**РУКОВОДСТВО ПОЛЬЗОВАТЕЛЯ**

Содержание

[1. Руководство программиста 3](#_Toc15081770)

[1.1 Общие сведения 3](#_Toc15081771)

[1.2 Функциональное назначение 3](#_Toc15081772)

[1.3 Описание логической структуры 3](#_Toc15081773)

[1.4 Используемые технические средства 7](#_Toc15081774)

[1.5 Входные данные 10](#_Toc15081775)

[1.6 Выходные данные 10](#_Toc15081776)

[2. Описание программы 11](#_Toc15081777)

# 1. Руководство программиста

## 1.1 Общие сведения

Программа имеет следующие свойства:

* Наименование основного файла: program.py
* Размер файла: 34,8 Кбайт
* Название продукта: РБФ-аппроксиматор
* Язык: русский

Для функционирования программы необходимо следующее программное обеспечение: 64-разрядная ОС Windows Vista или новее, либо 64-разрядная UNIX-подобная ОС, включая MacOS; желательно наличие установленного Python 3.6.

Программа реализована на языке Python; интерфейс реализован при помощи набора привязок графического фреймворка Qt PyQt5; Среда разработки – JetBrains PyCharm.

## 1.2 Функциональное назначение

Программа предназначена для аппроксимации и интерполяции сложных функциональных зависимостей многих переменных при помощи нейронных сетей. Алгоритмы, использующиеся для обучения нейронных сетей программы не подходят для решения задач классификации и кластеризации. Максимально допустимое количество элементов массива входных данных зависит от объема ОЗУ.

## 1.3 Описание логической структуры

Программа состоит из 8-ми модулей: rbf.py, содержащий классы РБФ-сетей; input.py, в котором реализована предобработка входных данных; program.py, в котором обрабатываются события интерфейса; newui.py, prederr.py, yesnodialog.py и predictwindow.py, код которых был автоматически сгенерирован при помощи PyQt5 и содержащих в себе описание окон программы и элементов в них. Ниже в таблицах 1-4 представлен список классов и публичных методов класса в каждом модуле, а также их описание.

Табл. 1. Содержание модулей, сгенерированных PyQt5

|  |  |  |
| --- | --- | --- |
| Классы | Методы | Назначение |
| Ui\_MainWindow (newui),  Ui\_Dialog (остальные) | setupUi | Создает окно и располагает на нем элементы интерфейса |
| retranslateUi | Изменяет названия элементов интерфейса |

Табл. 2. Содержание модуля input

|  |  |
| --- | --- |
| Функция | Назначение |
| readtrain | Открывает файлы, проверят их на правильность, пропускает указанное количество значений, убирает строки с пустыми значениями |
| \_getfile | Считывает данные из файла в зависимости от его расширения, проверяет заголовки столбцов на правильность |
| \_getnumbers | Возвращает количество входов и выходов сети |
| lincen | Генерирует указанное количество равноудаленных центров для каждого входа в диапазоне его значений |
| rancen | Генерирует указанное количество случайных центров для каждого входа в диапазоне его значений |
| \_getminmax | Возвращает минимальное и максимальное значение каждого столбца массива данных |

Табл. 3. Содержание модуля rbf

|  |  |  |
| --- | --- | --- |
| Класс | Метод | Назначение |
| RBF | \_gauss | возвращает значение функции Гаусса для входного вектора |
| \_laplace | возвращает значение функции Лапласа для входного вектора |
| \_cauchy | возвращает значение функции Коши для входного вектора |
| \_linear2 | возвращает значение кусочно-линейной функции для входного вектора |
| \_wigner | возвращает значение функции Вигнера для входного вектора |
| \_function\_decide | выбирает при расчетах функцию в зависимости от указанной |
| \_hidden\_layer | вычисляет значение матрицы выходов скрытого слоя сети |
| fit | обучение нейронной сети на основе выборок входных и выходных данных |
| predict | расчет значений на основе выборки входных данных |
| score | расчет погрешности рассчитанных значений по сравнению с исходными |
| TwoLayerRBF | Наследует методы RBF; перегружает методы fit, predict и \_hidden\_layer согласно альтернативной архитектуре | |
| \_sum\_layer | Возвращает декартово произведение элементов первого скрытого слоя |

Табл. 4. Содержание модуля program

|  |  |  |
| --- | --- | --- |
| Класс | Метод | Назначение |
| WarningWindow,  PredictedErr,  YesNo | \_\_init\_\_ | конструкторы классов |
| PredictWindow | getinfo | получает из основного окна программы конечный массив для сохранения в файл, заголовки столбцов и текущую используемую директорию |
| getotherinfo | получает из основного окна объект РБФ-сети, рабочие входную и выходную выборки, объект скейлера и флаги нормализации и денормализации |
| RBFInterface | \_rightdatacheck | проверяет загруженные данные на правильность, возвращает ошибку в случае неправильных данных |
| \_savedata | Сохраняет рассчитанные значения в файл |
| \_predict | Рассчитывает погрешность расчета; Если массив выходных значений не был загружен изначально, предлагает загрузить файл, где они есть |
| \_rightdatacheck | проверяет загруженные данные на правильность, возвращает ошибку в случае неправильных данных |
| \_changeimage | изменяет изображение РБФ при переключении типа РБФ в списке |
| \_changedir | изменяет директорию на последнюю открытую |
| \_gettrainfilenames | получает массив имен файлов для тренировочной выборки; блокирует или делает активной часть связанных элементов интерфейса |
| \_getworkfilenames | получает массив имен файлов для рабочей выборки; делает активной кнопку «Загрузить» |
| \_checkskip | активирует или отключает spinBox для пропуска данных обучающей выборки в зависимости от «галочки» |
| \_trainnorm | делает активным или неактивным выбор диапазона нормализации в зависимости от «галочки» |

Окончание таблицы

|  |  |  |
| --- | --- | --- |
| Класс | Метод | Назначение |
| RBFInterface | \_trainload | загружает, проверяет на правильность и обрабатывает обучающую выборку; делает активными или неактивными ряд элементов |
| \_workskip | активирует или отключает spinBox для пропуска данных рабочей выборки в зависимости от «галочки» |
| \_workload | Загружает, проверяет на правильность и обрабатывает рабочую выборку; делает активными или неактивными ряд элементов |
| \_centrainskip, \_centrand,  \_lincent | Делают активными или неактивными spinBox'ы при выборе способа генерации центров в зависимости от состояния соответствующих radioButton’ов |
| \_filecent | блокирует или делает активной элементы при выборе загрузки центров из файла |
| \_choosecentfiles | получает массив имен файлов для центров; делает активной кнопку «Загрузить» |
| \_centload | Загружает, проверяет на правильность и обрабатывает центры; делает активными или неактивными ряд элементов |
| \_fit | Если центры не были загружены из файла, генерирует центры; получает архитектуру, тип РБФ и заданный радиус; обучает ИНС |
| \_tranerr | рассчитывает погрешность обучения |
| \_save | Сохраняет в бинарном файле обученную РБФ и скейлеры |
| \_load | Загружает из бинарного файла объекты РБФ и скейлеров |
| \_predict | Рассчитывает значения на основе загруженной рабочей выборки и выводит их в отдельном окне. |

Вызов подпрограмм побочных модулей в управляющем модуле осуществляется стандартным способом путем подключения модулей в заголовке файла program и последующего вызова подпрограмм из них.

Основные вычисления в программе выполняются в модуле rbf, а все остальные предназначены лишь для удобства работы с ним. Поэтому имеет смысл рассмотреть алгоритмы работы только лишь самого расчетного модуля, приведенные далее на рис. 1 и 2.

## 1.4 Используемые технические средства

Программа предназначена для работы на персональном компьютере. Управление программой выполняется пользователем через графический интерфейс, отображаемый на экране дисплея. Манипуляция элементами интерфейса осуществляется компьютерной мышью и клавиатурой, входные и выходные данные загружаются и сохраняются на жестком диске, сменном носителе или других устройствах хранения информации.

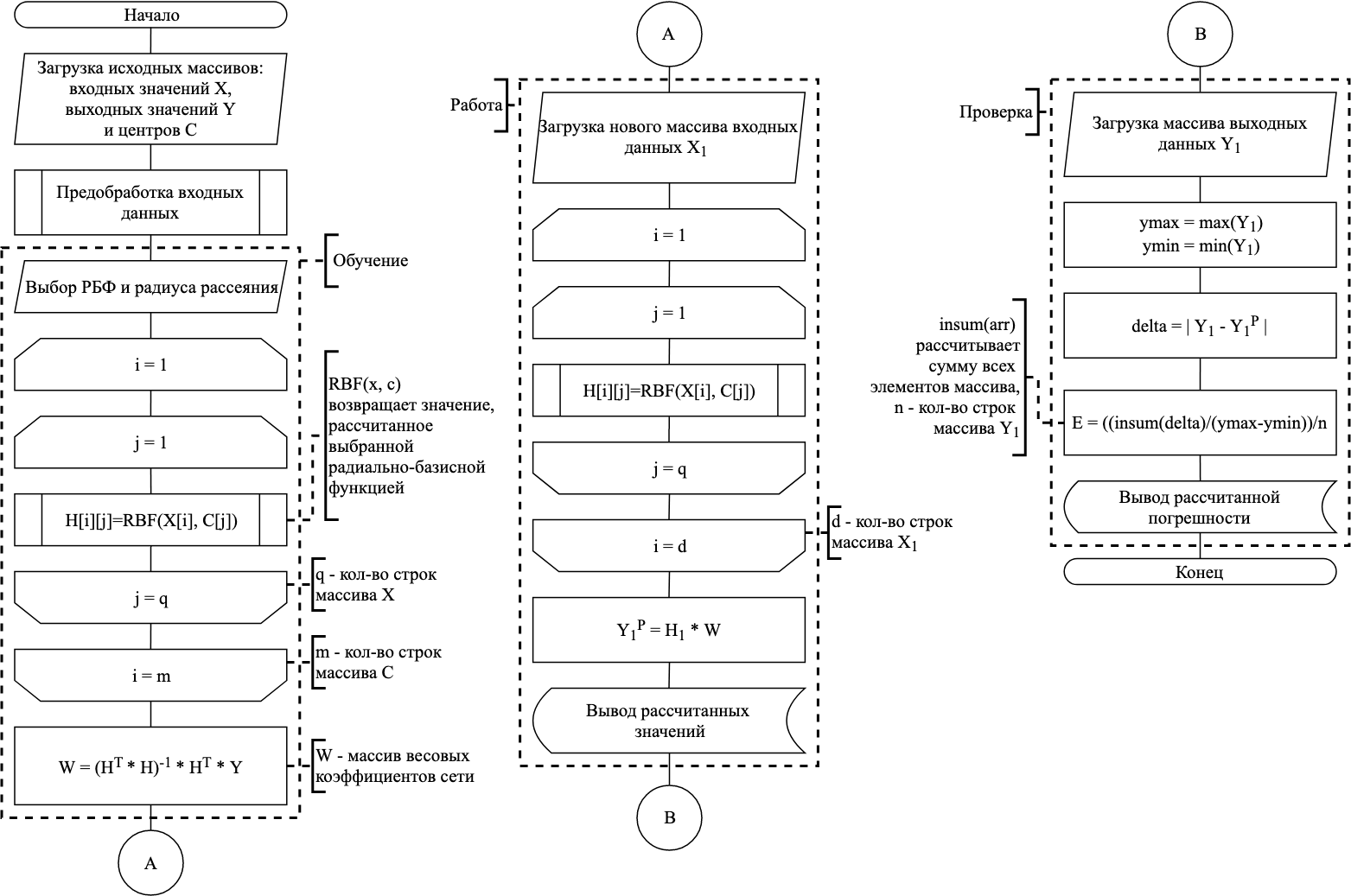


Рис. 1. Блок-схема работы классической РБФ-сети

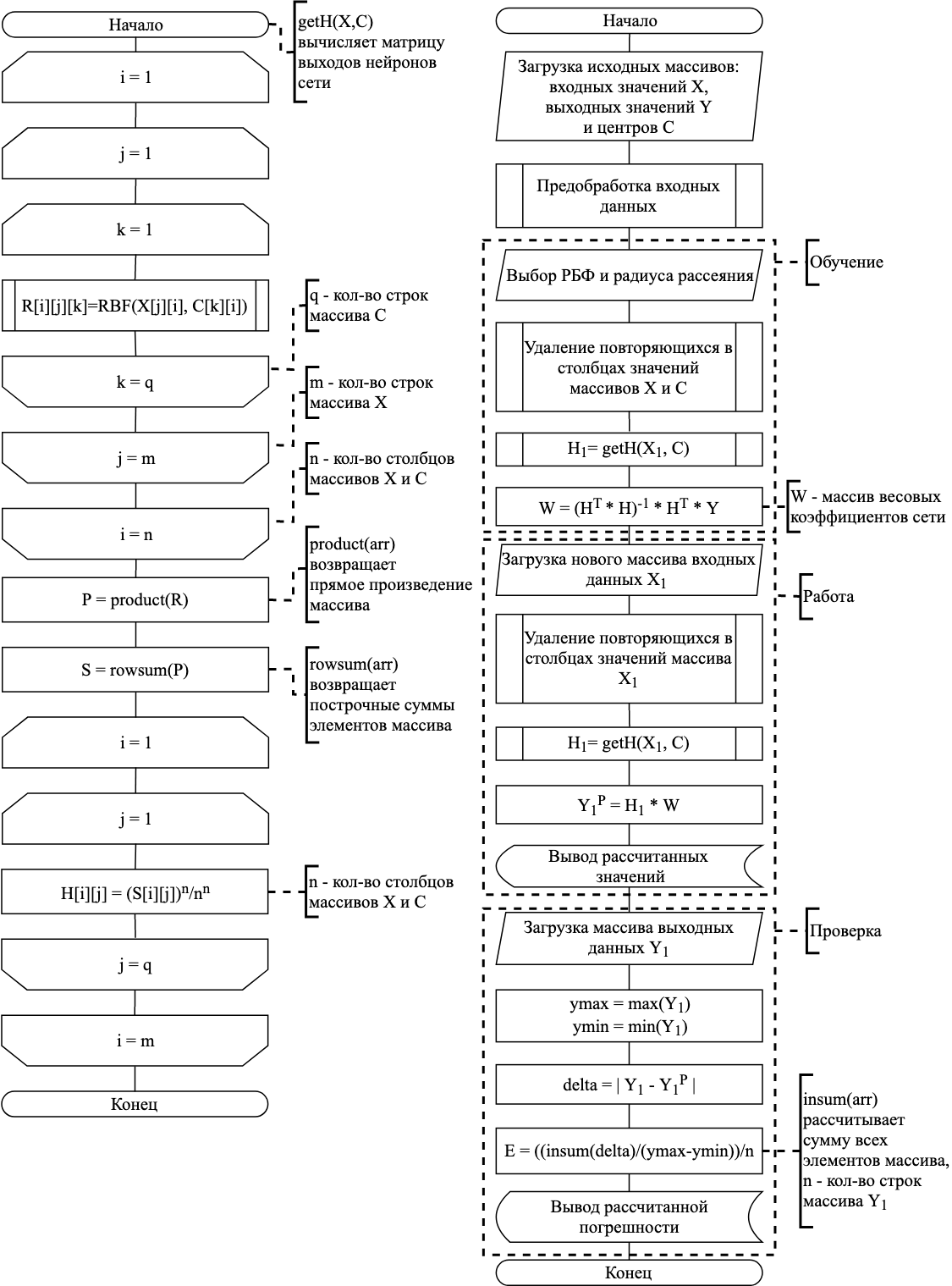


Рис. 2. Блок-схема работы альтернативной РБФ-сети

## 1.5 Входные данные

Входной файл представляет из себя таблицу, в которую сначала записаны независимые параметры, а затем – зависимые. Формат таблицы – файл с расширением .xls, .xlsx или .csv, либо текстовый файл формата .txt или .glo. В случае текстового файла, десятичным разделителем служит точка («.»), а числовым разделителем – знак табуляции. Столбцы параметров в файле должны быть озаглавлены: названия столбцов независимых переменных должны начинаться с латинской буквы «х», после которого идет порядковый номер столбца; аналогично озаглавливаются столбцы зависимых переменных, только начинаясь с латинской буквы «у». Пример записи входного файла приведен в табл. 5.

Табл. 5. Пример записи входного файла

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **x1** | **x2** | … | **xn** | **y1** | **y2** | … | **ym** |
| 0 | 0.1252 | … | -3.5162 | 15.25 | 0.311 | … | 2.2223 |
| 0 | 0.1252 | … | -1.087 | 15.25 | 0.311 | … | 2.4054 |
| 0 | 0.1252 | … | -0.3335 | 15.25 | 0.311 | … | 2.7113 |
| … | … | … | … | … | … | … | … |
| 1 | 1.1674 | … | -0.3335 | 6.6667 | 1 | … | 2.7113 |

## 1.6 Выходные данные

Выходные данные представляют из себя числовые результаты работы программы, выводимые сначала на экран персонального компьютера, которые затем имеется возможность сохранить на носитель в выбираемый каталог в одном из тех же форматов файлов, что и входные данные. Также имеется возможность сохранить обученную нейронную сеть и сопутствующие ей данные в бинарном файле с расширением .rbf.

# 2. Описание программы

Для более подробного ознакомления с функционалом программы рассмотрим сценарий работы с ней:

При запуске программы откроется главное окно, показанное ниже на рис. 3:

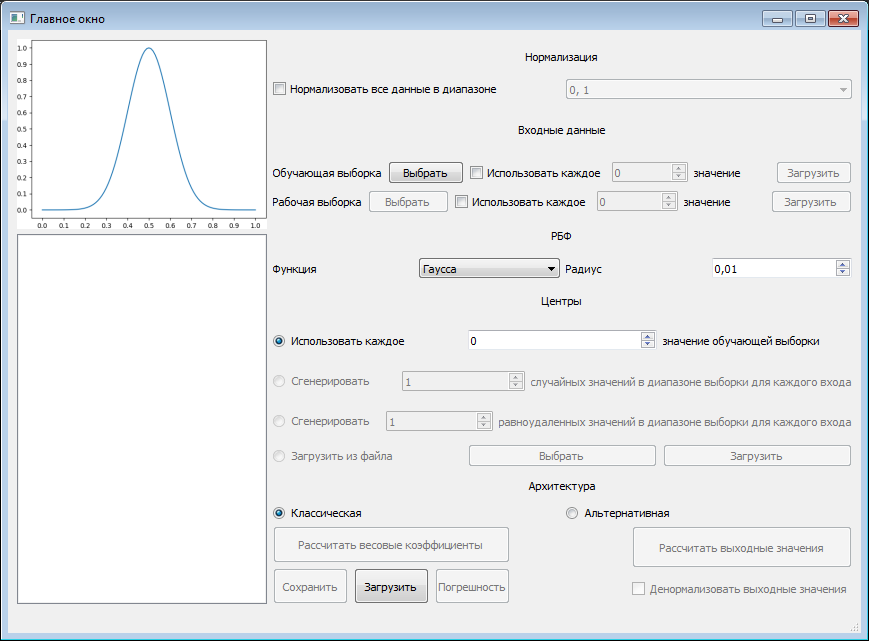


Рис. 3. Главное окно программы.

Поскольку на данный момент мы еще не загружали никаких данных, то большая часть элементов интерфейса недоступна для взаимодействия с пользователем.

Для загрузки входных данных рассмотрим поближе подгруппу элементов, с названием «Входные данные» (рис. 4).

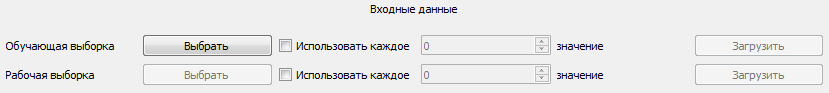


Рис. 4. Подгруппа «Входные данные».

Изначально у нас есть возможность выбрать файлы только для обучающей выборки. При нажатии на кнопку «Выбрать» откроется файловый менеджер операционной системы, в котором мы можем выбрать файлы с расширениями .glo, .txt, .xlsx, .csv. При этом в программе запоминается последняя директория, из которой пользователь выбрал файлы. Выбрав для загрузки интересующие нас файл или файлы, в поле в левом нижнем углу в строке состояния появится соответствующая запись (рис. 5):

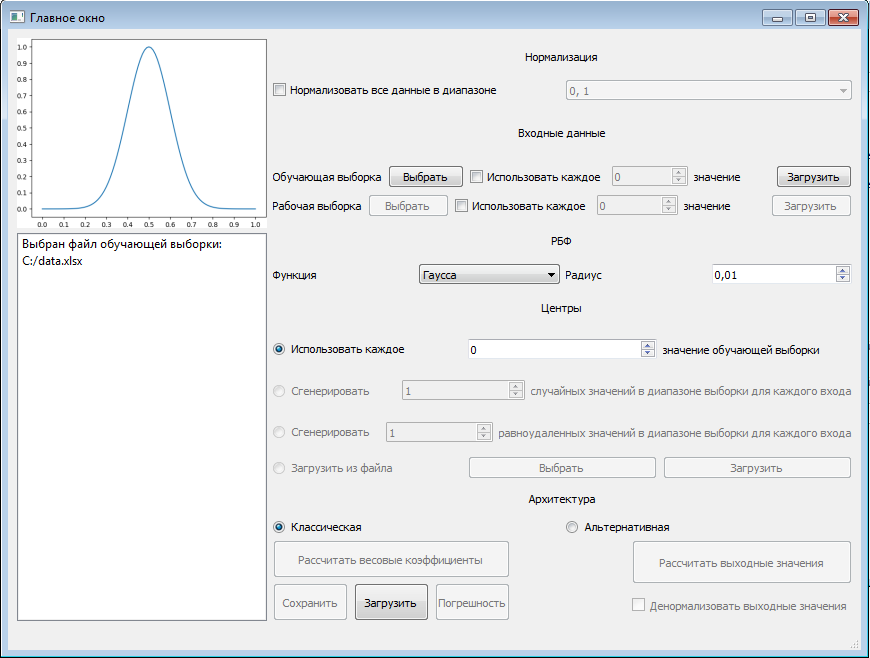


Рис. 5. Главное окно после выбора файла для обучения.

Перед загрузкой данных можно выбрать нормализовать загружаемые данные, отметив соответствующую галочку и выбрав диапазон нормализации (рис. 6):

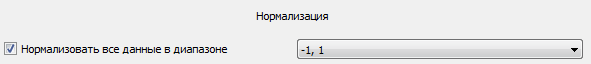


Рис. 6. Выбор нормализации данных.

Данные нормализуются в выбранный диапазон ([-1,1] или [0,1]) по выражению:

,

где – нормализованные входные и выходные значения, – i-е входное или выходное значение k-го примера выборки, *max* – максимальное значение диапазона нормализации (*max* = 1), *min* – минимальное значение диапазона нормализации (-1 или 0).

Также, если есть необходимость использовать только часть данных, то можно отметить галочку рядом с полем «Использовать каждое значение» и выбрать соответствующее число: 0 – используется каждое значение, 1 - значения берутся через одно, 2 – через 2 и т.д.

Определившись с необходимыми опциями по предобработке, можно жать кнопку «Загрузить». Программа обработает выбранные данные, и в случае успеха в строке состояния об этом будет уведомлено (рис. 7):



Рис. 7. Уведомление об удачной загрузке данных для обучения.

Если данные в выбранных файлах были оформлены не согласно указанным требованиям, то в строке состояния появится надпись «Ошибка», и выскочит окно, уточняющее природу ошибки. При этом данные загружены не будут и придется заново выбрать новые файлы для загрузки. Уведомления о возможных ошибках показаны на рис. 8 – 12.

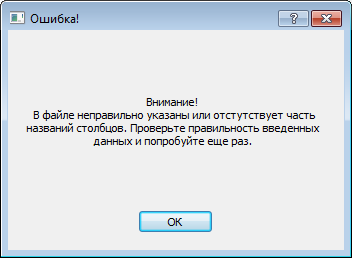


Рис. 8. Ошибка в названиях столбцов.

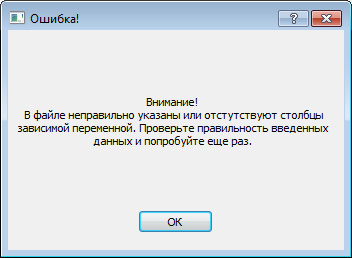


Рис. 9. Отсутствие зависимой переменной

в обучающей выборке.

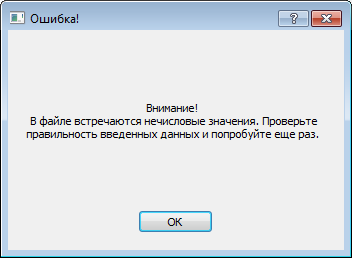


Рис. 10. Нечисловые значения в файле.

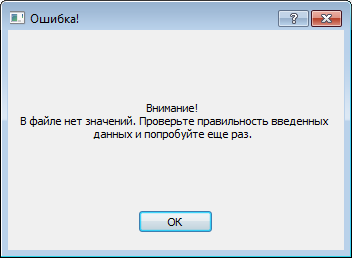


Рис. 11. Пустой файл.

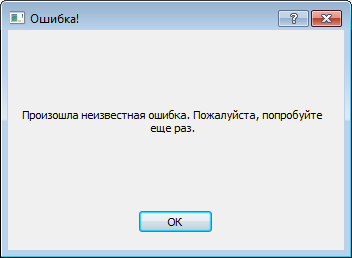


Рис. 12. Непредвиденная ошибка.

В файле могут содержаться пустые значения, однако при обработке поля с пустыми значениями будут удалены из конечной выборки.

Внимание! После нажатия кнопки «Загрузка», галочка для выбора нормализации станет более недоступной для пользовательского взаимодействия. При необходимости отменить, или наоборот выбрать нормализацию, придется снова выбрать файлы для загрузки и загрузить их повторно.

После успешной загрузки обучающей выборки станет доступной кнопка выбора файлов для рабочей выборки. Загрузка и обработка файлов здесь идентична описанному выше, правда есть дополнительная ошибка, связанная с новой выборкой (рис. 13):

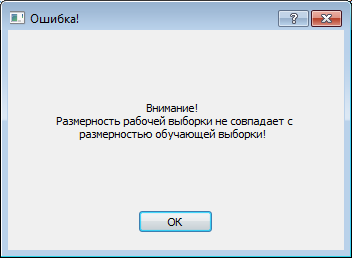


Рис. 13. Несовпадение размерностей выборок.

Загрузить рабочие данные при надобности можно и позже. После загрузки обучающей выборки можно перейти к разделу «РБФ» (рис. 14):

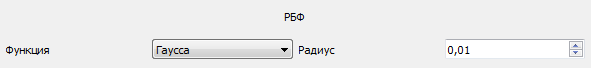


Рис. 14. Параметры функций активации.

Данная группа элементов отвечает за задание радиально-симметричных функций в нейронах среди функции Гаусса, Коши, кусочно-линейной, Вигнера и радиуса рассеяния r в диапазоне от 0,01 до 99,99. При смене функции из списка будет обновлен график РБФ в правом верхнем углу на график соответствующей функции (см. рис. 5).

Далее необходимо определиться со способом задания центров ИНС. Это можно сделать в соответствующем разделе (рис. 15):

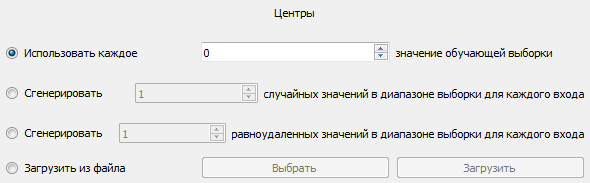


Рис. 15. Раздел «Центры».

При открытии программы доступен и активен только первый вариант (см. рис. 5), однако после загрузки обучающей выборки становятся доступными для взаимодействия остальные опции. Вместо ввода центров вручную было решено использовать загрузку из файла; загрузка центров из файла идентичная процессу загрузки остальных файлов, при этом есть дополнительный вариант ошибки (рис. 16):

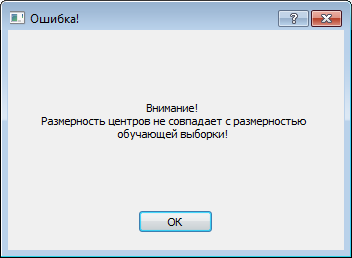


Рис. 16. Несовпадение размерности центров.

Далее идет выбор архитектуры ИНС (рис. 17); По умолчанию всегда выбирается классическая архитектура.



Рис. 17. Раздел «Архитектура».

Также после загрузки обучающей выборки становится доступна кнопка расчета весовых коэффициентов (рис. 18), при условии, что в качестве способа задания центров выбрана одна из первых трех опций; в случае выбора задания центров из файлов, кнопка «Расчет весовых коэффициентов» станет доступной только после успешной загрузки файла с центрами.

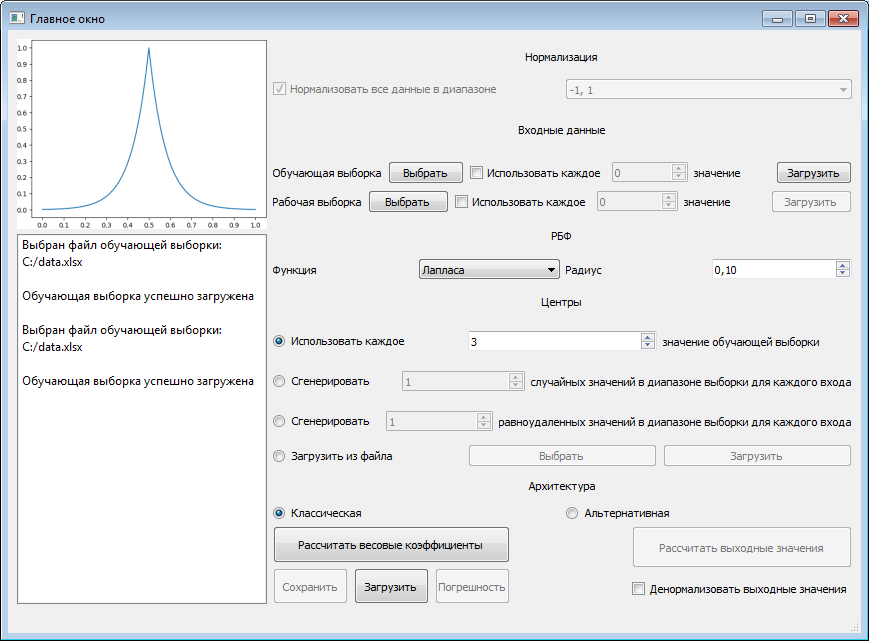


Рис. 18. Главное окно программы после загрузки обучающей выборки

и выбора центров.

После нажатия на эту кнопку начнется обучения нейронной сети с указанными параметрами. При больших объемах входных данных, программа на время расчетов может перестать отвечать, однако это не означает что она «зависла». После обучения, в строке состояния появится соответствующее уведомление, а также время обучения, в секундах или минутах (рис. 19):



Рис. 19. Уведомление об успешном обучении сети.

После этого станут доступными кнопки «Сохранить» и «Погрешность». При нажатии кнопки «Сохранить» откроется файловый менеджер и пользователь может указать путь для сохранения обученной нейронной сети, а также параметров обучения в файл формата .rbf. Нажатие кнопки «Погрешность» приведет к расчету погрешности обучения с выводом результата в строку состояния (рис. 20):



Рис. 20. Рассчитанная погрешность.

Погрешность рассчитывается по выражению:

,

где – погрешность одного выхода, , zi – значения функции обучающей выборки, z\*i – выходные значения нейронной сети, n – количество примеров выборки, m – количество выходов сети.

Кнопка «Загрузить», доступная изначально, позволяет открыть сохраненный ранее файл .rbf.

Для того, чтобы кнопка расчета выходных значений стала доступна, необходимо выполнение двух критериев: наличия рабочей выборки и обученной нейронной сети. Если также была выполнена нормализация данных, то станет доступна опция денормализации данных, выполняющейся по выражению:

.

После нажатия этой кнопки, откроется новое окно, содержащее в себе таблицу с выходными данными и кнопки «Сохранить результаты» и «Погрешность» (рис. 21). Также в главном окне будет уведомлено об успешном расчете и выведено время расчета.

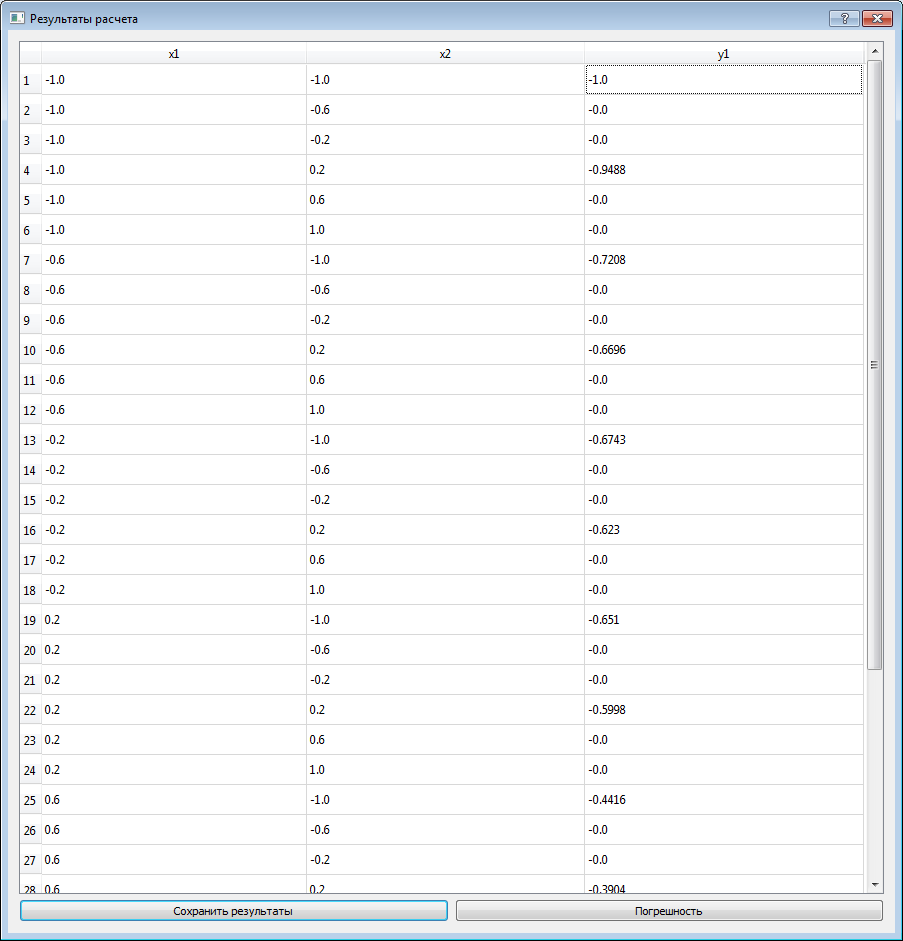


Рис. 21. Окно с результатами расчета.

Нажатие кнопки «Сохранить результаты» приведет к открытию файлового менеджера для сохранения выходных данных в форматах, указанных в п. 1. При нажатии кнопки «Погрешность» возможны 2 результата:

1. В файле рабочей выборки содержались зависимые переменные.

В этом случае погрешность расчета будет рассчитана сразу и результат будет выведен в новом окне (рис. 22):

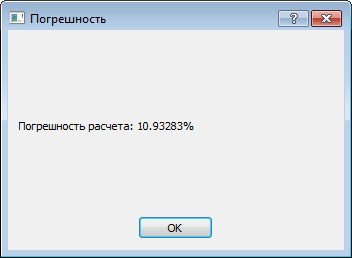


Рис. 22. Окно погрешности расчета.

1. В файле рабочей выборки не содержалось зависимых переменных.

При этом откроется окно, спрашивающее, хочет ли пользователь загрузить файл с расчетными переменными (рис. 23). Процесс загрузки идентичен всем предыдущим; в случае несоответствия загруженных данных с имеющимися, «выскочит» окно с ошибкой (рис. 24).

На этом функционал программы подходит к концу.

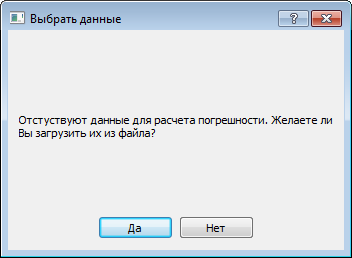


Рис. 23. Окно выбора файла с выходными параметрами.

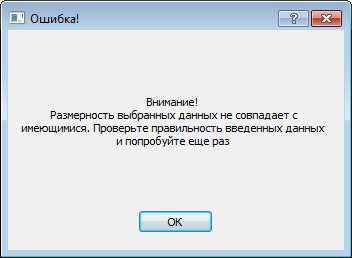


Рис. 24. Несоответствие файла с параметрами имеющимся данным.