

На правах рукописи

Микишев Илья Николаевич

**Исследование нейронных сетей для решения задач маршрутизации в
сетях передачи данных**

Направление подготовки
11.04.02 «Инфокоммуникационные технологии и системы связи»
направленность – Сети, системы и устройства телекоммуникаций
программа академической магистратуры

АВТОРЕФЕРАТ
магистерской диссертации
на соискание квалификации (степени) магистра

Екатеринбург 2020

Работа выполнена в Федеральном государственном бюджетном образовательном учреждении высшего образования «Сибирский государственный университет телекоммуникаций и информатики» Уральский технический институт связи и информатики (филиал) в г. Екатеринбурге (УрТИСИ СибГУТИ)

Научный руководитель, к.т.н., доцент

Н.В. Будылдина

Рецензент к.ф.-м.н., доцент

В.Т. Куанышев

Защита состоится «30» июня 2020 г. в 9:00 часов в Федеральном государственном бюджетном образовательном учреждении высшего образования «Сибирский государственный университет телекоммуникаций и информатики» Уральский технический институт связи и информатики (филиал) в г. Екатеринбурге (УрТИСИ СибГУТИ), г. Екатеринбург, ул. Репина, д. 15.

Секретарь Государственной аттестационной комиссии

О.А. Шумилова

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования и степень ее разработанности.

Современные сети передачи данных являются сложной структурой и основываются на транспортировке информации от отправителя к получателю. Чтобы информация была передана по оптимальному пути, проходя по маршруту с соответствующими определенными критериями, используются алгоритмы маршрутизации.

Маршрутизация – процесс передачи информации от отправителя к получателю через сеть промежуточных узлов. При этом при передаче сигнала на его маршруте будет, по крайней мере, не менее одного узла.

В процесс маршрутизации включены два основных компонента – определение оптимального маршрута и непосредственно транспортировка информации.

Определение маршрута для передачи пакета является сложной задачей и основывается на определенных параметрах или их комбинации. Маршрут может быть задан как изначально статическим – известен и не будет изменен в процессе передачи сигнала, так может быть и динамическим – изменяться в процессе передачи информации в зависимости от получаемой информации, в таком случае, сложность расчета маршрута возрастает.

На практике случаются ситуации, когда передать информацию необходимо в открытом виде, это означает, что не будут применены методы шифрования и криптографии. При таком типе передачи информации необходим выбор маршрута следования сигнала, при котором будет соблюдаться определенная степень защищенности, обеспечивающая невозможность подключения третьих лиц к физической среде передачи сигнала.

При маршрутизации трафика необходимо соблюдение заданных условий, причем необходимо выполнение не только требований, которые обеспечивают доставку информации при минимальных временных затратах, но и параметры, которые характеризуют безопасность маршрута.

В современных сетях существующие алгоритмы маршрутизации используют информацию о состоянии сети, в которой будет организована передача данных, информацию о ее полной структуре. При частых изменениях топологии сети, появлении или удалении соединений, изменениях в среде передачи данных, алгоритмы маршрутизации снижают эффективность своей работы и теряют способность оперативно поддерживать обмен информацией в сети. Для поддержания корректной работы сети передачи данных для доставки информации необходимо применение современных методов, которые способны решать задачу маршрутизации при неполной информации. К таким современным методам относятся нейронные сети.

Применение нейронных сетей позволит сохранять корректное функционирование сети передачи данных и передачу информации даже при частичной деградации или нарушении целостности сети. Таким образом, задача исследования нейронных сетей для решения задач маршрутизации в сетях передачи данных является актуальной и практически значимой. В рамках реализации проекта «Цифровая экономика», актуальность темы подтверждается

ростом внедрения нейронных сетей в условиях развития цифровых технологий в экономической, научной и социальной сфере.

Объект исследования – нейронные сети.

Предмет исследования – использование нейронных сетей для анализа параметров сетевого интерфейса и принятия решения о передаче информации.

Целью работы является исследование возможности нейронных сетей анализировать параметры сетевого интерфейса и принятия решения о передаче информации.

Для достижения поставленной цели необходимо решить следующие задачи:

- 1) провести анализ публикаций по теме исследования;
- 2) выполнить анализ алгоритмов маршрутизации, произвести их классификацию, а также проанализировать способы получения данных, которые необходимы для работы маршрутизирующей системы;
- 3) проанализировать параметры, влияющие на маршрутизацию трафика;
- 4) провести анализ архитектур нейронных сетей, которые способны решать задачи прогнозирования, управления и оптимизации;
- 5) произвести анализ и дать сравнительную характеристику функциям активации нейронов;
- 6) проанализировать методы обучения нейронных сетей;
- 7) выполнить анализ гиперпараметров нейронной сети;
- 8) разработать нейронную сеть, способную выполнять анализ параметров интерфейса сетевого оборудования и принимать решение о передаче информации.

Научная новизна работы заключается в исследовании возможности нейронной сети типа перцептрон выполнять анализ параметров сетевого интерфейса и принимать решение о последующей передаче информации.

Практическая значимость заключается в том, что продемонстрирована способность нейронных сетей анализировать параметры сетевого интерфейса и передавать информацию на основе данной оценки, а также позволяет:

- 1) повысить эффективность маршрутизации, передавая информацию по надежному маршруту;
- 2) повысить вероятность доставки информации от отправителя к получателю в условиях ухудшения сети связи.

Методология и методы исследования – разработка и написание кода в программной среде Sublime Text 3 на языке программирования Python с использованием библиотеки numpy.

Положения, выносимые на защиту:

- 1) графическое представление маршрутизации в сетях передачи данных;
- 2) архитектура нейронных сетей;
- 3) график сигмоидальной функции активации нейрона;
- 4) графическое представление работы алгоритма градиентного спуска;
- 5) графический результат влияния скорости обучения нейронной сети;
- 6) демонстрация работы нейронной сети.

Апробация результатов.

1) «Информационные технологии и когнитивная электросвязь» доклад по теме «Особенности построение нейронных сетей» // Материалы Межвузовского научного семинара – Екатеринбург, УрТИСИ СибГУТИ, 2019 г.;

2) «Цифровая экономика: взгляд студенчества» доклад по теме «Нейронные сети» // Материалы XXI научно-практической конференции студентов УрТИСИ СибГУТИ, г. Екатеринбург, 2019 г.;

3) «Исследование влияния количества нейронов в вычислительном слое нейронной сети на точность получаемого результата» // Международная научно-практическая конференция «ИНФОКОМ – 2020», г. Ростов-на-Дону, 2020 г.;

4) «Исследование скорости обучения нейронной сети при решении задачи нахождения глобального минимума функции» // VI Всероссийская научно-практическая конференция «Информационные технологии и когнитивная электросвязь», г. Екатеринбург, 2020 г.

Диссертационная работа включает введение, четыре главы, заключение, список литературы из 41 наименования. Объем диссертации 79 страниц, также включен 40 рисунок.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы магистерской диссертации, определены объект и предмет исследования, приведены цель и задачи работы, описаны научная новизна и практическая значимость.

В первой главе «Анализ публикаций» проведен анализ публикаций по теме исследования, который позволил определить степень проработанности вопроса использования нейронных сетей для анализа параметров сетевого интерфейса и принятия решения о передаче информации. Изученные материалы отражают популярные направления развития исследований в данной области.

В результате анализа трудов таких авторов, как А.И. Галушкина, С. Хайкина, Р. Гонсалеса, Р. Вудса, Дж. Ту, Р. Гонсалеса, Е.В. Бодянский, О.Г. Руденко, Н.И. Черняков, А.А. Евдокимов, И.Н. Лавриненко, А.В. Лавриненко, С. Осовский, Р. Каллан, В.А. Головкин, Л.Г. Комарцовой, А.В. Максимова, В.И. Комашинский и Смирнов Д.А. были сделаны выводы, подтверждающие актуальность темы магистерской диссертации.

Область исследования и применения нейронных сетей находится в состоянии активного развития после застоя, приходившегося на конец XX века. На данном этапе сформирован довольно существенный перечень фундаментальной литературы, описывающей принципы функционирования нейронных сетей, однако следует отметить, что вопрос применения нейронных сетей для решения задач маршрутизации в литературе не раскрыт.

Во второй главе «Технологии маршрутизации информации» рассматриваются алгоритмы маршрутизации и их классификация – рассмотрены достоинства и недостатки, а также проанализированы способы получения данных, которые необходимы для работы маршрутизирующей системы. Проанализированы параметры, влияющие на маршрутизацию трафика, среди которых:

- QoS;

- среднегодовое время простоя оборудования;
- наличие ошибок;
- утилизация;
- задержка при передаче сигнала;
- безопасность при передаче трафика.

Телекоммуникационные сети решают задачи транспортировки информации от отправителя к получателю, выполняя пересылку пакетов данных от узла к узлу по заданному маршруту, в большинстве случаев существует несколько маршрутов, по которым может происходить пересылка пакетов, на рисунке 1 представлено, что от отправителя к получателю ведет несколько маршрутов.

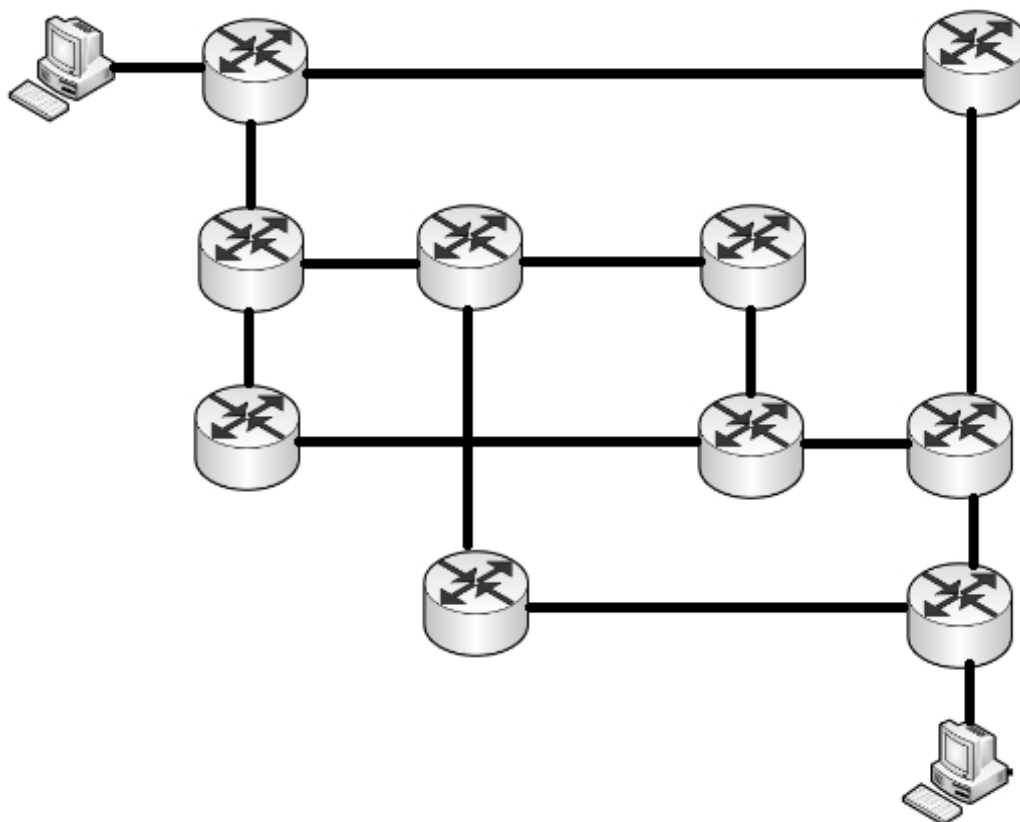


Рисунок 1 – Сеть с несколькими маршрутами

Для того чтобы выбрать оптимальный маршрут передачи данных, применяют алгоритмы маршрутизации, при этом маршрутизация включает в себя два процесса:

- 1) генерация таблицы маршрутизации – сведения о соединении с другим узлом, позволяя произвести оценку для последующей передачи данных;
- 2) управление маршрутами на основании данных, полученных из таблицы маршрутизации.

Средства, выполняющие алгоритм маршрутизации на основе построенной таблицы маршрутизации, получили название маршрутизаторы. Данные средства могут иметь как аппаратную, так и программную реализацию.

Для оценки маршрута, по которому будет передана информация используется критерий – метрика. Метрика характеризует приоритетность выбранного маршрута по отношению ко всем возможным маршрутам.

Выбор маршрута, по которому будет произведена передача информации, основывается на показателях качества, которые в свою очередь зависят от специфики работы сети.

Допустим, требуется выполнить передачу информации с минимальной задержкой, в таком случае, приоритетным показателем при оценке и поставке метрик будет являться расстояние, преодолеваемое сигналом, передаваемым к месту назначения. Алгоритмы маршрутизации оценивают следующие показатели сети: стоимость передачи информации, расстояние, преодолеваемое сигналом, защищенность канала связи, пропускную способность, задержку передачи. Важным критерием является защищенность канала связи, не допускающая перехвата информации третьими лицами, и гарантирующая доставку информации нужному адресату.

В настоящее время в большинстве случаев при построении телекоммуникационных сетей применяют адаптивные распределенные алгоритмы маршрутизации: алгоритмы состояния связей и дистанционно-векторные алгоритмы. Адаптивной, также именуемой динамической маршрутизацией называют систему, в которой любые изменения сети автоматически регистрируются в таблицах маршрутизации. Необходимость в создании алгоритмов такого типа возникает из-за наличия уязвимостей в существующих алгоритмах динамической маршрутизации. При распределенном построении сети каждый маршрутизатор является независимым узлом, выполняя обновление своей таблицы маршрутизации на основе данных, полученных о сети от других маршрутизаторов. Маршрутизатор собирает данные о сети, взаимодействует с другими маршрутизаторами, изменяет данные при получении новых. Такое построение телекоммуникационной сети обладает значительным преимуществом перед централизованным подходом, так как в телекоммуникационной сети с централизованным подходом присутствует только один маршрутизатор, который выполняет сбор информации о других маршрутизаторах, и выполняя в дальнейшем построение маршрута продвижения данных. При неисправностях в работе центрального маршрутизатора работа всей сети будет парализована, поэтому на практике такой подход построения телекоммуникационной сети не получил широкого распространения.

Используя дистанционно-векторный алгоритм (ДВА), маршрутизаторы регулярно производят обмен копиями таблиц маршрутизации. При таком взаимодействии маршрутизаторы регулярно обновляют свои таблицы маршрутизации, поддерживая данные о сети актуальными, в итоге, каждый маршрутизатор сети располагает полными данными о состоянии сети, о наличии других маршрутизаторов, и их таблицах маршрутизации. Построение маршрута основывается на использовании вектора расстояния, который характеризует необходимое количество переходов для достижения конечного узла передачи информации. В таблице маршрутизации присутствует суммарный вектор, который

указывает, на каком расстоянии от заданного узла находится соответствующий узел или сеть.

Для построения и обеспечения каждого маршрутизатора информацией точного графа сети используются алгоритмы состояния сети (АСС). Функционирование сети, построенной на основе АСС связано с поддержанием информации, содержащей информации о топологии сети. В основе функционирования используется информация о одном графе сети, что повышает надежность и устойчивость процесса маршрутизации к каким-либо изменения конфигурации. В отличие от АСС, ДВА маршрутизация не содержит информации об удаленных сетях и маршрутизаторах.

Для построения таблицы маршрутизации применяется процедура, содержащая в себе два этапа.

Процесс генерации базы данных о топологии связей в сети. Формируется структура сети, в которой вершинами являются маршрутизаторы, а ребрами – связи между маршрутизаторами. Происходит процесс обмена информации с соседними маршрутизаторами о графе сети, которой каждый владеет на текущий момент времени, при этом маршрутизаторы не вносят корректировок, сохраняя информацию, в результате обмена данными все маршрутизаторы получают идентичные сведения о состоянии и топологии сети. Вся полученная информация упорядочивается, формируя логическую топологию в виде древовидной структуры сети, в такой структуре корнем дерева является текущий маршрутизатор, а ветвями – возможные маршруты передачи информации. При изменении конфигурации процесс генерации базы данных о топологии связей в сети повторяется, выполняя построение нового графа сети.

Процесс нахождения оптимальных маршрутов передачи информации и генерация таблицы маршрутизации. После выполнения процесса построения дерева связи, необходимо выполнить нахождение оптимального маршрута для дальнейшей передачи информации. В протоколах маршрутизации, основывающихся на АСС (например, OSPF – Open Shortest Path First), используется итеративный алгоритм Дейкстры для поиска оптимального графа сети. Каждый маршрутизатор выполняет рассылку пакетов до всех известных ему подсетей, при этом, в каждом найденном пути запоминается исключительно первый шаг, только он и заносится в последующем в таблицу маршрутизации.

В современных сетях используемые маршрутизаторы, например такие как Cisco для передачи информации используются следующие процессы [29]:

1) процессы маршрутизации, на основе которых работают сетевые протоколы, например IS-IS (Intermediate System-to-Intermediate System), EIGRP (Enhanced Interior Gateway Routing Protocol), BGP (Border Gateway Protocol) или OSPF (Open Shortest Path First);

2) таблица маршрутизации, которая получает сведения от интерфейсов маршрутизатора о состоянии канала связи и производит обмен с другими маршрутизаторами аналогичными таблицами маршрутизации, получая сведения о всей сети;

3) процесс переадресации, который обращается к таблице маршрутизации для принятия решения о переадресации информации на другой интерфейс текущего маршрутизатора.

Помимо описанных процессов существуют и другие параметры, которые могут быть учтены при маршрутизации.

Задержка до передаваемого узла оказывает значительное влияние при передаче информации, которая должна передаваться потоково, например, при видеотрансляции, либо видеоконференции.

Задержка зависит от двух факторов – расстояния и времени обработки на промежуточных узлах. Чем больше расстояние, тем больше будет задержка, тем большее количество промежуточных узлов потребуется пройти пакету до конечного хоста. На промежуточных узлах пакет обрабатывается, проверяется его целостность, определяется дальнейшее направление передачи, помимо этих процессов может возникнуть очередь на передачу, вызванная, например, малой производительностью оборудования, переутилизацией, недоступностью одного из промежуточных хостов.

При передаче пакетов необходимо учитывать, чтобы интерфейсы, через которые передаются пакеты, не были переутилизированы – это означает, что полоса пропускания была бы свободной для передачи пакетов.

При передаче трафика необходимо учитывать важный параметр – приоритизацию передаваемого трафика, данный сервис в технологическом режиме обеспечивает QoS (Quality of Service – качество обслуживания). QoS используется для контроля передаваемого трафика, например, если на маршрутизаторе установить чтобы приоритетным трафиком был VoIP, а FTP был низкоприоритетным, в таком случае, при передаче трафика приоритет будет у VoIP пакетов, а FTP могут быть отброшены.

Для анализа среды передачи используется ведение логов, в частности, запись о наличии ошибок. Ошибки бывают двух видов – вызванные неполадками на линии связи, либо ошибки, вызванные устройством.

При наличии ошибок следует избегать передачи информации по участкам, где они наблюдаются, ибо при передаче информация может быть повреждена, либо вовсе не быть переданной.

Помимо ошибок в логах оборудования имеется информация о времени работы интерфейса, называется – аптайм. Если сигнал регулярно появляется и пропадает, либо среднее время простоя в течение года велико – такой интерфейс для передачи пакетов не является надежным.

Передавая информацию, необходимо учитывать безопасность маршрута, по которому она будет передана, ибо если информация является конфиденциальной, то в таком случае необходимо обеспечить невозможность ее получения третьими лицами, и чтобы информация была доставлена конкретному хосту.

Анализ параметров интерфейса сетевого оборудования позволяет сформировать оценку о надежности интерфейса и в дальнейшем учитывать данную информацию для принятия решения о передаче информации. Выбор надежного маршрута способствует передаче информации предназначенному хосту без ее утери или кражи.

В третьей главе «Искусственные нейронные сети» рассмотрены основы построения нейронных сетей, произведен анализ архитектур построения нейронных сетей, способных выполнять анализ данных, оптимизацию, прогнозирование, а также принятие решений и управление сетью. Рассмотрены способы обучения нейронной сети: обучение с учителем, обучение без учителя, обучение с подкреплением. Рассмотрены функции активации нейронов: ступенчатая, линейная, сигмоидальная, гиперболический тангенс, ReLu. Выполнен анализ гиперпараметров нейронной сети: скорость обучения, влияние нейронов смещения, количество слоев нейронной сети, а также количество нейронов в каждом слое. Рассмотрен популярный метод обучения – метод обратного распространения ошибки для наиболее встречаемой архитектуры – перцептрона.

Искусственная нейронная сеть (далее применяется «нейронная сеть») – интерпретация работы человеческого мозга посредством машинного кода. Нейронная сеть состоит из нейронов и связей между ними – синапсами.

Нейрон – вычислительная единица нейронной сети, выполняет элементарные вычисления над получаемыми данными. Нейроны делятся на слои, формируя уровни взаимодействия, существуют следующие типы слоев – входной, промежуточный (так же именуется как вычислительный, внутренний или скрытый) и выходной. Входной слой предназначен для получения данных и дальнейшей их передаче по нейронной сети, промежуточный слой выполняет функции вычисления и обработки, выходной слой нейронов выполняет вывод результата.

Синапс – соединение между двумя нейронами. Синапс имеет один важный показатель – это вес, благодаря ему передаваемая информация изменяется, и от этого зависит, какой нейрон будет активирован в последующем.

Функционирование нейронных сетей невозможно без их предварительного обучения, поэтому по отношению к нейронным сетям применяется машинное обучение, которое также влияет на вес синапса, значение которого изначально задается случайным образом, но в дальнейшем вычисляется и корректируется для влияния на функционирование нейронной сети, выполняя ее настройку.

Машинное обучение – процесс, при котором происходит обучение нейронной сети при решении множества однотипных задач, позволяя нейронной сети использовать полученную информацию для решения задач, для которых создавалась сеть.

Пример строения нейронной сети представлен на рисунке 2.

Сеть радиально-базисных функций. Данный тип нейронных сетей использует в качестве активационных функций радиально-базисные, также именуемые RBF-сетями, на выходе которой является линейная комбинация параметров нейрона и базисных функций.

Нейронные сети радиально-базисных функций нашли применение в прогнозировании временных рядов, системах управления и классификации данных.

Нейронные сети радиально-базисных функций характеризуются следующими особенностями:

1) имеют только один скрытый слой;

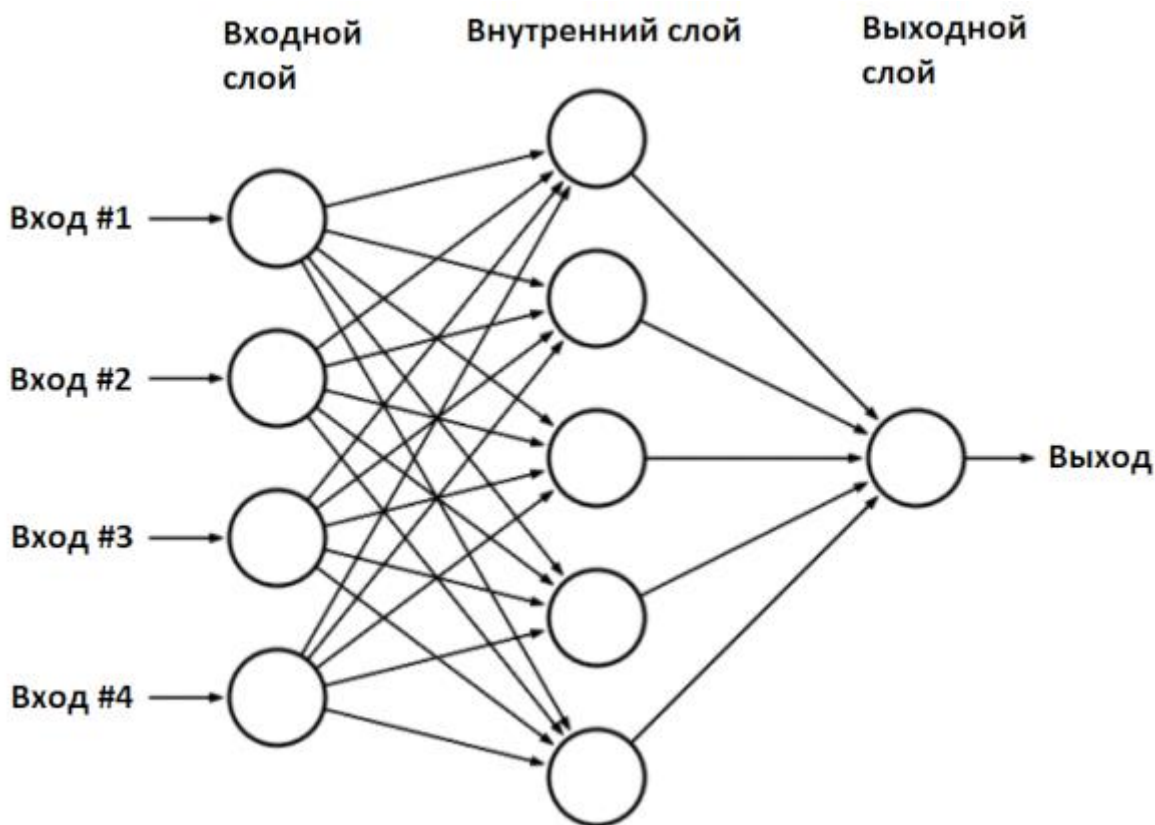


Рисунок 2 – Пример строения нейронной сети

2) нелинейную активационную функцию имеют исключительно нейроны, находящиеся в скрытом слое;

3) веса синаптических связей входных и выходных слоев всегда равняются единице.

В качестве аргумента радиально-базисные функции передачи принимают расстояние между заданным центром активационной функции и входным вектором. Чем ближе к центру входной вектор, тем значения этой функции больше.

Перцептрон. Нейронная сеть с архитектурой типа перцептрон состоит из трех типов элементов:

1) датчики, также называются S-элементы, либо сенсорный элемент, которые реагируют на поступающую информацию;

2) ассоциативные элементы, называются А-элементы, производят ассоциации на основании данных обучения;

3) реагирующие элементы, именуемые R-элементами, реагируют на основе информации, полученной от ассоциативных элементов.

Однослойная модель перцептрона на практике не получила широкого распространения из-за малой масштабируемости и ограниченного применения, но позволила развить идею перцептрона и модернизировать ее. В настоящее время применяется модель многослойного перцептрона.

Многослойный перцептрон. Отличительной особенностью многослойного перцептрона является наличие более одного скрытого слоя, также можно отнести наличие функции активации нелинейного типа, как правило сигмоидальной.

Поступающие на вход нейронам сигналы не бинарные, кодируются десятичными числами, но при этом нормализуются, чтобы принимали значения от нуля до единицы, это необходимо так как применяется сигмоидальная функция активации.

Показателем эффективности работы перцептрона является не реакция на поступающие сигналы (ибо можно получить реакцию любого типа и перцептрон будет выдавать результат, но он может быть неверным), а прогресс в выработке таких реакций, приближение к правильному ответу.

Рекуррентные нейронные сети. Нейронные сети передают сигнал последовательно, в направлении от нейронов, расположенных на входе (первом слое) нейронной сети, к ее выходу (нейронам, располагающимся на последнем слое), проходя все промежуточные вычислительные слои – такая структура нейронной сети имеет устойчивое поведение при эксплуатации, но существует особый класс нейронных сетей, называемых рекуррентными. Особенностью рекуррентных нейронных сетей является наличие обратной связи. Нейронные сети с наличием обратной связи содержат в своей структуре соединения, связи которых идут от наиболее дальних нейронов (относительно первого слоя) к наиболее близким нейронам. Такие нейронные сети могут быть неустойчивы, динамичны и иметь сложную динамику поведения. Нейронные сети, построенные на основе рекуррентных сетей, имеют большой интерес в сфере исследования, однако при решении практических задач не нашли широкого применения, уступая место традиционным типам нейронных сетей с прямой передачей данных.

Применение нейронной сети для решения задачи маршрутизации на телекоммуникационной сети устанавливает ограничения в виде задач, которые потребуются решать нейронной сети в процессе эксплуатации, а именно: анализ данных, оптимизация, прогнозирование, а также принятие решений и управление сетью. Рассмотренные архитектуры нейронных сетей подходят под заданные условия для их дальнейшего применения на телекоммуникационной сети. Сеть радиально-базисных функций состоит исключительно из одного скрытого слоя, моделируя произвольную нелинейную функцию, что в свою очередь утилизирует вопрос о количестве слоев при использовании многослойного перцептрона. Применяя методы линейного моделирования в сети радиально-базисных функций, можно оптимизировать параметры линейной комбинации в выходном слое, способствуя быстрому нахождению глобального минимума функции, а не локальных, которые мешают процессу обучения многослойного перцептрона. По этой причине сеть радиально-базисных функций быстрее обучается, нежели многослойный перцептрон, но при этом, прежде чем применять линейную оптимизацию в выходном слое сети радиально-базисных функций, необходимо определить положение центров радиальных элементов и их количество. Подобного рода алгоритмы производительнее алгоритмов, производящих обучение многослойного перцептрона, но меньше подходят для поиска оптимальных решений.

Опытным путем установлено, что для правильного моделирования типичной функции сети радиально-базисных функций, необходимо большее количество элементов. По этой причине модель сети радиально-базисных функций будет

работать медленнее, требовать больше памяти по сравнению с многослойным перцептроном.

Отличительной особенностью нейронных сетей является имитация структуры биологической нейронной сети и ее качества – обучение. Нейронные сети состоят из двух компонентов: нейронов и соединяющих их компонентов, называемых синапсами. У синапса имеется значимый для функционирования нейронной сети параметр, который называется вес. Наличие в нейронной сети синапсов позволяет информации передаваться от нейрона к нейрону, при этом параметр вес определяет, будет ли передана данная информация последующему нейрону.

Чтобы применить нейронную сеть для решения задачи, ее, как и человека, необходимо первоначально обучить, для этого применяются методы машинного обучения. Машинное обучение – процесс, при котором происходит обучение нейронной сети при решении множества однотипных задач.

Существуют следующие типы обучения нейронных сетей:

- обучение с учителем;
- обучение без учителя;
- обучение с подкреплением.

Обучение нейронной сети с наличием учителя означает присутствие тренировочного датасета. Датасет (обучающий набор) – набор примеров с истинными значениями, среди которых могут выступать размеченные данные, показатели, классы, наличие тэгов. Обучающий набор состоит из наблюдений и наборов данных, для которых приведены значения как входных, так и выходных переменных.

Точный размер датасета, необходимый для обучения нейронной сети, определяется на основе ряда эвристических правил, которые связывают количество необходимых наблюдений с размером нейронной сети, одно из таких правил гласит – количество наблюдений в датасете должно превышать количество связей в сети в десять раз.

Количество наблюдений напрямую коррелирует со сложностью решаемой задачи нейронной сетью, чем сложнее задача, тем большее количество наблюдений должен содержать в себе датасет. Также необходимо учитывать количество анализируемых переменных, с ростом количества переменных количество требуемых наблюдений растет нелинейно, поэтому даже при небольшом числе переменных может потребоваться большое множество наблюдений. Данная особенность известна как проклятие размерности.

Большинство решаемых задач требуют наличия в датасете несколько сотен или тысяч наблюдений, но сложные задачи будут требовать значительно большее число наблюдений для корректного обучения нейронной сети. При этом даже для решения тривиальной задачи потребуется не меньше сотни наблюдений.

Обучая нейронную сеть обучающим набором, в котором недостаточно количества наблюдений, в результате работы нейронная сеть будет предоставлять неправильные ответы, не выполняя своих функций.

При решении реальных задач нейронная сеть будет работать с недостоверными данными, которые могут быть искажены, зашумлены, либо быть неполными, либо некоторые переменные могут вовсе отсутствовать. Благодаря

широкой выборке среди наблюдений нейронные сети могут показывать эффективную работу при работе с шумами, однако, можно столкнуться с так называемыми выбросами, которые могут исказить получаемый результат обучения. Выбросы – значения, находящиеся далеко за пределами области нормальных значений некоторой переменной.

Для достижения необходимого результата нейронную сеть, как и человека требуется обучить. Для определенного рода задач, например, как классификация используется обучение с учителем, первоначально предоставляя нейронной сети тренировочный датасет, позволяющий выполнить анализ данных и формируя фундаментальные установки в работе нейронной сети.

Если нейронной сети продемонстрировать сначала датасет с изображениями животных, то она сможет запомнить и классифицировать животных по определенным типам, например, найти животное на изображении, определить его вид, определить окрас, либо другие признаки. В дальнейшем при анализе изображения нейронная сеть будет опираться на те данные, которые она получила в ходе обучения. Но при таком типе обучения для получения наиболее достоверных результатов требуется, чтобы датасет не был однородным. Нейронная сеть, анализируя данные, старается найти закономерность и в дальнейшем использует полученные данные при анализе новой информации, как и человек. Известен случай, когда нейронную сеть обучали распознавать танки на изображении, но когда продемонстрировали другой сет, то нейронная сеть отмечала неверные изображения, где отсутствовали танки, в процессе анализа было выяснено – все фотографии из тренировочного датасета были выполнены в одинаковых тонах и ландшафте, на данном примере проявляется свойство нейронной сети – всеохватность, нейронная сеть может не только выполнять свою задачу по классификации, но и научиться распознаванию других объектов. Но у данного метода обучения есть один значимый минус – наличие обширного датасета, для его создания требуется большое количество информации, времени и ресурсов.

Когда требуется найти заранее неизвестные ответы, то применяют обучение нейронной сети без учителя. При таком обучении демонстрируется набор данных, и нейронная сеть самостоятельно анализирует, находит общие признаки и черты.

Если не имеется обширного датасета, то подойдет метод обучения с частичным привлечением учителя. Существует метод генеративно-сопоставительной сети или GAN. При обучении данным методом существуют две нейронные сети – генератор и дискриминатор. Генератор создает и пытается имитировать информацию, представленную в обучающей выборке. Дискриминатор получаемую информацию из генератора сравнивает с информацией из датасета, если находит отличие – отбрасывает. Данная сеть все время тренируется, и генератор старается сделать идеальную подделку, а дискриминатор лучше отделять подделки от оригинальных данных.

Как и говорилось ранее, нейронные сети подобны биологическому мозгу и также могут выполнять определенные функции в виде анализа, прогнозирования, дополнительно они могут, как и биологический мозг работать, опираясь на положительный полученный опыт, такие нейронные сети обучаются с подкреплением.

Суть подобного метода обучения сводится к обратной связи в виде положительных отметок при обучении, если нейронная сеть предоставила правильный ответ, то такой ответ отмечается как истинный, и в будущем нейронная сеть будет ориентироваться на те элементы и признаки, которые позволили получить истинный ответ. Нейронная сеть в последующих итерациях будет пытаться найти оптимальный способ достижения ранее полученных истинных ответов. При работе такой сети будет производиться как поиск новых способов достижения истинных значений, так и воспроизведение действий, которые уже позволяли достичь истинных значений – нейронная сеть будет учитывать обратную связь.

Наличие нейронов в сети позволяет выполнять важную задачу при функционировании нейронной сети – выполнять расчет. Нейрон считает взвешенную сумму на своих входах, и полученный результат обрабатывает функция активации, после чего определяется, будет ли активирован нейрон, и вследствие передана информация, либо нейрон не будет активирован, и информация не будет передана дальше по синаптическим связям.

Функция активации позволяет определить выходное значение нейрона после выполнения обработки полученного результата взвешенной суммы в зависимости от порогового значения активации. Функция активации позволяет определить, будет ли активирован нейрон, либо останется инактивным.

Для того чтобы определить границы активации, нужно установить пороговое значение, при котором будет считаться, что нейрон активирован, в ином случае будет считаться, что нейрон не был активирован.

Используя ступенчатую функцию, активация будет происходить при значении больше нуля, если значение равно нулю либо меньше, то нейрон будет считаться не активированным.

Ступенчатая функция активации – простейшая функция, позволяющая функционировать нейронной сети, но у нее имеются недостатки – отсутствует возможность создания стэка слоев нейронной сети, а также способна решать только задачи, где имеется два варианта ответа.

Если для решения задачи требуется чтобы нейронная сеть предоставляла более двух ответов, были промежуточные ответы, то выполнение данных условий сможет обеспечить линейная функция активации.

Линейная функция представляет собой прямую линию и пропорциональна взвешенной сумме, полученной на конкретном нейроне.

Преимущество данной функции перед ступенчатой заключается в наличии промежуточных значений. Если активированными оказываются более одного нейрона, то в таком случае в качестве решения выбирается значение того нейрона, у которого значение функции является максимальным.

Но у данной функции имеется существенный недостаток, который влияет на процесс обучения, в частности, существует метод обучения обратного распространения ошибки, который для корректировки значения веса синапса использует метод, называемый градиентный спуск (более подробно будет изложено в дальнейшем), суть метода заключается в нахождении минимума функции, при этом задействуется метод частных производных, при этом, если

вычислить производную, то она будет постоянной. В таком случае значение градиента также будет всегда постоянным, корректировки, выполняемые методом обратного распространения, также будут постоянными и не зависящими от изменений, которые были выполнены при предыдущих итерациях.

Помимо постоянной производной и корректировки веса синапса на постоянное значение, существует сложность стэкирования нейронной сети – создание нейронной сети, включающей в свою архитектуру более одного вычислительного слоя.

Функция активации, получившая свое название из-за своей формы – сигмоидальная функция активации. Особенность сигмоидальной функции активации заключается так же, как и у линейной функции, в том, что они обе не бинарные, имеют промежуточные значения, что означает – функция аналоговая, в отличие от ступенчатой функции активации, характерен гладкий градиент.

Сигмоида подходит для решения задач классификации, стремится привести итоговое значение к одному из краев кривой, данная особенность функции активации позволяет находить четкие границы при решении задач.

График сигмоидальной функции активации показан на рисунке 3.

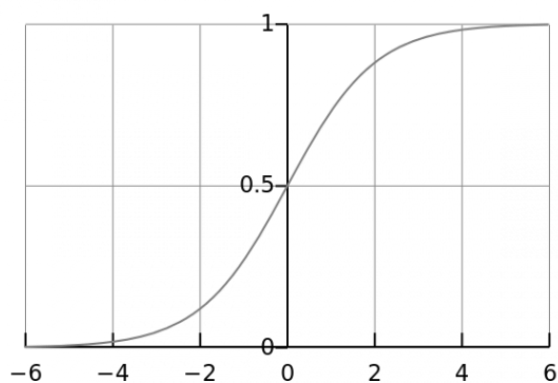


Рисунок 3 – График сигмоидальной функции активации

В современное время сигмоида является наиболее используемой функцией активации при построении нейронных сетей, но она не лишена недостатков, которые оказывают значительное влияние на функционирование сети.

При приближении к концам сигмоиды значения по оси ординат слабо реагируют на изменения, происходящие на оси абсцисс, следовательно, градиент в данных областях имеет также незначительные изменения.

Когда градиент слабо изменяется, либо может иметь крайне малые значения при изменении (также называют исчезновением градиента), то в таком случае нейронная сеть медленно обучается, либо перестает обучаться вовсе.

Функция активации аналогичная по форме сигмоидальной функции называется гиперболическим тангенсом.

Функция гиперболический тангенс имеет одинаковые с сигмоидой свойства – нелинейная, поэтому подходит для создания многослойных нейронных сетей. Но стоит учитывать существенное отличие – градиент значительно изменяет свои

значения по сравнению сигмной. Выбор функции гиперболического тангенса или сигмной зависит от типа решаемой задачи и требований к амплитуде градиента. Аналогично сигмной гиперболический тангенс имеет и недостаток, заключающийся в исчезновении градиента.

Рассмотрим функцию активации типа ReLu. У функции активации типа ReLu имеется свойство – разреженность активации. Предположим, что имеется большая по размеру нейронная сеть. При использовании в качестве функции активации гиперболический тангенс или сигмную потребуется выполнять большое количество вычислений, нагружая сеть, затрачивая временной ресурс на выполнение расчетов, такое явление получило название – плотная активация, подобное функционирование нейронной сети является затратным. Для того чтобы снизить затраты, необходимо, чтобы часть нейронов не была активирована, тогда бы активация стала разреженной и эффективной. Функция активации типа ReLu нашла свое применение в глубоких нейронных сетях, позволяя значительно упростить вычисления и экономя временной ресурс.

Помимо свойства обучения, у нейронных сетей имеется также одна уникальная особенность, которая позволяет значительно влиять на результат, управляя работой сети, такая особенность – это наличие гиперпараметров. Гиперпараметры также имеют важное свойство – подбираются вручную и основываются на эмпирическом методе проб и ошибок, опыте, полученном ранее разработчиком нейронной сети. К гиперпараметрам относятся:

- скорость обучения;
- количество вычислительных слоев;
- количество нейронов, расположенных в каждом слое;
- наличие или отсутствие нейронов смещения.

Вычислительные слои можно стэкировать, создавая глубокие нейронные сети, способные глубже изучить признаки.

Для того чтобы определить, какая архитектура нейронной сети требуется, следует исходить из того, какие данные будет сеть обрабатывать. Например, при изучении английского языка, в частности его грамматики, требуется знать большое количество понятий – в таком случае будет актуальным использование широкой нейронной сети, которая сможет проанализировать большое количество значений.

Если перед нейронной сетью поставлена задача по изучению преобразования Фурье, то необходимо учитывать, что понятий и параметров немного, но каждое из них требуется глубоко понимать.

При построении архитектуры нейронной сети необходимо руководствоваться начальными условиями, входными параметрами и типом задач, которые должна будет выполнять нейронная сеть. Архитектура нейронной сети определяет будущие возможности сети, а также корректность полученных результатов после ее работы.

В нейронных сетях, помимо стандартных нейронов, применяется особый тип – нейроны смещения. Нейрон смещения имеет отличительную особенность – его вход и выход всегда равняется единице, а также отсутствуют входящие синаптические связи. Нейроны смещения могут либо присутствовать в нейронной сети, при этом находясь в количестве одного на каждом слое, либо вовсе

отсутствовать. Синаптические связи нейронов смещения так же, как и у обычных нейронов – со всеми нейронами следующего уровня, при этом отсутствует синаптическая связь между двумя нейронами смещения. Из описанных правил построения нейронной сети, в составе которой имеются нейроны смещения, можно сделать вывод, что присутствие нейронов смещения возможно на всех слоях нейронной сети, кроме выходного, по причине отсутствия необходимости формирования связи с нейронами следующего уровня. Присутствие нейронов смещения в структуре нейронной сети обусловлено наличием необходимости сдвига полученного результата в область допустимых значений, либо значений, которые необходимо получать по результату функционирования нейронной сети.

В области нейросетевых технологий применяется метод машинного обучения, называемый методом обратного распространения ошибки.

Принцип работы метода основан на прохождении сигнала от входа нейронной сети до ее выхода, последующем вычислении ошибки и прохождении обратного сигнала от выхода нейронной сети к ее входу, выполняя корректировку весов с учетом вычисленной ошибки при прямом прохождении сигнала. Метод обратного распространения нашел применение в нейронных сетях, называемых перцептрон.

Используя градиентный спуск, можно решить задачи нахождения минимального значения функции потерь, поиск минимума функции повышает точность модели. Одновременно при увеличении точности модели меняются параметры – вес и смещение. На рисунке 4 показан пример работы алгоритма градиентного спуска.

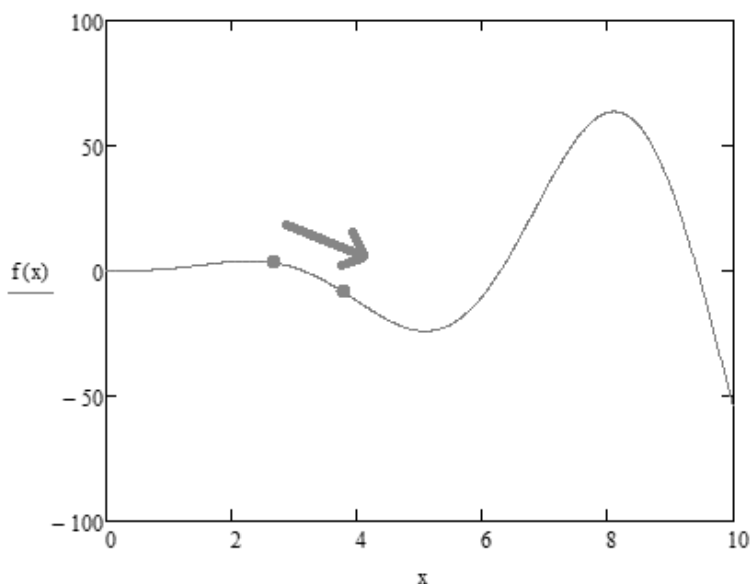


Рисунок 4 – Пример работы алгоритма градиентного спуска

Помимо методов машинного обучения, которые оказывают значительное влияние на процесс обучения нейронных сетей, существует также гиперпараметр, который называется скоростью обучения. Скорость обучения – коэффициент, который устанавливается и подбирается опытным путем вручную. Пример влияния скорости обучения нейронной сети показан на рисунке 5.

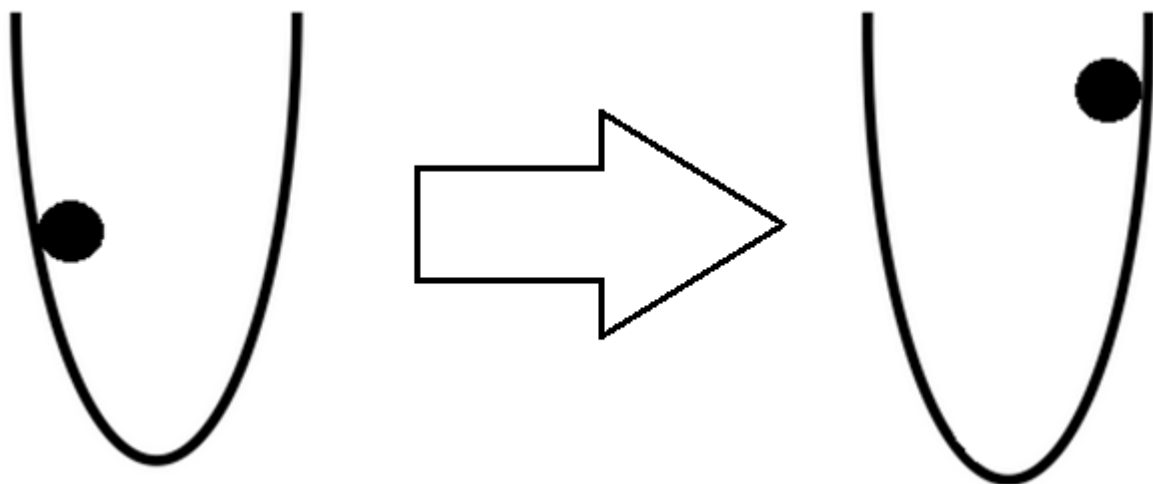


Рисунок 5 – Пример влияния скорости обучения нейронной сети

При подборе значения скорости обучения следует учитывать, что если указать слишком низкое значение, то поиск глобального минимума функции может занять большое количество времени, либо он вовсе не будет найден, либо будет найден неверный минимум функции, снижающий точность работы нейронной сети. При большом значении скорости обучения значения функции будут быстро меняться, пропуская значения глобального минимума.

В четвертой главе «Постановка эксперимента» представлено описание нейронной сети, способной выполнять анализ параметров интерфейса сетевого оборудования и принимать решение о передаче информации на основе надежности маршрута.

Для реализации исследования и разработки нейронной сети была выбрана программная среда Sublime Text 3, язык программирования Python версии 3.8.2, с поддержкой библиотеки `numpy`.

Параметры нейронной сети:

- 1) значение скорости обучения – 10;
- 2) тип функции активации – сигмоида;
- 3) архитектура нейронной сети – перцептрон;
- 4) количество вычислительных слоев нейронной сети – один;
- 5) количество нейронов в входном слое – шесть;
- 6) количество нейронов в вычислительном слое – восемь;
- 7) количество нейронов в выходном слое – один;
- 8) метод обучения – метод обратного распространения;
- 9) количество итераций – сорок тысяч.

Наличие шести нейронов в входном слое обеспечивает необходимый ввод данных из массива датасета. Экспериментально было установлено, что наличие восьми нейронов в вычислительном слое позволяет достичь глобального минимума для решаемой задачи, увеличение количества слоев или нейронов в вычислительном слое негативно повлияло на получаемый результат, нейронная сеть начинала ошибаться – было достигнуто переобучение нейронной сети, при котором происходил поиск отсутствующих закономерностей.

Нейронная сеть анализировала данные, основываясь на допустимом значении выбранного параметра (превышало допустимый порог или нет). Параметры, которые были выбраны для анализа нейронной сетью:

- QoS;
- наличие ошибок;
- утилизация выбранного интерфейса;
- безопасность канала связи;
- время бесперебойной работы.

Первоначально нейронная сеть тестировалась с значением скорости значения от 0.01 до 100, было выявлено:

- при скорости обучения равной 10 нейронная сеть имела наибольшую точность после обучения;

- при заданных параметрах точность работы нейронной сети составила 0.00077, разработанная нейронная сеть однозначно определяла надежность выбранного интерфейса для маршрутизации трафика;

- при исследовании использовался упрощенный датасет для определения возможности нейронной сети к анализу информации.

На рисунке 6 демонстрируется работа нейронной сети.

```
Значение ошибки после 0 итераций: 0.7515022563135849
Значение ошибки после 10000 итераций: 0.001604046596759651
Значение ошибки после 20000 итераций: 0.0011124450246059578
Значение ошибки после 30000 итераций: 0.0008996860901764239
Значение ошибки после 40000 итераций: 0.000774371461528227
Выбранный маршрут является надежным
```

Рисунок 6 – Демонстрация работы нейронной сети

Можно сделать вывод о возможности применения датасета с конкретными значениями представленных параметров, которые были использованы для анализа.

Было выявлено, что нейронные сети способны анализировать информацию однозначно, не допуская ошибок, основываясь на предыдущем опыте. Поэтому нейронные сети способны решать задачи оценки интерфейса и принимать решение о маршрутизации трафика.

ЗАКЛЮЧЕНИЕ

В рамках магистерской диссертации была исследована возможность нейронной сети типа перцептрон выполнять анализ параметров сетевого интерфейса и принимать решение о последующей передаче информации.

В ходе диссертационного исследования были решены следующие задачи.

Проведен анализ публикаций по теме исследования – рассмотрены труды таких авторов, как А.И. Галушкин, С. Хайкин, Р. Гонсалес, Р. Вудс, Дж. Ту, Р. Гонсалес, Е.В. Бодянский, О.Г. Руденко, Н.И. Черняков, А.А. Евдокимов, И.Н. Лавриненко, А.В. Лавриненко, С. Осовский, Р. Каллан, В.А. Головкин, Л.Г. Комарцова, А.В. Максимов, В.И. Комашинский и Смирнов Д.А.

Выполнен анализ алгоритмов маршрутизации, рассмотрена их классификация – рассмотрены достоинства и недостатки, а также проанализированы способы получения данных, которые необходимы для работы маршрутизирующей системы.

Проанализированы параметры, влияющие на маршрутизацию трафика, среди которых:

- QoS;
- среднегодовое время простоя оборудования;
- наличие ошибок;
- утилизация;
- задержка при передаче сигнала;
- безопасность при передаче трафика.

Проведен анализ архитектур нейронных сетей, которые способны решать задачи прогнозирования, управления и оптимизации.

Произведен анализ и дана сравнительная характеристика функциям активации нейронов, рассмотрены преимущества наиболее используемых функций.

Проанализированы методы обучения нейронных сетей, рассмотрены условия, при которых необходимо выбирать определенный метод обучения.

Выполнен анализ гиперпараметров нейронной сети, рассмотрено влияние каждого параметра на функционирование нейронной сети и конечный результат.

Разработана нейронная сеть, способная выполнять анализ параметров интерфейса сетевого оборудования и принимать решение о передаче информации на основе надежности маршрута. Было выявлено, что нейронные сети возможно применять для оценки параметров интерфейса сетевого оборудования и принимать решение о маршрутизации трафика однозначно, не допуская ошибок, основываясь на опыте, полученном при обучении и дальнейшем функционировании сети.

СПИСОК РАБОТ, ОПУБЛИКОВАННЫХ АВТОРОМ ПО ТЕМЕ ДИССЕРТАЦИИ

1 «Информационные технологии и когнитивная электросвязь» доклад по теме «Особенности построение нейронных сетей» // Материалы Межвузовского научного семинара – Екатеринбург, УрТИСИ СибГУТИ, 2019 г.;

2 «Цифровая экономика: взгляд студенчества» доклад по теме «Нейронные сети» // Материалы XXI научно-практической конференции студентов УрТИСИ СибГУТИ, г. Екатеринбург, 2019 г.;

3 «Исследование влияния количества нейронов в вычислительном слое нейронной сети на точность получаемого результата» // Международная научно-практическая конференция «ИНФОКОМ – 2020», г. Ростов-на-Дону, 2020 г.;

4 «Исследование скорости обучения нейронной сети при решении задачи нахождения глобального минимума функции» // VI Всероссийская научно-практическая конференция «Информационные технологии и когнитивная электросвязь», г. Екатеринбург, 2020 г.