МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования

«Нижегородский государственный университет   
им. Н.И. Лобачевского»

**Институт информационных технологий, математики и механики**

**Кафедра информатики и автоматизации научных исследований**

Направление: Прикладная информатика

**МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ**

Тема:

**«Нейросетевой модуль гибридной интеллектуальной системы»**

**Выполнил:**

студент группы 381207 - 1

Алявдин Александр Владимирович

Подпись\_\_\_\_\_\_\_\_\_\_

**Научный руководитель:**

к.т.н., доцент Басалин П.Д.

Подпись\_\_\_\_\_\_\_\_\_\_

**Зав. Кафедры**:

д.т.н, профессор Прилуцкий М.Х.

Подпись\_\_\_\_\_\_\_\_\_

Нижний Новгород  
2018

Оглавление

[Введение 3](#_Toc515530878)

[1. Постановка задачи 5](#_Toc515530879)

[2. Представление знаний о предметной области на инфологическом уровне 6](#_Toc515530880)

[2.1. Дерево решений как формализм описания знаний на инфологическом уровне 6](#_Toc515530881)

[2.2. Граф решений 7](#_Toc515530882)

[3. Гибридная система интеллектуальной поддержки процессов принятия решений 9](#_Toc515530883)

[3.1. Концепция системы, основанной на знаниях 9](#_Toc515530884)

[3.2. Нейросетевой подход к построению интеллектуальной системы 10](#_Toc515530885)

[3.3. Концепция гибридной системы интеллектуальной поддержки и её основные компоненты 12](#_Toc515530886)

[4. Нейронная сеть как основа механизма вывода гибридной системы интеллектуальной поддержки 15](#_Toc515530888)

[4.1. Многослойный персептрон 15](#_Toc515530889)

[4.2. Активационные функции нейронов 17](#_Toc515530890)

[5. Нейросетевой модуль гибридной системы интеллектуальной поддержки 21](#_Toc515530891)

[5.1. Основные компоненты нейросетевого модуля 21](#_Toc515530892)

[5.2. Структура и топология нейронной сети 22](#_Toc515530893)

[5.2.1. Сенсорный слой 22](#_Toc515530894)

[5.2.2. Скрытые слои 23](#_Toc515530895)

[5.2.3. Моторный слой 23](#_Toc515530896)

[5.3. Алгоритм построения нейронной сети по графу решений 23](#_Toc515530897)

[5.4. Пример построения нейронной сети по графу решений из области диагностики неполадок сетевого соединения 25](#_Toc515530898)

[6. Программная реализация 26](#_Toc515530899)

[6.1. Общее описание программной реализации 26](#_Toc515530900)

[6.2. Архитектура программной системы 26](#_Toc515530901)

[6.2.1. Диаграмма классов компонента нейронной сети 26](#_Toc515530902)

[6.2.2. Диаграмма классов компонента построения нейронной сети 27](#_Toc515530903)

[6.2.3. Диаграмма классов компонента дообучения нейронной сети 27](#_Toc515530904)

[6.3. Интеллектуальный интерфейс 28](#_Toc515530905)

[6.4. Компонент построения нейронной сети по графу решений 28](#_Toc515530906)

[6.5. Компонент нейросетевого вывода 29](#_Toc515530907)

[7. Полученные результаты 30](#_Toc515530908)

[7.1. Тестовый базис 30](#_Toc515530909)

[7.1.1. Фрагмент графа решений для области диагностики неполадок сетевого подключения 30](#_Toc515530910)

[7.1.2. Результаты работы нейросетевого модуля 31](#_Toc515530912)

[8. Сравнение результатов работы нейросетевого и аналитического вывода 32](#_Toc515530913)

[Заключение 33](#_Toc515530914)

[Список литературы 34](#_Toc515530915)

[Приложение 35](#_Toc515530916)

Введение

Во многих областях своей деятельности человеку приходится сталкиваться с задачами трудноформализуемого и неформального плана, алгоритмы (сценарии) решения которых заранее не известны.

К таким задачам относятся задачи, математически формулируемые, но не имеющие априори известного алгоритма решения, гарантирующего получение приемлемого результата в реальном масштабе времени, особенно в случае их высокой размерности (NP-трудные проблемы). К таким задачам относятся, в частности, задачи принятия решений в предметных областях, связанных с диагностикой, прогнозированием процессов, управлением сложными системами.

Помочь человеку в решении этих задач способна *система интеллектуальной поддержки*, объединяющая возможности компьютера со знаниями и опытом эксперта в такой форме, что она способна предложить «разумный» совет или осуществить «разумное» решение поставленной задачи. Такая система должна уметь рассуждать при сомнительных, неполных данных, объяснять ход своих рассуждений понятным пользователю способом, самообучаться и адаптироваться к конкретным условиям применения.

Концепция системы, основанной на знаниях (СОЗ), позволяет добиться этой цели, обеспечивает открытость и масштабируемость системы, однако аналитический механизм получения решения в таких системах достаточно громоздкий и не всегда срабатывает в реальном масштабе времени.

Для решения данной проблемы предлагается использовать нейронную сеть, что значительно повышает производительность системы, кроме того, появляется возможность интегрировать аналитический модуль системы с нейросетевым, используя результаты работы аналитического модуля для дообучения нейронной сети.

*В первом разделе* данной работы поставлены основные цели и задачи.

*Во втором разделе* рассмотрены концепции представления знаний на инфологическом уровне, их преимущества и недостатки.

*В третьем разделе* описана концепция гибридной системы интеллектуальной поддержки, концепция системы, основанной на знаниях, которая является базовой для гибридной системы, а также основные компоненты гибридной системы.

*В четвертом разделе* рассмотрен многослойный персептрон - структура нейронной сети, на базе которой предлагается создать нейронную сеть для нейросетевого модуля гибридной системы. Также в этом разделе рассмотрены основные типы активационных функций нейронов.

*В пятом разделе* описан подход создания нейросетевого модуля гибридной системы интеллектуальной поддержки, его основные компоненты, а также механизм построения и структура нейронной сети, которые лежат в его основе. Также в этом разделе рассмотрены концепции дообучения сети.

*В шестом разделе* описана программная реализация модуля, его архитектура, диаграммы классов и детали реализации.

*В седьмом разделе* приведены результаты работы системы на тестовом базисе и сравнение результатов работы нейросетевого модуля с аналитическим модулем на тестовых примерах.

1. Постановка задачи

Основная цель данной работы - создать модуль, позволяющий принимать решения в различных предметных областях на основании нейросетевого механизма вывода и, при необходимости, использовать результаты работы аналитического модуля для дообучения нейронной сети.

Для достижения данной цели поставлены следующие задачи:

1. **Разработать подход к созданию нейросетевого модуля гибридной системы интеллектуальной поддержки.** В рамках данной задачи, необходимо рассмотреть концепцию создания системы, основанной на знаниях и гибридной интеллектуальной системы. Разработать алгоритм преобразования знаний в эквивалентную по «рассуждениям» нейронную сеть, а также сформировать структуру нейронной сети и способ дообучения нейронной сети
2. **Разработать архитектуру нейросетевого модуля.** В рамках данной задачи необходимо описать основные компоненты нейросетевого модуля гибридной системы интеллектуальной поддержки, разработать программную архитектуру, диаграммы классов.
3. **Создать программную реализацию.** В рамках данной задачи необходимо определиться со средой разработки, языком, используемыми форматами файлов и библиотеками, написать и протестировать программу, создать тестовую инфраструктуру.
4. **Сравнить результаты работы нейросетевого и аналитического модулей гибридной системы.** В рамках данной задачи нужно сравнить полученные результаты двух модулей и проанализировать их.

2. Представление знаний о предметной области на инфологическом уровне

Знания системы могут формироваться на основе:

* рекомендаций эксперта;
* аналогий;
* примеров;
* наблюдений, открытий и экспериментов;
* умозаключений на базе глубинных представлений.

Неавтоматизированный сбор знаний экспертов – это трудоемкий процесс, плохо структурированный и носящий циклический характер. Одна его итерация состоит в получении от эксперта максимального количества знаний о каком-либо фрагменте предметной области и анализ их на непротиворечивость.

Для ускорения процесса накопления знаний и придания ему комфортного характера необходима разработка специального программного интерфейса приобретения знаний от эксперта. При создании такого интерфейса важно правильно определиться с языком описания знаний на инфологическом (содержательном) уровне. С одной стороны, этот язык должен быть достаточно формален, нагляден и удобен для эксперта; он должен позволять эксперту использовать свою терминологию на естественном языке. С другой стороны, этот язык не должен создавать проблем компьютеру с его распознаванием и интерпретацией, как, например, естественный язык.

2.1. Дерево решений как формализм описания знаний на инфологическом уровне

Анализ действий экспертов, связанных с диагностикой состояний сложных объектов и систем, показывает, что профессионал проводит экспертизу вполне целенаправленно, придерживаясь определенной стратегии. При этом он руководствуется множеством правил (эвристик), которые могут быть представлены в форме

ЕСЛИ *<посылка>* ТО  *<заключение>,*

называемой *продукционным правилом*.

Правило может быть задано с *коэффициентом уверенности*, определяющим положительным значением меньше или равном единице.

Указанным выше требованиям при описании знаний продукционного типа на инфологическом уровне в достаточной степени удовлетворяет традиционное *дерево решений*. Это бинарная древовидная структура с вершинами двух типов:

* *вершины условий* (*типа «овал»*), из которых могут исходить только две ветви (одна – по «Да», другая – по «Нет»);
* *вершины выводов* (*типа «прямоугольник»*), имеющие единственный выход (обязательно на вершину условия) для промежуточных выводов и не имеющие выхода (не продолжающие логику рассуждений) в случае окончательного вывода.

Каждая вершина условия содержит некоторое высказывание, которое может принимать значения «Истина» или «Ложь» (соответствует выходам по «Да» или «Нет»). Вершина вывода содержит одно или несколько предложений, описывающих некоторое промежуточное или окончательное заключение в виде набора фактов или ссылок на определенные действия. Корень дерева обозначается как вершина условия (овал), содержащая высказывание, с которого начинается процесс логических рассуждений.

Таким образом, дерево решений представляет собой структуру, описывающую логику рассуждений в предметной области. Для описания структурированных (с целью сужения пространства поиска) знаний может использоваться не одно, а несколько деревьев решений.

Данная структура дерева решений обладает следующими недостатками. Во-первых, она не ориентирована на описание нечеткой логики рассуждений эксперта. Во-вторых, заставляет дублировать фрагменты дерева решений в случаях прихода к ним с разных направлений.

2.2. Граф решений

*Граф решений* представляет собой ориентированный помеченный граф с вершинами двух типов:

* *вершины условий* (*типа «овал»*), из которых могут исходить несколько ребер (по «Да» или «Нет» со своими коэффициентами уверенности, помечающими эти ребра);
* *вершины выводов* (*типа «прямоугольник»*), имеющие один вход и один выход (обязательно на вершину условия) для промежуточных выводов и не имеющие выхода (не продолжающие логику рассуждений) в случае окончательного вывода.

Каждая вершина условия (вершина-овал) содержит некоторое высказывание, которое может принимать значения «Истина» или «Ложь» с различными коэффициентами уверенности (каждому коэффициенту уверенности соответствует помеченный им выход).

Вершина вывода содержит одно или несколько предложений, описывающих некоторое промежуточное или окончательное заключение в виде набора фактов или ссылок на определенные действия. Вершина вывода помечается (например, в нижнем правом углу) *коэффициентом уверенности*, определяющим положительным значением меньше единицы степень адекватности (правдоподобия) вывода условиям его активации.

Граф имеет одну корневую вершину-овал, содержащую высказывание, с которого начинается описание логики рассуждений. Некоторые из вершин-овалов могут иметь несколько входных ребер с разных направлений, не приводящих к появлению обратных связей (исключением является корневая вершина-овал, не имеющая ни одного входного ребра).

Таким образом, граф решений позволяет описать более глубокую, нечеткую логику рассуждений эксперта в предметной области. Для описания структурированных (с целью сужения пространства поиска) знаний может использоваться не один, а несколько графов решений.

3. Гибридная система интеллектуальной поддержки процессов принятия решений

3.1. Концепция системы, основанной на знаниях

Под *интеллектуальной системой* (ИС, или IS — Intellectual System)принято понимать программно-техническую систему, способную решать задачи, традиционно считающиеся творческими, принадлежащие конкретной предметной области, знания о которой хранятся в памяти интеллектуальной системы.

*Экспертная система* (ЭС, или ES — Expert System) – система, объединяющая возможности компьютера со знаниями и опытом эксперта в такой форме, что система может предложить «разумный совет» или осуществить «разумное решение» поставленной задачи. При этом система способна пояснить «ход своих рассуждений» в понятной для пользователя форме.

По определению Эдварда Фейгенбаума (Станфордский университет), *система, основанная на знаниях* (СОЗ, или на английском knowledge-based system), – это интеллектуальная компьютерная программа, использующая знания и процедуру вывода для решения проблем, которые настолько сложны, что требуют привлечения эксперта.

С точки зрения функционального назначения, понятия СОЗ, ИС и ЭС можно рассматривать как синонимичные. Однако в определении СОЗ явно прослеживается базовый принцип организации системы *декларативного* (не предписывающего) типа. Именно этот принцип, предполагающий четкое отделение друг от друга базы знаний и механизма вывода, максимально обеспечивает модульный принцип построения, открытость системы, возможность создания оболочек экспертных систем (empty expert systems), настраиваемых через формализм базы знаний на различные предметные области. В остальном СОЗ присущи все особенности, характерные для ИС (ЭС):

* ограниченность определенной областью экспертизы;
* качественный характер выходных результатов;
* способность рассуждать при сомнительных, неполных данных;
* способность объяснять ход своих рассуждений понятным пользователю способом;
* способность самообучаться и адаптироваться к конкретным условиям применения и многое другое.

Целесообразность создания и использования СОЗ в конкретной предметной (проблемной) области определяется следующими критериями:

* отсутствие строгих алгоритмов и существование эвристических методов (приемов) решения задач;
* наличие эксперта, способного решать задачи и объяснять ход их решения;
* надежность и статичность имеющихся знаний (статичность понимается в том плане, что есть сложившийся, устоявшийся костяк знаний);
* сомнительный характер доступных данных;
* в основе решения задач лежит метод логических рассуждений.

*Автономная* СОЗ традиционно призвана оказывать консультации и использует в основном эвристические приемы решения задач, не привлекая формальные методы моделирования, анализа и синтеза.

*Неавтономная* СОЗ выступает как *средство интеллектуальной поддержки процессов принятия решений* в различных предметных (проблемных) областях. Она может быть организована в виде интеллектуальной надстройки над прикладными программами, а может и интегрироваться с последними, придавая им интеллектуальность.

Основой для создания как автономных, так и неавтономных экспертных систем может служить *оболочка СОЗ* (empty knowledge-based system), реализующая определенные формы представления знаний, механизмы их приобретения и интерпретации с использованием процедурного анализа и метапроцедур, лежащих в основе интеллектуальной деятельности человека.

3.2. Нейросетевой подход к построению интеллектуальной системы

*Искусственная нейронная сеть* (ИНС) – это существенно параллельно распределенный процессор, обладающий способностью к приобретению, сохранению и репрезентации опытного знания и сходный с мозгом в двух аспектах:

* знание приобретается сетью в процессе обучения;
* для сохранения знания используются силы межнейронных связей (*синаптические веса*).
* Составными элементами ИНС являются спецпроцессоры двух видов:
* *искусственные нейроны* (или просто *нейроны*), суммирующие поступающие на их входы сигналы и преобразующие смещенную на величину порога сумму в соответствии с заданной активационной функцией нейрона;
* *связи между нейронами*, реализующие межнейронные взаимодействия в виде сигналов, умножаемых на синаптические веса связей.

Схематическое представление модели нейрона с входными связями изображено на рис. 1, где использованы следующие обозначения:

**** — сигналы, поступающие на входные связи нейрона;

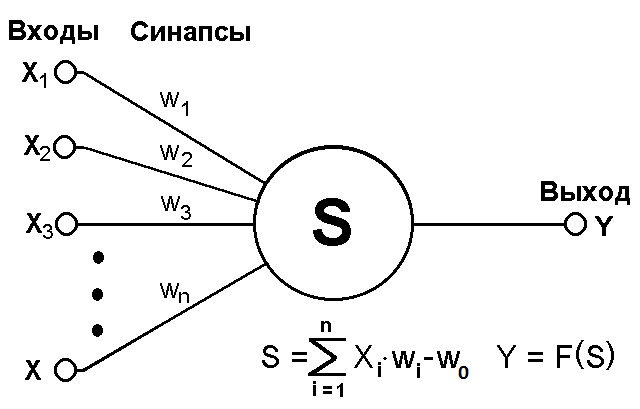
**** — синаптические веса входных связей;

**** — порог срабатывания нейрона;

****;

**** — активационная функция нейрона, реализующая тот или иной вид преобразования суммы взвешенных входных сигналов, смещенной на величину порога активационной функции ****.

Рис.1. Схематическое представление нейрона с входными связями



В нейронной сети между нейронами **** и **** могут быть установлены направленные связи **** и ****, каждой из которых присваиваются, соответственно, веса **** и **** (*синаптические веса*) (рис. 2).

Рис. 2. Межнейронные связи

...

...

(i, j)

*wji*

*wij*

(j, i)

Таким образом, объединяя нейроны между собой направленными связями, можно строить различные топологии (структуры) нейронных сетей и обучать их решению конкретных задач. *Обучить ИНС* – это значит подобрать такие значения синаптических весов ее межнейронных связей и, возможно, порогов и параметров активационных функций нейронов, при которых сеть приобретает способность выполнять желаемый вид преобразования.

3.3. Концепция гибридной системы интеллектуальной поддержки и её основные компоненты

В составе оболочки *гибридной* системы интеллектуальной поддержки процессов принятия решений предусмотрены следующие основные компоненты (рис. 3):

* исходно пустая база знаний;
* механизм вывода;
* подсистема приобретения знаний;
* механизм обучения нейронной сети;
* подсистема объяснения;
* интеллектуальный интерфейс.

*База знаний* включает две традиционные составляющие. Одна из них (*база правил*) объединяет в себе долговременные знания о предметной области в виде набора продукционных правил, получаемых от подсистемы приобретения знаний на этапе настройки оболочки системы на определенную предметную область. Другая составляющая базы знаний (*рабочая память*) представляет ее динамическую часть, в которой хранятся факты (оперативные данные), описывающие текущую ситуацию (состояние процесса принятия решений).

*Механизм вывода* представляет собой аналитическое ядро системы, выполняющее функцию построения прямой логической цепочки рассуждений с использованием стратегий поиска в глубину и в ширину.

*Подсистема приобретения знаний* в качестве главного механизма получения знаний от эксперта использует граф решений. Принимая на вход описание графа решений, она формирует в требуемом формате набор продукционных правил, который передает в базу правил.

Кроме этого подсистема приобретения знаний должна генерировать нейронную сеть, эквивалентную (по логике «рассуждений») исходному графу решений.

Достаточно обученная нейронная сеть при необходимости способна заменить механизм вывода и осуществлять существенно параллельный *нейровывод*. Широтой охвата пространства поиска нейровыводом можно управлять путем селективной активизации входных сигналов сети.

*Подсистема объяснения* призвана представить по требованию пользователя в понятном для него виде цепочку рассуждений, приведшую к тому или иному заключению (заключениям). При работе системы в аналитическом режиме рассуждений обеспечение такой функции не вызывает проблем. Сложнее дело обстоит с нейровыводом, хотя какие-то механизмы раскрытия логики «рассуждений» сети, очевидно, могут быть реализованы (об этом говорит пунктирная связь нейронной сети с подсистемой объяснения, отмеченная на рис. 3).

*Интеллектуальный интерфейс* объединяет в себе лингвистические, информационные и программные средства взаимодействия пользователя, инженера по знаниям и эксперта с соответствующими компонентами системы. Они должны быть ориентированы на неподготовленного пользователя, обладать способностью настраиваться на его терминологию и создавать по возможности комфортные условия для работы в системе.

Рабочая память

База правил

Подсистема приобретения знаний

Интеллектуальный интерфейс

Пользователь

Инженер по знаниям

Эксперт

Подсистема объяснения

Механизм вывода

База знаний

Рис.3. Схема взаимодействия компонент оболочки гибридной системы интеллектуальной поддержки

Алгоритм формирования продукционных правил

Алгоритм формирования нейронной сети

Граф решений

Решения нейронной сети

Решения механизма вывода

Механизм обучения нейронной сети

Ошибка

Нейронная сеть

Набор продукционных правил

4. Нейронная сеть как основа механизма вывода гибридной системы интеллектуальной поддержки

4.1. Многослойный персептрон

*Многослойный персептрон* представляет собой сеть прямого распространения с одним *входным* (*сенсорным*)слоем, одним *выходным* (*моторным*) слоем и одним или несколькими *внутренними* (*скрытыми, связующими*) слоями нейронов. Характерной его чертой является *прямонаправленность*: информация, преобразуясь, передается от входного слоя через скрытых слоев к выходному слою.

В стандартной (регулярной) топологии персептрона каждый нейрон   слоя   непосредственно воздействует с синаптическими весами  на все нейроны   следующего слоя  и ни на какие другие (– соответственно, число нейронов в слое  и в слое ;  и  обозначают, соответственно, входной и выходной слои). Схема многослойного персептрона стандартной топологии представлена на рис. 4.

**..................**

**.......................**

**................**

**.......................**

*х1*

*х2*

*хN*

*y1*

*y2*

*yM*

Входной

слой

Скрытые слои

Выходной слой

Рис. 4. Многослойный персептрон стандартной топологии

**.................................**

Модифицированные версии многослойного персептрона могут иметь нерегулярные связи между слоями (какие-то связи могут отсутствовать) и прямые (непосредственные) связи между нейронами несмежных слоев.

Входной слой персептрона выполняет функции приема и ретрансляции входных сигналов  на нейроны первого скрытого слоя. Основное нелинейное преобразование сигналов происходит в скрытых слоях. Нейроны выходного слоя осуществляют суперпозицию взвешенных сигналов последнего скрытого слоя (имеют линейную активационную функцию) или выполняют нелинейное преобразование, как и нейроны скрытых слоев.

*Обучение персептрона* – это итеративный целенаправленный процесс изменения значений весов синаптических связей (и, возможно, порогов активационных функций нейронов), реализуемый «внешней» процедурой (*алгоритмом обучения*) с использованием тренировочных шаблонов (*обучение с супервизором*) до тех пор, пока сеть не приобретет желаемые свойства.

*Тренировочный шаблон* представляет собой пару векторов , один из которых  – вектор известных входных сигналов, другой  – вектор желаемых выходных сигналов. В процессе обучения векторы  из тренировочного набора  последовательно подаются на вход ИНС и для каждого из них оценивается ошибка между фактическим  и желаемым  откликом (выходом) сети . Затем определяется общая ошибка , на основании которой (а возможно, и с использованием данных о предыдущих итерациях) алгоритм обучения осуществляет модификацию значений настроечных параметров сети, направленную на уменьшение ошибки. Как вариант, модификация значений варьируемых параметров сети может осуществляться после оценки действия каждого очередного шаблона, т.е. по «локальной» ошибке . Процесс обучения повторяется до тех пор, пока сеть не приобретет способность выполнять желаемый тип преобразования, заданный тренировочным набором шаблонов .

В результате обучения сеть достигает способности правильно реагировать не только на шаблоны множества , но и на другие допустимые возмущения, с которыми она никогда ранее не имела дела. В этом состоит уже упоминавшееся *свойство ИНС обобщать*. Ошибка в обобщении, которая может иметь место, определяется двумя причинами:

* ограниченностью размеров сети, ведущей к недостаточному качеству аппроксимации;
* ограниченностью объема обучающей выборки.
  1. Активационные функции нейронов

Работа и обучение нейронной сети зависят не только от её топологии, но и от конфигурации каждого отдельного нейрона, в частности, от активационных функций нейронов и их параметров.

**** — активационная функция нейрона, реализующая тот или иной вид преобразования суммы взвешенных входных сигналов, смещенной на величину порога активационной функции ****.

Графические представления некоторых из возможных видов активационных функций приведены на рис. 5.

Рис. 5а представляет *линейную* активационную функцию

.

На рис. 5б изображена *линейная с насыщением* активационная функция

,

имеющая настроечный параметр .

На рис. 5в представлена *ступенчатая* активационная функция (ее часто обозначают )

.

*a*

*s*

*f*(*s*)

- *a*

- *a*

*a*

*б)*  Линейная с насыщением активационная функция

Рис. 12. Некоторые из возможных видов активационных функций

0

*s*

*f*(*s*)

0

1

*s*

*f*(*s*)

*в)*  Ступенчатая активационная функция

*а)*  Линейная активационная функция

0

*f*(*s*)

*s*

1

*г)*  Сигмоидальная активационная функция

0,5

*-σ*

1

*f*(*s*)

*s*

*a*

*д)*  Радиально-симметричная активационная функция

0

0

*K*

*s*

*f*(*s*)

*е)*  Пороговая активационная функция

*σ*

Рис. 12 (продолжение). Некоторые из возможных видов активационных функций

*Сигмоидальная* активационная функция, описываемая выражением ,

изображена на рис. 5*г*. При малых значениях настроечного параметра  функция является более пологой, нежели при больших, приближающих ее к единичной ступеньке.

Рис. 5*д* представляет *радиально-симметричную* активационную функцию

,

с настроечным параметром **.**

Один из самых распространенных видов активационной функции, используемый для многослойных персептронов – это сигмоидальная функция. Её преимущество в том, что сигмоидальная функция дифференцируема на всей оси абсцисс и диапазон. Непрерывность первой производной позволяет обучать сеть градиентными методами (например, методом обратного распространения ошибки).

Сигмоид сужает диапазон изменения так, что значение выхода нейрона лежит между нулем и единицей. Его дополнительное преимущество состоит в контроле усиления. Для слабых сигналов выходная кривая имеет наклон, дающий большее усиление. Когда величина сигнала становится больше, усиление падает. Так большие сигналы воспринимаются сетью без насыщения, а слабые сигналы проходят по сети без сильного ослабления.

1. Нейросетевой модуль гибридной системы интеллектуальной поддержки

5.1. Основные компоненты нейросетевого модуля

Нейросетевой модуль гибридной системы должен содержать:

* Интеллектуальный интерфейс для эксперта или иной компонент для получения знаний в том виде, который пригоден для построения нейронной сети
* Интерфейс для взаимодействия с пользователем, получения от него данных о состоянии системы и вывода решения
* Компонент построения нейронной сети по входному графу решений в заданном формате
* Нейронная сеть, построенная по графу решений, сохраняемая на носителе
* Компонент дообучения нейронной сети, который способен использовать результаты, генерируемые аналитическим модулем системы

Высокоуровневая архитектура основных компонентов и их связь и взаимодейтсвие с компонентами аналитического модуля отражена на рис. 6.

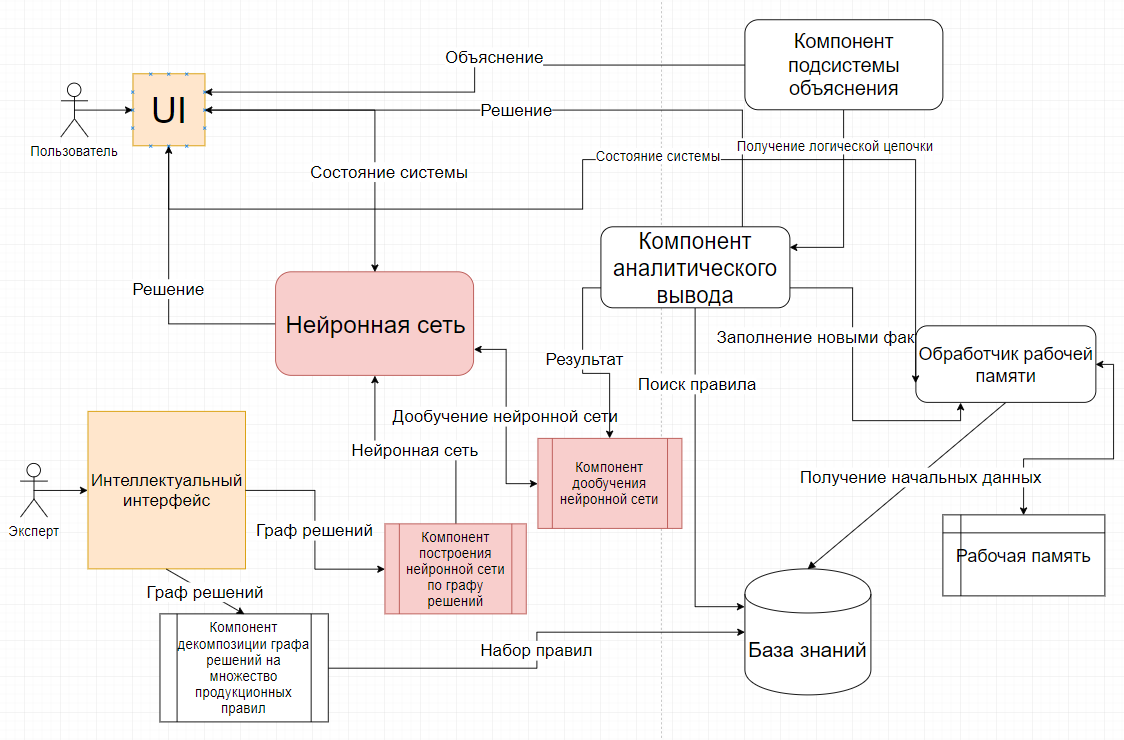


Рис. 6 – Высокоуровневая архитектура компонентов гибридной системы интеллектуальной поддержки. Выделенные компоненты – основные в нейросетевом модуле

* 1. Структура и топология нейронной сети

Нейронная сеть, которая строится по графу решений, должна быть эквивалентна по логике «рассуждений» исходному графу решений. Значит необходимо сопоставить какие-то элементы топологии графа, или его структурные элементы, некоторым элементам нейронной сети.

Нейронная сеть, которая строится по графу решений, представляет собой многослойный персептрон нерегулярной структуры. Это значит, что каждый нейрон каждого слоя может быть связан только с последующими слоями, но не каждый нейрон очередного слоя связан с каждым нейроном следующего слоя. Данная топология нейронной сети выбрана как наиболее отражающая логику «рассуждений» исходного графа.

* + 1. Сенсорный слой

Абсолютно логично поставить в соответствие вершинам-условиям (овалам) графа решений нейроны сенсорного слоя нейронной сети, т.к. вершины-условия используются для получения информации о текущем состоянии системы.

Каждая исходящая связь из вершины-условия должна породить собственный нейрон для реализации нечеткой логики, чтобы у пользователя была возможность получить рассуждения сразу по нескольким фактам, в случае сомнений.

Логично использовать коэффициенты уверенности каждой связи, исходящей из вершины-условия как синаптический вес соответствующей исходящей связи из нейрона. Однако просто использовать коэффициенты уверенности как синаптические веса нельзя, т.к. в таком случае рассуждения, связанные с коэффициентами уверенности близкими к нулю проводиться не будут, а это делает итоговое решение неполным и исключает из него часть рассуждений (например факт «Нет искры» – это «Есть искра» с коэффициентом уверенности 0)

В связи с этим предлагается задавать синаптические веса исходящих связей нейронов сенсорного слоя равными 1. А для решения проблемы, связанной с просчетом нужных направлений рассуждений, использовать специальную активационную функцию у нейронов сенсорного слоя:

*,* где *–* порог активационной функции, равный коэффициенту уверенности эксперта на соответствующей исходящей связи вершины-условия*, s –* коэффициент уверенности факта, который задает пользователь

* + 1. Скрытые слои

Скрытые слои нейронной сети – промежуточное звено, предназначенное для «рассуждений». Логично поставить в соответствие нейронам скрытых слоёв сети вершины промежуточных выводов, а также вершины окончательных выводов, сделав нейроны моторного слоя ретранслирующими.

Связи между нейронами распределяются соотвественно связям в графе.

Активационная функция нейронов скрытых слоёв – сигмоидальная функция c фиксированным порогом. Такой фиксированный порог необходим для того, чтобы функция выдавала корректные результаты в диапазоне [0,1], а не в диапазоне [-0.5,0.5]. С помощью настроечного параметра a можно управлять крутизной сигмоиды, настраивая работу сети.

*,* где *a* – настроечный параметр функции

* + 1. Моторный слой

Выходной слой сети является ретранслирующим, то есть активационная функция каждого нейрона – линейная с параметром 1. Каждый нейрон моторного слоя должен соответствовать каждой вершине окончательного вывода в исходном графе решений. Значения, полученные в результате работы сети – это и есть решение, проанализировав его, пользователь может сделать выбор.

* 1. Алгоритм построения нейронной сети по графу решений

Предлагается следующий алгоритм построения нейронной сети по графу решений:

**ШАГ 1** Пройти по всем вершинам графа, если это эллипс - найти все связи, выходящие из него, добавить нейрон с порогом активационной функции, равным коэффициенту уверенности связи, на сенсорный слой. Записать id узла равным id вершины в графе. **ПЕРЕЙТИ НА ШАГ2**

**ШАГ 2** Для каждого нейрона сенсорного слоя выполнить:

**ШАГ 2.0** Найти вершину графа с id равным id текущего нейрона. Найти все вершины, связанные с этой вершиной. Для каждой из них **ВЫПОЛНИТЬ РЕКУРСИВНУЮ ФУНКЦИЮ TRAVERSE** с параметрами index = 0, weigth = вес соответствующей связи, sender = текущий нейрон сенсорного слоя, current = текущая вершина.

**Функция traverse**:

Функция принимает на вход *index* – номер текущего скрытого слоя нейронной сети, *sender* – нейрон, для которого новая связь будет исходящей. *weight* - вес связи между *sender* и очередным нейроном, *current – вершина, для которой выполняется обход*

**ШАГ 2.1** Найти вершину графа с id равным id нейрона sender. Если тип этой вершины – эллипс, то **ПЕРЕЙТИ НА ШАГ 2.2**, иначе проверить количество скрытых слоев сети, если оно меньше index + 1, то добавить новый скрытый слой. **ПЕРЕЙТИ НА ШАГ 2.1.0**

**ШАГ 2.1.0** Определить количество связей найденной вершины. Если количество связей равно нулю, то **ПЕРЕЙТИ НА ШАГ 2.1.1**, иначе **ПЕРЕЙТИ НА ШАГ 2.1.2**

**ШАГ 2.1.1** Проверить, нет ли на скрытых слоях нейрона с id равным (id вершины + 'motor') если есть, то **ПЕРЕЙТИ НА ШАГ 2.1.1.1**, иначе **ПЕРЕЙТИ НА ШАГ 2.1.1.2**

**ШАГ 2.1.1.1** Создать нейрон на моторном слое, создать нейрон на скрытом слое с номером index. Создать связь между нейроном sender и нейроном скрытого слоя, присвоить связи вес weight. Создать связь между нейроном скрытого слоя и нейроном моторного слоя. Присвоить связи вес, который соответствует коэффициенту найденной на предыдущих шагах вершины.

**ШАГ 2.1.1.2** Найти нейрон скрытого слоя с id равным id вершины. Создать связь между нейроном, который вызвал функцию traverse и нейроном скрытого слоя, присвоить связи вес weight.

**ШАГ 2.1.2** Проверить, нет ли на скрытых слоях нейрона с id равным id вершины. Если нет, то **ПЕРЕЙТИ НА ШАГ 2.1.2.1**, иначе **ПЕРЕЙТИ НА ШАГ 2.1.2.2**

**ШАГ 2.1.2.1** Создать нейрон на скрытом слое с индексом index, создать связь между нейроном sender и нейроном скрытого слоя, присвоить связи вес weight. Для всех вершин, связанных с текущим, вызвать функцию traverse, передав в неё параметр index + 1 и weight = коэффициент узла-прямоугольника, sender = нейрон скрытого слоя.

**ШАГ 2.1.2.2** создать связь между нейроном sender и нейрономом скрытого слоя, присвоить связи вес weight

**ШАГ 2.2** Для всех вершин, связанных с найденной вершиной, вызвать функцию traverse, передав в неё параметр index, weight, sender равные текущим.

* 1. Пример построения нейронной сети по графу решений из области диагностики неполадок сетевого соединения

1. Программная реализация

6.1. Общее описание программной реализации

Программная реализация нейросетевого модуля гибридной интеллектуальной системы написана на языке Python 3.5.6, использована среда разработки PyCharm Community Edition, использованы только стандартные модули-библиотеки языка, такие как “ElementTree”, “json”.

Разработка велась с использованием системы контроля версий Git, использовался репозиторий GitHub, средство для управления репозиторием – SourceTree.

Исходные коды основных модулей системы доступны в приложении и в репозитории, ссылка так же в приложении.

* 1. Архитектура программной системы
     1. Диаграмма классов компонента нейронной сети

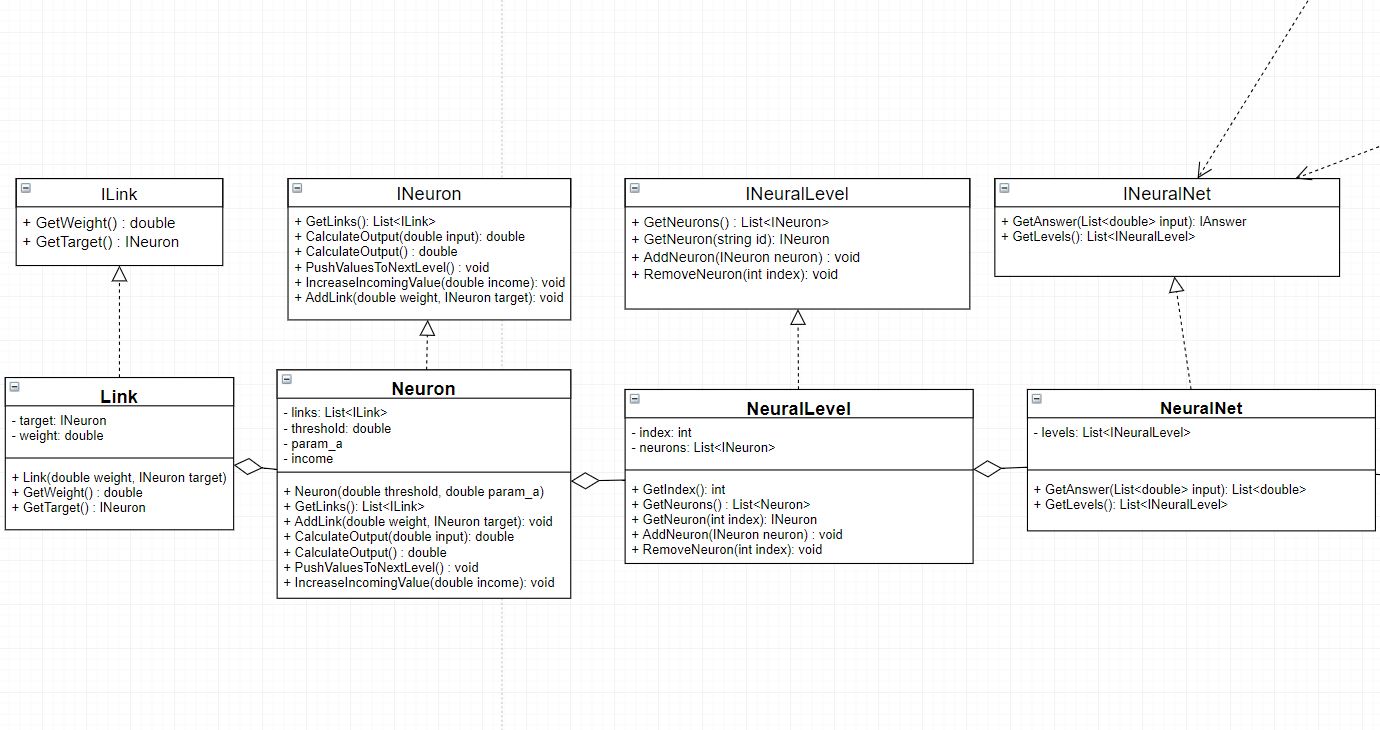


Рис. 7 – диаграмма классов компонента нейронной сети

*Основные компоненты*: NeuralNet – класс нейронной сети, агрегирует класс NeuralLevel – класс слоя нейронной сети - агрегирует Neuron – класс нейрона – агрегирует Link – класс связи нейронной сети

* + 1. Диаграмма классов компонента построения нейронной сети

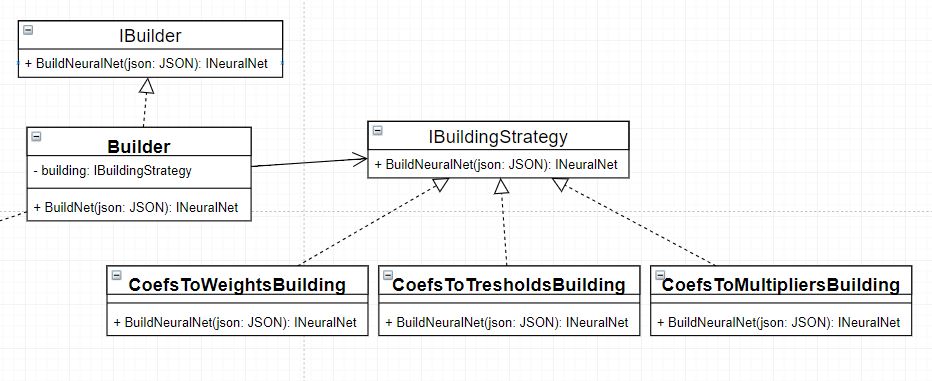


Рис. 8 – диаграмма классов компонента построения нейронной сети

Основные компоненты: Builder – модуль обучения нейронной сети, с помощью паттерна стратегия позволяет настраивать стратегию построения нейронной сети(IBuildingStrategy) – CoefsToThreshholdsBuilding– класс построения с помощью перевода коэффициентов уверенности в пороги активационных функций, CoefsToValuesBuilding – класс построения нейронной сети с помощью перевода коэффициентов уверенности в синаптические веса нейронов, CoefsToWeightsBuilding – класс построения нейронной сети с помощью перевода коэффициентов уверенности в пороги активационных функций.

* + 1. Диаграмма классов компонента дообучения нейронной сети

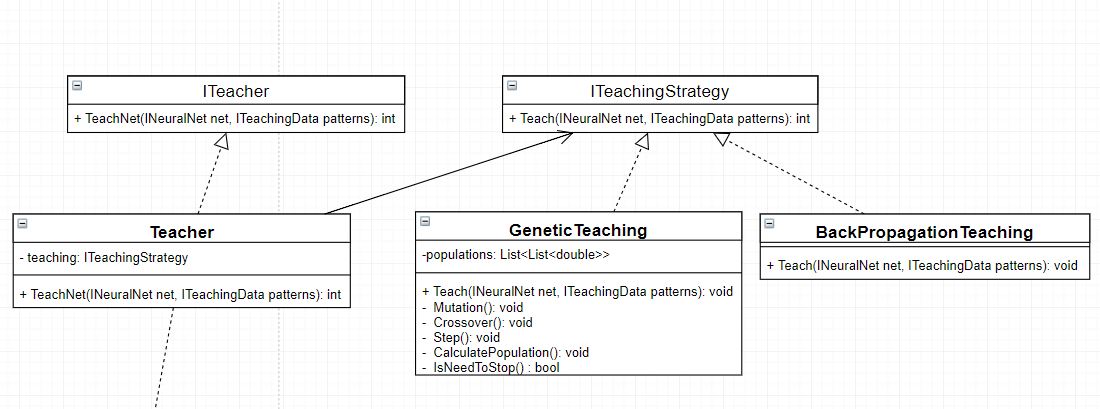


Рис. 9 – диаграмма классов компонента построения нейронной сети

Основные компоненты: Teacher – модуль обучения нейронной сети, с помощью паттерна стратегия позволяет настраивать стратегию обучения(ITeachingStrategy) – GeneticTeaching – класс обучения с помощью генетического алгоритма, или BackPropagationTeaching – класс обучения с помощью алгоритма обратного распространения ошибки

* 1. Интеллектуальный интерфейс

Интеллектуальный интерфейс реализован в приложении в виде парсера графа в формате XML, который преобразует это представление графа в json-представление, которое принимается компонентом построения нейронной сети. Для генерации XML с визуальным представлением графа решений, а также для рисования графа решений используется бесплатный сервис draw.io

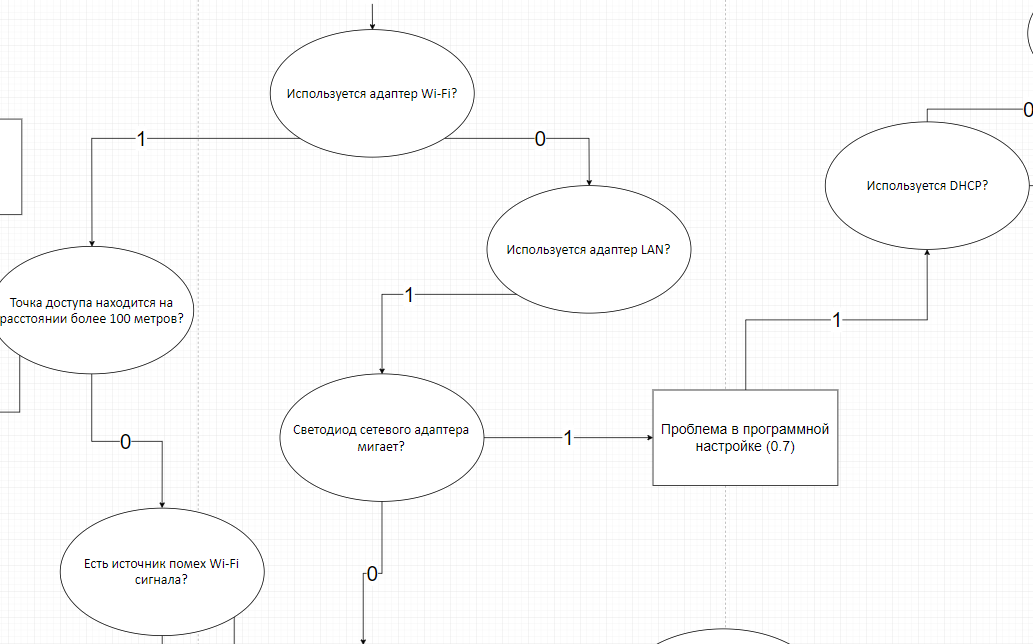


Рис. 10 – рисование графа решений, который преобразуется в json-представление

* 1. Компонент построения нейронной сети по графу решений

Данный компонент принимает на вход json заданного формата: корневой узел должен имееть два аттрибута: **nodes** – массив описаний узлов и **links** – массив описаний связей

* 1. Компонент нейросетевого вывода

1. Полученные результаты

7.1. Тестовый базис

7.1.1. Фрагмент графа решений для области диагностики неполадок сетевого подключения

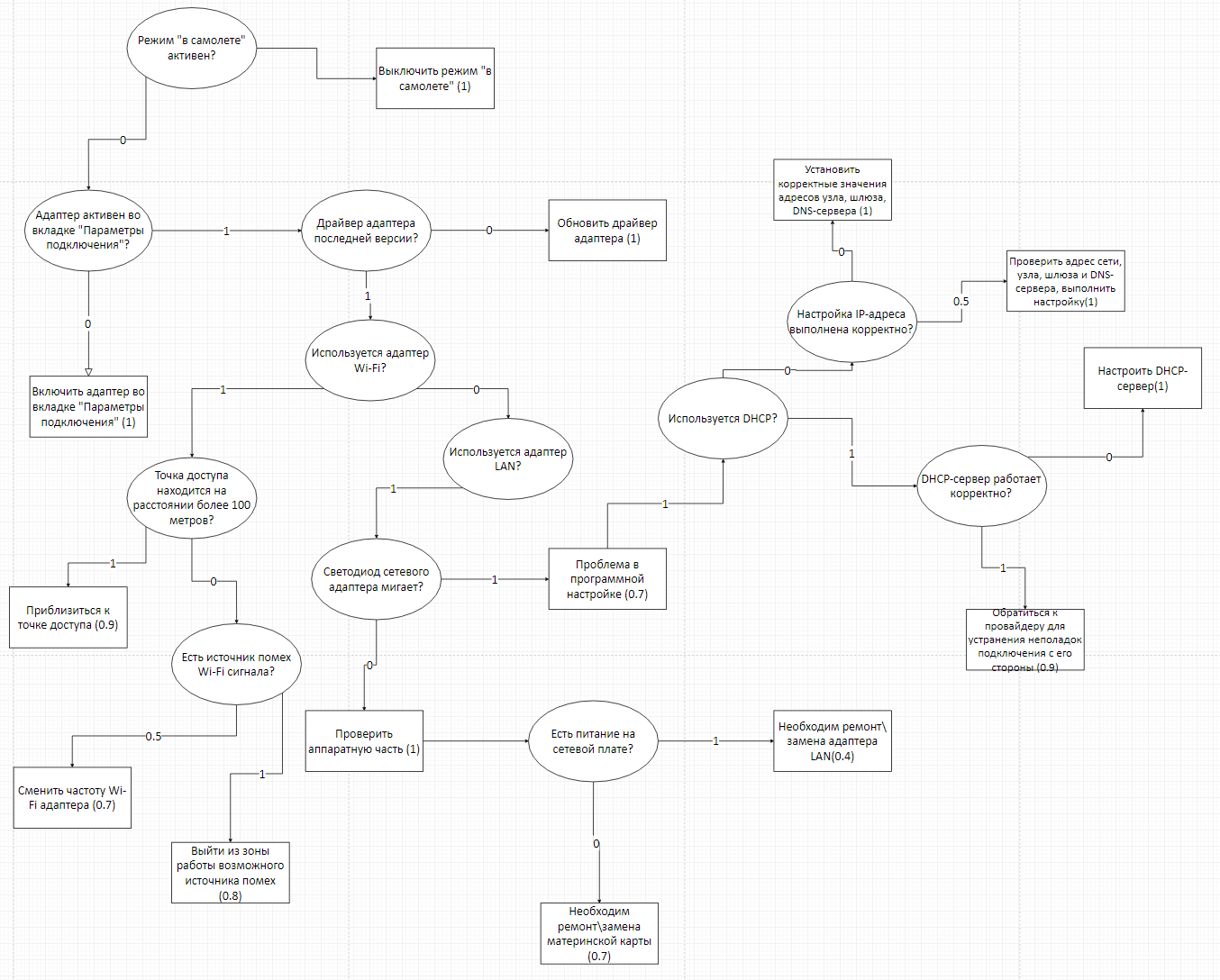


Рис. 11 – Фрагмент графа решений для области диагностики неполадок сетевого подключения

7.1.2. Результаты работы нейросетевого модуля

8. Сравнение результатов работы нейросетевого и аналитического вывода

Заключение

Список литературы

Приложение