1.3)$$

где характерные точки S_L , K_L , K_R и S_R определены в формулах (1.2) и (1.3).

Однако нечеткое число будет полностью определяться (1.3) только в том случае, когда известен вид ФП. При создании компьютерного приложения использовались ФП LR-типа, определяемые следующим выражением [38]:

$$\mu_A(x) = \begin{cases} f_L(x), & x \in [S_L, K_L]; \\ 1, & x \in [K_L, K_R]; \\ f_R(x), & x \in (K_R, S_R]; \\ 0, & x \notin [S_L, S_R]. \end{cases} \quad (1.4)$$

Выражение (1.4) позволяет получать разнообразные ФП, среди них наиболее известны треугольные, трапецеидальные и кусочно-непрерывные полиномиальные.

В данной диссертационной работе треугольную функцию принадлежности нечеткого числа $N = \langle A, B, C \rangle$ предлагается задавать в виде:

$$\mu(x) = f_L(x) \cdot H(x - A) \cdot H(B - x) + f_R(x) \cdot H(x - B) \cdot H(C - x),$$

где $f_L(x)$, $f_R(x)$ – функции, соответственно, левой и правой частей функции принадлежности, представляющие собой полиномы второго порядка, $H(x)$ – единичная функция Хевисайда [39].

В случае когда $f_L(x)$, $f_R(x)$ являются полиномами второго порядка и $f_L'(x) = 0$ и $f_R'(x) = 0$ в характеристических точках нечеткой величины, форму ФП, определяемой выражением (1.4), можно изменять, налагая на функции $f_L(x)$ и $f_R(x)$ одно из четырех дополнительных условий:

$$\begin{cases} f_L'(A) = 0; \\ f_R'(B) = 0. \end{cases} \quad (1.5)$$

$$\begin{cases} f_L'(B) = 0; \\ f_R'(B) = 0. \end{cases} \quad (1.6)$$

$$\begin{cases} f_L'(A) = 0; \\ f_R'(C) = 0. \end{cases} \quad (1.7)$$

$$\begin{cases} f_L'(B) = 0; \\ f_R'(C) = 0. \end{cases} \quad (1.8)$$

Внешний вид треугольных ФП нечетких чисел с учетом условий (1.5) – (1.8) представлен на рисунке 1.14

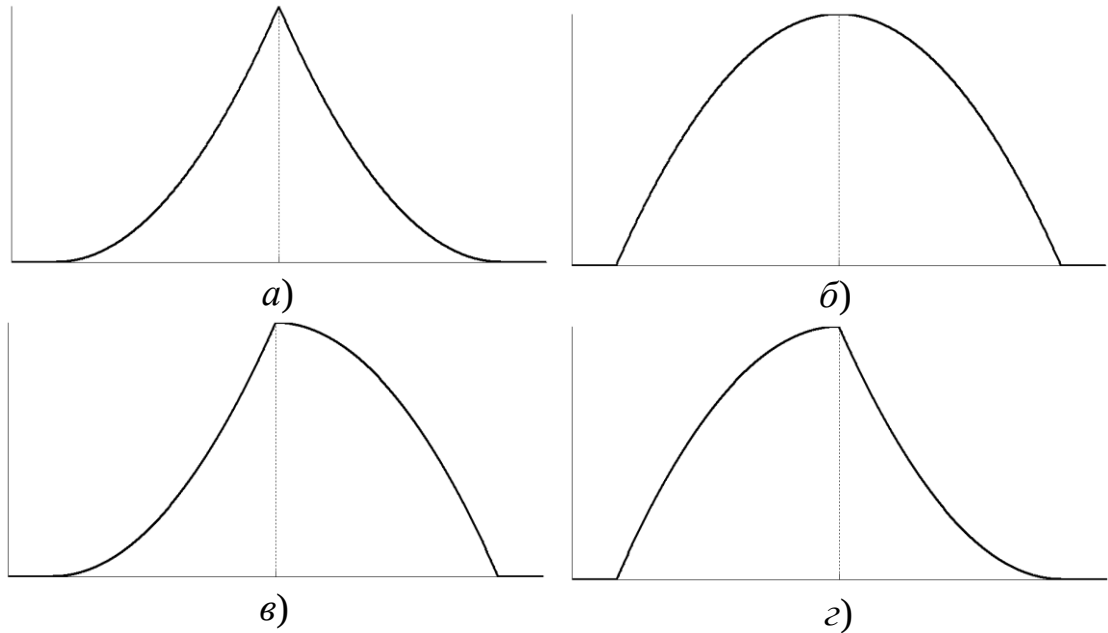


Рисунок 1. 14. Формы ФП с учетом дополнительных условий: а) условия (1.5); б) условия (1.6); в) условия (1.7); г) условия (1.8).

Данные ФП предлагается использовать в качестве функций активации нейронов и на основе этих ФА предлагается создание новых моделей нечетких нейронов и новых моделей нечетких нейронных сетей. Очевидно, что эти модели нечетких нейронных сетей отличаются от моделей ННС, рассмотренных выше в подразделе 1.1 и изучаемых в [13], поэтому будем именовать ННС, изучаемые в [13], ННС первого типа, а ННС, предлагаемые нами с моделями нечетких нейронов, использующих в качестве функций активации ФП, представленные на рисунке 1.13 а) – г), будем называть ННС второго типа. Отметим, используя структурные представления, что в ННС первого типа реализуются нечеткие отношения между нейронами как элементами нейронной сети, т.е. ННС первого типа реализуют

системы нечеткого логического вывода методами теории нейронных сетей. В ННС второго типа нечеткость является атрибутом нейрона как элемента нейронной сети. В данной трактовке ВНС, отмеченные нами выше в подразделе 1.1 так же является ННС второго типа.

В диссертации исследуются возможности применения полиномиальных кусочно-непрерывных функций, представленных на рис. 1.13 а) – з), в качестве функций активации нейронной сети, на основе которых создаются новые модели нечетких нейронов, новые модели нечетких нейронных сетей (ННС второго типа) и на их основе решаются типовые задачи обнаружения закономерностей в наборах данных.

1.4 Выводы по первой главе

1. Дано общее введение в искусственные нейронные сети. Отмечены основные достижения и проблемы теории и практики ИНС. Проанализированы процессы обучения нейронных сетей: с учителем и без учителя. Наиболее известными и чаще всего применяемым алгоритмами обучения ИНС являются методы: обратного распространения ошибки; сопряженного градиента; евенберга – Марквардта.

2. ИНС является некоторой моделью естественной нейронной сети. На функциях активации ИНС базируется решение о том, передавать сигнал или нет. Выполнен анализ различных типов ФА ИНС и предложено применение в ИНС новых ФА, базирующихся на ФП нечетких чисел LR - типа. Данный тип ННС предложено именовать ННС 2-го типа. На основе данного типа ННС решаются следующие типовые задачи обнаружения закономерностей в данных:

- 1) прогнозирование временных рядов со стохастическим трендом;
- 2) определение параметров детерминированного сигнала на фоне белого шума;

- 3) классификация интенсивности марковского случайного потока событий.
- 4) классификация объектов и их состояний.

Перечисленные задачи последовательно решаются в следующих главах:

- Задачи 1 и 2 решаются во 2-й главе;
- Задача 3 решается в 3-й главе;
- Задача 4 решается в 4-й главе.

2 МНОГОСЛОЙНЫЕ ОДНОНАПРАВЛЕННЫЕ НЕЙРОННЫЕ СЕТИ С НОВЫМИ ФУНКЦИЯМИ АКТИВАЦИИ ДЛЯ ПРОГНОЗА И ОПРЕДЕЛЕНИЯ ПАРАМЕТРОВ СИГНАЛОВ

2.1 Использование новых функций активации в многослойной однаправленной нейронной сети для прогноза значений временных рядов

2.1.1 Временной ряд

Временной ряд имеет разные определения [40 - 46]. Но в целом, временным рядом обычно называют последовательность наблюдений (данных, точек), упорядоченную по времени, но возможно упорядочение данных и по другому параметру [47]. Фактически, это совокупность данных о функционировании некоторой системы, таких как цена акций, данные о трафике и т.п. Временные ряды могут использоваться двумя способами для разных целей:

- Ретроспектива - использование исторических данных для анализа предыдущего поведения системы. Применения включают диагностику или распознавание отказа машины, или заболевания человека.
- Перспектива - использование данных для прогнозирования будущего поведения системы. Приложения включают прогнозирование запасов или цен, прогноз спроса на рынке и прогнозирование естественных данных.

Анализ временных рядов базируется на методах, которые исследуют закономерности поведения таких временных рядов для составления прогнозов. Прогнозирование временных рядов - это использование модели для прогнозирования будущих событий или будущих значений данных на основе известных прошлых событий. Это процесс, который производит набор выходов по заданному набору исторических переменных. Прогнозирование предполагает, что будущие события основаны на прошлых или нынешних событиях, в которых

некоторые аспекты прошлых моделей будут продолжаться и в будущем. Прошлые отношения могут быть затем обнаружены посредством изучения и наблюдения. Другими словами, прогнозирование временных рядов открывает связь между настоящими, прошлыми и будущими наблюдениями. Согласно Пламмеру [48], целью прогнозирования временных рядов является наблюдение или моделирование существующих рядов данных, которые могут быть в разных формах, например, серии финансовых данных (запасы, индексы, обменные курсы и т. д.), физически наблюдаемые серии данных (солнечные пятна, погода и т. д.), а также серии математических данных (последовательность Фибоначчи, интегралы дифференциальных уравнений и т.д.). Временной ряд представляет собой упорядоченную последовательность, характер которой меняется во времени t . Прогнозирование временных рядов принимает существующую серию данных $X_{t-n}, \dots, X_{t-2}, X_{t-1}, X_t$, и прогнозирует значения данных $X_{t+1}, X_{t+2}, \dots$. Теоретически эти ряды можно рассматривать как непрерывную функцию переменной времени t . Для практических целей время обычно рассматривается в терминах дискретных временных шагов. Размер временного интервала зависит от проблемы и может быть от миллисекунды до часов, дней или даже нескольких лет.

2.1.2 Создание нейронной сети с использованием метода скользящего окна для прогноза временных рядов

Для прогнозирования временных рядов используется метод скользящего окна. Размер окна и сегмента увеличивается до тех пор, пока мы не достигнем наименьшей погрешности [49]. После выбора первого сегмента следующий сегмент выбирается [50] с конца первого сегмента. Процесс повторяется до тех пор, пока все данные временных рядов не будут сегментированы. Процесс скользящего окна показан на рисунке 2.1 с размером окна 5. Скользящее окно накапливает исторические данные временных рядов [51]. Каждое число (1, 2, 3, ... ,10) представляет собой ежедневное наблюдение данных временных рядов за 1 день, 2

дня, 3 дня, ... ,10 дней соответственно. Первоначально окно покрывало числа от 1 до 5, что означает, что данные за 5 дней используются для прогнозирования данных на следующий день, затем окно скользит с правой стороны на один день, чтобы покрыть еще 5 дней (от 2 до 6) наблюдений, чтобы прогнозировать на следующий день. Процесс будет продолжен до тех пор, пока не будут исчерпаны данные временного ряда определенного периода времени, рассмотренного для экспериментальной цели. Выбор размера скользящего окна в начальный период определяется опытным путем, а затем исходя из набора данных.



Рисунок 2.1. Процесс скользящего окна

Структура многослойной однонаправленной сети в нашей работе состоит из входного слоя, одного скрытого слоя, и выходного слоя. Данная НС представлена на рисунке 2.2 [52].

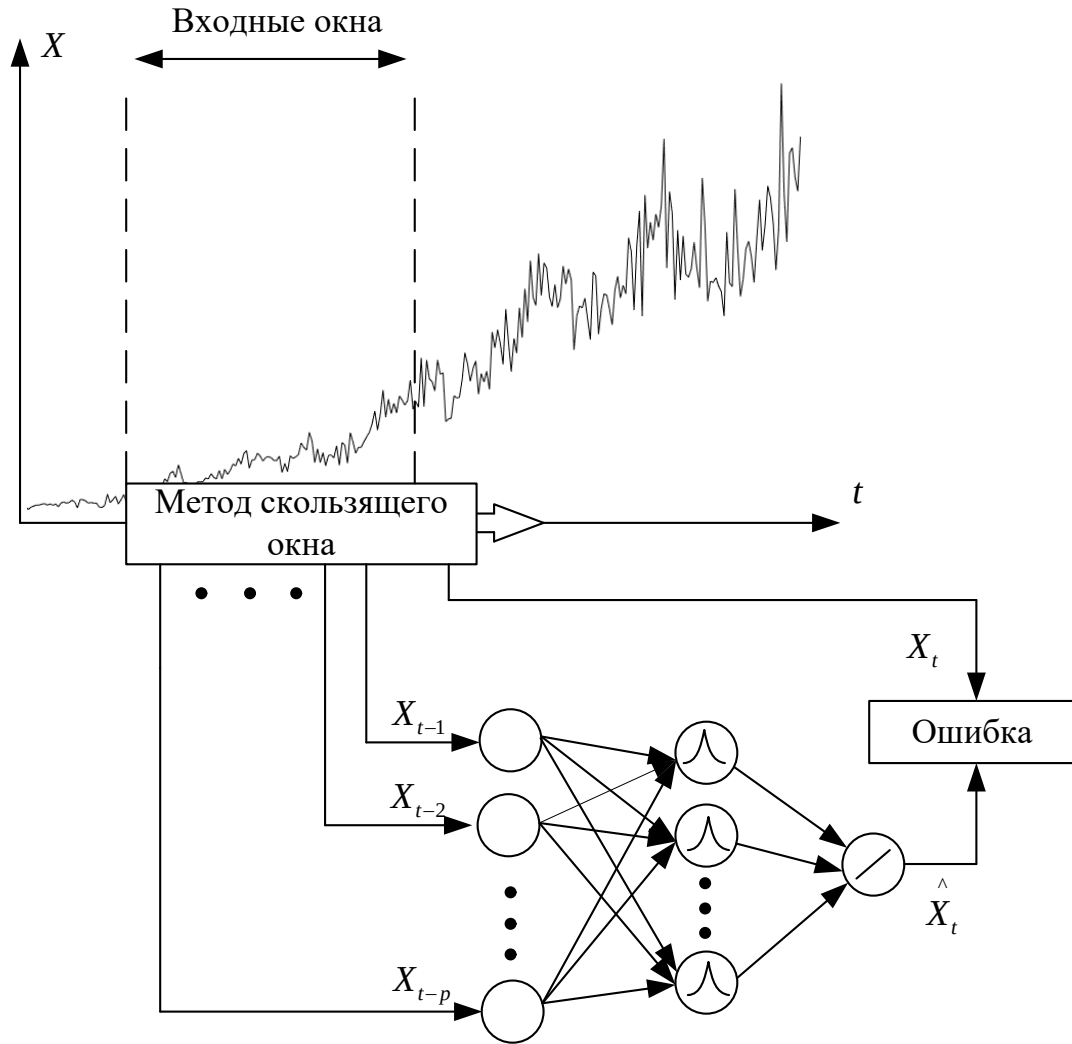


Рисунок 2.2. Метод скользящего окна в многослойной однонаправленной нейронной сети

На рисунке 2.2 показан метод скользящего окна в многослойной однонаправленной нейронной сети. Математическая модель сети может быть выражена как

$$\hat{X}_t = w_0 + \sum_{j=1}^Q w_j \cdot f \left(w_{0j} + \sum_{i=1}^P w_{i,j} \cdot X_{t-i} \right), \quad (2.1)$$

где P и Q - количество входных и скрытых нейронов соответственно, $w_j, j=0,1,\dots,Q$ - элементы весовых векторов от скрытого слоя до выхода,

$w_{i,j}, i=1,2,...,P \& j=1,2,...,Q$ - элементы весовых векторов от входа к скрытым нейронам.

Для согласованности сети и сокращения времени вычисления, входные данные должны быть нормализованы. Следовательно, использованные данные текущего исследования нормализуются до начала обработки. Нормализацию можно выполнить следующим образом:

$$X_{\text{норм}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}.$$

Для процесса обучения данные распределяются случайным образом с использованием 70% для обучения и 30% для тестирования. Алгоритм обратного распространения Левенберга - Марквардта (ОРЛМ) используется для обучения. Алгоритм ОРЛМ был разработан, чтобы аппроксимировать производную второго порядка, без необходимости вычислять матрицу Гессииана, тем самым увеличивая скорость обучения. Матрица Гессииана может быть аппроксимирована, как показано в уравнении (2.2), и градиент можно вычислить, как описано в уравнении (2.3):

$$H = J^T J, \quad (2.2)$$

$$g = J^T e, \quad (2.3)$$

где H - матрица Гессииана, g - градиент, J - матрица Якобиана, которая содержит первые производные сетевых ошибок относительно весов, и e - вектор сетевых ошибок во всех образцах обучения [53]. Для оценки матрицы Якобиана используется стандартный алгоритм обратного распространения для аппроксимации матрицы Гессииана [54].

Обновление веса вычисляется по формуле [55, 56]:

$$W_{k+1} = W_k - (J^T J + \mu I)^{-1} J^T e, \quad (2.4)$$

где W - вектор, содержащий все веса, I - единичная матрица и μ - параметр размера шага обучения, значение которого уменьшается после каждого успешного

шага.

Следует отметить, что этот алгоритм использует матрицу Якобиана для вычислений, считая, что функция производительности является средним значением суммы квадратов ошибок. Следовательно, сети должны использовать среднеквадратичную ошибку (СКО) δ для сетевого обучения:

$$\delta = \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}, \quad (2.5)$$

где y_i обозначает i -ую выборку, $\hat{y}_i$ - приближенные данные, полученные сетью для значения y_i , и n - количество выборок данных для сетевого обучения.

Задача прогнозирования поведения сложных динамических систем, особенно в экономике и социальной сфере является трудноформализуемой. Прогноз таких систем должен опираться почти исключительно на выявлении скрытых закономерностей в накопленных данных. Предсказание финансовых временных рядов – типичный пример такого рода задач. Практика применения указанных методов показала, что они эффективны лишь в редких случаях [57].

НС, изображенная на рисунке 2.2, обучена для прогнозирования временного ряда. Значения весов обновляются в процессе обучения по уравнению (2.4). Временные ряды обобщенных координат, используемых для обучения, показаны на рисунке 2.3. Используем 3 набора данных временных рядов, которые дают ежемесячный Австралийский импорт из Японии от июля 65-го года до октября 93-го года (Набор данных 1) [58], обменный курс австралийского доллара (доллар Австралии за 1 доллар США) от июля 69-го года до августа 95-го года (Набор данных 2) [59] и экспорт квартальных австралийских национальных счетов от сентября 59-го года до июня 95-го года (Набор данных 3) [60]. Для каждого набора данных проводим 5 экспериментов, в которых используем обычную (сигмоид) многослойную однонаправленную нейронную сеть (МОНС), МОНС с ФА вида a ,

МОНС с ФА вида *б*, МОНС с ФА вида *в*, МОНС с ФА вида *г*. Здесь и далее ФА вида *а* соответствует ФА, представленной на рисунке 1.13.*а*, ФА вида *б* соответствует ФА, представленной на рисунке 1.13.*б*, ФА вида *в* соответствует ФА, представленной на рисунке 1.13. *в*, ФА вида *г* соответствует ФА, представленной на рисунке 1.13.*г*, то есть всего проводится 15 экспериментов. В этих экспериментах используются одинаковые структуры для каждого набора данных, т.е. для каждого набора данных структуры МОНС имеют одинаковое количество нейронов в скрытом слое и размеры скользящего окна не отличаются.

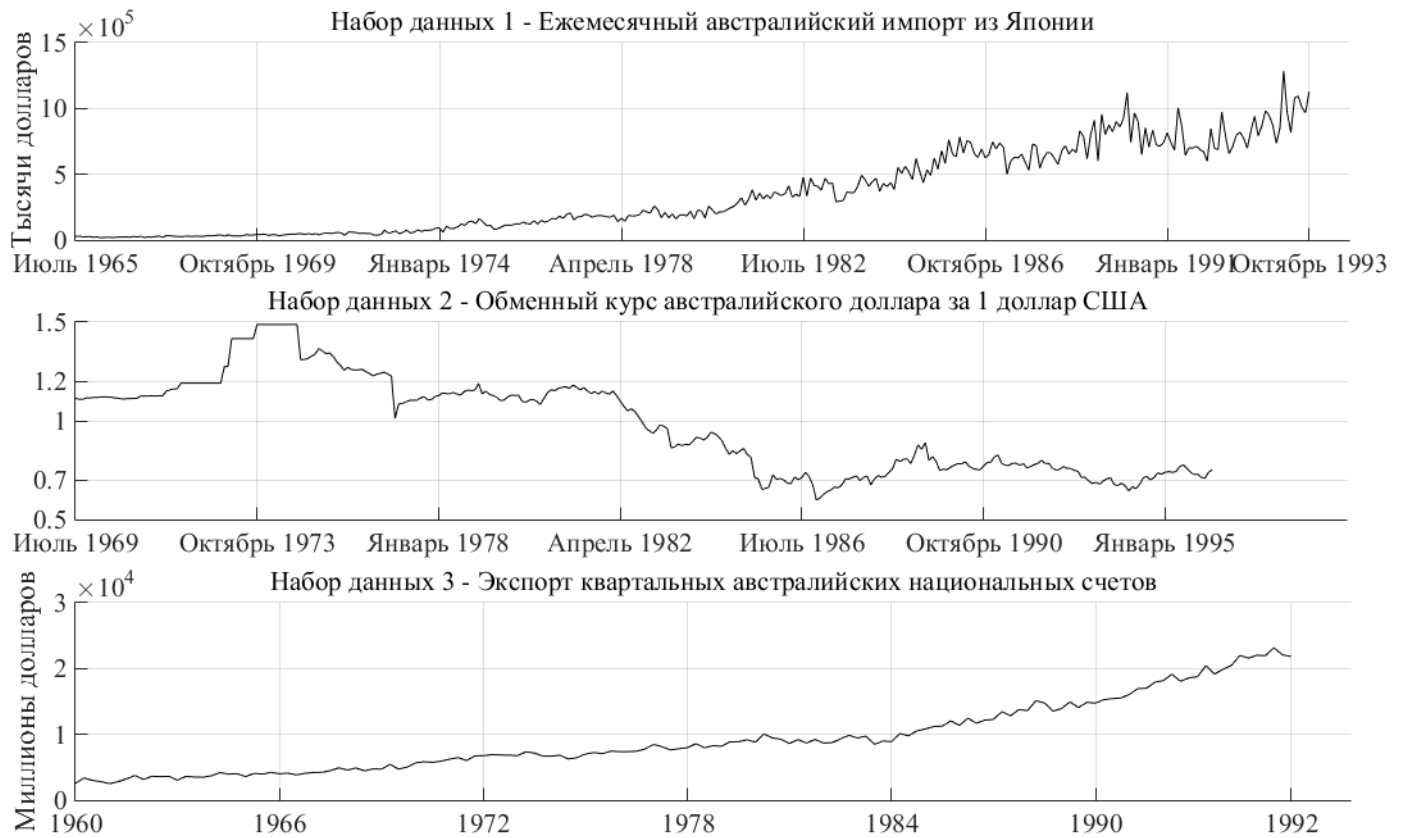
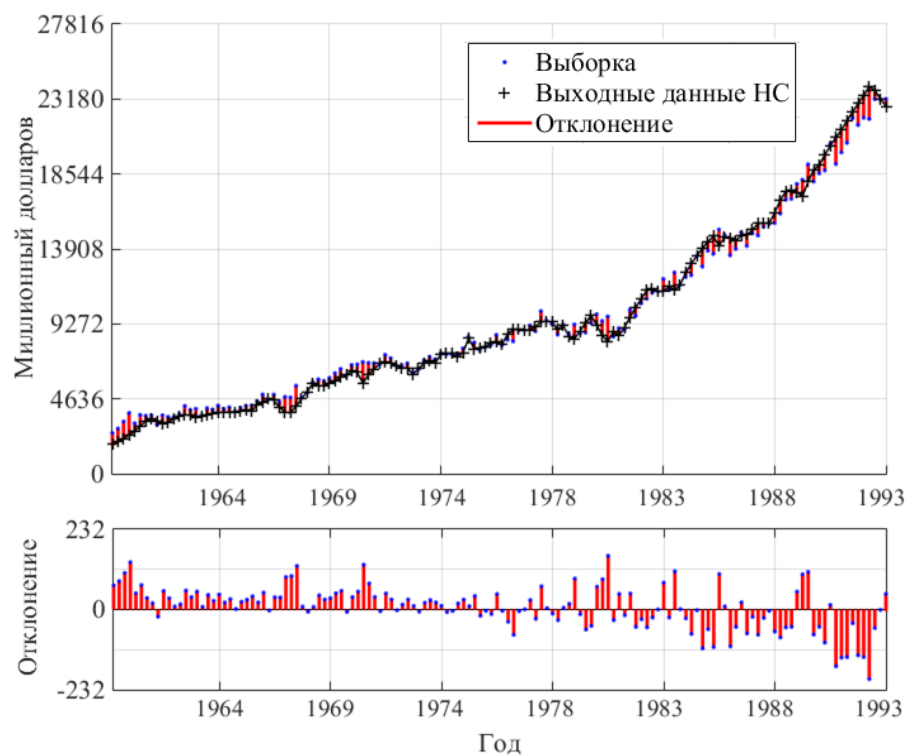
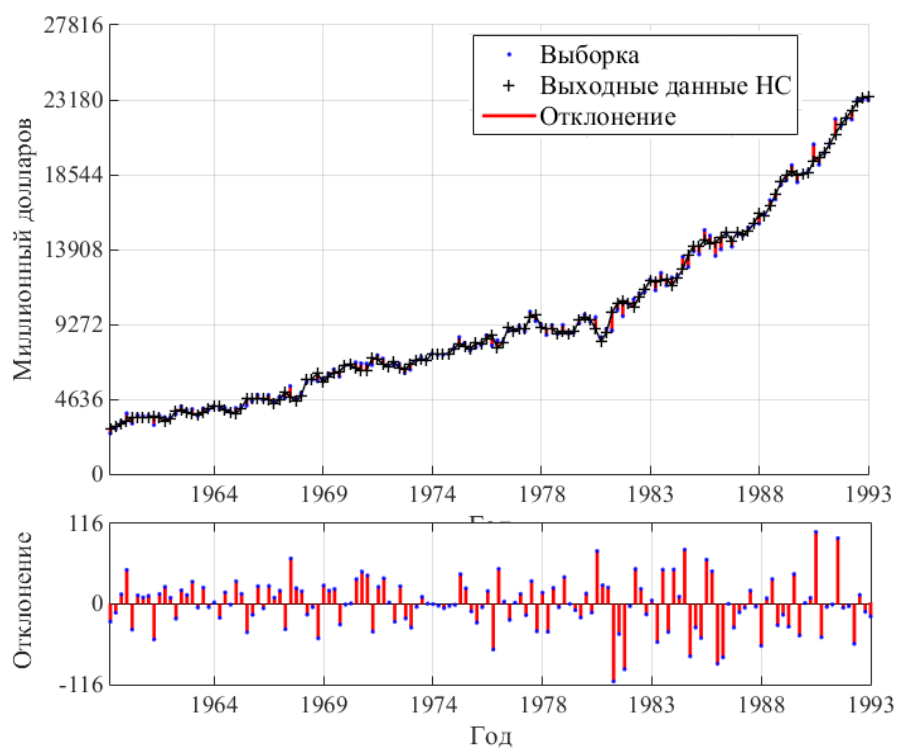


Рисунок 2.3 Наборы данных

На рисунке 2.4 показаны выборка и выходные данные после обучения МОНС с ФА вида *а*, *б*, *в*, *г* соответственно и отклонение между ними для набора данных 3.

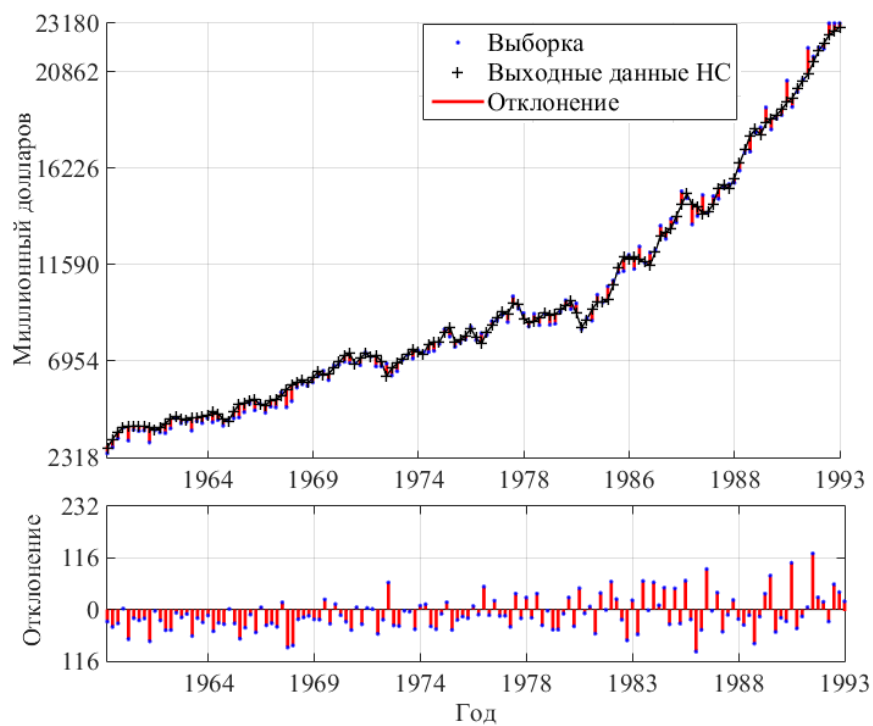


а)

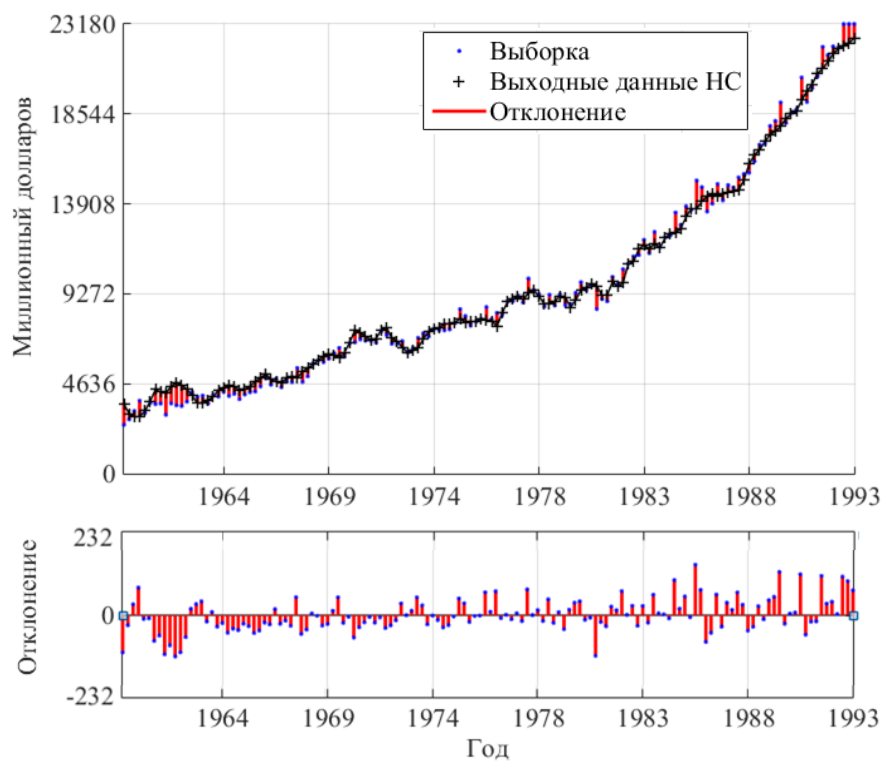


б)

Рисунок 2.4. Результаты тестирования с использованием НС с ФА вида *а*, *б*, *в*, *г* соответственно (начало)



в)



г)

Рисунок 2.4. Результаты тестирования с использованием НС с ФА вида *а*, *б*, *в*, *г* соответственно (окончание)

В таблице 2.1 приведены СКО вычисленные по формуле (2.5), для определения лучшей модели МОНС для прогнозирования временных рядов.

Таблица 2.1. СКО каждого эксперимента

	Обычная МОНС	МОНС с ФА вида <i>a</i>	МОНС с ФА вида <i>б</i>	МОНС с ФА вида <i>в</i>	МОНС с ФА вида <i>г</i>
Набор данных 1	0.0023	0.0025	0.0021	0.0023	0.0024
Набор данных 2	0.0018	0.0017	0.0016	0.0016	0.0021
Набор данных 3	2.7399×10^{-4}	7.0090×10^{-4}	2.3366×10^{-4}	2.8933×10^{-4}	4.3826×10^{-4}

Из результатов тестирования следует, что разрабатываемые нами модели нечетких нейронов и ННС второго типа вполне успешно решают задачу прогнозирования временных рядов. СКО для прогнозируемых данных в МОНС с ФА вида *a*, *б*, *в*, *г* не значительно отличаются от СКО для прогнозируемых данных в обычной МОНС. Эксперименты показывают, что модель НС с ФА вида *б* наиболее эффективна. Модели нейронной сети с ФА вида *в* также показывают хорошие результаты, но они незначительно отличаются от обычной МОНС. В наборах данных на рисунках 2.3 ННС прогнозирует тренды, которые можно интерпретировать как вероятностные (стохастические) сигналы на фоне случайных помех. Разрабатываемые ННС 2-го типа (особенно с ФА вида *б*) эффективно решают подобные задачи.

В следующем подразделе мы разрабатываем и исследуем НС с использованием ФА вида *a*, *б*, *в*, *г* для определения основных параметров детерминированного сигнала на фоне белого шума.

2.2 Определение амплитуды детерминированного сигнала на фоне белого шума с помощью ННС второго типа

Известно большое число прикладных задач, среди которых и задачи диагностирования технических систем, в которых необходимо определить параметры сигнала на фоне случайных помех. При постановке подобных задач часто используется аддитивная модель смеси детерминированного сигнала и белого шума.

Белый шум - это случайный сигнал, имеющий одинаковую интенсивность на разных частотах, что придает ему постоянную спектральную плотность мощности [61, 62]. Термин используется в том же или аналогичном значении во многих научных и технических дисциплинах, включая физику, акустическую инженерию, телекоммуникации и статистическое прогнозирование. Белый шум относится к статистической модели сигналов и источников сигналов, а не к какому-либо конкретному сигналу. Самый простой белый шум является гауссовском - значения, которые принимать шумовой сигнала, распределены по гауссовскому закону распределения [63, 64]. На рисунках 2.5, 2.6 представлены примеры гауссовского белого шумового сигнала и аддитивной смеси сигнала и белого шума. Белый шум получил свое название от белого света [65], хотя свет, который кажется белым, обычно не имеет плоской спектральной плотности мощности в видимой полосе.

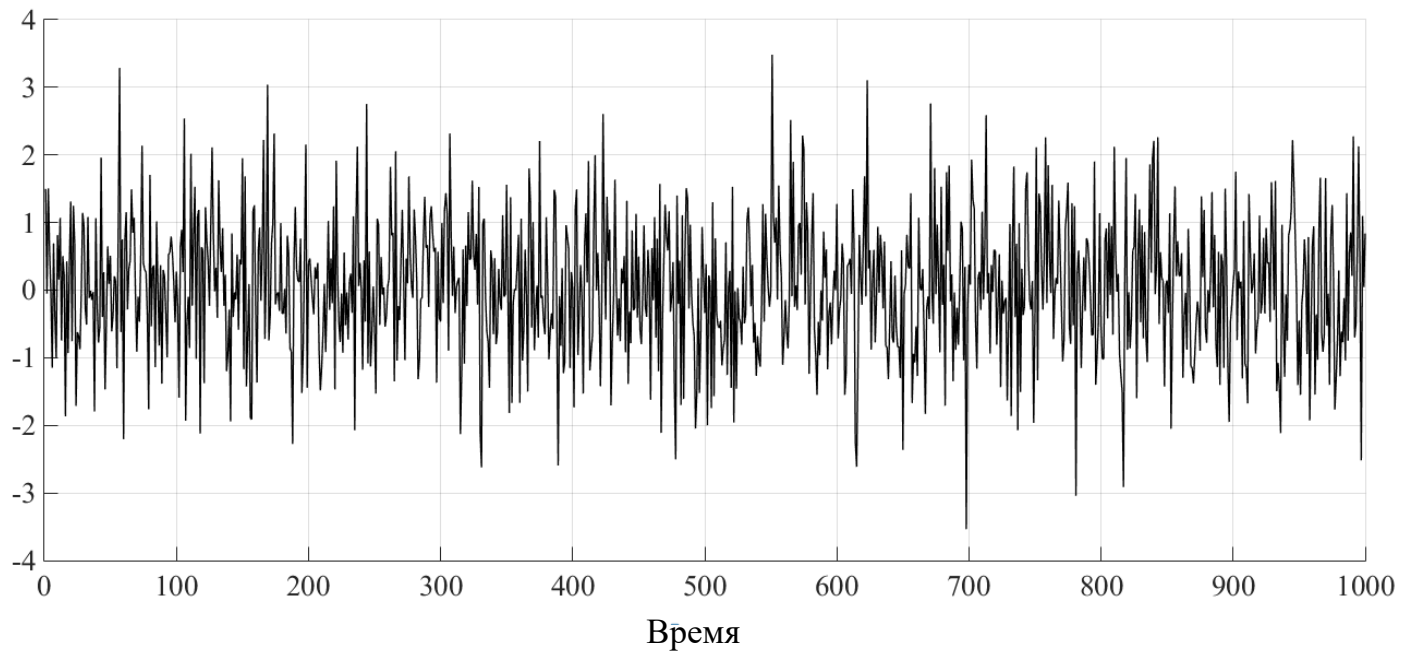


Рисунок 2.5. График гауссовского белого шумового сигнала

Выборки сигнала белого шума могут быть последовательными во времени или расположены вдоль одной или нескольких пространственных координат. При цифровой обработке изображений пиксели изображения с белым шумом обычно располагаются в виде прямоугольной сетки и предполагаются как независимые случайные величины с равномерным распределением вероятности по некоторому интервалу. Концепция может быть определена также для сигналов, распространяющихся по более сложным областям, таким как сфера или тор.

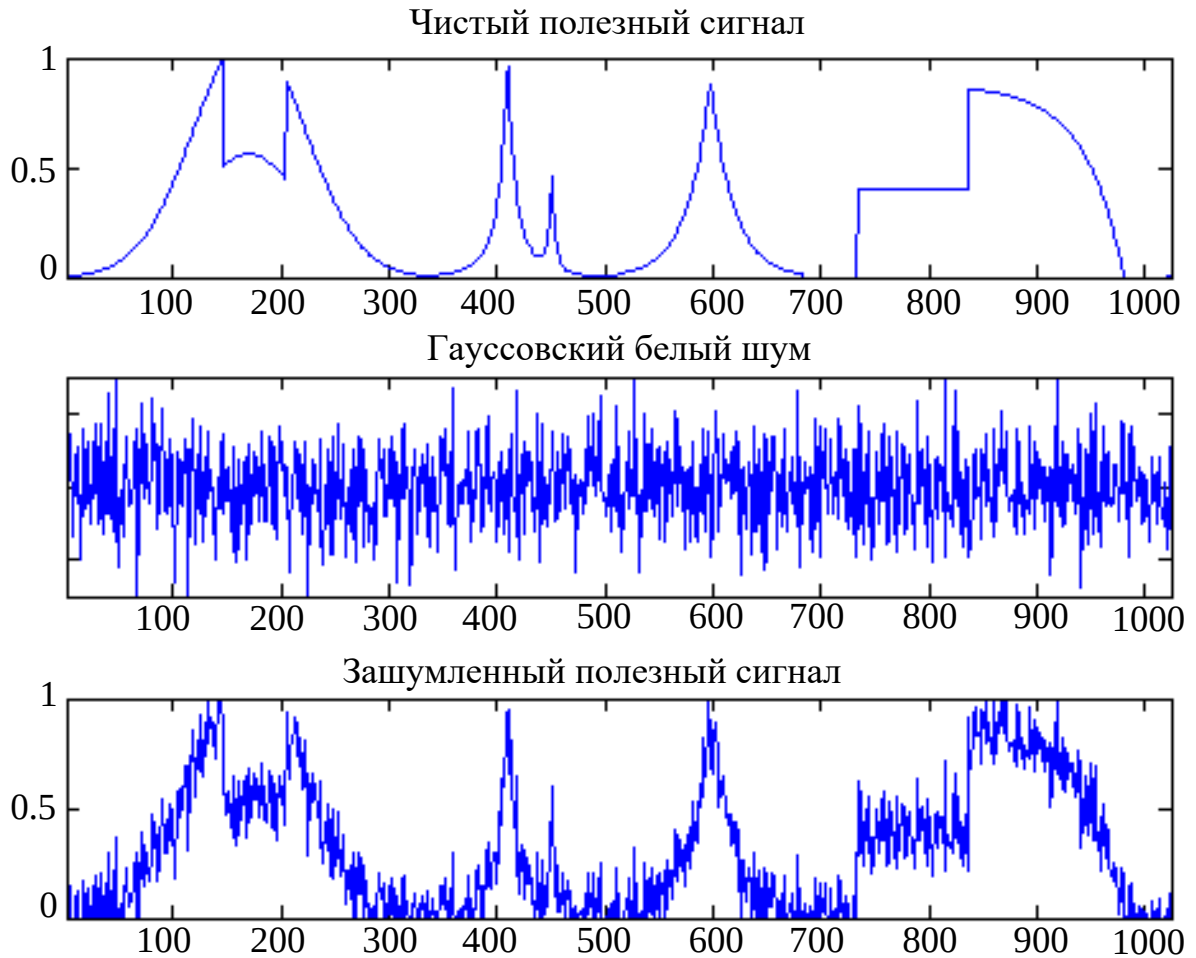


Рисунок 2.6. Аддитивная модель гауссовского шума

В качестве модели используем аддитивную смесь сигнала y и белого шума ε . В качестве сигнала используем сигнал в виде синусоиды:

$$y = x \times \sin(2\pi ft), \quad (2.6)$$

где x – амплитуда основного сигнала, f – основная частота сигнала. В качестве белого шума ε используем $N(0, \sigma^2)$, т.е. белый шум, является нормально распределенными случайными величинами с нулевым математическим ожиданием и дисперсией σ^2 .

Для эксперимента выбираем амплитуды из интервала $[1, 1,5]$; дискретное время t : 500 эквидистантных значений от 0 до 0,5 (с); и задаем $f = 60$ Гц. Затем добавляется аддитивный гауссовский белый шум с отношением сигнал /шум = 0,05

дБ [66]. Таким образом, на вход ННС поступает временной ряд, представляющий собой упорядоченную последовательность значений, характер которой меняется во времени t . Задачей ННС является прогноз значения x – амплитуды сигнала на входе ННС. Подобные задачи возникают во многих приложениях ННС. На рисунке 2.7 пунктиром показан чистый оригинальный сигнал – синусоида по функции (2.6) с амплитудой 1,5 и этот же сигнал после добавления к нему аддитивного гауссовского белого шума с отношением сигнал/шум = 0,05дБ.

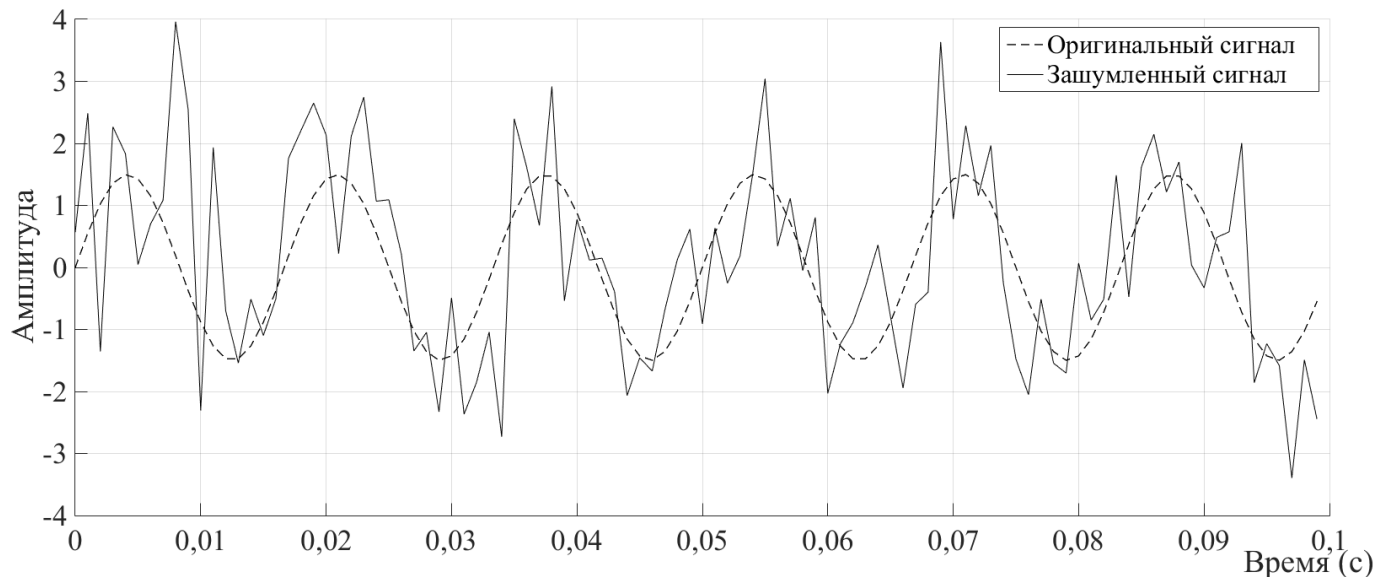


Рисунок 2.7. Сигнал для эксперимента

Итак, входными значениями ННС является аддитивная смесь гармонического сигнала и белого шума, которая моделируется временным рядом, а выходными значениями ННС являются амплитуда сигнала, который определяется формулой (2.6).

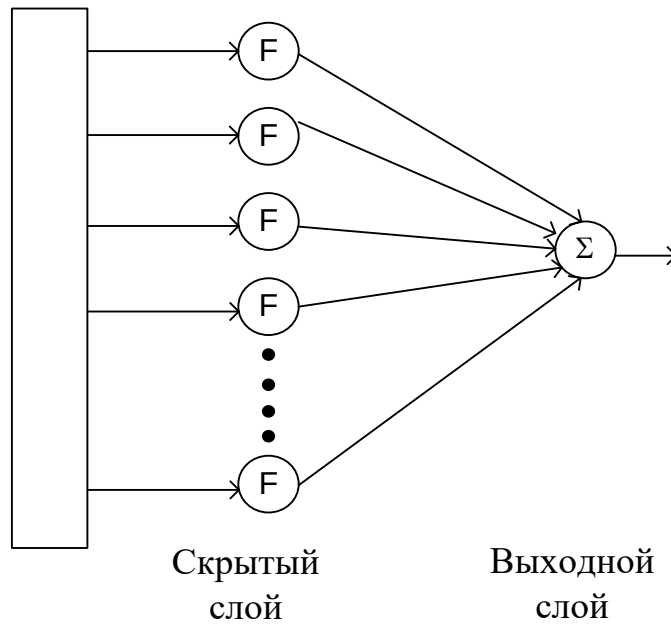


Рисунок 2.8. Структура сети

Структура сети, используемая в программе, показана на рисунке 2.8. Её можно представить в виде многослойной однонаправленной сети. Структура состоит из входного слоя, одного скрытого слоя, и выходного слоя. Скрытый слой состоит из 10 нейронов с нечеткими ФА вида a , b , v , z . В выходном слое используется линейная ФА. Разделяем все входные значения на секции: 70% для обучения и 30% для проверки. В процессе обучения использован метод Левенберга-Марквардта.

Для тестирования, выбираем 5 амплитуд от 1,1 до 1,5 и с каждой из которой создаем 20 зашумленных сигналов. Мы также проводим 5 экспериментов, в которых используем обычную (сигмоид) МОНС для определения амплитуды этих 5 амплитуд. В таблице 2.2 приведены СКО вычисленные по формуле (2.5), для определения лучшей модели МОНС для определения амплитуды сигнала на фоне белого шума.

Таблица 2.2. СКО результаты эксперимента

Амплитуда Тип НС	1,1	1,2	1,3	1,4	1,5
МОНС с ФА вида <i>a</i>	0.0230	0.0043	0.0023	0.0053	0.0157
МОНС с ФА вида <i>б</i>	0.0129	0.0028	0.0041	0.0082	0.0339
МОНС с ФА вида <i>в</i>	0.0088	0.0016	4.9653×10^{-4}	1.9299×10^{-4}	4.2292×10^{-4}
МОНС с ФА вида <i>г</i>	0.0164	0.0025	2.1084×10^{-4}	0.0100	0.0217
Обычная МОНС	0.0278	0.0074	0.0034	0.0063	0.0315

Из результатов тестирования следует, что разрабатываемые нами модели нечетких нейронов и ННС второго типа вполне успешно решают задачу определения амплитуды детерминированного сигнала на фоне белого шума. СКО в МОНС с ФА вида *a*, *б*, *в*, *г* в большинстве экспериментов показывают лучшее качество определения. Эксперименты показывают, что модель НС с ФА вида *в* наиболее эффективна. Модели нейронной сети с ФА вида *a*, *б*, *г* также показывают хорошие результаты, но они незначительно отличаются от обычной НМОНС.

2.3 Выводы по второй главе

1. Разработана структура МОНС с использованием метода скользящего окна для прогноза временных рядов с ФА вида *a*, *б*, *в*, *г*. Процесс обучения реализуется с помощью алгоритма Левенберга - Марквардта обратного распространения. Реальные наборы данных используются для тестирования.

2. Выполнена разработка структура МОНС с ФА вида *a*, *б*, *в*, *г* для определения амплитуды детерминированного сигнала на фоне белого шума. Алгоритм Левенберга - Марквардта обратного распространения тоже используется для обучения. Синусоидальные сигналы с добавлением к ним гауссовского белого шума с отношением сигнал/шум = 0,05 дБ используются в качестве наборов данных тестирования.

3. СКО используется для определения лучшей модели МОНС. В экспериментах прогнозирования временных рядов, СКО в МОНС с ФА вида *a*, *б*, *в*, *г* незначительно отличаются от СКО в обычной МОНС. Эксперименты показывают, что модель НС с ФА вида *б*, *в* наиболее эффективна. Тем не менее, тестирование разработанных МОНС на аддитивной смеси синусоидального сигнала с гауссовским белым шумом (отношением сигнал/шум = 0,05 дБ), их сравнительный анализ с известным МОНС с сигмоидом ФА показали значительно лучшие результаты по эффективности и точности определения, особенно МОНС с ФА вида *в*.

3 НЕЧЕТКАЯ НЕЙРОСЕТЕВАЯ КЛАССИФИКАЦИЯ ИНТЕНСИВНОСТИ МАРКОВСКОГО СЛУЧАЙНОГО ПОТОКА СОБЫТИЙ

3.1 Марковские случайные процессы

Случайный процесс называется марковским процессом если для каждого момента времени t_0 вероятность любого состояния системы в будущем (при $t > t_0$) зависит только от её состояния в настоящем (при $t = t_0$) и не зависит от того, когда и каким образом система пришла в это состояние [67] (т.е. как развивался процесс в прошлом). Пусть S техническое устройство, которое характеризуется некоторой степенью изношенности S . Нас интересует, как оно будет работать дальше. В первом приближении характеристики работы системы в будущем (частота отказов, потребность в ремонте) зависят от состояния устройства в настоящий момент и не зависят от того, когда и как устройство достигло своего настоящего состояния. Представим формульную модель функционирования подобной системы [68].

Пусть $\lambda(t)$ – кусочно-постоянный случайный процесс с двумя состояниями $\lambda(t) = \lambda_1$ либо $\lambda(t) = \lambda_2$ ($\lambda_1 < \lambda_2$). Случайный процесс $\lambda(t)$ принципиально не наблюдаем. Время пребывания процесса $\lambda(t)$ в i -м состоянии является случайной величиной, распределенной по экспоненциальному закону с параметрами λ_i ($i = 1, 2$), т.е. согласно [69] рассматриваемый случайный процесс моделирует МАР-поток (Markovian Arrival Process) событий с экспоненциальной функцией распределения:

$$F_i(t) = 1 - e^{-\lambda_i t}, \quad i = 1; 2. \quad (3.1)$$

Марковские случайные процессы названы по имени выдающегося русского математика А.А.Маркова (1856-1922), впервые начавшего изучение вероятностной связи случайных величин и создавшего теорию, которую можно назвать «динамикой вероятностей» [70]. МАР-поток событий часто используются в теории массового обслуживания [71] (в англоязычной литературе – Queueing Theory (теория очередей).

Основоположником исследований в данной области считается датский ученый А.К. Эрланг, опубликовавший в 1909–1917 г.г. работы, посвященные обслуживанию требований, поступающих на телефонную станцию. Развитие теории массового обслуживания связано с её применением в системах управления запасами, транспортных системах, системах связи, компьютерных сетях [72]. Частные случаи МАР-потоков называют также МС-потоками (Markov Chain) [73, 74]. Эти потоки можно представить в виде последовательно идущих друг за другом кусков простейших (пуассоновских) потоков, каждый из которых характеризуется своей интенсивностью λ_i прихода событий. Интенсивность λ_i потоков событий является важной характеристикой, модели МС-потока, которая хорошо аппроксимирует входящий поток событий реальной системы массового обслуживания.

В работе [75] предложен алгоритм оценивания числа состояний и значений интенсивностей МС-потока событий по выделенным в реализации потока интервалам постоянства интенсивности. Приведен пример работы алгоритма с использованием имитационной модели МС-потока, что позволило сравнить результаты работы алгоритма с исходными данными и результатами работы имитационной модели. Кроме того, предложена процедура оценивания интенсивностей состояний МС - потока в тех случаях, когда количество состояний должно быть не более некоторого наперед заданного числа. В работе [76] А.М.Горцев, Л.А. Нежелская, А.А.Соловьев предлагают алгоритм оптимальной оценки состояний МАР-потока событий в условиях непродлевающегося мертвого времени. Это, в свою очередь, позволяет изменять режимы работы системы обслуживания в зависимости от того или иного состояния МАР-потока (адаптироваться к изменениям интенсивности входящего потока событий). Выражения апостериорных вероятностей для оценки состояний потока получены в явном виде, что позволяет производить вычисления без привлечения численных методов. Сам же алгоритм оценки состояния потока обеспечивает минимум безусловной (полной) вероятности ошибки вынесения решения.

В этой главе решается задача исследования возможности применения НС, структура которой может быть представлена в виде многослойной однонаправленной сети (рисунок 3.1) с использованием ФА вида a , b , v , z , которые представлены на рисунке 1.13, для классификации интенсивности λ_i потоков событий.

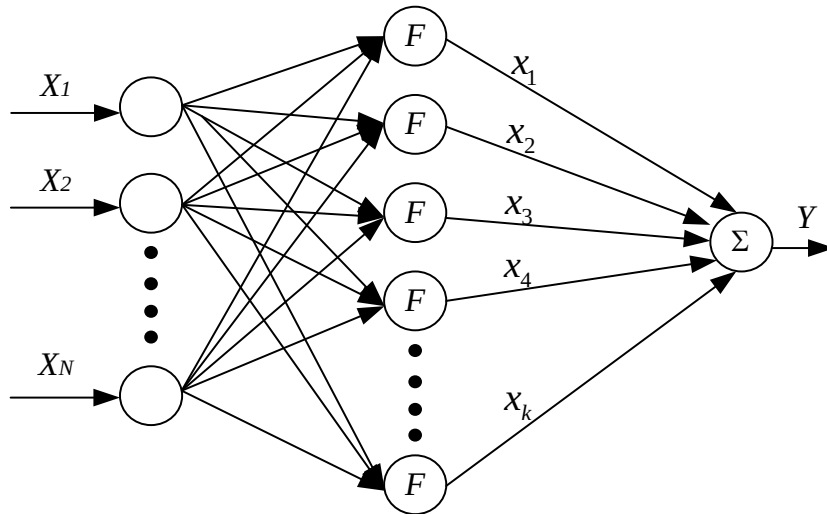


Рисунок 3.1. Структура сети

ННС состоит из входного слоя, одного скрытого слоя и выходного слоя. Скрытый слой состоит из нейронов с нечеткими ФА, виды которых для нечетких нейронных сетей второго типа представлены в [77, 78]. В выходном слое используется линейная ФА. На входы ННС поступают марковские случайные процессы с экспоненциальной функцией распределения (3.1).

3.2 Настройка функций активации

Выше отмечалось, мы используем четыре вида нечетких ФА второго типа, исследованных в [77, 78] и представленных на рисунках 1.13.

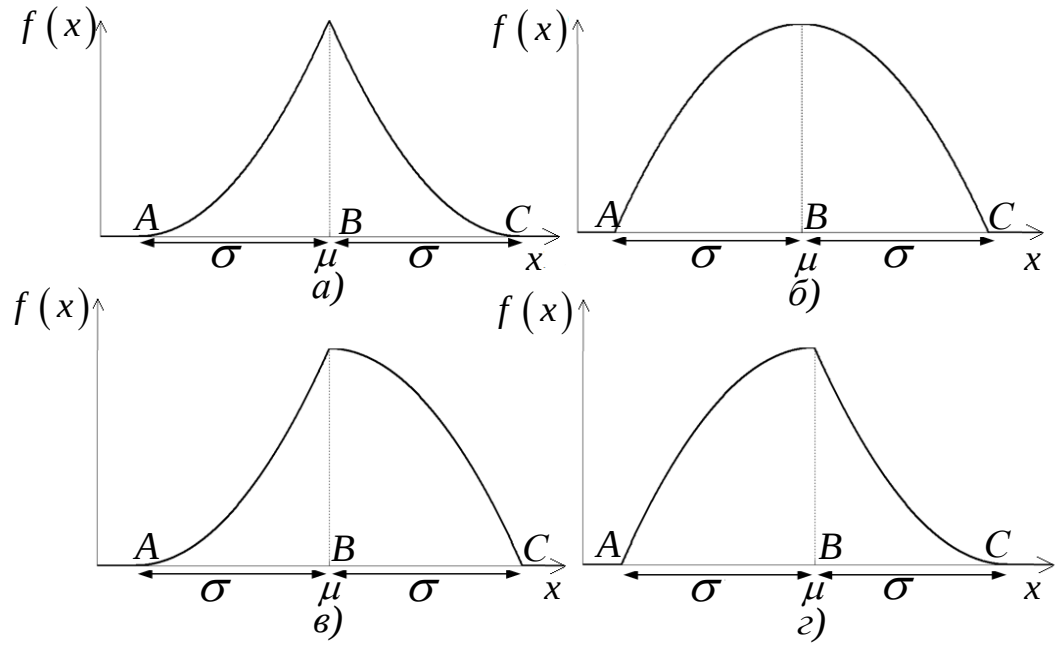


Рис. 3.2 – Функции активации ФА второго типа.

Набор точек $\langle A, B, C \rangle$ на рисунках 3.2 являются характеристическими точками ФА.

Известно, что $f_L(x)$, $f_R(x)$ являются полиномами второго порядка и $f_L'(x) = 0$, $f_R'(x) = 0$ в характеристических точках, т.е.

$$f_L(x) = \sum_{i=0}^2 a_i x^i \quad f_R(x) = \sum_{i=0}^2 b_i x^i \quad (3.2)$$

Для ФА вида a (рисунок 3.2), коэффициенты полиномов $f_L(x)$ и $f_R(x)$ задаются следующими выражениями [79]:

$$\begin{aligned} a_0 &= \frac{A^2}{(A-B)^2}; \quad a_1 = \frac{-2A}{(A-B)^2}; \quad a_2 = \frac{1}{(A-B)^2}, \\ b_0 &= \frac{C^2}{(B-C)^2}; \quad b_1 = \frac{-2C}{(B-C)^2}; \quad b_2 = \frac{1}{(B-C)^2}. \end{aligned} \quad (3.3)$$

Зададим $AB = BC = \sigma$; и $B = \mu$. Тогда $f_L(x)$ и $f_R(x)$ определяются следующим образом:

$$\begin{aligned} f_L(x) &= \frac{1}{\sigma^2} x^2 + \frac{-2(\mu - \sigma)}{\sigma^2} x + \frac{(\mu - \sigma)^2}{\sigma^2}, \mu - \sigma \leq x \leq \mu, \\ f_R(x) &= \frac{1}{\sigma^2} x^2 + \frac{-2(\mu + \sigma)}{\sigma^2} x + \frac{(\mu + \sigma)^2}{\sigma^2}, \mu \leq x \leq (\mu + \sigma), \end{aligned} \quad (3.4)$$

где μ и σ называются центром и шириной функций активации.

Пусть $a = (\mu - \sigma)$ если $\mu - \sigma \leq x \leq \mu$ и $a = (\mu + \sigma)$ если $\mu \leq x \leq (\mu + \sigma)$, тогда получим общее описание (обобщенная) ФА вида a :

$$f(x) = \frac{1}{\sigma^2} x^2 + \frac{-2a}{\sigma^2} x + \frac{a^2}{\sigma^2}. \quad (3.5)$$

Аналогично, для ФА вида b , $f_L(x)$ и $f_R(x)$ определяются следующим образом:

$$\begin{aligned} f_L(x) &= \frac{-1}{\sigma^2} x^2 + \frac{2\mu}{\sigma^2} x + \frac{(\mu - \sigma)^2 - 2\mu(\mu - \sigma)}{\sigma^2}, \mu - \sigma \leq x \leq \mu, \\ f_R(x) &= \frac{-1}{\sigma^2} x^2 + \frac{2\mu}{\sigma^2} x + \frac{(\mu + \sigma)^2 - 2\mu(\mu + \sigma)}{\sigma^2}, \mu \leq x \leq (\mu + \sigma). \end{aligned} \quad (3.6)$$

На этой основе получим общее описание (обобщенная) ФА вида b :

$$f(x) = \frac{-1}{\sigma^2} x^2 + \frac{2\mu}{\sigma^2} x + \frac{a^2 - 2\mu a}{\sigma^2}, \quad (3.7)$$

где $a = (\mu - \sigma)$ если $\mu - \sigma \leq x \leq \mu$ и $a = (\mu + \sigma)$ если $\mu \leq x \leq (\mu + \sigma)$.

Для ФА вида v , $f_L(x)$ и $f_R(x)$ определяются следующим образом:

$$\begin{aligned} f_L(x) &= \frac{1}{\sigma^2} x^2 + \frac{-2(\mu - \sigma)}{\sigma^2} x + \frac{(\mu - \sigma)^2}{\sigma^2}, \mu - \sigma \leq x \leq \mu, \\ f_R(x) &= \frac{-1}{\sigma^2} x^2 + \frac{2\mu}{\sigma^2} x + \frac{(\mu + \sigma)^2 - 2\mu(\mu + \sigma)}{\sigma^2}, \mu \leq x \leq (\mu + \sigma). \end{aligned} \quad (3.8)$$

На этой основе получим общее описание (обобщенная) ФА вида *в*:

$$f(x) = k_1 \left(\frac{1}{\sigma^2} x^2 + \frac{-2a}{\sigma^2} x + \frac{a^2}{\sigma^2} \right) + k_2 \left(\frac{-1}{\sigma^2} x^2 + \frac{2\mu}{\sigma^2} x + \frac{a^2 - 2\mu a}{\sigma^2} \right), \quad (3.9)$$

где $a = (\mu - \sigma)$, $k_1 = 1$, $k_2 = 0$ если $\mu - \sigma \leq x \leq \mu$ и $a = (\mu + \sigma)$, $k_1 = 0$, $k_2 = 1$ если $\mu \leq x \leq (\mu + \sigma)$.

Аналогично, для ФА вида *г*, общее описание функций имеет следующее выражение:

$$f(x) = k_1 \left(\frac{1}{\sigma^2} x^2 + \frac{-2a}{\sigma^2} x + \frac{a^2}{\sigma^2} \right) + k_2 \left(\frac{-1}{\sigma^2} x^2 + \frac{2\mu}{\sigma^2} x + \frac{a^2 - 2\mu a}{\sigma^2} \right), \quad (3.10)$$

где $a = (\mu + \sigma)$, $k_1 = 0$, $k_2 = 1$ если $\mu - \sigma \leq x \leq \mu$ и $a = (\mu - \sigma)$, $k_1 = 1$, $k_2 = 0$ если $\mu \leq x \leq (\mu + \sigma)$.

Выражение (3.5), (3.7), (3.9), (3.10) являются функциями активации, которые используются в ядре нейрона скрытого слоя ННС на рисунке 3.1.

3.3 Обучение ННС

Процесс обучения ННС состоит в выборе трех наборов параметров: центр μ и ширина σ для каждого из нейронов ННС и матрицы выходных весов между нейронами ННС и выходными узлами.

Как правило, центр и распределение функций активации нейронов ННС должны иметь характеристики, подобные данным. Одним из подходов к созданию интеллектуального выбора центров является выполнение кластеризации *К-средних* на учебном наборе и использование кластерных центров в качестве μ . Алгоритм кластеризации *К-средних* является простым повторяющимся алгоритмом кластеризации, разделяющий определенный набор данных на заданное пользователем число кластеров k [80]. Это неиерархичный метод неконтролируемого обучения. Он позволяет разделить произвольный набор данных на заданное число кластеров так, что объекты внутри одного кластера были достаточно близки друг к

другу, а объекты разных не пересекались. Другими словами, цель этого алгоритма — объединить в группы сходные данные по некоторым заданным критериям. Чаще всего при кластеризации используются меры расстояния. По умолчанию данный алгоритм использует квадрат Евклидова расстояния [81].

Алгоритм K -средних в нашем случае рассматривает распределение каждого случайного процесса как точку в функциональном пространстве. Базовый алгоритм K -средних произвольно определяет число кластерных центров в многомерном функциональном пространстве. Каждое распределение в наборе данных присваивается кластеру, произвольный средний вектор которого является самым близким. Процедура продолжается до тех пор, пока не произойдет существенного изменения положения средних векторов классов между последовательными итерациями алгоритмов.

Однако алгоритм K -средних очень чувствителен в начальных отправных точках. Алгоритм K -средних генерирует начальный кластер случайным образом. Когда случайные начальные точки близки к окончательному решению, алгоритм K -средних имеет высокую вероятность обнаружить центр кластера. В противном случае это приведет к некорректным результатам кластеризации [82].

Каждый кластер соответствует ядру функций активации. Краткое описание алгоритма K -средних, используемого в данной работе, состоит в следующем [83]:

Шаг 1: случайно выбираются количество начальных центров кластера из представленных выше наборов данных $\mu_1, \mu_2, \dots, \mu_k$, где k — количество начальных центров кластера, чем больше k , тем точнее результаты и продолжительнее время обработки, определяется $\mu_a = \{x_{b1}, x_{b2}, \dots, x_{bN}\}$ где $a = 1; 2; \dots; k$ и b — случайное число от 1 до L , где L — количество входных данных, N — количество дискретных мониторов времени наблюдения входных данных.

Шаг 2: определяются Евклидово расстояние от центра распределения x_i , ($i = 1; 2; \dots; L$) до кластера, C_j ($j = 1; 2; \dots; k$) если $\|x_i - z_j\| < \|x_i - z_p\|$, $p = 1; 2; \dots; k$,

$j \neq q$ тогда x_i находится в кластере C_j .

Шаг 3: вычисляются новые кластерные центры μ_i^* по формуле:

$$\mu_i^* = \frac{1}{n_i} \sum_{x_j \in C_i} x_j, \quad (j=1; 2; \dots; k), \quad (3.11)$$

где n_i является количеством распределений, принадлежащих кластеру C_i .

Шаг 4: если $\mu_i^* = \mu_i$ ($i=1; 2; \dots; k$), тогда останов. В противном случае следует перейти к шагу 2.

Если имеются ограничения по времени и если процесс кластеризации не завершается на шаге 4, тогда он выполняется для максимального фиксированного числа итераций. Чем больше числа итераций, тем точнее определяются центры кластеров.

Настройка ширины ядра является труднорешаемой проблемой при построении ННС. Если ширина ядра слишком велика, то оцененная плотность вероятности сглаживается. Наоборот, когда она слишком мала, может быть чрезмерная адаптация к определенному набору данных. Ширина ядра устанавливается как среднее расстояние между данными соответствующего кластера [84]:

$$\sigma_i = \sqrt{\frac{1}{p} \sum_{j=1}^p (X_j - \mu_i)^2}, \quad (3.12)$$

где σ_i является шириной i -ого нечеткого скрытого нейрона, μ_i является центром i -ого нечеткого скрытого нейрона, p - количество функциональных данных в данном кластере. X_j - j -ые данные в i -ом кластере. Ширина ядра (3.12) гарантирует, что отдельные функции активации не слишком острые или/и не слишком плоские.

Конечный набор параметров для тренировки - это выходные веса. При построении ННС, задача корректировки весов сети между скрытым слоем и выходным слоем может быть отнесена к задачам оптимизации, обычно решаемыми методом градиентного спуска (также известный как реализация метода наименьших

квадратов [85]). В данной работе, веса ННС между скрытым слоем и выходным слоем рассчитываются с использованием метода псевдообращения [86-88] Этот алгоритм преодолевает многие проблемы в традиционных алгоритмах градиента, таких как остановка критерия, скорость обучения, количество эпох и локальные минимумы. Благодаря более короткому времени обучения и способности к обобщению, он подходит для приложений реального времени.

На рисунке 3.1 X представляют собой входные наборы данных о пространственных входных признаках. Выход y ННС в наблюдаемый момент времени j ($j=1; 2; \dots; N$, N – количество наблюдаемых моментов времени) имеет следующий вид:

$$y_j = \sum_{i=1}^k w_i f(x_i), j=1; 2; \dots; N, \quad (3.13)$$

где w_i – вес между выходным нейроном и i -м скрытым нейроном; $f(x)$ – обобщенная ФА; k – число нейронов скрытого слоя.

Уравнение (3.13) можно записать в виде [89]

$$Y = \Phi w, \quad (3.14)$$

где

$$\Phi(\mu, \sigma, X) = \begin{bmatrix} \varphi_1(\mu_1, \sigma_1, X_1) & \dots & \varphi_K(\mu_K, \sigma_K, X_1) \\ \dots & \dots & \dots \\ \varphi_1(\mu_1, \sigma_1, X_v) & \dots & \varphi_K(\mu_K, \sigma_K, X_v) \\ \dots & \dots & \dots \\ \varphi_1(\mu_1, \sigma_1, X_N) & \dots & \varphi_K(\mu_K, \sigma_K, X_N) \end{bmatrix}, v=1; 2; \dots; N. \quad (3.15)$$

φ – функция активации. Матрица Φ (размерность $N \times K$) называется выходной матрицей скрытого слоя нейронной сети; i -я строка Φ является i -м скрытым выходным нейроном относительно входов $X_1, X_2, \dots, X_N$.

Пусть T – матрица выходных данных обучения. Если $\Phi w = T$, то выходные веса w вычисляются по формуле

$$w = (\Phi^T \Phi)^{-1} \Phi^T T, \quad (3.16)$$

где $(\Phi^T \Phi)^{-1} \Phi^T$ является обобщенной псевдообратной [90] выходной матрицей Мура–Пенроуза скрытого слоя.

Итак, основные параметры ННС определяются путем вычисления центров и ширины с использованием алгоритма кластеризации K -средних и определения весов выходного слоя ННС с использованием метода псевдообращения.

3.4 Эксперименты по классификации интенсивности потоков событий

Марковские случайные потоки событий, поступающие на входы ННС, представляются в виде двух последовательно идущих друг за другом кусков простейших (пуассоновских) потоков, каждый из которых характеризуется своей интенсивностью λ_i прихода событий.

В данной главе представлены результаты исследования по классификации параметров случайных потоков на основе четырёх экспериментов:

- первый эксперимент: марковские случайные потоки событий имеет 2 интенсивности $\lambda_1 \in (0; 0,5]$, $\lambda_2 \in (0,5; 1]$, параметры λ_1 , λ_2 не фиксированы и не известны;
- второй эксперимент: марковские случайные потоки событий имеет 2 интенсивности $\lambda_1 \in (0; 0,5]$, $\lambda_2 \in (0,5; 1]$, параметры λ_1 , λ_2 фиксированы, но не известны;
- третий эксперимент: марковские случайные потоки событий имеет 2 интенсивности $\lambda_1 = 0,25$, $\lambda_2 = 0,75$, то есть параметры λ_1 , λ_2 фиксированы и известны.
- четвертый эксперимент: марковские случайные потоки событий имеет 10 разных интенсивностей λ_i ($i = 1, 2, \dots, 10$) ($\lambda_1 < \lambda_2 < \dots < \lambda_{10}$)

В первом эксперименте, входными данными являются потоки, состоящие из 1000 случайных процессов с параметрами $\lambda(t)$, принимающими значения

$\lambda = 0,001; 0,002; \dots; 1$. То есть, входные данные можно представить в виде матрицы размерности 1000×1000 . Выходные данные можно представить в виде матрицы-вектора размерности 1000×1 , элементы которой имеют только два значения 0 или 1. Значение 0 обозначает, что классифицируемый поток имеет распределения с параметрами λ_1 из интервала $(0; 0,5]$, значение 1 обозначает, что классифицируемый поток имеет распределения с параметрами λ_2 из интервала $(0,5; 1]$.

Во втором эксперименте, входными данными являются потоки, состоящие из 500 случайных процессов с параметром $\lambda_1 \in (0; 0,5]$ и 500 случайных процессов с параметром $\lambda_2 \in (0,5; 1]$. Выходные данные представляются матрицей-вектором размерности 1000×1 , состоящей из значений 0 и 1. Значение 0 обозначает, что классифицируемый поток имеет распределение с параметром λ_1 из интервала $(0; 0,5]$, значение 1 обозначает, что классифицируемый поток имеет распределение с параметром λ_2 из интервала $(0,5; 1]$.

В третьем эксперименте, входными данными являются 200 случайных процессов с параметром $\lambda_1 = 0,25$ и 200 случайных процессов с параметром $\lambda_2 = 0,75$. Выходные данные также являются матрицей-вектором размерности 400×1 , которая состоит из значений 0 и 1. Значение 0 обозначает, что поток имеет распределение параметром $\lambda_1 = 0,25$; значение 1 обозначает, что поток имеет распределение с параметром $\lambda_2 = 0,75$.

В четвертом эксперименте, входными данными являются потоки, состоящие из 500 случайных процессов с параметром $\lambda_i (i = 1, 2, \dots, 10)$. Данные каждого процесса получены из 1000 моментов времени. То есть, входные данные можно представить в виде матрицы размерности 5000×1000 . Выходные данные также являются матрицей-вектором размерности 5000×1 , состоящей из значений 1, 2, ..., 10 соответственно с параметрами $\lambda_1, \lambda_2, \dots, \lambda_{10}$.

Результаты тестирования

Эксперименты проводились на базе процессора Core i5-4460 3,2 ГГц с 8 ГБ ОЗУ под MATLAB R2013a. Графики генерировались с использованием MATLAB.

На рисунке 3.3 представлена зависимость качества кластеризации от количества нейронов в скрытом слое в момент времени $t = 1000$ с использованием ФА вида б в первом эксперименте.

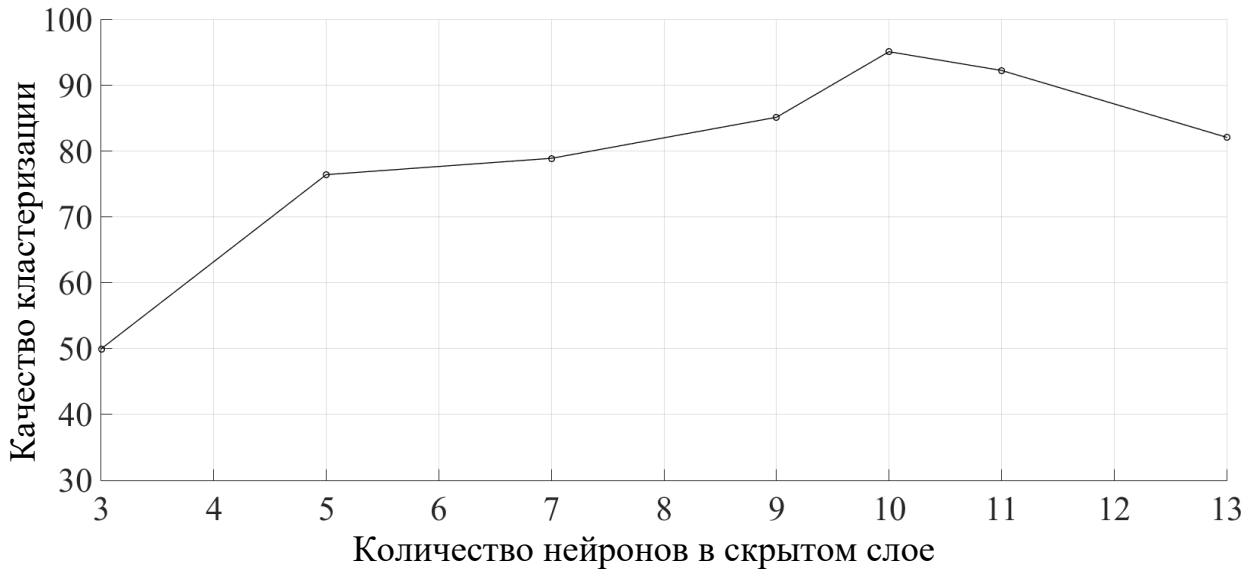


Рисунок 3.3. Зависимость качества кластеризации от количества нейронов в скрытом слое

Из рисунка 3.3 следует, что использование 10 нейронов в скрытом слое обеспечивает самое высокое качество кластеризации. Поэтому в первом эксперименте, ННС использует с 10 нейронов с нечеткими ФА, виды которых для нечетких нейронных сетей второго типа представлены в [77, 78] в скрытом слое.

В первом эксперименте наборы данных для тестирования состояли из 400 случайных процессов $\lambda^*(t)$, $\lambda^* = 0,001; 0,001; \dots; 1$. Результаты классификации с помощью ННС процессов по интенсивностям λ_1 , λ_2 представлены на рисунке 3.4, 3.5, 3.6, 3.7.

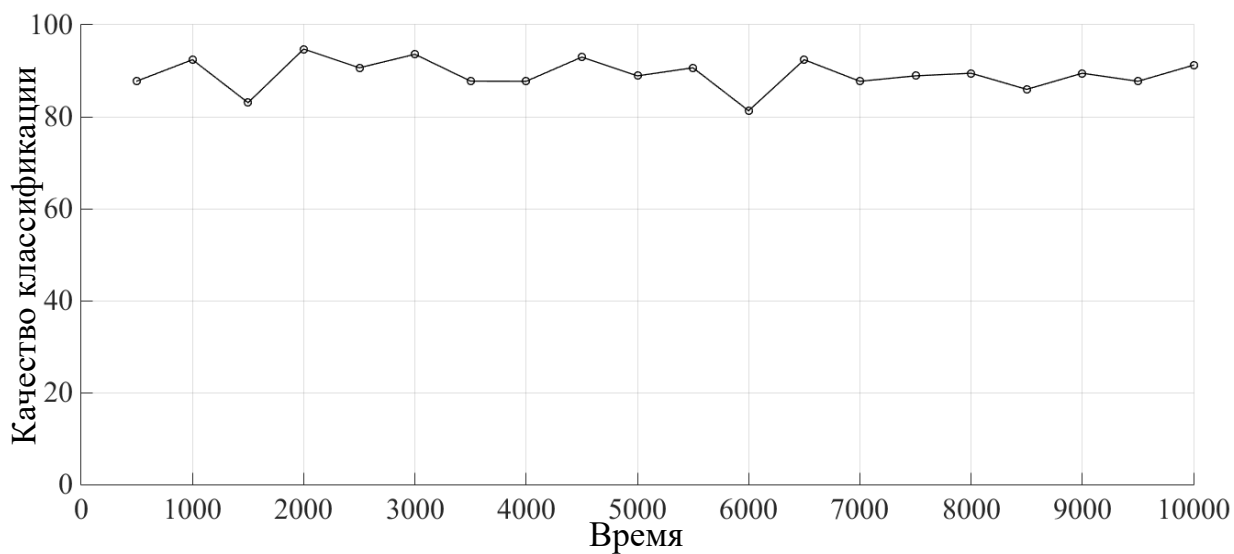


Рисунок 3.4. Эффективность классификации по интенсивности марковских случайных потоков событий в первом эксперименте с ФА вида *a*

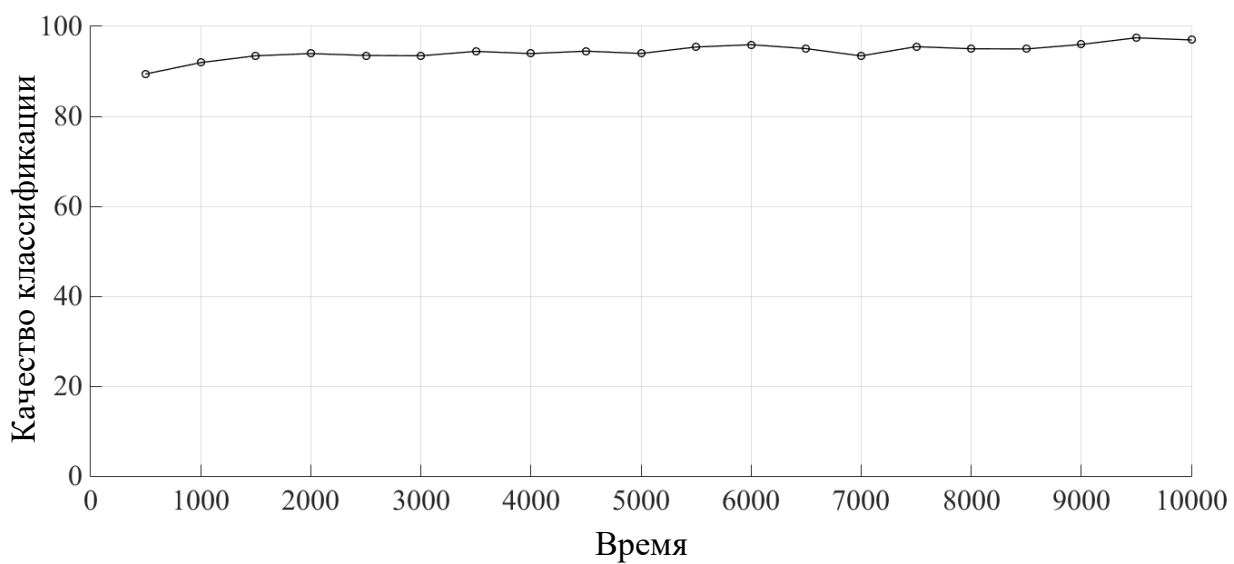


Рисунок 3.5. Эффективность классификации по интенсивности марковских случайных потоков событий в первом эксперименте с ФА вида *б*

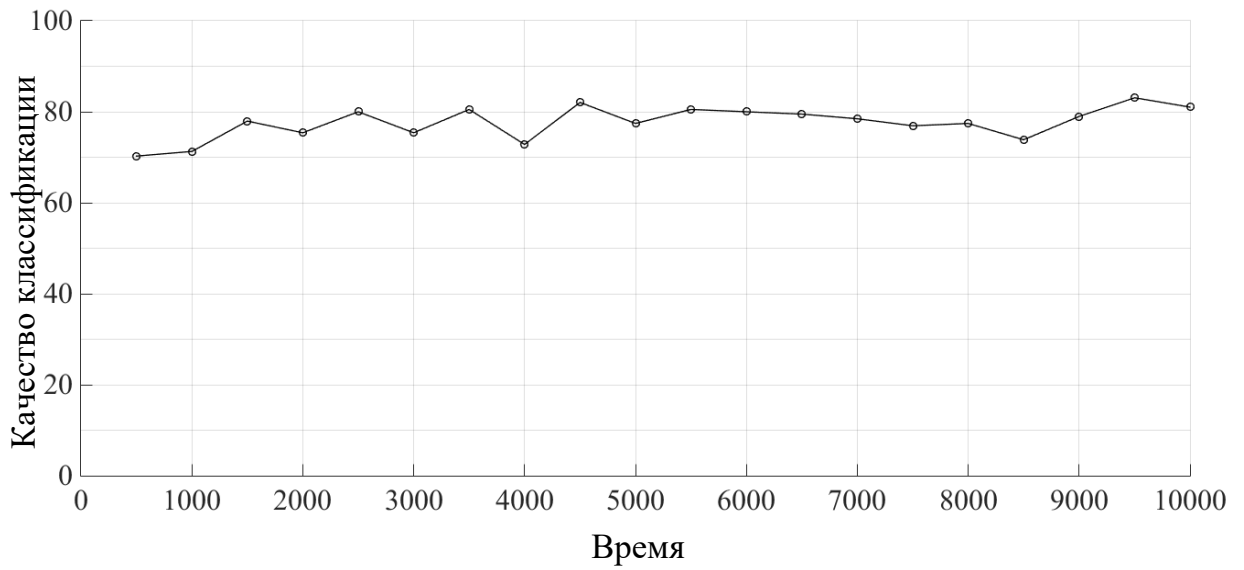


Рисунок 3.6. Эффективность классификации по интенсивности марковских случайных потоков событий в первом эксперименте с ФА вида *б*

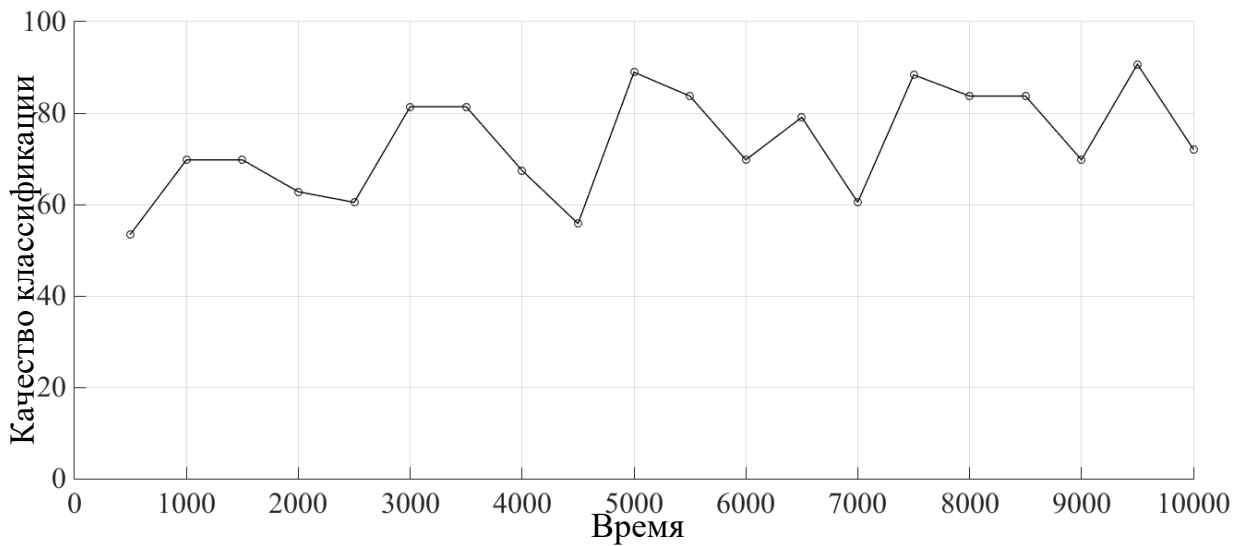


Рисунок 3.7. Эффективность классификации по интенсивности марковских случайных потоков событий в первом эксперименте с ФА вида *г*

Из этого эксперимента следует, что использование ФА вида *б* обеспечивает более высокую эффективность классификации интенсивности марковских случайных потоков событий. Время наблюдения не влияет на результаты

классификации, или, ФА вида δ можно приняться во всех времени наблюдения. ННС с ФА вида α тоже показывает хорошие результаты классификации, которые не зависят от времени наблюдения. Результаты с использованием ФА вида ϵ и ФА вида δ довольно похожие. Однако, ННС с ФА вида δ имеет нестабильные результаты классификации. Эти два вида ФА можно рекомендовать для использования больших времени наблюдения.

В рисунке 3.8 легко заметить, что разница между количеством процессов с интенсивностями λ_1 , λ_2 в первом эксперименте с использованием ФА вида α и δ количеством процессов с интенсивностями λ_1 , λ_2 в наборе данных не значительна.

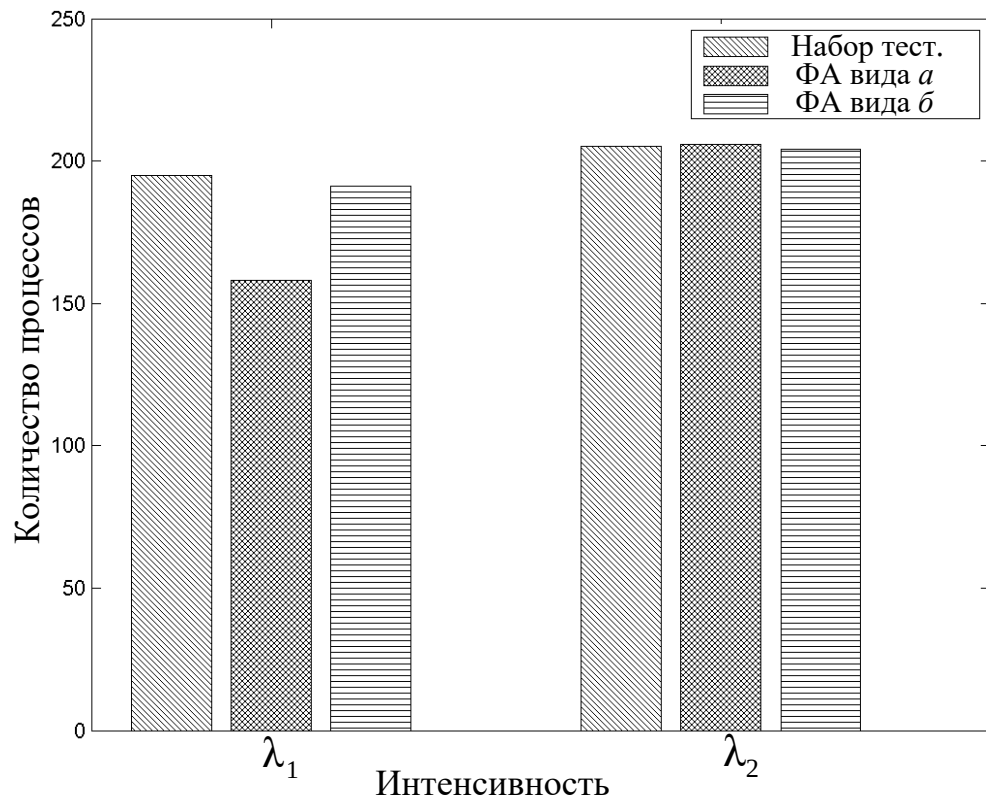


Рисунок 3.8. Количество процессов с интенсивностями λ_1 , λ_2 в момент времени $t = 500$ в наборе данных тестирования и количество правильных решений при определении параметров процессов в первом эксперименте

Во втором эксперименте набор данных для тестирования состояли из 400 случайных процессов с параметром $\lambda_1 = 0,25$ и 400 случайных процессов с

параметром $\lambda_2 = 0,75$. Эффективность классификации этих процессов по интенсивности λ_1, λ_2 во втором эксперименте практически равна 100%.

В третьем эксперименте наборы данных для тестирования состояли из 400 случайных процессов с параметром $\lambda_1 = 0,25$ и 400 случайных процессов с параметром $\lambda_2 = 0,75$. Эффективность классификации этих процессов по интенсивности λ_1 , в третьем эксперименте также практически равна 100%.

На рисунке 3.9 представлена зависимость качества кластеризации от количества нейронов в скрытом слое в момент времени $t = 10000$ с использованием ФА вида б в четвертом эксперименте.

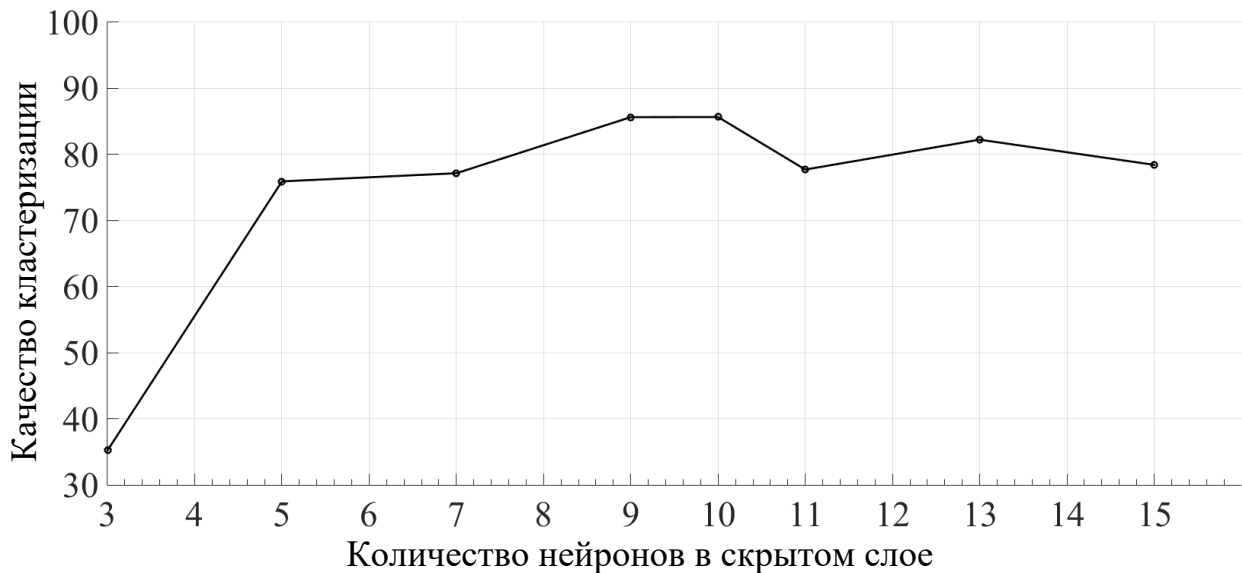


Рисунок 3.9. Зависимость качества кластеризации от количества нейронов в скрытом слое

Из рисунка 3.9 следует, что использование 10 нейронов в скрытом слое обеспечивает самое высокое качество кластеризации. Поэтому в четвертом эксперименте, ННС использует с 10 нейронов с нечеткими ФА, виды которых для нечетких нейронных сетей второго типа представлены в [77, 78] в скрытом слое.

В четвертом эксперименте, результаты классификации с помощью ННС процессов по интенсивностям λ_1, λ_2 представлены на рисунке 3.10, 3.11, 3.12, 3.13.

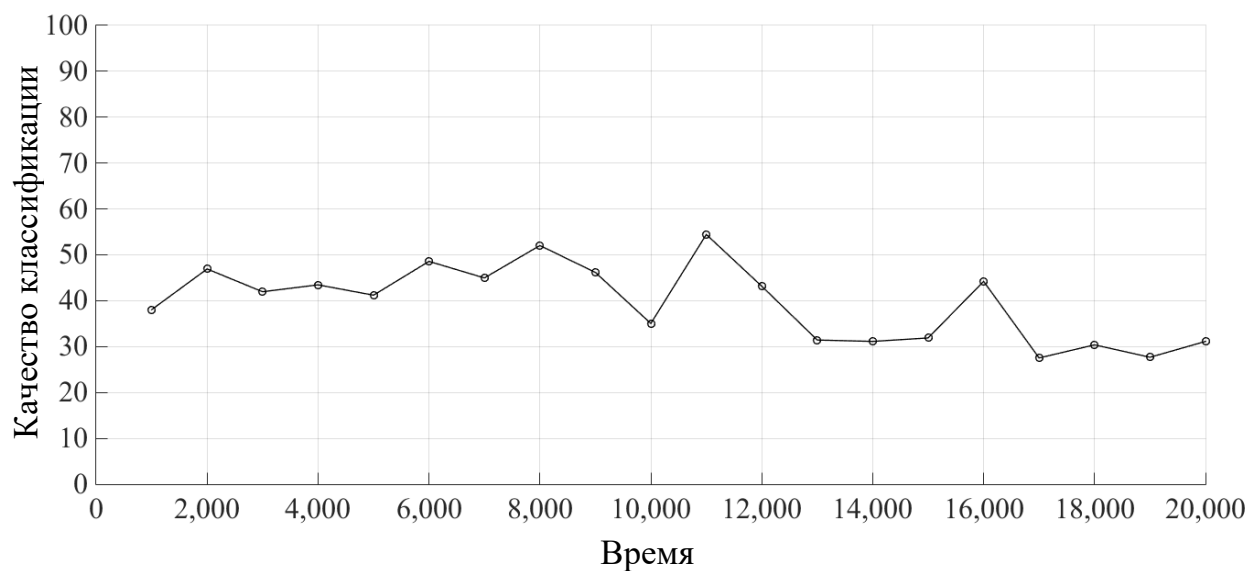


Рисунок 3.10. Эффективность классификации по интенсивности марковских случайных потоков событий в четвертом эксперименте с ФА вида *a*

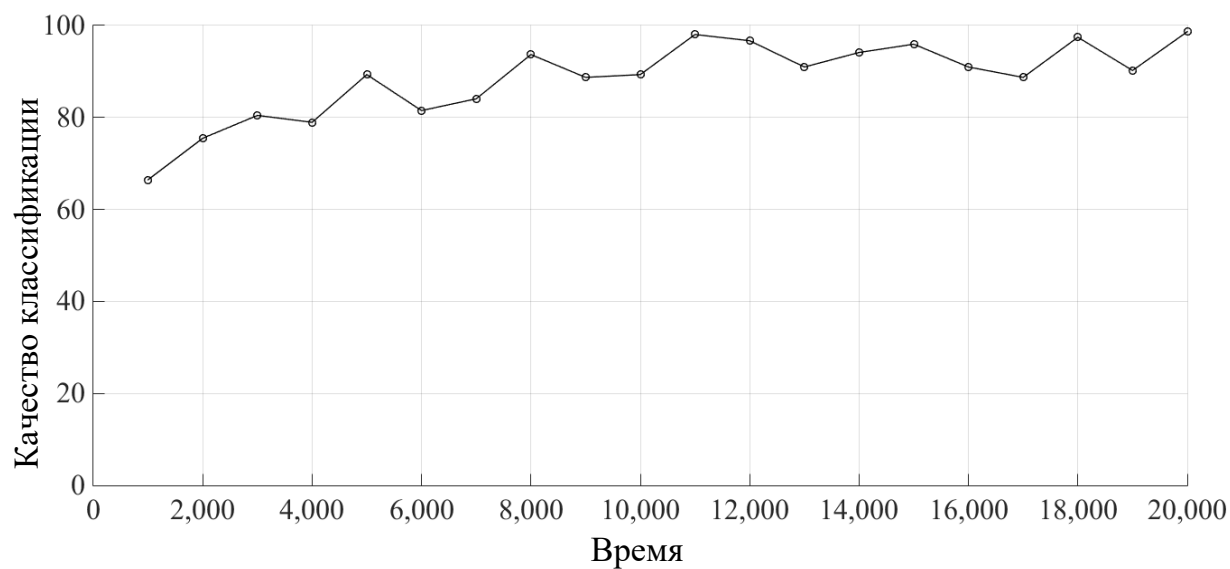


Рисунок 3.11. Эффективность классификации по интенсивности марковских случайных потоков событий в четвертом эксперименте с ФА вида *b*

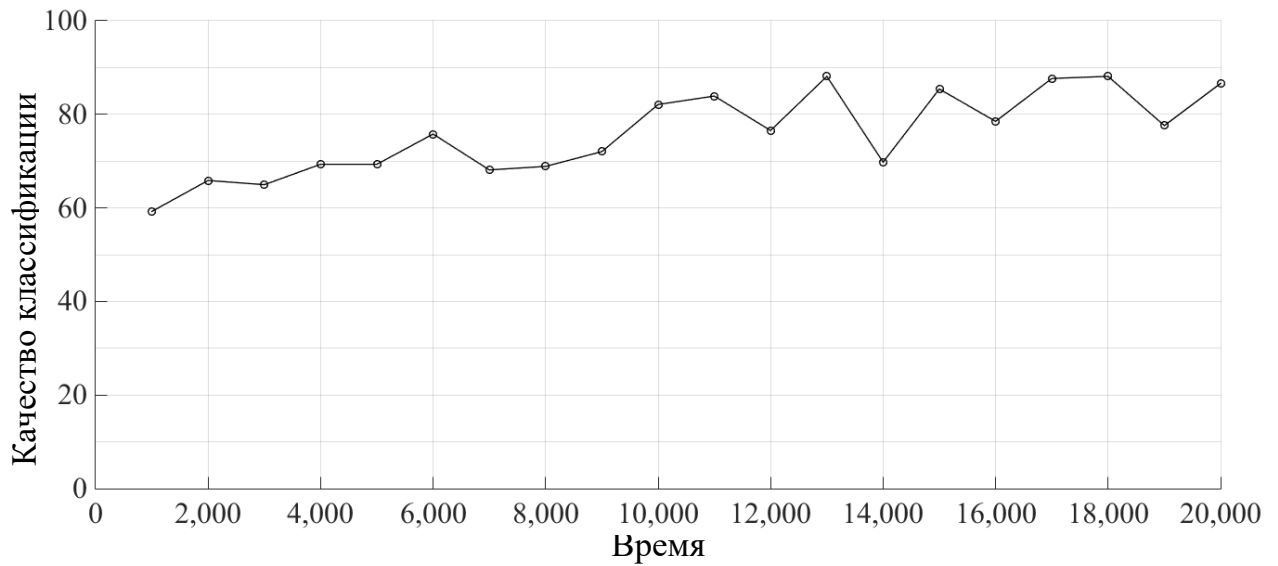


Рисунок 3.12. Эффективность классификации по интенсивности марковских случайных потоков событий в четвертом эксперименте с ФА вида v

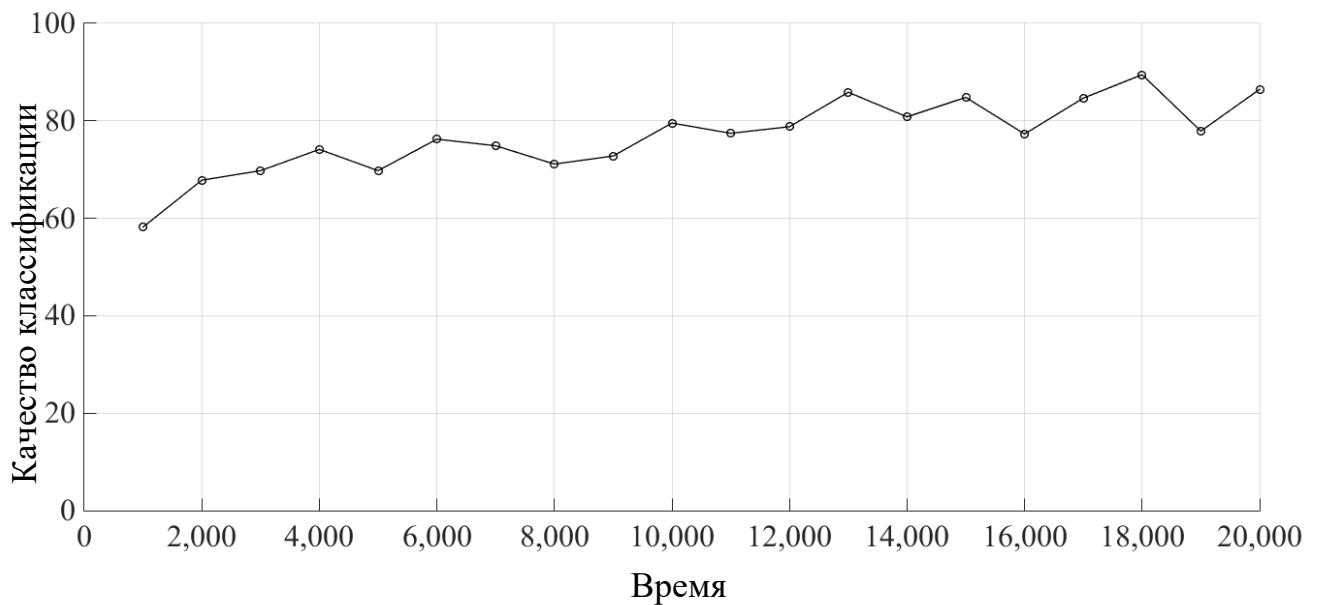


Рисунок 3.13. Эффективность классификации по интенсивности марковских случайных потоков событий в четвертом эксперименте с ФА вида z

Так же, как в первом эксперименте, в четвертом эксперименте, ННС ФА вида b имеет самые лучшие результаты. Однако, в этом эксперименте, время наблюдения

влияет на эффективность классификации: чем больше время наблюдения, тем лучше результаты классификации.

3.5 Выводы по третьей главе

1. Поставлена задача исследования возможности применения ННС 2-го типа для классификации интенсивности потоков событий на основе теории марковских случайных процессов.

2. Создана структура ННС 2-го типа в виде многослойной однонаправленной сети, скрытый слой которой состоит из 10 нейронов с нечеткими ФА. Нечеткие ФА определены через параметры μ и σ - центр и ширина функций активации.

3. Процесс обучения ННС 2-го типа состоит в выборе трех наборов параметров: для каждого из нейронов ННС и матрицы выходных весов между нейронами ННС и выходными узлами. Алгоритм кластеризации К-средних используется для определения центра μ . Ширина σ устанавливается на среднее расстояние между данными соответствующего кластера. Определение весов выходного слоя ННС выполняется с использованием метода псевдообращения.

4. Анализ результатов экспериментов показывает что, разрабатываемые нами модели нечетких нейронных сетей вполне успешно решают задачи классификации по интенсивности марковских случайных потоков событий. При этом на всех тестовых наборах ННС с ФА вида $\tilde{b}$ обеспечивает лучший результат во всех экспериментах. Заметим, что в известных нам исследованиях НС не ставились и не решались подобные задачи. В научной литературе по теорий массового обслуживания [72] теория марковских случайных процессов применяется для определения параметров интенсивности потоков событий. Очевидно, что эти задачи отличаются от решаемых нами задач классификации.

4 НЕЙРО-НЕЧЕТКАЯ КЛАССИФИКАЦИЯ ОБЪЕКТОВ И ИХ СОСТОЯНИЙ

Нейронная сеть представляет собой сложные отношения между входами и выходами [91, 92]. На основе объединения результатов теории нейронных сетей и теории нечетких множеств возможно создание оригинальных интеллектуальных систем [92].

Нечеткие системы, построенные на нечетких правилах, успешно применяются для различных задач классификации [93-99]. Эти системы зависят от лингвистических правил, которые предоставляются экспертами, или правил, извлекаемых из набора обучающих данных различными методами, такими как методы кластеризации. Однако, нечеткие системы, как правило, не учатся и не приспособляются [100], в то время как нейронные сети обладают способностью к самообучению и самоорганизации. По этим причинам вполне логично создание нейро-нечетких систем, которые представляют собой комбинацию систем нечеткого вывода и нейронной сети [101-105]. Нейро-нечеткая система представляет собой нечеткую нейронную сеть, которая включает в себя систему нечетких выводов, преодолевает некоторые ограничения нейронных сетей и расширяет пределы нечетких систем. Нейро-нечеткие классификаторы определяют распределения классов и показывают отношения ввода-вывода [100-107], тогда как нечеткие системы описывают системы с использованием естественного языка. Нейронные сети используются для настройки или обучения системных параметров в нейро-нечетких приложениях.

В настоящей главе решаются задачи автоматической классификации объектов и их состояний. Общий подход к задачам автоматической классификации объектов изложен в [108]. Согласно этому подходу эмпирические данные представляются в виде матрицы данных Z , строки которой соответствуют различным наблюдаемым объектам, а столбцы – параметрам, описывающим состояние каждого объекта. Если

параметры имеют различный физический смысл, то осуществляется преобразование матрицы данных Z в стандартизованную матрицу данных X . Формулы преобразования Z в X приведены в [94]. Матрице данных ставится в соответствие корреляционная матрица R , элементами которой являются выборочные коэффициенты корреляции r_{ij} , пропорциональные скалярному произведению двух векторов-столбцов матрицы данных: скалярное произведение i -го вектор-столбца и j -го вектор-столбца. Коэффициент корреляции является удобным показателем «близости» или «связи» между параметрами. На этой основе разработаны многие структурные методы обработки эмпирических данных, в том числе методы автоматической классификации объектов, образов и изображений [108]. Задача классификации объектов трактуется как геометрическая задача о выделении в пространстве классификационных признаков «плотных» скоплений точек. Для её решения необходимо задание правила отнесения любого объекта к тому или иному классу. Подобные задачи сформулированы и решены в многочисленных отечественных и зарубежных публикациях. Однако поиск эффективных решений задач классификации объектов и их состояний продолжается. В настоящей работе для решения задач классификации объектов и их состояний разрабатывается и исследуется нечеткая НС, базирующаяся на нечетких ФА, предложенных нами выше в главе 1 (рисунок 1.13), и алгоритме масштабированного сопряженного градиента (МСГ) [109].

В известных нечетких нейронных системах – нейро-нечетких классификаторах (ННК) каждый узел сети описывается параметризованной колоколообразной функцией активации (ФА) или гауссовой ФА. В настоящей главе предлагаются новые модели нечетких НС (ННС), использующих импульсные нечеткие ФА нейронов ННС. Поставим задачу разработки ННК на основе этих нейронов, обучения ННК алгоритмом МСГ и исследования возможности ННК для решения задач классификации объектов и их состояний. Для тестирования ННК выберем

классические наборы данных: ирисы Фишера и шкалы баланса. Ирисы Фишера — это набор данных для задачи классификации, на примере которого Рональд Фишер в 1936 году продемонстрировал эффективность разработанного им метода дискриминантного анализа [110]. Этот набор данных стал уже классическим и часто используется для иллюстрации эффективности работы различных статистических алгоритмов. Шкалы баланса также являются классическим набором данных для задачи классификации. Этот набор данных предложен Зиглером [111] и описывает процесс взвешивания с использованием шкалы баланса. Исследование проведено с помощью средств компьютерного моделирования MATLAB [112].

4.1 Нейро-нечеткий классификатор

Задача нечеткой классификации - это задача разбиения пространства классификационных признаков на нечеткие классы. Можно описать пространство классификационных признаков с нечеткими областями и управлять каждой областью с помощью нечетких правил [113]. Типичное нечеткое правило классификации P_i , которое демонстрирует связь между пространством входных данных и классами, выглядит следующим образом:

$$P_i: \text{если } x_{s1} - A_{i1} \text{ и } \dots x_{sj} - A_{ij} \dots \text{ и } x_{sn} - A_{in}, \text{ то класс} - C_k,$$

где x_{sj} представляет j -й входной элемент s -ой выборки; A_{ij} обозначает нечеткое множество j -го входного элемента в i -м правиле; C_k представляет собой метку k -го класса. A_{ij} определяется соответствующей функцией принадлежности.

Для достижения этой цели нечеткий параметр функции принадлежности [93] и нечеткий кластер должны быть оптимизированы. Когда правила нечеткой классификации строятся как сеть, то их параметры могут быть адаптированы к нейронной сети. В результате системы с нечеткой классификацией и нейронные сети могут быть объединены с сохранением их положительных свойств. Комбинированная система называется нейро-нечетким классификатором, который

представляет собой адаптивную сеть с несколькими входами и несколькими выходами [114].

ННК представляет собой многослойную однонаправленную сеть, состоящую из следующих слоев: входной слой, слой нечетких ФА, слой фаззификации, слой дефаззификации, слой нормализации и выходной слой. Классификатор имеет несколько входов и несколько выходов. На рисунке 4.1 изображен ННК с двумя классификационными признаками и тремя классами. Каждый вход определяется тремя лингвистическими переменными. Таким образом, существует девять нечетких правил.

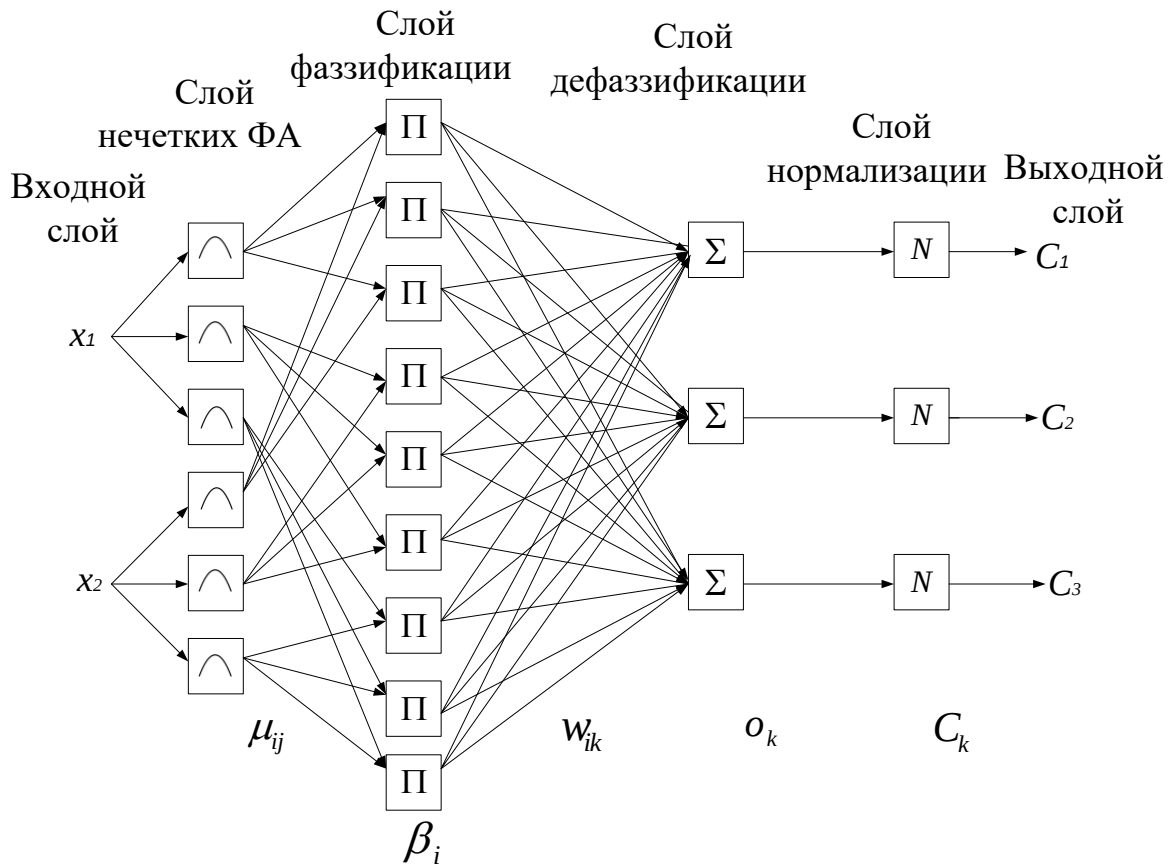


Рисунок 4.1. Структура нейро-нечеткого классификатора

Слой нечетких ФА: функция активации для каждого входа идентифицируется в этом слое. В данной главе используется четыре вида нечетких ФА. Внешний вид этих ФА представлен на рисунке 1.13 а, б, в, г, а также на рисунке 3.2 а, б, в, г.

Набор точек $\langle A, B, C \rangle$ (см. рисунок 3.2) являются характеристическими

точками ФА. Задаем $AB = BC = \sigma$; и $B = \mu$. Пусть $a = (\mu - \sigma)$, $k_1 = 0$, $k_2 = 1$ если $\mu - \sigma \leq x \leq \mu$ и $a = (\mu + \sigma)$, $k_1 = 1$, $k_2 = 0$ если $\mu \leq x \leq (\mu + \sigma)$, тогда получим четыре ФА соответственно [115]:

$$a) \begin{cases} f_{ij}(x_{sj}) = \frac{1}{\sigma_{ij}^2} x_{sj}^2 + \frac{-2a_{ij}}{\sigma_{ij}^2} x_{sj} + \frac{a_{ij}^2}{\sigma_{ij}^2}, \\ f_{ij}(x_{sj}) = 0, \text{ иначе,} \end{cases}$$

$$б) \begin{cases} f_{ij}(x_{sj}) = \frac{-1}{\sigma_{ij}^2} x_{sj}^2 + \frac{2\mu_{ij}}{\sigma_{ij}^2} x_{sj} + \frac{a_{ij}^2 - 2\mu_{ij}a_{ij}}{\sigma_{ij}^2}, \\ f_{ij}(x_{sj}) = 0, \text{ иначе,} \end{cases}$$

в)

$$\begin{cases} f_{ij}(x_{sj}) = k_1 \left(\frac{1}{\sigma_{ij}^2} x_{sj}^2 + \frac{-2a_{ij}}{\sigma_{ij}^2} x_{sj} + \frac{a_{ij}^2}{\sigma_{ij}^2} \right) + k_2 \left(\frac{-1}{\sigma_{ij}^2} x_{sj}^2 + \frac{2\mu_{ij}}{\sigma_{ij}^2} x_{sj} + \frac{a_{ij}^2 - 2\mu_{ij}a_{ij}}{\sigma_{ij}^2} \right), \\ f_{ij}(x_{sj}) = 0, \text{ иначе,} \end{cases}$$

г)

$$\begin{cases} f_{ij}(x_{sj}) = k_1 \left(\frac{-1}{\sigma_{ij}^2} x_{sj}^2 + \frac{2\mu_{ij}}{\sigma_{ij}^2} x_{sj} + \frac{a_{ij}^2 - 2\mu_{ij}a_{ij}}{\sigma_{ij}^2} \right) + k_2 \left(\frac{1}{\sigma_{ij}^2} x_{sj}^2 + \frac{-2a_{ij}}{\sigma_{ij}^2} x_{sj} + \frac{a_{ij}^2}{\sigma_{ij}^2} \right), \\ f_{ij}(x_{sj}) = 0, \text{ иначе.} \end{cases}$$

Слой фаззификации: каждый узел в этом слое генерирует сигнал, соответствующий степени выполнения нечеткого правила для выборки. Этот сигнал можно назвать степенью возбуждения (the firing strength) нечеткого правила в отношении объекта, подлежащего классификации. Степень возбуждения по этому правилу заключается в следующем:

$$\beta_{is} = \prod_{j=1}^N f_{ij}(x_{sj}), \quad (4.1)$$

где N - количество классификационных признаков.

Слой дефаззификации: в этом слое вычисляются веса выходов; каждое правило влияет на каждый класс в соответствии с их весами. Если правило управляет определенной областью класса, вес между этим выходом правила и конкретным классом будет больше, чем другие веса. В противном случае веса классов малы. Вес выхода для s -й выборки, относящейся к k -му классу, вычисляется следующим образом:

$$o_{sk} = \sum_{i=1}^U \beta_{is} w_{ik}, \quad (4.2)$$

где w_{ik} обозначает степень принадлежности к k -му классу, который контролируется i -м правилом и U представляет собой количество правил; o_{sk} обозначает весовой выход для s -й выборки, принадлежащего k -му классу.

Слой нормализации: выходы сети должны быть нормализованы, так как суммирование весов может быть больше чем 1, в некоторых случаях

$$h_{sk} = \frac{o_{sk}}{\sum_{l=1}^K o_{sl}} = \frac{o_{sk}}{\delta_s}, \quad \delta_s = \sum_{l=1}^K o_{sl}, \quad (4.3)$$

где h_{sk} представляет нормированное значение s -ого выборки, которая относится к k -му классу; и K - количество классов. После этого метка класса s -й выборки определяется максимальным значением как

$$C_s = \max_{k=1,2,\dots,K} \{h_{sk}\}, \quad (4.4)$$

где C_s обозначает метку класса для s -й выборки.

4.2 Обучение нейро-нечеткого классификатора

Антецедентные параметры сети $\theta = \{S_{M \times N}, C_{M \times N}, W_{M \times K}\}$ [116] могут быть адаптированы любым методом оптимизации S и C – матриц, содержащих значения σ и μ , соответственно; W представляет весовую матрицу соединений от слоя

фаззификации до слоя дефаззификации; M , N , K - количество правил, признаков и классов соответственно; метод кластеризации K - средних используются для получения исходных параметров матриц S и C . Метод кластеризации K - средних предназначен для разделения входного пространства признаков на несколько кластеров, в которых каждая точка данных принадлежит кластеру с ближайшим средним значением. Это приводит к разделению пространства данных. Для данного набора данных этот метод может оценить количество кластеров и центров кластеров.

На рисунке 4.2 показано пространство признаков с двумя входами $\{x_1, x_2\}$. Предположим, что каждый вход делится на три нечетких множества с помощью метода кластеризации K - средних. Каждый нечеткий набор характеризуется соответствующей функцией принадлежности; В результате каждый вход имеет три функции принадлежности. Правило нечеткой классификации описывает связь между входным пространством объектов и классами. Формирование нечетких правил если-то показано на рисунке 2. Каждый вход представлен в виде трех функций принадлежности; Таким образом, у нас есть девять нечетких правил.

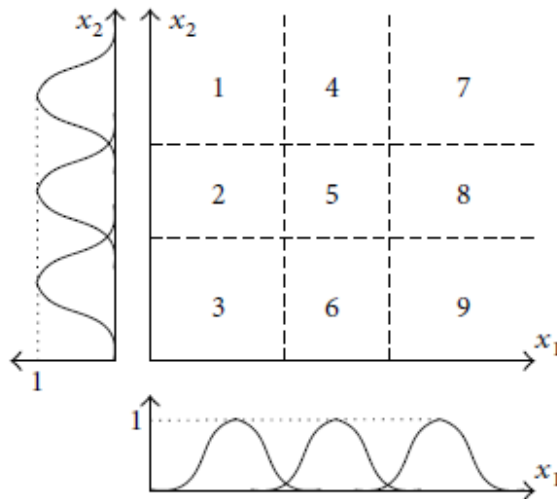


Рисунок 4.2. Разделение функционального пространства с двумя входами и тремя функциями принадлежности для каждого входа

Метод МСГ используется для адаптации параметров сети. МСГ - это алгоритм сопряженного градиента второго порядка [114], который минимизирует целевые

функции нескольких переменных. Теоретическая основа этого алгоритма была доказана Моллером. Этот алгоритм относится к методам первого порядка, использующих первые производные. МСГ использует механизм масштабирования размера шага, который позволяет избежать трудоемкого поиска линий на каждую итерацию обучения, что повышает его быстродействие. Основанный на Моллере [109], метод МСГ показывает суперлинейную конвергенцию [107] при решении многих задач.

Параметр w_{ik} в формуле (4.2) также можно адаптировать с помощью метода МСГ. Однако при обучении w_{ik} может быть больше 1. В этих случаях значения весов могут быть потеряны для кластеров из этого же класса. По этой причине либо w_{ik} должен быть ограничен, либо w_{ik} определяется из отношения числа выборки k -го класса в i -й области нечеткого правила относительно общего числа выборок k -го класса:

$$w_{ik} = \frac{D_i}{D_k}, \quad (4.5)$$

где D_i - число выборок k -го класса, которые принадлежат к i -й нечеткой области правил и D_k - число всех выборок k -го класса.

Функция стоимости, используемая в методе МСГ, определяется из наименьших средних квадратов по разности цели и вычисленного значения класса.

$$E = \frac{1}{N} \sum_{s=1}^N E_s, \quad (4.6)$$

$$E_s = \frac{1}{2} \sum_{k=1}^K (t_{sk} - h_{sk})^2, \quad (4.7)$$

где N – количество выборок, t_{sk} и h_{sk} являются целевыми и вычисленными значениями s -й выборки, принадлежащей k -му классу, соответственно. Если s -й выборки принадлежит k -му классу, то целевое значение t_{sk} устанавливается равным

1, а в противном случае - равным 0.

Цель алгоритма МСГ - найти оптимальный или близкий к оптимальному параметр θ^* из функции стоимости $E(\theta)$. В алгоритме МСГ следующий ближайший вектор обновления, θ_{t+1} , к текущему вектору θ_t определяется как

$$\theta_{t+1} = \theta_t - g_t H^{t-1} \quad (4.8)$$

где $g_t = E'(\theta_t)$ и $H_t = E''(\theta_t)$ являются вектором градиента и гессенской матрицей $E(\theta_t)$ соответственно. Произведение $-g_t H^{t-1}$ называется шагом Ньютона; направление Ньютона обозначено знаком минус. Если матрица Гессеана положительно определена, а $E(\theta_{t+1})$ квадратична, метод Ньютона непосредственно достигает локального минимума за один шаг [93]. Однако для достижения локального минимума обычно требуется больше итераций. Моллер [109] ввел вектор временных параметров $\theta_{m,t}$, который находится между θ_{t+1} и θ_t и определяется как

$$\theta_{m,t} = \theta_t + \gamma_t d_t, \quad 0 < \gamma_t \ll 1, \quad (4.9)$$

где γ_t - размер короткого шага и $d_t = -g_t$ - сопряженный вектор направления вектора временного параметра на t -й итерации. Фактическое обновление параметра рассчитывается как

$$\theta_{t+1} = \theta_t + \alpha_t d_t, \quad (4.10)$$

где θ_{t+1} - вектор обновления следующего параметра; θ_t - вектор текущего параметра; и α_t - фактический размер шага обновления параметра и рассчитывается следующим образом:

$$\alpha_t = \frac{-d_t^T E'(\theta_t)}{d_t^T s_t}, \quad s_t = E''(\theta_t) d_t \approx \frac{E'(\theta_{m,t}) - E'(\theta_t)}{\gamma_t}, \quad (4.11)$$

где s_t - информация второго порядка, а α_t обозначает базовый размер длинного шага. Для расчета α_t , информация второго порядка s_t должна быть получена из градиентов первого порядка.

В алгоритме МСГ два разных градиента вектора параметров вычисляются на любой итерации. Градиент d_t вектора параметров весов $\theta_{m,t}$ рассчитывается с использованием размера γ_t короткого шага; и градиент фактического обновления параметра θ_{t+1} рассчитывается с использованием размера длинного шага α_t , который получается с использованием $\theta_{m,t}$.

Частную производную от E , относящуюся к μ_{ij} , можно вычислить:

$$\frac{\partial E}{\partial \mu_{ij}} = \sum_{s=1}^N \frac{\partial E}{\partial E_s} \left(\sum_{k=1}^K \frac{\partial E_s}{\partial h_{sk}} \cdot \frac{\partial h_{sk}}{\partial o_{sk}} \cdot \frac{\partial o_{sk}}{\partial \beta_{is}} \cdot \frac{\partial \beta_{is}}{\partial f_{ijs}} \cdot \frac{\partial f_{ijs}}{\partial \mu_{ij}} \right). \quad (4.12)$$

Для функции активации вида b :

$$\begin{cases} f_{ij}(x_{sj}) = \frac{-1}{\sigma_{ij}^2} x_{sj}^2 + \frac{2\mu_{ij}}{\sigma_{ij}^2} x_{sj} + \frac{a_{ij}^2 - 2\mu_{ij}a_{ij}}{\sigma_{ij}^2}, \\ f_{ij}(x_{sj}) = 0, \text{ иначе,} \end{cases} \quad (4.13)$$

где $a = (\mu - \sigma)$ если $\mu - \sigma \leq x \leq \mu$ и $a = (\mu + \sigma)$ если $\mu \leq x \leq (\mu + \sigma)$; $f_{ij}(x_{sj})$ представляет собой степень принадлежности i -го правила и j -й признак; x_{sj} представляет собой s -й выборку и j -й признак входной матрицы.

Из (4.1) - (4.13) получим:

$$\frac{\partial E}{\partial \mu_{ij}} = \frac{1}{N} \sum_{s=1}^N \sum_{k=1}^K \left((h_{sk} - t_{sk}) \left(\frac{1 - h_{sk}}{\delta_s} \right) w_{ik} \beta_{is} \frac{2x_{sj} - 2\mu_{ij}}{-x_{sj}^2 + 2\mu_{ij}x_{sj} + \sigma_{ij}^2 - \mu_{ij}^2} \right).$$

Аналогично

$$\frac{\partial E}{\partial \sigma_{ij}} = \frac{1}{N} \sum_{s=1}^N \sum_{k=1}^K \left((h_{sk} - t_{sk}) \left(\frac{1 - h_{sk}}{\delta_s} \right) w_{ik} \beta_{is} \frac{2(x_{sj} - 2\mu_{ij})^2}{\sigma_{ij}(-x_{sj}^2 + 2\mu_{ij}x_{sj} + \sigma_{ij}^2 - \mu_{ij}^2)} \right).$$

Аналогично получаются частные производные от E , относящиеся к μ_{ij} для остальных ФА.

ННС обучается методом оптимизации МСГ с использованием частных

производных E -отношения по приведенным выше параметрам.

4.3 Тестирование нейро-нечеткого классификатора

Разработанный выше ННК с нашими функциями активации и алгоритмом МСГ исследован на задаче классификации набора данных ирисы Фишера для нахождения отображения между четырьмя входными переменными (длина наружной, ширина наружной, длина внутренней и ширина внутренней) и тремя классами (Setosa, Versicolor и Virginica) [110].

В наборе данных имеется 150 образцов, 75 из них используются как данные для обучения ННК, а остальные 75 для тестирования ННК. ННК обучается на наборе учебных данных за 100 эпох. Затем скорректированная сеть ННК исследуется на наборе данных тестирования. Желаемые и рассчитанные выходы соответствуют всем 75 тестовым данным.

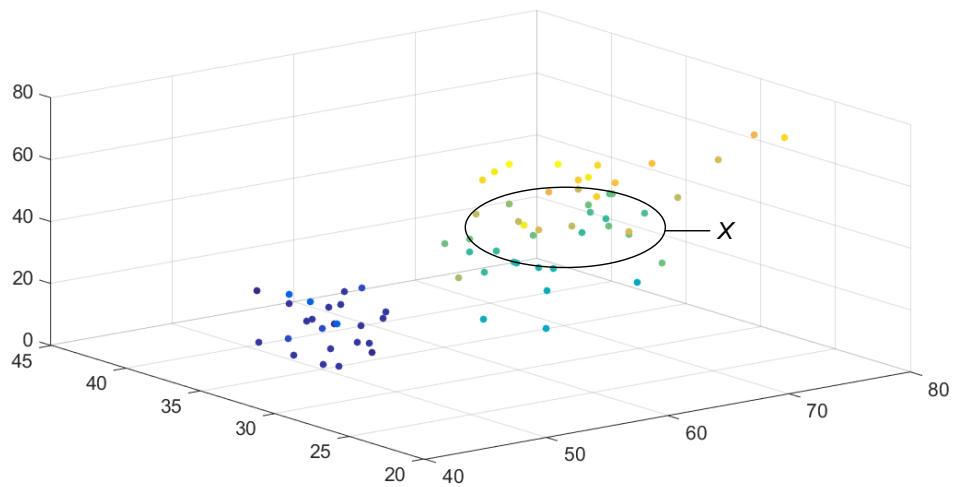


Рисунок 4.3. Набор данных ирисы Фишера после использования метода кластеризации К – средних

После использования метода кластеризации К – средних, набор данных ирисы Фишера имеет довольно четкую классификацию. Однако, в области X мы не можем классифицировать данные.

Эксперимент с ННК по классификации состояний объекта проведен на наборе данных шкалы баланса, который получен из UC Irvine Machine Learning Repository и

был сгенерирован для моделирования результатов психологических экспериментов, проведенных Зиглером [111]. Примеры в наборе данных соответствуют одному из трех классов: шкала справа, слева или сбалансирована. Набор данных состоит из 625 примеров и 5 атрибутов: имя класса (слева, баланс, справа), левый вес (1, 2, 3, 4, 5), левое расстояние (1, 2, 3, 4, 5), правый вес (1, 2, 3, 4, 5) и правое расстояние (1, 2, 3, 4, 5). Например, элемент набора данных «2, 5, 2, 1» (левый вес = 2, левое расстояние = 5, правый вес = 2, правое расстояние = 1) следует классифицировать по классу «слева». Для обучения используем 313 образцов, а для тестирования – 312.

Результаты тестирования ННК представлены в таблице 4.1. Эффективность работы ННК при идентификации теста будем оценивать процентным отношением между количеством правильных результатов тестирования и общим количеством данных тестирования. Чем больше процентное отношение, тем выше эффективность работы ННК на этом тесте. Данное исследование характеризует эффективность работы ННК, т.е. эффективность классификации. СКО характеризует точность результатов экспериментов с ННК. Чем меньше эта ошибка, тем выше точность работы ННК на этом тесте. То есть это исследование характеризует точность классификации.

Таблица 4.1. Результаты классификации набора данных ННК с различными ФА

Тип ФА	Эффективность ННК на тесте (%)		СКО	
	Ирисы Фишера	Шкала баланса	Ирис Фишера	Шкала баланса
Гауссоида	96	88.78	$2.297 \cdot 10^{-7}$	0.13
Вида <i>a</i>	96	87.82	$5.031 \cdot 10^{-6}$	0.12
Вида <i>б</i>	97.33	89.10	$6.361 \cdot 10^{-4}$	0.15
Вида <i>в</i>	97.33	77.88	0.0418	0.31
Вида <i>г</i>	96	60.25	0.041	0.61

Из таблицы 4.1 следует, что результаты классификации с использованием предлагаемых нами ФА не превосходят результат классификации с использованием гауссовой ФА (СКО ННК с гауссовой ФА на тесте ирисы Фишера самая малая), однако использование ННК с ФА вида a и b вполне приемлемо на разных тестах. Эффективность ННК с ФА вида b и c на тесте ирисы Фишера самая высокая.

Проведем сравнительный анализ результатов классификации ННК с предлагаемыми ФА и классификации с использованием радиальной базисной нейронной сети (РБНС). РБНС является популярной альтернативой известному многослойному персептрону, поскольку имеет простую структуру и более высокую скорость обучения [117]. РБНС использует радиальные базисные функции в качестве функций активации. В таблице 4.2 представлены результаты эксперимента по классификации предлагаемым ННК и РБНС для выбранных тестов ирисы Фишера и шкалы баланса.

Таблица 4.2 Результаты классификации ННК и РБНС

Метод использования	Эффективность на тестах (%)	
	Ирисы Фишера	Шкала баланса
ННК	96 – 97.33	60.25 – 89.10
РБНС	65.33	54.91

Из таблицы 4.2 следует, что предлагаемый ННК эффективнее РБНС в решаемой задаче классификации.

На рисунке 4.4 показана динамика СКО ННК на тесте ирисы Фишера в зависимости от учебных эпох.

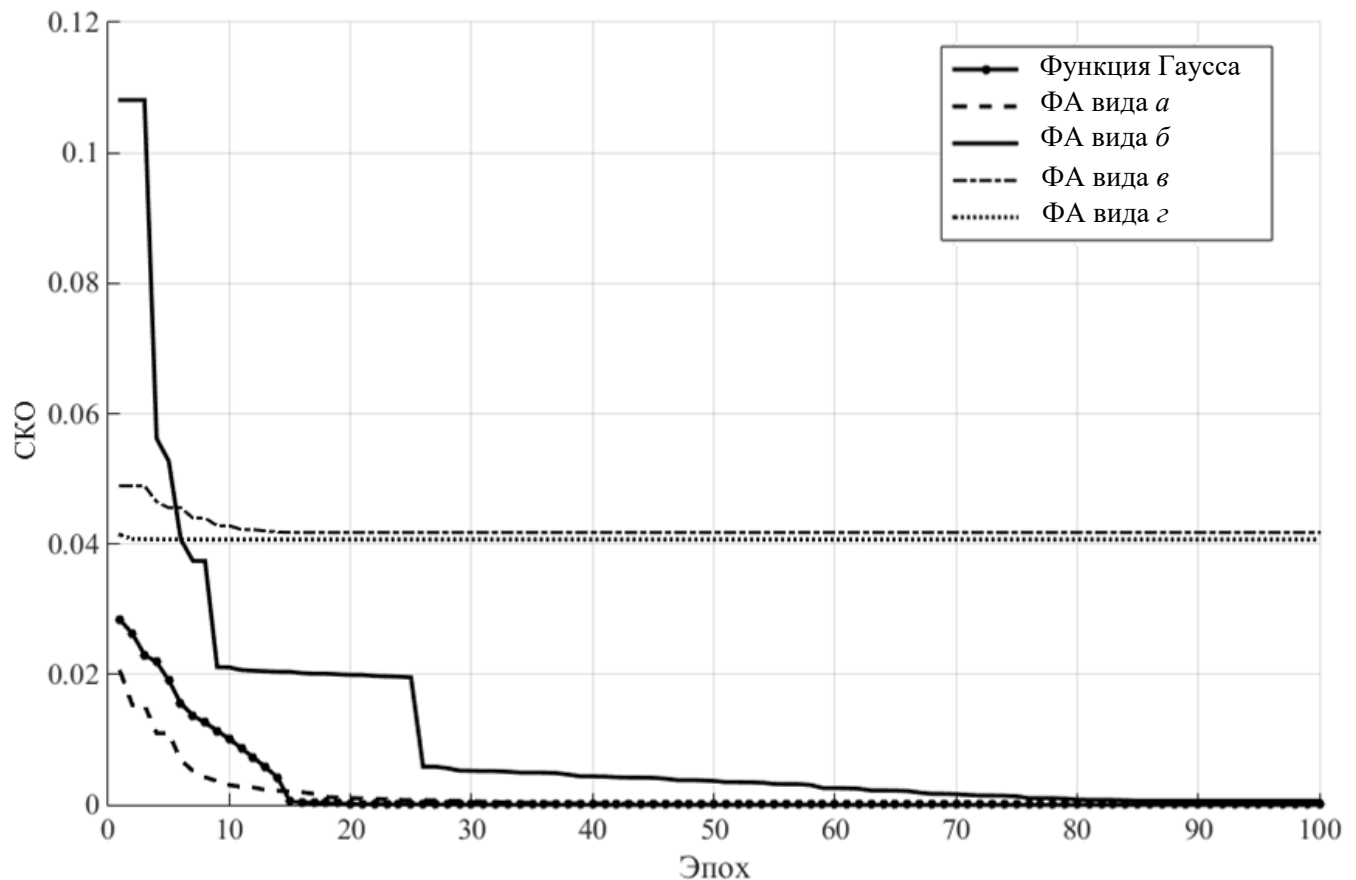


Рисунок 4.4. СКО ННК на тесте ирисы Фишера

Из рисунка 4.4 следует, что обучение ННК с ФА ϵ и ζ на основе алгоритма СМГ не эффективно. СКО (точность работы ННК с этими ФА) почти не изменяются и слишком велики, поэтому для ННК этого типа необходим поиск других методов обучения. Относительно обучения ННК с ФА a и b на основе алгоритма СМГ следует сделать более оптимистичный вывод и утверждать, что ННК данных типов эффективно решают задачи классификации объектов и их состояний.

4.4 Выводы по четвертой главе

1. Разработана структура ННК классификатора, приведено описание функции и структуры слоев ННК.
2. Антецедентные параметры сети определяются методом кластеризации К – средних. Разработан алгоритм МСГ для обучения ННК. На любой итерации вычисляются два разных градиента вектора параметров сети. На этой основе

определяются частные производные функции стоимости ННК.

3. Тестирование разработанных ННК на наборах данных ирисы Фишера и шкалы баланса, их сравнительный анализ с известным ННК с гауссовой ФА и РБНС показали конкурентоспособность двух из четырех разработанных ННК по эффективности и точности классификации. Для двух неконкурентоспособных ННК необходим поиск других методов обучения и, в частности, возможно применение метода Нелдера – Мида [118], который успешно зарекомендовал себя при обучении НС, входящей в структуру системы автоматического управления.

ЗАКЛЮЧЕНИЕ

В диссертации решена актуальная научно-техническая задача алгоритмического обеспечения нейронной сети нового типа с полиномиальными кусочно-непрерывными функциями активации для обнаружения закономерностей в наборах данных. Основные научные и практические результаты, полученные в диссертационной работе, состоят в следующем.

1. Проведён анализ процессов обучения НС для различных типов функций активации ФА НС. Предложено применение в НС новых ФА (ФА вида a , b , v , z), базирующихся на полиномиальных кусочно-непрерывных функциях принадлежности нечетких чисел LR - типа. Новый тип НС с полиномиальными кусочно-непрерывными функциями активации предложено именовать нечеткими НС (ННС) второго типа. В ННС первого типа реализуются нечеткие отношения между нейронами как элементами нейронной сети, т.е. ННС первого типа реализуют системы нечеткого логического вывода методами теории нейронных сетей. В ННС второго типа нечеткость является атрибутом нейрона как элемента нейронной сети. На основе ННС второго типа создано алгоритмическое обеспечение нейронных систем обработки информации для обнаружения закономерностей в наборах данных и решения следующих задач:

- прогнозирование временных рядов со стохастическим трендом;
- определение параметров детерминированного сигнала на фоне белого шума;
- классификация интенсивности марковского случайного потока событий;
- классификация объектов и их состояний.

2. Разработана структура МОНС с использованием метода скользящего окна для прогноза временных рядов с ФА вида a , b , v , z . Процесс обучения реализуется с помощью алгоритма Левенберга - Марквардта обратного распространения. Реальные наборы данных используются для тестирования.

Среднеквадратичная ошибка (СКО) используется для определения лучшей модели МОНС. СКО в МОНС с ФА вида a , b , v , z не значительно отличаются от СКО в обычной МОНС. Эксперименты показывают, что модель НС с ФА вида b , v наиболее эффективна.

3. Выполнена разработка структуры МОНС с ФА вида a , b , v , z для определения амплитуды детерминированного сигнала на фоне белого шума. Алгоритм обратного распространения Левенберга - Марквардта используется для обучения. Синусоидальные сигналы с добавлением к нему аддитивного гауссовского белого шума с отношением сигнал/шум = 0,05 дБ используются в качестве наборов данных для тестирования. Лучшие результаты показали в МОНС с ФА вида a , b , v , z (особенно МОНС с ФА вида v) по сравнению с результатами в обычной МОНС.

4. Исследована возможность применения ННС нового типа для классификации интенсивности потоков событий (марковских случайных процессов). Создана структура ННС второго типа в виде многослойной однонаправленной сети с использованием ФА вида a , b , v , z . Процесс обучения ННС второго типа состоит в выборе трех наборов параметров: центра μ и ширины σ для ФА вида a , b , v , z нейронов ННС и матриц выходных весов между нейронами ННС и выходными узлами ННС. С помощью алгоритма кластеризации К-средних получен набор параметров центра μ . Набор параметров ширины σ определяется по среднему расстоянию между данными соответствующего кластера. Метод псевдообращения матриц используется для определения матрицы весов выходного слоя ННС. Результаты экспериментов доказывают что, разрабатываемые нами модели нечетких нейронных сетей, успешно решают задачи классификации интенсивности марковских случайных потоков событий. ННС с ФА вида b обеспечивает лучший результат во всех экспериментах.

5. Выполнена разработка нейро-нечеткого классификатора (ННК). ННК представляет собой многослойную однонаправленную сеть, состоящую из

следующих слоев: входной слой, слой нечетких ФА, слой фаззификации, слой дефаззификации, слой нормализации и выходной слой. Антецедентные параметры сети определяются методом кластеризации К – средних. Алгоритм масштабированного сопряженного градиента (МСГ) используется для обучения ННК. Для тестирования используются типовые наборы данных: ирисы Фишера и шкала баланса. Проведен сравнительный анализ разработанных ННК с известными ННК: ННК с гауссовой ФА и радиальной базисной нейронной сетью (РБНС). Сравнительный анализ доказывает конкурентоспособность двух из четырех разработанных ННК по эффективности и точности классификации. ННС с ФА вида *a* и *б* показывают лучшие результаты по классификации объектов и их состояний.

Отметим перспективы развития темы диссертации по следующим направлениям:

- возможно использование в качестве ФА НС полиномиальных кусочно – непрерывных функций третьего и более высокого порядка, а также сплайн – функций высоких порядков.

- исследованные нами ННС именуются ННС второго типа, однако представляют также научный интерес ННС третьего типа, в которых нечеткость является как атрибутом нейронов (ФА нейронов), так свойствами отношений (связей) между нейронами (элементами ННС).

СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ

ВНС – Вероятностная нейронная сеть.

ДНС – Детерминированная нейронная сеть.

ЕИ – Естественный интеллект.

ИИ – Искусственный интеллект.

ИНС – Искусственная нейронная сеть.

ЛМОР – Левенберга - Марквардта обратного распространение.

МОНС – Многослойная однонаправленная нейронная сеть.

МП – Многослойный персептрон.

МСТ – Масштабированный сопряженный градиент.

НАНС – Нелинейная авторегрессионная нейронная сеть.

ННК – Нейро-нечеткий классификатор.

ННС – Нечеткая нейронная сеть.

НС – Нейронная сеть.

РБНС – Радиальная базисная нейронная сеть.

РБФ – Радиальная базисная функция.

РБФНС – Радиально-базисная функциональная нейронная сеть.

СКО – Среднеквадратичная ошибка.

ФА – Функция активации.

ФП – Функция принадлежности.

ANFIS – Adaptive Network Based Fuzzy Inference System.

MAP – Markovian Arrival Process.

MC – Markov Chain.

$$\|X\| = \sqrt{\sum_{i=1}^n x_i^2} \text{ – норма матрицы } X.$$

A^{-1} – обратная матрица, $A^{-1}A = \mathbf{I}$, где $\mathbf{I}$ – единичная матрица.

$\mathbf{x}^T = (x_1, x_2, \dots, x_n) = (x_i, i = \overline{1, n})$ – вектор строка, T – символ транспонирования.

$\sup(X)$ – супремум, или точной (наименьшей) верхней гранью (границей) множества X .

СПИСОК ЛИТЕРАТУРЫ

1. McCulloch, W. S. A logical calculus of the ideas imminent in nervous activity / W. S. McCulloch, W. A. Pitts // Bull. Math. Biophys. – 1943. – Vol. 5. – P.115-133
2. Hebb, D. O. The Organization of Behavior / D. O. Hebb. – New York: Wiley, 1949. – 335p.
3. Minsky, M. Perceptrons / M. Minsky, S. Papert. – Cambridge, MIT Press, 1969.
4. Hopfield, J. J. Neural networks and to physical systems with emergent collective computational abilities / J. J. Hopfield // Proc. Natl. Acad. Sci. – 1982. – Vol. 79. – P. 2554-2558.
5. Kohonen, T. Self-organization and Associative Memory / T. Kohonen. – New York: Springer-Verlag, 1984. – 326 p.
6. Rumelhart, D. E. Parallel Distributed Processing: Explorations in the Microstructure of Cognition / D. E. Rumelhart, J. L. McClelland and the PDP Research Group. – Cambridge, MIT Press/Bradford Books, 1986. – 547 p.
7. Спицын, В.Г. Применение искусственных нейронных сетей для обработки информации: Методические указания к лабораторным работам / В.Г. Спицын, Ю.Р. Цой. – Томск: Изд-во ТПУ, 2007. – 32 с.
8. Broomhead, D. Radial basis functions, multi-variable functional interpolation and adaptive networks / D. Broomhead, D. S. Lowe. – Technical report, DTIC Document, 1988. – 39 p.
9. Лабораторная работа №2 по дисциплине «Искусственные нейронные сети» на тему «Нейронные сети с радиальными базисными функциями». Кафедра информатики, Физико-математический факультет, Амурский гуманитарно-педагогический государственный университет, 2007.

10. Nyanteh, Y.D. Application of artificial intelligence to stator winding fault diagnosis in Permanent Magnet Synchronous Machines / Y. D. Nyanteh, S. K. Srivastava, C. S. Edrington, D. A. Cartes // Electric Power Syst. Res. – 2013. – Vol. 103. – P. 201-213.
11. López, M. Application of SOM neural networks to short-term load forecasting: The Spanish electricity market case study / M. López // Electric Power Syst. Res. – 2012. – Vol. 91. – P. 18–27.
12. Ibrahim, M. Nonlinear autoregressive neural network in an energy management strategy for battery/ultra-capacitor hybrid electrical vehicles / M. Ibrahim, S. Jemei, G. Wimmer, D. Hissel // Electric Power Syst. Res. – 2016. – Vol. 136. – P. 262–269.
13. Ярушкина, Н. Г. Нечеткие нейронные сети с генетической настройкой . Лекции по нейроинформатике. Часть 1/ Н. Г. Ярушкина. – М.: Изд-во МИФИ, 2004. – С. 151- 199.
14. Нечеткие нейронные сети [Электронный ресурс]. – Режим доступа: https://studbooks.net/1783599/informatika/nechetkie_neyronnye_seti
15. Астафуров, В. Г. Программная система для автоматической классификации текстур перистой облачности на основе нейро-нечеткой сети / В. Г. Астафуров, Т. В. Евсюткин // Доклады ТУСУР. – 2015. – № 2(36). – С. 125–129.
16. Астафуров, В. Г. Классификация многоярусной облачности по данным MODIS с использованием технологии нейронных сетей и методов нечеткой логики / В. Г. Астафуров, А. В. Скороходов // Современные проблемы дистанционного зондирования Земли из космоса. – 2015. – Т. 12, №6. – С. 162-173.
17. Гафаров, Ф.М. Искусственные нейронные сети и приложения: учеб. пособие / Ф.М. Гафаров, А.Ф. Галимянов. – Казань: Изд-во Казан. ун-та, 2018. –121с.

18. Калан, Р. Основные концепции нейронных сетей.: Пер. с англ / Р. Калан – М.: Издательский дом «Вильямс», 2001. – 287 с.: ил. – Парал. тит. англ.
19. Комарцова, Л. Г. Нейрокомпьютеры: Учебное пособие для вузов / Л. Г. Комарцова, А. В. Максимов. – 2-е изд., перераб. и доп. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2004. – 400 с.: ил. – (Информатика в техническом университете.)
20. Алгоритм обратного распространения ошибки [Электронный ресурс]. - Режим доступа: <http://www.aiportal.ru/articles/neural-networks/back-propagation.html>
21. Метод сопряженных градиентов — математический аппарат [Электронный ресурс]. – Режим доступа: <https://www.intuit.ru/studies/courses/88/88/lecture/20539>
22. Федосин, С.А. Анализ и сравнение методов обучения нейронных сетей / С.А. Федосин, Д.А. Ладяев, О.А. Марьина // Вестник Мордовского университета. – 2010. - № 4. – С. 79-88.
23. Осовский, С. Нейронные сети для обработки информации. – Пер. с польск. И.Д. Рудинского / С. Осовский. – М.: Финансы и статистика, 2002. – 344 с.
24. Скороходов, А.В. Сравнительный анализ градиентных методов минимизации / А.В. Скороходов, А.В. Тунгусова // Доклады ТУСУРа. – 2011. – № 2 (24). – С. 98-102.
25. Марьина, О.А. Методы обучения многослойного персептрона. Попытки оптимизации задачи поиска глобального минимума функции энергии [Электронный ресурс] / О.А. Марьина, Д.А. Ладяев. - Режим доступа: http://fetmag.mrsu.ru/2009-1/pdf/Marjina_Ladjaev_Global_minimum.pdf
26. Стохастические методы обучения нейронных сетей. [Электронный ресурс]. – Режим доступа: <https://www.intuit.ru/studies/courses/88/88/lecture/20539>
27. Отмани, Л. Нейронные байесовские сети в применении к распознаванию

речи / Л. Отмани, А. Бенъету // International Conference on Electrical Engineering and Automatic Control 2013. [Электронный ресурс]. – Режим доступа: http://masters.donntu.org/2014/fknt/brynza/library/otmani_2013.html

28. Molecular Biology of the Cell, Second Edition (Appendix to Chapter 17: Cells of the Adult Human Body: A Catalog) / B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, J.D. Watson. – Garland Publishing Inc., New York, 1989. – pp. 995-997.

29. Fiala, J.C. Dendrite Structure / J.C. Fiala, K.M. Harris . - Oxford: Oxford Press, 1999. — pp. 1-34.

30. Нейросети и нейрокомпьютеры. [Электронный ресурс]. – Режим доступа: http://de.ifmo.ru/bk_netra/page.php?index=44&layer=2&tutindex=25

31. Горбачевская, Е.Н. Классификация нейронных сетей / Е.Н. Горбачевская // Вестник Волжского университета им. В.Н. Татищева. – 2012. – № 2. – С. 128–134.

32. Галушкин, А. И. Нейронные сети: основы теории / А.И. Галушкин. – М.: Горячая линия – Телеком, 2010. – 496 с

33. Ефремов, А. А.. О применении кусочно-непрерывных функций к заданию функций принадлежности нечетких чисел (L-R)-типа [Электронный ресурс] / А. А. Ефремов, А. М. Кориков // Вестник науки Сибири. - 2011 - №. 1 (1). - С. 340-343. - Режим доступа: <http://sjs.tpu.ru/journal/issue/view/2/showToc/sect/4>

34. Ефремов, А. А. Вычисление нечеткой вероятности безотказной работы систем с нечеткими параметрами моделей надежности / А. А. Ефремов // Доклады Том. гос ун-та систем упр. и радиоэлектроники. – 2015. – № 2(36). – С. 136–140.

35. Kwang, H. Lee. First Course on Fuzzy Theory and Applications/ Lee H. Kwang; Advanced Institute of Science and Technology, Repulic of South Korea, 2005. – 334 p.

36. Harris, J. Fuzzy Logic Applications in Engineering Science/ J. Harris;

Published by Springer, The Netherlands, 2006. – 216 p

37. Основные термины и определения нечеткого множества [Электронный ресурс]. – Режим доступа: http://www.life-prog.ru/1_36943_osnovnie-termini-i-opredeleniya.html, свободный (дата обращения: 02.05.2015).
38. Нгуен, А.Т. Компьютерное моделирование показателей нечеткой надежности / А.Т. Нгуен, А.А. Ефремов // Доклады ТУСУР. – 2016. – Т. 19, № 1. – С. 57-62. DOI: 10.21293/1818-0442-2016-19-1-57-62.
39. Половко, А. М. Основы теории надежности: учебное пособие / А. М. Половко, С. В. Гуров. – СПб: БХВ-Петербург, 2008. – 702 с.
40. Лоскутов, А.Ю. Анализ временных рядов: Курс лекций / А.Ю. Лоскутов. – М.: Физический факультет МГУ. – 113с.
41. Афанасьев, В.Н. Анализ временных рядов и прогнозирование: Учебник / В.Н. Афанасьев, М.М. Юзбашев. – М.: Финансы и статистика, 2001. – 228 с.
42. Кизбикенов, К.О. Прогнозирование и временные ряды : учебное пособие / К.О. Кизбикенов. – Барнаул : АлтГПУ, 2017. – 113с.
43. Методы и модели анализа временных рядов : метод. указания к лаб. работам / сост. С.И. Татаренко. – Тамбов : Изд-во Тамб. гос. техн. ун-та, 2008. – 32 с.
44. Газизов, Д. И. Обзор методов статистического анализа временных рядов и проблемы, возникающие при анализе нестационарных временных рядов /Д.И. Газиров // Научный журнал. – 2018. – № 3 (4). – С. 9-14.
45. Петрова, М.В. Декомпозиция и анализ временного ряда на примере индекса потребительских товаров и розничной торговли RTSCR [Электронный ресурс] / М.В. Петрова // Сборник статей участников IV Международного научного студенческого конгресса «Россия и ВТО: экономические, правовые и социальные аспекты». - -2014. - Режим доступа: <https://docplayer.ru/80016544-Dekompoziciya-i-analiz-vremennogo-ryada-na-primere-indekса-potrebitelskih-tovarov-i-roznichnoy->

torgovli-rtscr.html

46. Черемушкин, С.В. Прогнозирование экономических временных рядов на основе анализа интервенций и определения вероятностных интервалов / С.В. Черемушкин // Финансовый менеджмент, 2011. – № 5. – с. 94-116.
47. Андерсон, Т. Статистический анализ временных рядов / Т. Андерсон – М.: Изд-во «Мир», 1976. – 757с.
48. Plummer, E. A. Time series forecasting with feed-forward neural networks: guidelines and limitations: Thesis Master of Science / Plummer Eric A. – 2000. – 60p.
49. Mozaffari, L. Vehicle speed prediction via a sliding-window time series analysis and an evolutionary least learning machine: A case study on San Francisco urban roads / L. Mozaffari, A. Moxaffari, N. L. Azad // Engineering science and technology, an international journal. – 2015. – Vol.18. – p.150-162.
50. Vafaeipour, M. Application of sliding window technique for prediction of wind velocity time series / M. Vafaeipour, O. Rahbari, M. A. Rosen, F. Fazelpour, P. Ansarirad // International journal of Energy and environmental engeering (springer). – 2014. – Vol. 5. – p. 105-111.
51. Yu,Y. Time Series Outlier Detection Based on Sliding Window Prediction [Электронный ресурс] / Y. Yu, Y. Zhu, S. Li, D. Wan // Mathematical Problems in Engineering. – 2014. – Режим доступа: <https://www.hindawi.com/journals/mpe/2014/879736/>
52. Нгуен, А.Т. Нейронные сети с полиномиальными кусочно-непрерывными функциями активации для поиска закономерностей в данных / А.Т. Нгуен, А.М. Корилов // Доклады ТУСУР. – 2019. – Т. 22, № 1. – С. 71-76. doi: 10.21293/1818-0442-2019-22-1-71-76.
53. Матрица Якоби и якобиан. [Электронный ресурс]. – Режим доступа: <http://pmpu.ru/vf4/algebra2/dets/jacobian>

54. Hagan, M. T. Training feed forward networks with the Marquardt algorithm / M. T. Hagan, M. B. Menhaj // IEEE Trans. Neural Netw. – 1994. – Vol. 5. – p. 989–993
55. Nguyen, A. T. Application of Artificial Neural Networks to Predict Dynamic Responses of Wing Structures due to Atmospheric Turbulence/ A. T. Nguyen, J. H. Han, A. T. Nguyen. International Journal of Aeronautical and Space Sciences. – 2017. – Vol. 18 (3). – с. 474-484.
56. Marquardt, D.W. An Algorithm for Least-Squares Estimation of Nonlinear Parameters / D.W. Marquardt. – Journal of the Society for Industrial and Applied Mathematics. – Vol. 11, No. 2, 1963. – pp. 431–441
57. Мазуров, М.Е. Прогнозировании финансовых временных рядов с помощью метода самоорганизованной критичности / М.Е. Мазуров // Экономика, Статистика и Информатика. – 2014. - № 3. – с.153-157.
58. Monthly Australian imports from Japan: thousands of dollars. Jul 65 – Oct 93. [Электронный ресурс]. – Режим доступа: <https://datamarket.com/data/set/22qx/monthly-australian-imports-from-japan-thousands-of-dollars-jul-65-oct-93#!ds=22qx&display=line>
59. Exchange rate of Australian dollar: \$A for 1 US dollar. Monthly average: Jul 1969 – Aug 1995. [Электронный ресурс]. – Режим доступа: <https://datamarket.com/data/set/22wv/exchange-rate-of-australian-dollar-a-for-1-us-dollar-monthly-average-jul-1969-aug-1995#!ds=22wv&display=line>
60. Quarterly Australian national accounts exports: millions of dollars at 1989/90 prices. Sep 59 – Jun 95. [Электронный ресурс]. – Режим доступа: <https://datamarket.com/data/set/22my/quarterly-australian-national-accounts-exports-millions-of-dollars-at-198990-prices-sep-59-jun-95#!ds=22my&display=line>
61. Carter, M. Op Amps for Everyone / M. Carter, R. Bruce // Texas Instruments. – 2009. – P. 10–11.
62. Исаков, В.Н. Статистическая теория радиотехнических систем (курс

лекций) [Электронный ресурс]. - Режим доступа: <http://strts-online.narod.ru/files/lec3.pdf>

63. Нормальный (гауссовский) закон распределения. [Электронный ресурс]. - Режим доступа: <https://scicenter.online/teoriya-veroyatnostey-scicenter/normalnyi-gaussovskiy-zakon-145639.html>

64. Альшанский, М.А. Гауссовский белый шум с траекториями в пространстве $S'(H)$ / М.А. Альшанский . – Изв. вузов. Матем., 2011. – с. 3-11.

65. Stein, M. L. Interpolation of Spatial Data: Some Theory for Kriging / M. L. Stein . – Springer Series in Statistics. Springer, 1999. – P. 40.

66. Хоровиц, П., Хилл, У. Искусство схемотехники: В 3-х томах: Т. 2. Пер. с англ. / П. Хоровиц, У. Хилл. – 4-е изд., перераб. и доп. – М.: Мир, 1993. - с. 51.

67. Алиев, Т.И. Основы моделирования дискретных систем: Учебное пособие / Т.И. Алиев . - СПб: СПбГУ ИТМО, 2009. – 363 с.

68. Вентцель, Е.С. Исследование операций / Е.С. Вентцель. – М.: Советское радио, 1972. – 552 с.

69. Дудин, А. Н. Системы массового обслуживания с коррелированными потоками / А. Н. Дудин, В. И. Кли-менок. – Минск: Изд-во БГУ, 2000. – 175 с.

70. Костин, В.Н. Статистические методы и модели: Учебное пособие / В.Н. Костин, Н.А. Тишина. – Оренбург: ГОУ ОГУ, 2004. – 138 с.

71. Назаров, А.А. Асимптотически пуассоновские мар-потoki / А.А. Назаров, И.Л. Лапатин. // Вестник Томского Государственного Университета. – 2010. – № 4 (13). – С. 72-78.

72. Гнеденко, Б. В. Введение в теорию массового обслуживания / Б. В. Гнеденко, И. Н. Коваленко. – Изд. 3-е, испр. и доп. – М.: КомКнига, 2005. – 400 с.

73. Горцев, А. М. О связи МС-потоков и МАР-потоков событий / А. М.

Горцев, Л. А. Нежелская // Вестник Том. гос. ун-та. Управление, вычислительная техника и информатика. – 2011. – № 1 (14). – С. 13–21.

74. Беккерман, Е. Н. Аппроксимация МС-потокком реального потока событий / Е. Н. Беккерман, С.Г. Катаев, С. С. Катаева, Д. Ю. Кузнецов // Вестник Том. гос. ун-та. Приложение. – 2005. – № 14. – С. 248–253.

75. Беккерман, Е. Н. Алгоритм оценивания числа состояний и значений интенсивностей МС-потока событий / Е. Н. Беккерман, С. С. Катаева // Вестник Том. гос. ун-та. Приложение. – 2009. – № 3(8). – С. 19–26.

76. Горцев, А. М. Оптимальная оценка состояний мар-потока событий в условиях непродлевающегося мертвого времени / А. М. Горцев, Л. А. Нежелская, А. А.Соловьев // Вестник Том. гос. ун-та. Приложение. – 2010. – № 3(12). – С. 54–64.

77. Нгуен, А. Т. Модель нейронной сети с нечеткими функциями активации для прогнозирования временного ряда / А. Т. Нгуен, А. М. Корилов // Доклады ТУСУРа. – 2016. – Т. 19, № 4. – С. 50–52.

78. Nguyen, A. T. Models of neural networks with fuzzy activation functions [Электронный ресурс] / A. T. Nguyen, A. M. Korikov // IOP Conference Series: Materials Science and Engineering. – 2017. Vol. 177. – Режим доступа: <http://iopscience.iop.org/article/10.1088/1757-899X/177/1/012031>, свободный (дата обращения: 01.03.2017).

79. Корилов, А.М. Нечеткая нейросетевая классификация интенсивности марковского случайного потока событий/ А.М. Корилов, А.Т. Нгуен // Доклады ТУСУРа. – 2017. – Т. 20, № 2. – С. 79–83. DOI: 10.21293/1818-0442-2017-20-2-79-83.

80. Пирсон, Л. Наука данных для чайников. – М.: Джон Уайли & сыновья, 2015. – 408с.

81. Осипова, Ю.А. Применение кластерного анализа методом k-средних для

классификации текстов научной направленности / Ю.А. Осипова, Д.Н. Лавров // Математические структуры и моделирование. – 2017. – № 3(43). С. 108–121.

82. Wu, Z. An Optimal Graph Theoretic Approach to Data Clustering: Theory and Its Application to Image Segmentation / Z. Wu, R.M. Leahy // IEEE Transactions on pattern analysis and machine intelligence. – 1993. – Vol. 15. – № 11. – P. 1101-1113.

83. Nguyen, A. Tu. Fuzzy Neural Network Technology Support Decision-Making / A.Tu Nguyen, A.M. Korikov, A.Tuan Nguyen // Advances in Computer Science Research. — 2017. — Vol. 72 : Information technologies in Science, Management, Social sphere and Medicine (ITSMSSM 2017). — P. 128-131. DOI: <https://doi.org/10.2991/itsmssm-17.2017.27>

84. Нгуен, А.Т. Задачи обучения нечетких нейронных сетей / А.Т. Нгуен, А.М. Корилов // Сборник материалов международной научно-практической конференции «Электронные средства и системы управления». – Томск: В-Спектр, 2017. – С. 65- 68

85. Численные методы в физике и радиофизике: Учебно-методическое пособие / Г.М. Тептин, О.Г. Хуторова, Ю.М. Стенин, А.А. Журавлев, В.Р. Ильдиряков, В.Е. Хуторов, К.В. Скобельцын. – Казань: КФУ, 2013. – 38 с.

86. Wettscherech, D. Improving the Performance of Radial Basis Function Networks by Learning Center Locations / D. Wettscherech, T. Dietterich // Electronic Proceedings of the Neural Information Processing Systems Conference. – 1991. – № 4. – P. 1133–1140.

87. Сараев, П.В. Использование псевдообращения в задачах обучения искусственных нейронных сетей // Электронный журнал «Исследовано в России». – 2001. – № 29. – С. 308–317.

88. Кудин, Г. Применение методов псевдообращения к задачам оптимизации структур / Г. Кудин // International Journal “Information Theories and Applications”. –

2013. – Vol. 20, N 4. – p.360-366.

89. Нгуен, А.Т. Нечеткая нейросетевая кластеризация интенсивности МС - потоков событий / А.Т. Нгуен, А.М. Корилов // Сборник материалов международной научно-практической конференции студентов, аспирантов и молодых ученых «Молодежь и современные информационные технологии». – 11.2017. – С. 53-54.

90. Зубов, Н.Е. О вычислении псевдообратной матрицы. Общий случай / Н.Е. Зубов, В.Н. Рябченко // Вестник Московского государственного технического университета им. Н.Э. Баумана. Серия «Естественные науки». – 2018. – № 1. – С. 16-25.

91. Рассел, С. Искусственный интеллект: современный подход / С. Рассел, П. Норвиг. – М.: Вильямс, 2006. – 1408 с.

92. Ata, R. An adaptive neuro-fuzzy inference system approach for prediction of tip speed ratio in wind turbines / R. Ata, Y. Kocyigit // Expert Systems with Applications. – 2010. – Vol. 37, N 7. – P. 5454–5460.

93. Jang, J. S. R. Neuro-fuzzy and soft computing / J. S. R. Jang, C. T. Sun, E. Mizutani. – Upper Saddle River: Prentice Hall, 1997. – 614p.

94. Joshi, A. On neurobiological, neuro-fuzzy, machine learning, and statistical pattern recognition techniques / A. Joshi, N. Ramakrishnan, E.N. Houstis, J.R. Rice // IEEE Transactions on Neural Networks. – 1997. – Vol. 8. – P. 18-31.

95. Lin, C. T. Support vector based fuzzy neural network for pattern classification / C.T. Lin, C.M. Yeh, S.F. Liang, J.F. Chung, N. Kumar // IEEE Transactions on Fuzzy Systems. – 2006. – Vol. 14(1). – P. 31–41.

96. Mitra, S. Knowledge-based fuzzy MLP for classification and rule generation / S. Mitra, R. K. De, S. K. Pal // IEEE Transactions on Neural Networks. – 1997. – Vol. 8(6). – P. 1338–1350.

97. Nozaki, K. Adaptive fuzzy rule-based classification systems / K. Nozaki, H. Ishibuchi, H. Tanaka // IEEE Transactions on Fuzzy Systems . – 1996. – Vol. 4(3). – p.238-250.
98. Simpson, P. K. Fuzzy min–max neural networks. I. Classification / P. K. Simpson // IEEE Transactions on Neural Networks. – 1992. – Vol. 3(5). – P. 776-786
99. Zadeh, L. A. Fuzzy sets and their application to pattern classification and clustering analysis / L. A. Zadeh // Classification and Clustering. Proceedings of an Advanced Seminar Conducted by the Mathematics Research Center, the University of Wisconsin, Madison. – 1976. – P. 251–299.
100. Chatterjee, A. A PSO-aided neuro-fuzzy classifier employing linguistic hedge concepts / A. Chatterjee, P. Siarry // Expert Systems with Applications. – 2007. – Vol. 33(4). – p. 1097-1109.
101. Jang, J. S. R. ANFIS: Adaptive network based fuzzy inference systems / J. S. R. Jang // IEEE Transactions on Systems Man and Cybernetics. – 1993. – Vol. 23(3). – P. 665-685.
102. Juang, C. F. An online self-constructing neural fuzzy inference / C. F. Juang, C. T. Lin // IEEE Transactions on Fuzzy Systems. – 1998. – Vol.6. - Issue 1. – P. 12-32.
103. Kasabov, N. K. Evolving fuzzy neural networks for supervised/unsupervised online knowledge-based learning / N. K. Kasabov // IEEE Transactions on Systems Man and Cybernetics, Part B. – 2001. – Vol. 31(6). – P. 902–918.
104. Kasabov, N. K. DENFIS: Dynamic evolving neural-fuzzy inference system and its application for time-series prediction / N. K. Kasabov, Q. Song // IEEE Transactions on Fuzzy Systems. – 2002. – Vol. 10(2). – P. 144–154.
105. Nauck, D. Neuro-fuzzy systems for function approximation / D. Nauck, R. Kruse // Fuzzy Sets and Systems. – 1999. – Vol. 101(2). – P. 261–271.

106. Marin-Blazquez, J. G. From approximative to descriptive fuzzy classifiers / J. G. Marin-Blazquez, Q. Shen // IEEE Transactions on Fuzzy Systems. – 2002. – Vol. 10(4). – P. 484–497.
107. Nauck, D. Fuzzy data analysis with NEFCLASS / D. Nauck // International Journal of Approximate Reasoning. – 2003. – Vol. 32(2–3). – P. 103–130.
108. Браверман, Э. М. Структурные методы обработки эмпирических данных / Э. М. Браверман, И.Б Мучник. – М.: Наука, 1983. – 464 с.
109. Moller, M. F. A scaled conjugate gradient algorithm for fast supervised learning / M. F. Moller // Neural Networks. – 1993. – Vol. 6, N 4. – P. 525–533.
110. Fisher, R. A. The Use of Multiple Measurements in Taxonomic Problems / R. A. Fisher // Annals of Eugenics. – 1936. Vol. 7. –P. 179-188.
111. Siegler, R. S. Three aspects of cognitive development / R. S. Siegler // Cognitive Psychology. – 1976. – Vol. 8, N 4. – P. 481-520.
112. Attaway, S. MATLAB: A Practical Introduction to Programming and Problem Solving. – Third Edition / S. Attaway. – Burlington, MA: Elsevier Inc., 2013. – 539 p.
113. Chakraborty, D. Neuro-fuzzy scheme for simultaneous feature selection and fuzzy rule-based classification / D. Chakraborty, N. R. Pal // IEEE Transactions On Neural Networks. – 2004. – Vol. 15, N 1. – P. 110–123.
114. Методы решения задачи минимизации квадратичной функции. Проблемы сходимости: Метод. указания / сост. К. В. Григорьева; СПб. гос. архит.-строит. ун-т. – СПб., 2009. – 36 с.
115. Корилов, А. М. Нейро-нечеткая классификация объектов и их состояний / А.М. Корилов, А.Т. Нгуен //Научный вестник НГТУ. – 2018. – № 3 (72). – С. 73–86. – doi: 10.17212/1814-1196-2018-3-73-86. DOI: 10.17212/1814-1196-2018-3-73-86.
116. Abe, S. Support vector machines for pattern classification. – NewYork, NY:

Springer, 2010. – 473 p.

117. Wu Y. Using Radial Basis Function Networks for Function Approximation and Classification / Y. Wu, H. Wang, B. Zhang, K. L. Du // *ISRN Applied Mathematics*. – 2012. – Article ID 324194. – 34 p.

118. Игумнов, И. В. Метод Нелдера – Мида в обучении искусственной нейронной сети, входящей в состав ШИМ-элемента / И. В. Игумнов, Н. Н. Куцый // *Научный вестник НГТУ*. – 2017. – № 3. – С. 22–30.

Приложение А. Акт внедрения в учебный процесс



УТВЕРЖДАЮ

Проректор по учебной работе ТУСУР

П.Е. Троян

« 8 » апреля 2019 г.

Акт

внедрения в учебный процесс

результатов диссертации Нгуен Ань Ту «Алгоритмическое обеспечение нейронной сети с полиномиальными кусочно-непрерывными функциями активации для обнаружения закономерностей в данных», представленной на соискание ученой степени кандидата технических наук

Комиссия в составе:

председатель – Кориков А.М. (зав. кафедрой АСУ, д.т.н.),

члены комиссии

- Мицель А.А. (профессор кафедры АСУ, д.т.н.),

- Алферов С.М. (доцент кафедры АСУ, к.т.н.),

установила, что:

1) Результаты второй, третьей и четвертой глав диссертации используются при проведении практических занятий по дисциплине «Интеллектуальные системы» для магистрантов направления 09.04.01 Информатика и вычислительная техника (профили: «Программное обеспечение вычислительных машин, систем и компьютерных сетей» и «Автоматизированные системы обработки информации и управления в экономике»).

2) Результаты диссертации используются при выполнении практических и лабораторных занятий по дисциплине «Моделирование систем» для бакалавриата по направлению 09.03.01 Информатика и вычислительная техника (профиль: «Программное обеспечение средств вычислительной техники и автоматизированных систем»)

Председатель комиссии

/ А.М. Кориков/

Члены комиссии:

/ С.М. Алферов/

/ А.А. Мицель/

Приложение Б. Акт внедрения в компании Midas Core Pte Ltd



Research Implementation Report

Midas Core Pte Ltd (Registration No. 201812525R, Singapore) is a FinTech company that provides blockchain solutions, trading tools software, crypto wallet and crypto exchange. Models of fuzzy neural network with activation functions formed in polynomial piecewise continuous in Nguyen Anh Tu's thesis are applied to the process of improving product quality of the company. Results of the second chapter of the thesis are used for determining state of the financial securities market our company.

Nguyen Anh Tu demonstrates exceptional critical thinking, possesses great analytical skills to diagnose problems and devise viable solutions. Anh Tu would be a valuable asset to any organization and has my highest recommendation.

Please feel free to reach out to me if you need further information.

20 June 2019

Director of Midas Core Pte. Ltd



Phillip Phung

Приложение В. Акт внедрения в компании Center of Education and Technology Transfer VIET4C

RESEARCH IMPLEMENTATION REPORT

We confirm that the results of Nguyen Anh Tu's dissertation research on the topic: "Algorithmic support for neural networks with polynomial piecewise continuous activation functions to detect data patterns" are of practical interest and have been used in the development of our business.



26 June 2019

Mai Kim Binh,

Director

Microtech VietNam Joint Stock Company

Tel: +84 915 570 122

Email: info@microtechvn.net

Website: <http://microtechvn.net>