Министерство науки и высшего образования Российской Федерации

ФГБОУ ВО «АЛТАЙСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

Институт цифровых технологий, электроники и физики

Кафедра вычислительной техники и электроники (ВТиЭ)

Лабораторная работа №1

**Знакомство с нейросетевым анализом**

|  |
| --- |
| Выполнил: студент гр. 5.306М:  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Лаптев А. В.  Проверил: доц. каф. ВТиЭ  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Шайдуров А. А.  Оценка \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  «\_\_\_\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_2024 г. |

**Цель:** проведение нейросетевого анализа сложных данных. В рамках данной работы необходим программный продукт «Deductor Studio».

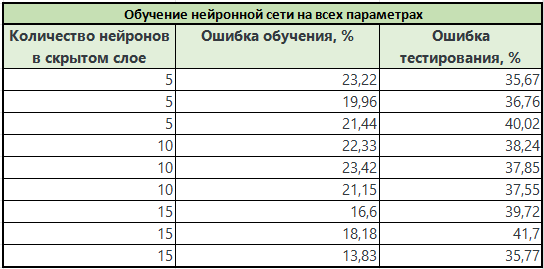
**Задачи:** в рамках данной лабораторной работы необходимо провести исследования с применением нейросетевого и статистического анализа.

1. На первом этапе вам необходимо обучить нейронные сети для решения задачи классификации.
2. На втором этапе вы должны с применением статистического анализа сократить количество входных параметров. И обучить новые нейронные сети на основе сокращенной обучающей выборки.
3. На третьем этапе вы должны сформировать итоговую таблицу и построить графики зависимости усредненных ошибок от количества нейронов в скрытом слое.

**Ход работы:**

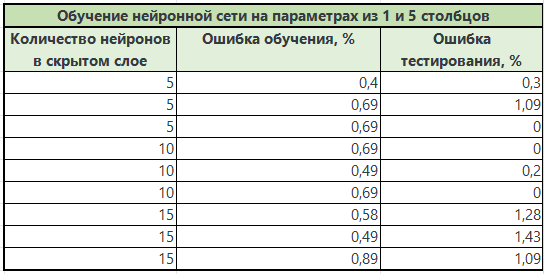
1. Для решения задачи классификации при обучении использовались нейронные сети с 5, 10 и 15 нейронами в скрытом слое. В результате обучения были получены следующие значения ошибок:

*Таблица 1. Обучение нейронной сети на всех параметрах*

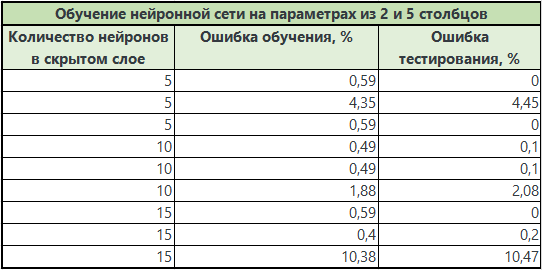


1. После того, как было произведено обучение с использованием всех входных параметров было произведено сокращение исходного количества столбцов до 5 (первые 5 столбцов имели наибольшее влияние на итоговые результаты обучения моделей). И для каждой следующей пары столбцов: 1 и 5, 2 и 5, 3 и 5, 4 и 5 －были обучены новые нейронные сети для такого же числа нейронов в скрытом слое. В результате были получены следующие значения ошибок:

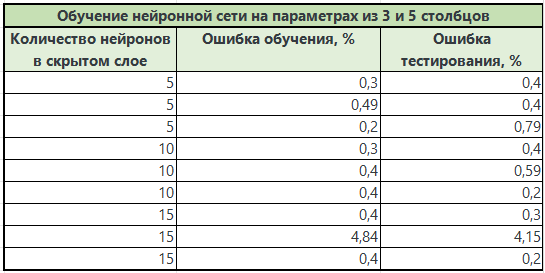
*Таблица 2. Обучение нейронной сети на параметрах 1 и 5 столбцов*



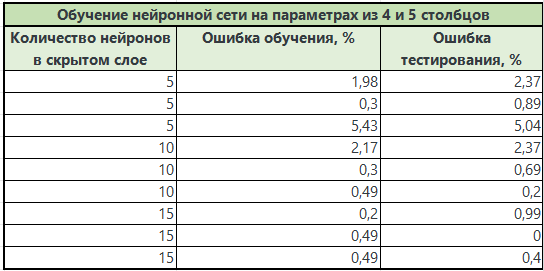
*Таблица 3. Обучение нейронной сети на параметрах 2 и 5 столбцов*



*Таблица 4. Обучение нейронной сети на параметрах 3 и 5 столбцов*



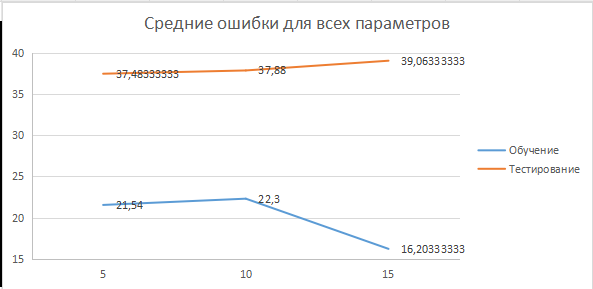
*Таблица 5. Обучение нейронной сети на параметрах 4 и 5 столбцов*



1. В результате все средние значения были сведены в общую таблицу и для указанных значений был построен ряд графиков:

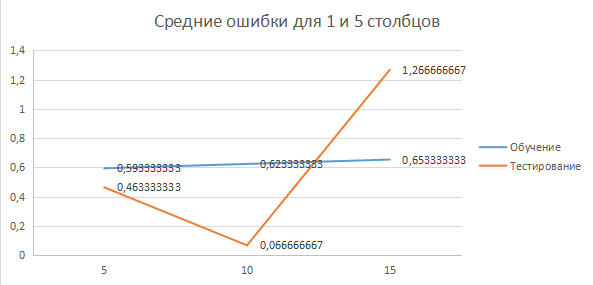
*Таблица 6. Сводная таблица средних ошибок для каждого тестирования*



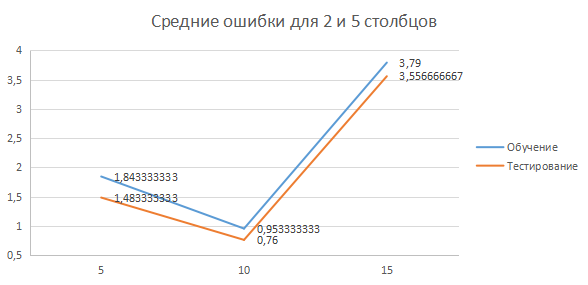


*Рис. 1. Средние ошибки при использовании всех параметров.*

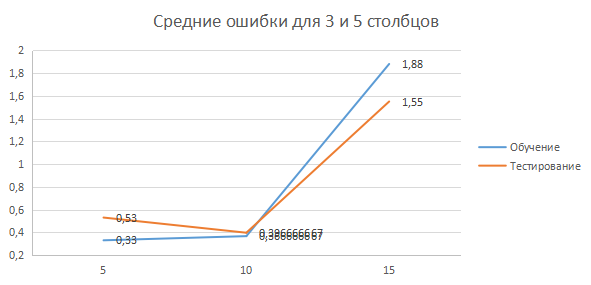
На данном графике видно, что для 5 и 10 нейронов в скрытом слое средние ошибки обучения и тестирования примерно одинаковы и разница в процентном отношении также сохраняется, но для 15 нейронов четко видно резкое снижение ошибки при обучении, но ее увеличение для тестовой выборки, что говорит о том, что на таком количестве нейронов модель начинает переобучаться, подстраиваясь под обучающие значения, что негативно влияет на предсказания и нежелательно.



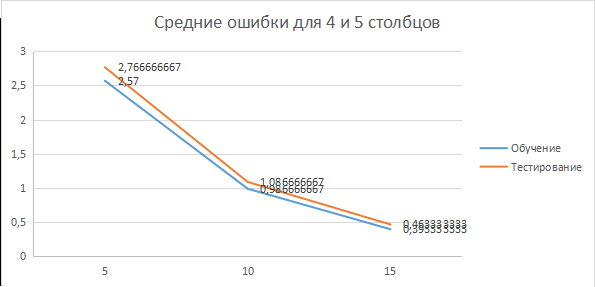
*Рис. 2. Средние значения ошибок для 1 и 5 столбцов.*



*Рис. 3. Средние значения ошибки для 2 и 5 столбцов.*



*Рис. 4. Средние значения ошибки для 3 и 5 столбцов.*



*Рис. 5. Средние значения ошибки для 4 и 5 столбцов.*

На графиках, которые представлены выше показано то, как изменяются средние значения ошибки для конкретных пар столбцов со значимыми входными параметрами. Можно заметить, что на всех графиках, кроме последнего лучше всего модель показывает себя на количестве нейронов равном 10, а для 15 нейронов либо вырастает среднее значение ошибки (рис. 3 и рис. 4), либо наблюдается переобучение (рис. 2).

Для графика на рис. 5 наблюдается практически линейное снижение ошибки при увеличении числа нейронов в скрытом слое, но, поскольку, для 10 нейронов все равно среднее значение ошибки значительно ниже, чем для 5 нейронов, то в целом, для всей системы, можно назвать оптимальным количеством именно 10 нейронов в скрытом слое.

Также, можно заметить, что для после построения нейронной сети для пар столбцов значение ошибки сильно меньше, чем для модели, которая использует все параметры. Это связано с тем, что часть начальных параметров были малоинформативны и создавали шум в данных, что затрудняло обучение модели, что в свою очередь также позволило модели лучше обобщить их и найти закономерности.

**Вывод:** в ходе выполнения работы был проведен нейросетевой анализ сложных данных; проведено обучение модели на разном количестве нейронов в скрытом слое для разного количества входных параметров; в ходе анализа применялся статистический анализ для сокращения числа параметров; а также рассчитаны средние значения ошибок для обучающей и тестовой выборок; построены графики, которые иллюстрируют зависимости средних значений ошибок от числа нейронов.