

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«САМАРСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИМЕНИ АКАДЕМИКА С.П. КОРОЛЕВА»

Отчет по лабораторной работе №4  
по курсу  
«Интеллектуальные системы управления»

Выполнил: Иванов И.И., группа 6132

Проверил:  
к.т.н., доцент Солдатова О.П.

САМАРА 2023

Цель работы

Целью данной лабораторной работы является демонстрация способности нейронной сети решать задачи классификации при неполных или недостоверных данных. Сеть необходимо обучить классификации по пяти классам по 10 ‑ 20 количественным или качественным признакам.

Используемая модель

Многослойный персептрон

Модель многослойного персептрона представлена на рисунке 1.

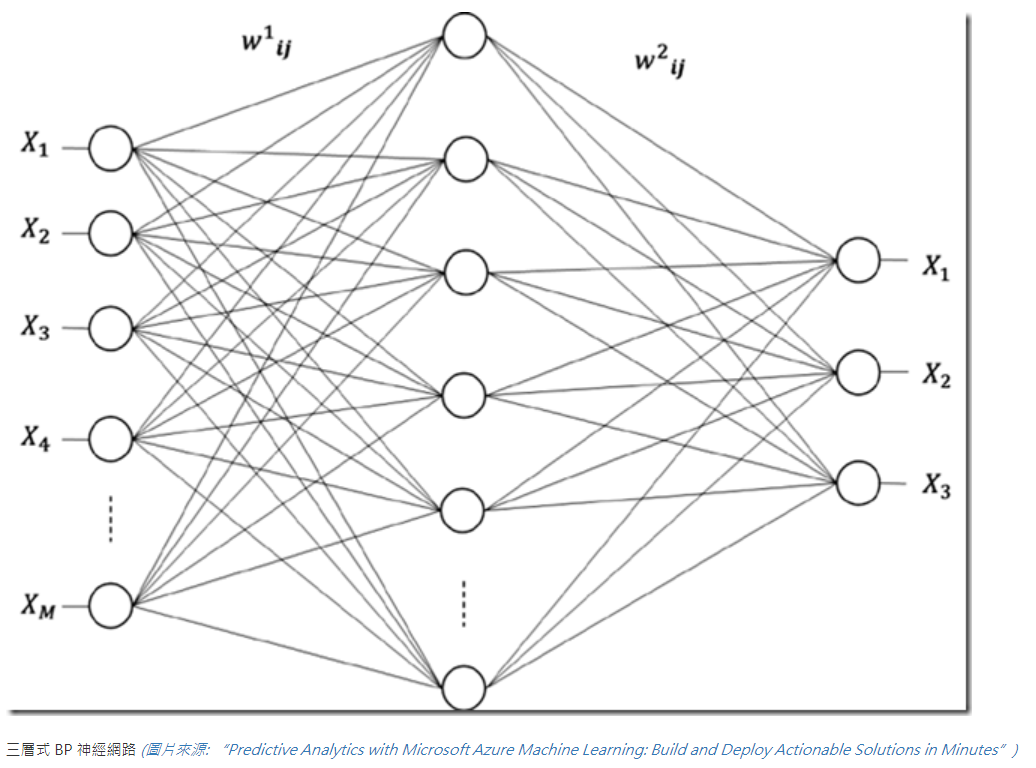


Рисунок 1 – Модель многослойного персептрона

Порядок выполнения работы

Данная лабораторная работа выполняется в несколько этапов:

1. Необходимо выбрать предметную область, отобрать числовые характеристики объектов и задать их диапазоны. Примеры типов предметных областей для классификации перечислены ниже.
2. Сгенерировать обучающую выборку размерностью от 30 до 50 примеров для каждого класса.
3. Написать программу, имитирующую работу многослойного персептрона и провести обучение сети по методу наискорейшего спуска с использованием алгоритма обратного распространения ошибки. Выбрать начальное значение числа скрытых слоёв, равным единице, число нейронов в скрытом слое – равным полусумме входов и выходов сети.
4. Провести обучение многослойного персептрона по алгоритму наискорейшего спуска с учётом моментов (1 занятие).
5. Продемонстрировать работу многослойного персептрона преподавателю (2 занятие).
6. Исследовать зависимость погрешности обучения для каждого алгоритма от значения коэффициента обучения.
7. Исследовать зависимость погрешности обучения и классификации для каждого алгоритма от объёма обучающей выборки.
8. Исследовать зависимость погрешности обучения и классификации для каждого алгоритма от числа нейронов скрытого слоя.
9. Исследовать зависимость погрешности классификации от числа итераций обучения.
10. Составить отчёт, который должен содержать постановку задачи, обучающие выборки, результаты исследований по пунктам 6 - 9 и сдать его преподавателю (3 занятие).

Предметная область

В качестве предметной области была выбрана классификация графических образов букв.Таким образом, нейронная сеть должна уметь распознавать и классифицировать образы черных латинских букв на белом фоне.

В качестве обучаемой выборки были выбраны 26 печатный латинских букв и 26 арабских латинских букв.

Пример латинских букв печатных и арабских печатных представлены на рисунке 2 и 3 соответственно.



Рисунок 2 – Печатная латинская буква



Рисунок 3 – Арабская печатная латинская буква

Обучающие выборки

В качестве обучающей выборки были заготовлены файлы .txt со списком пикселей латинских печатных и арабских букв. Перед обучением загружаются эти файлы и на основе данных выборок производится обучение нейронной сети.

Обучение нейронной сети

В качестве функции активации была выбрана сигмоидальная функция или сигмоид [2]:

|  |  |
| --- | --- |
|  | (1) |

В качестве алгоритма обучения были использованы методы обратного распространения ошибки (BackPROP), RPROP и iRPROP+.

1. Метод обратного распространения ошибки

Обратное распространение ошибки – это метод вычисления градиента, который используется при обновлении весов многослойного персептрона.

Цели обратного распространения просты:

* Отрегулировать каждый вес пропорционально тому, насколько он способствует общей ошибке;
* В результате итеративного уменьшения ошибки каждого веса получить ряд весов, которые дают хорошие результаты распознавания.

В основе алгоритма обратного распространения ошибки лежит применение метода градиентного спуска для поиска глобального минимума целевой функции [2]:

|  |  |
| --- | --- |
|  | (2) |

где – yj – эталонное значение нейрона в выходном слое,

dj – текущее значение нейрона в выходном слое.

Формула взвешенной суммы выглядит следующим образом [2]:

|  |  |
| --- | --- |
|  | (3) |

где – веса между нейронами,

– значения нейронов.

Формула активационного состояния нейрона выглядит следующим образом [2]:

|  |  |
| --- | --- |
|  | (4) |

Формула подсчета ошибок скрытого слоя выглядит следующим образом [2]:

|  |  |
| --- | --- |
|  | (5) |

Формула пересчета весов выглядит следующим образом [2]:

|  |  |
| --- | --- |
|  | (6) |

где – коэффициент скорости обучения.

Алгоритм обучения нейронной сети методом обратного распространения ошибки можно разделить на 2 этапа:

1. Расчёт выходных сигналов для заданного тренировочного примера;
2. Сравнение рассчитанных выходных сигналов с эталоном и обновление весовых коэффициентов связей между узлами на основе найденных отличий.

Тренировочный пример – это последовательность данных, которыми оперирует нейронная сеть.

1. Алгоритмы обучения RPROP и iRPROP+

В отличие от стандартного алгоритма BackPROP, RPROP использует только знаки частных производных для подстройки весовых коэффициентов. Алгоритм использует так называемое «обучение по эпохам», когда коррекция весов происходит после предъявления сети всех примеров из обучающей выборки.

|  |  |
| --- | --- |
|  | (7) |

Для определения величины коррекции используется следующее правило [3]:

Если на текущем шаге частная производная по соответствующему весу wij поменяла свой знак, то это говорит о том, что последнее изменение было большим, и алгоритм проскочил локальный минимум, и, следовательно, величину изменения необходимо уменьшить на η– и вернуть предыдущее значение весового коэффициента: другими словами, необходимо произвести «откат» [3]:

|  |  |
| --- | --- |
|  | (8) |

Если знак частной производной не изменился, то нужно увеличить величину коррекции на η+ для достижения более быстрой сходимости. Зафиксировав множители η− и η+, можно отказаться от глобальных параметров настройки нейронной сети, что также можно рассматривать как преимущество рассматриваемого алгоритма перед стандартным алгоритмом Backprop.

Рекомендованные значения для η− = 0.5, η+ = 1.2, но нет никаких ограничений на использование других значений для этих параметров.

Для того, чтобы не допустить слишком больших или малых значений весов, величину коррекции ограничивают сверху максимальным Δmax и снизу минимальным Δmin значениями величины коррекции, которые по умолчанию, соответственно, устанавливаются равными 50 и 1.0E-6.

Начальные значения для всех Δij устанавливаются равными 0.1. Опять же, это следует рассматривать лишь как рекомендацию, и в практической реализации можно задать другое значение для инициализации.

Для вычисления значения коррекции весов используется следующее правило [3]:

|  |  |
| --- | --- |
|  | (9) |

Если производная положительна, т.е. ошибка возрастает, то весовой коэффициент уменьшается на величину коррекции, в противном случае – увеличивается.

Затем подстраиваются веса [3]:

|  |  |
| --- | --- |
|  | (10) |

Сразу после появления, алгоритм RProp завоевал популярность благодаря простоте и высокой скорости и эффективности обучения. Однако, нужно отметить, что данный алгоритм использует только информацию о локальном градиенте ошибки (точнее, только градиент ошибки по данному весу). В целом ряде работ были попытки улучшить скорость и эффективность обучения нейронной сети за счет учета глобального изменения ошибки. Одна из простейших 24 модификаций базового алгоритма оказалась на редкость эффективной. Суть отличия iRPROP+ от RPROP – если изменился знак градиента, производить откат только в случае, если общая ошибка классификации увеличилась с предыдущего шага [6].

Схема алгоритма обучения RPROP представлена на рисунке 4. Схема алгоритма обучения iRPROP+ представлена на рисунке 5.

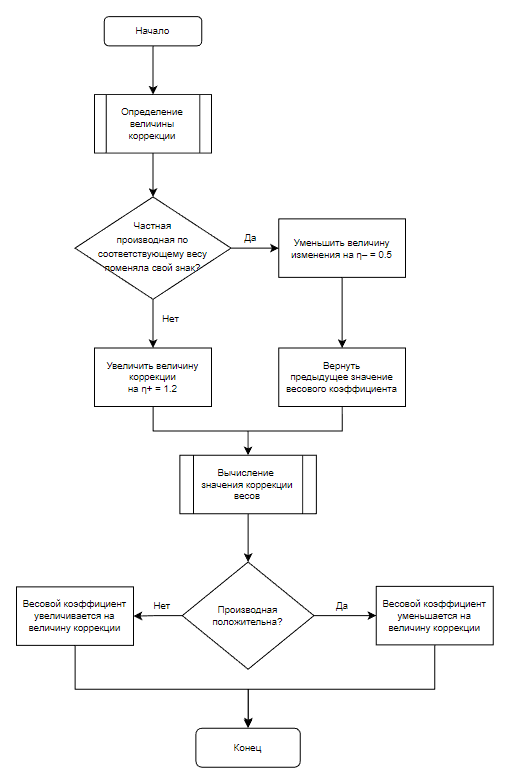


Рисунок 4 – Схема алгоритма обучения RPROP

Схема алгоритма обучения iRPROP+ представлена на рисунке 5.

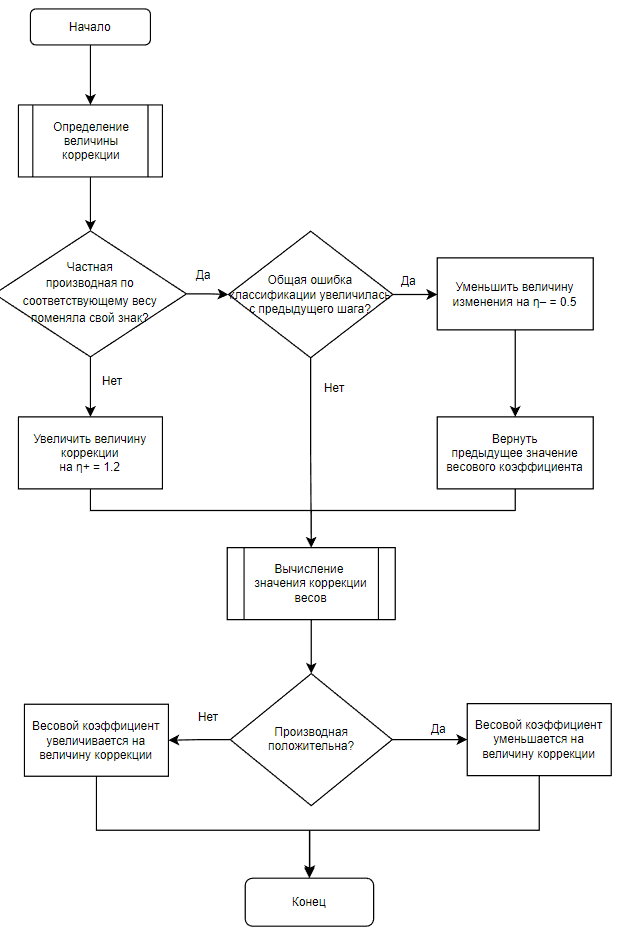


Рисунок 5 – Схема алгоритма обучения iRPROP+

Исследования

1. Исследовать зависимость погрешности обучения для каждого алгоритма от значения коэффициента обучения.
2. Исследовать зависимость погрешности обучения и классификации для каждого алгоритма от объёма обучающей выборки.
3. Исследовать зависимость погрешности обучения и классификации для каждого алгоритма от числа нейронов скрытого слоя.
4. Исследовать зависимость погрешности классификации от числа итераций обучения.

В исследованиях в качестве погрешностей обучения и прогнозирования считается среднеквадратическое отклонение:

*,*

где – объем выборки,

– полученное значение,

- эталонное значение.

Синей линией обозначена погрешность обучения печатных букв. Красной линией обозначения погрешность обучения печатных латинских букв.

1. Исследовать зависимость погрешности обучения для каждого алгоритма от значения коэффициента обучения.

1 Алгоритм обратного распространения ошибки

Для исследования использована обучающая выборка объемом 52 буквы. На рисунке 6 изображена зависимость погрешности обучения от значения коэффициента обучения алгоритма обратного распространения ошибки. Для подсчета погрешности были выбраны 100 нейронов скрытого слоя и 10 итераций обучения.

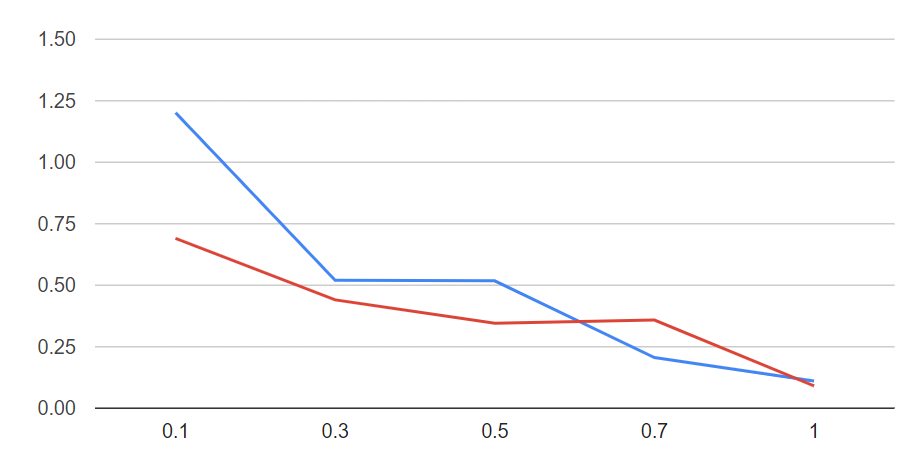


Рисунок 6 - Зависимость погрешности обучения от значения коэффициента обучения алгоритма обратного распространения ошибки.

2 Алгоритм RPROP

Для исследования использована обучающая выборка объемом 52 буквы. На рисунке 7 изображена зависимость погрешности обучения от значения коэффициента обучения алгоритма RPROP. Для подсчета погрешности были выбраны 100 нейронов скрытого слоя и 10 итераций обучения.

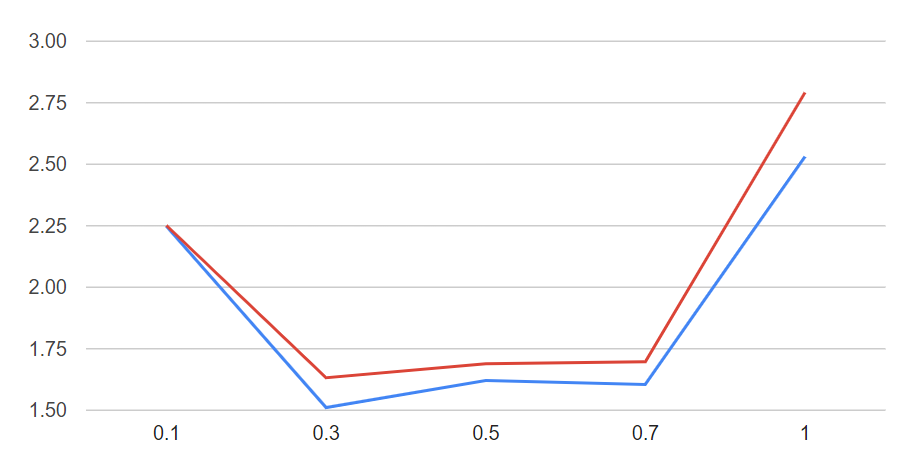


Рисунок 7 - Зависимость погрешности обучения от значения коэффициента обучения алгоритма RPROP

3 Алгоритм iRPROP+

Для исследования использована обучающая выборка объемом 52 буквы. На рисунке 8 изображена зависимость погрешности обучения от значения коэффициента обучения алгоритма iRPROP+. Для подсчета погрешности были выбраны 100 нейронов скрытого слоя и 10 итераций обучения.

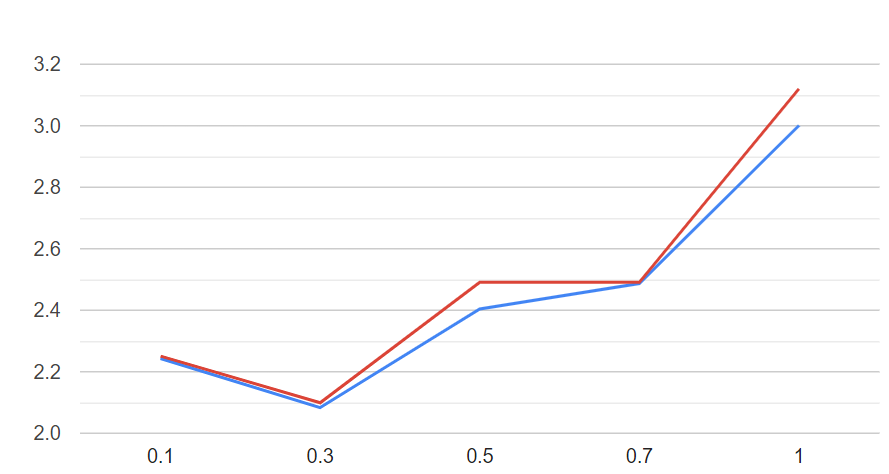


Рисунок 8 - Зависимость погрешности обучения от значения коффициента обучения алгоритма iRPROP+

2 Исследовать зависимость погрешности обучения и классификации для каждого алгоритма от объёма обучающей выборки

1 Алгоритм обратного распространения ошибки

На рисунке 9 изображена зависимость погрешности обучения от объема обучающей выборки алгоритма обратного распространения ошибки. Для подсчета погрешности были выбраны 100 нейронов скрытого слоя, коэффициент обучения 0.1 и 10 итераций обучения.

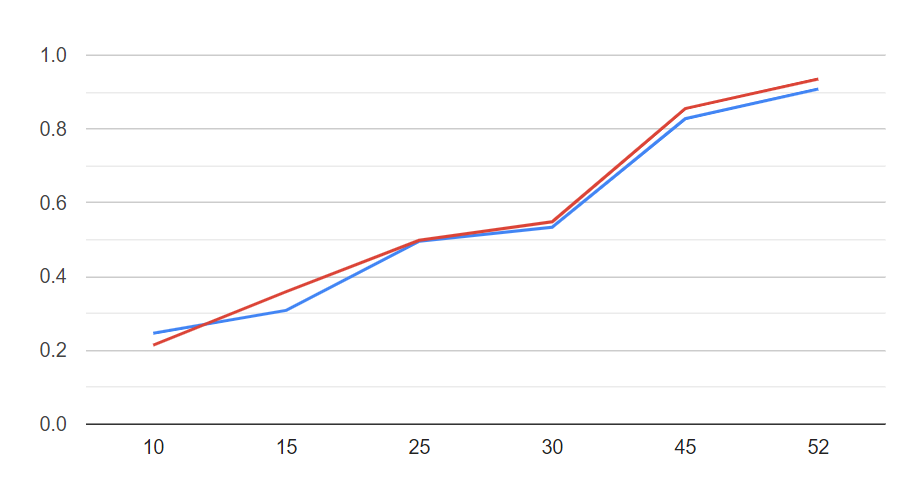


Рисунок 9 - Зависимость погрешности обучения от объема обучающей выборки

2 Алгоритм RPROP

На рисунке 10 изображена зависимость погрешности обучения от объема обучающей выборки алгоритма RPROP. Для подсчета погрешности были выбраны 120 нейронов скрытого слоя, коэффициент обучения 0.3 и 10 итераций обучения.

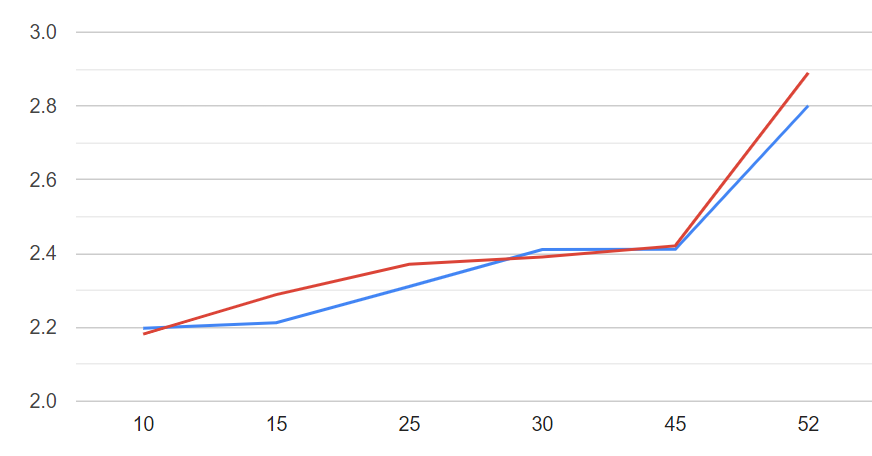


Рисунок 10 - Зависимость погрешности обучения от объема обучающей выборки

3 Алгоритм iRPROP+

На рисунке 11 изображена зависимость погрешности обучения от объема обучающей выборки алгоритма iRPROP+. Для подсчета погрешности были выбраны 120 нейронов скрытого слоя, коэффициент обучения 0.3 и 10 итераций обучения.

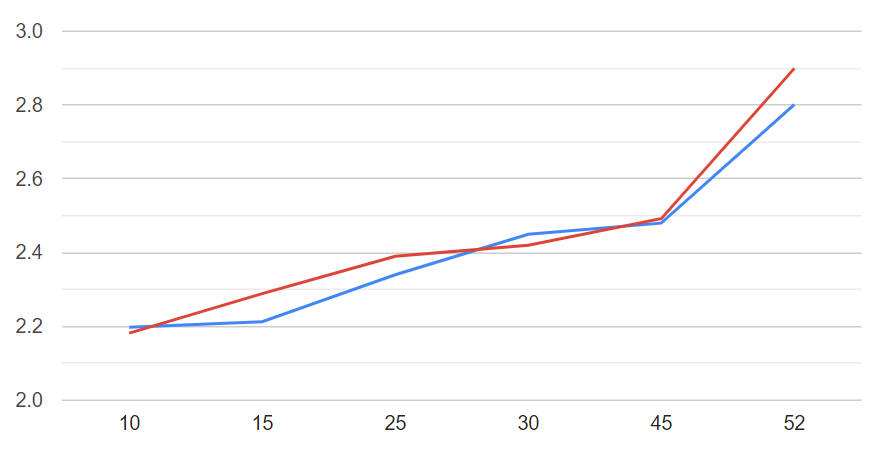


Рисунок 11 - Зависимость погрешности обучения от объема обучающей выборки

3 Исследовать зависимость погрешности обучения и классификации для каждого алгоритма от числа нейронов скрытого слоя

1 Алгоритм обратного распространения ошибки

Для исследования использована обучающая выборка объемом 52 буквы. На рисунке 12 изображена зависимость погрешности обучения от числа нейронов скрытого слоя обучения алгоритма обратного распространения ошибки. Для подсчета погрешности были выбраны коэффициент обучения 0.5 и 10 итераций обучения.

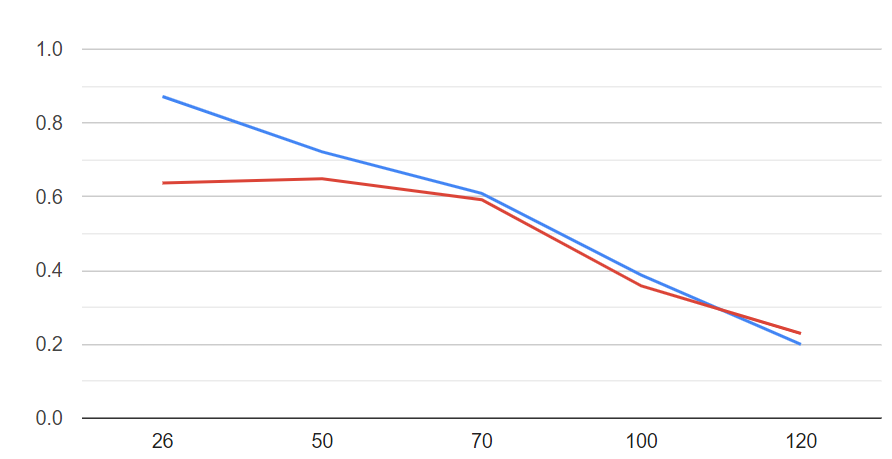


Рисунок 12 - Зависимость погрешности обучения от числа нейронов скрытого слоя

2 Алгоритм обучения RPROP

Для исследования использована обучающая выборка объемом 52 буквы. На рисунке 13 изображена зависимость погрешности обучения от числа нейронов скрытого слоя обучения алгоритма RPROP. Для подсчета погрешности были выбраны коэффициент обучения 0.3 и 10 итераций обучения.

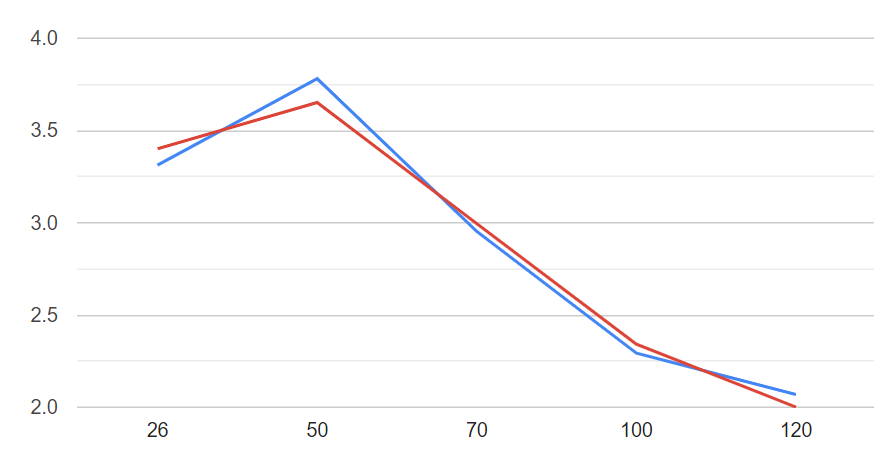


Рисунок 13 - Зависимость погрешности обучения от числа нейронов скрытого слоя

3 Алгоритм обучения iRPROP+

Для исследования использована обучающая выборка объемом 52 буквы. На рисунке 14 изображена зависимость погрешности обучения от числа нейронов скрытого слоя обучения алгоритма iRPROP+. Для подсчета погрешности были выбраны коэффициент обучения 0.3 и 10 итераций обучения.

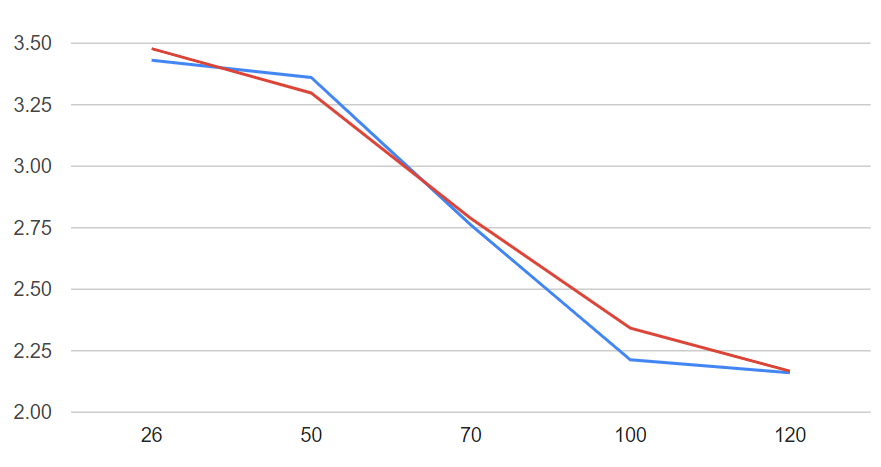


Рисунок 14 - Зависимость погрешности обучения от числа нейронов скрытого слоя

4 Исследовать зависимость погрешности классификации от числа итераций обучения

Для исследования использована обучающая выборка объемом 52 буквы. На рисунке 15 изображена зависимость погрешности обучения от числа итераций обучения алгоритма обратного распространения ошибки. Для подсчета погрешности были выбраны коэффициент обучения 0.5 и 100 нейронов скрытого слоя.

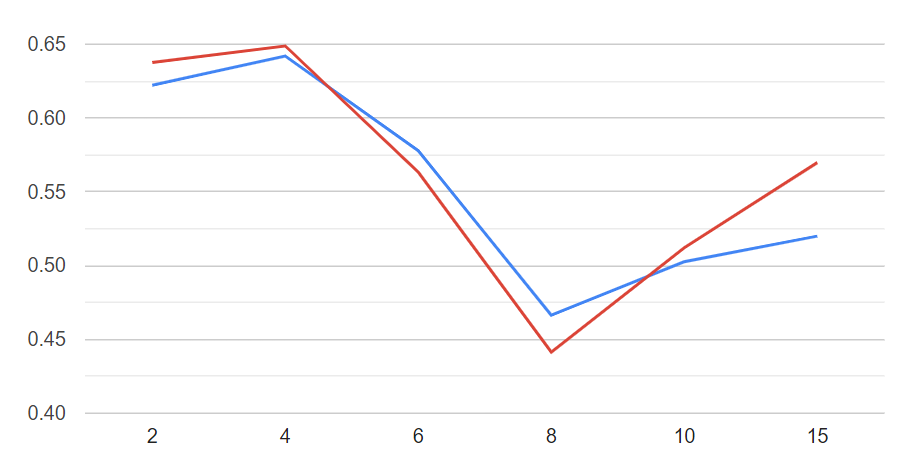


Рисунок 15 - Зависимость погрешности обучения от числа итераций обучения

Для исследования использована обучающая выборка объемом 52 буквы. На рисунке 16 изображена зависимость погрешности обучения от числа итераций обучения алгоритма RPROP. Для подсчета погрешности были выбраны коэффициент обучения 0.3 и 120 нейронов скрытого слоя.

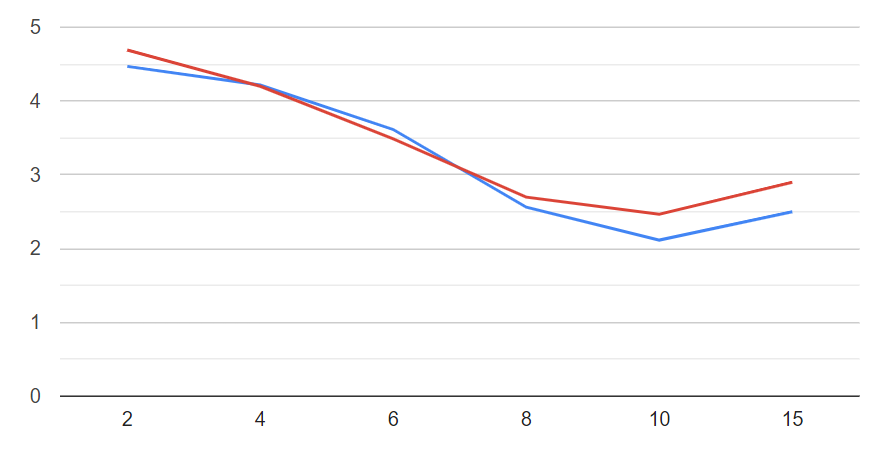


Рисунок 16 - Зависимость погрешности обучения от числа итераций обучения

Для исследования использована обучающая выборка объемом 52 буквы. На рисунке 17 изображена зависимость погрешности обучения от числа итераций обучения алгоритма iRPROP+. Для подсчета погрешности были выбраны коэффициент обучения 0.3 и 120 нейронов скрытого слоя.

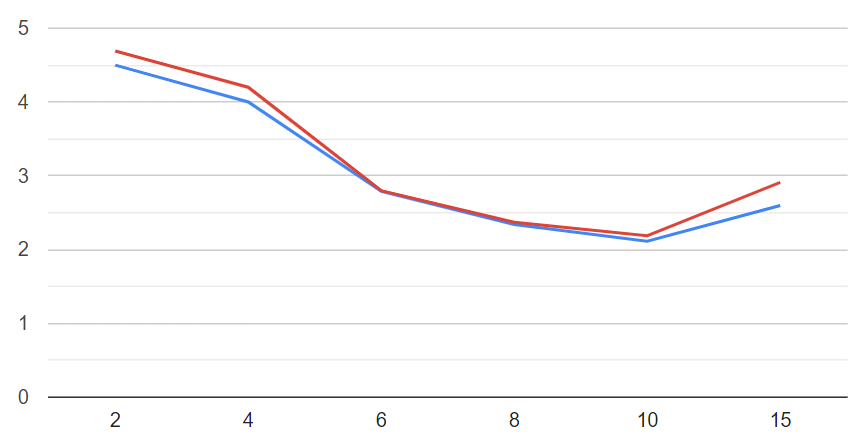


Рисунок 17 - Зависимость погрешности обучения от числа итераций обучения

ВЫВОДЫ

В результате работы была написана программа, имитирующая работу многослойного персептрона и проведено обучение сети по алгоритмам обратного распространения ошибки, RPROP и iRPROP+. Также проведены исследования зависимости погрешностей обучения от разных параметров.

Из полученных результатов можно сделать вывод, что для достижения наилучшего процента распознавания коэффициент обучения для алгоритма обучения обратного распространения ошибки должен быть 1; для алгоритма обучения RProp и iRProp+ - 0.3; а количество нейронов в скрытом слое для алгоритма обучения обратного распространения ошибки должно быть 120; для алгоритмов обучения RProp и iRProp+ - 120. Во всех трех алгоритмах, чем больше объем обучающих данных, тем больше погрешность.

Список использованной литературы

1. Искусственная нейронная сеть [Электронный ресурс]/ – Электрон. текстовые дан., Режим доступа: http://www.machinelearning.ru/wiki/index.php?title=Искусственная\_нейронная\_сеть, свободный;
2. Метод обратного распространения ошибки [Электронный ресурс]. – https://ru.wikipedia.org/wiki./Метод\_обратного\_распространения\_ошибки
3. Алгоритм RPROP [Электронный ресурс] – Электрон. Текстовые дан., Режим доступа: <https://basegroup.ru/community/articles/rprop>, свободный.
4. Гуисов, М.И. Применение нейронных сетей для оценки позиции в игре Го [Текст]/ Гуисов, М.И. <http://is.ifmo.ru/diploma-theses/guisov/guisov.pdf>
5. Рашид, Т. Создаем нейронную сеть [Текст]/ Т. Рашид; пер. с англ. Гузикевича, А.Г., ред. Гузикевича, А.Г. – М: Диалектика, 2017 – 273 с.