****

Институт информационных и вычислительных технологий

Кафедра управления и интеллектуальных технологий

**Отчет по лабораторной работе №1**

**По курсу «Нейрокомпьютеры и их применение»**

**«Изучение основных понятий теории ИНС и приемов работы с пакетом Statistica Neural Networks»**

Выполнили студенты: Михайловский Михаил, Озеров Сергей

Группа: А-03-21

Проверил: Колпинский Сергей Викторович

**Москва 2024**

Оглавление

[Введение 2](#_Toc159761705)

[Логическая функция AND 2](#_Toc159761706)

[Линейная функция активации 3](#_Toc159761707)

[Пороговая функция активации 4](#_Toc159761708)

[Логистическая функция активации 5](#_Toc159761709)

[Логическая функция OR 6](#_Toc159761710)

[Линейная функция активации 7](#_Toc159761711)

[Пороговая функция активации 7](#_Toc159761712)

[Логистическая функция активации 8](#_Toc159761713)

[Логическая функция XOR 9](#_Toc159761714)

[Линейная функция активации 9](#_Toc159761715)

[Пороговая функция активации 10](#_Toc159761716)

[Логистическая функция активации 11](#_Toc159761717)

[Модель достаточная для предсказания XOR 11](#_Toc159761718)

Введение

Исследуем структуру простой нейросети, имеющей два входа, один выход и не имеющей скрытых слоёв (рис. 1).

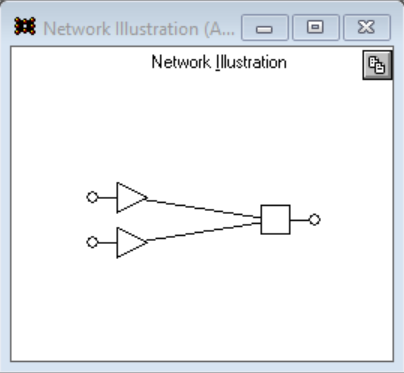


Рис. 1. Исследуемая структура нейросети.

Попробуем на основе этой структуры создать модели, которые будут предсказывать значения нескольких основных логических функций (AND, OR, XOR).

Логическая **функция AND**

Создадим модель для предсказания значений функции AND. Её таблица истинности представлена на рис. 2.

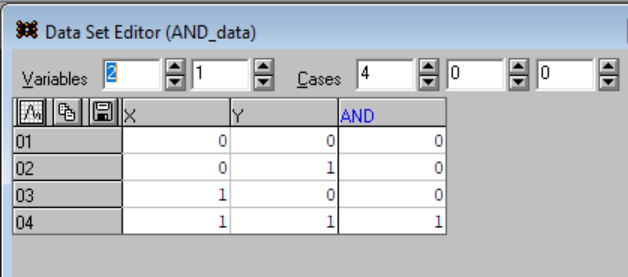


Рис. 2. Таблица истинности логической функции AND.

Установим выходную функцию активации линейной (рис. 3). При графическом представлении в декартовой плоскости Oxy значения данной функции являются линейно разделимыми. Поэтому проводить классификацию соответствующих меток должно быть не сложно.

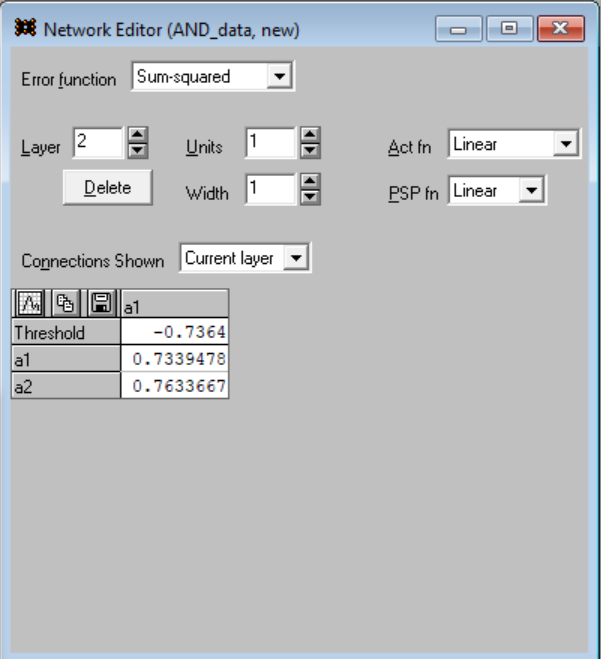


Рис. 3. Окно настроек слоёв нейросети.

Для обучения модели используем метод обратного распространения ошибки (back propagation). Использованные параметры представлены на рис. 4. Далее будем создавать модели с различными выходными фукнциями активации.

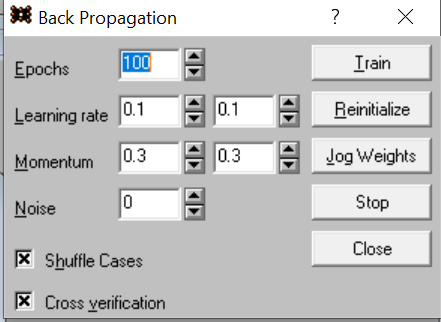


Рис. 4. Окно настройки обучения по back propagation.

Линейная функция активации

Обучим модель с линейной выходной функцией активации. Как видим график очень быстро приходит к установившемуся значению. Однако, ошибка в результате получается ненулевой (порядка 0.25), это связано с тем, что наша модель предсказывает не бинарные значения, а вещественные.

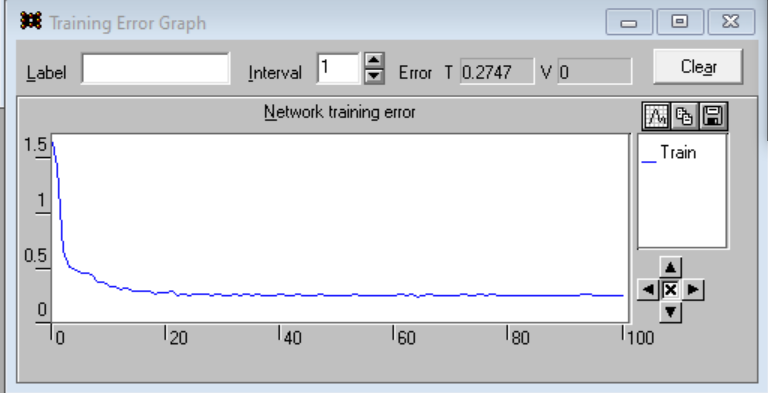


Рис. 5. График ошибки в течение обучения модели с линейной функцией активации.

В результате прогона нашей таблицы данных для функции AND получаем результаты, представленные на рис. 6. Как мы видим, модель предсказывает вещественные значения, довольно далекие от истинных. Однако, можно отметить, что каждое из них находится ближе к истинному значению.

Таким образом, если добавить на выход, например, пороговую функцию активации, то в результате модель будет в точности предсказывать все значения функции AND.

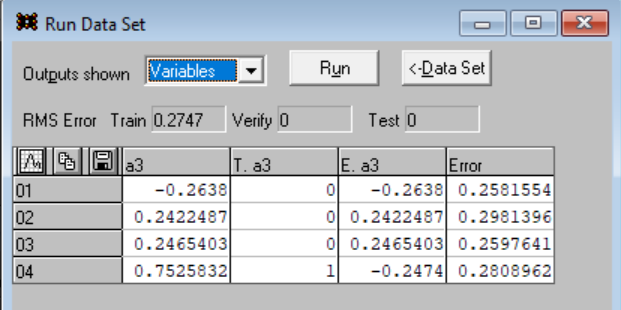


Рис. 6. Предсказания модели значений AND по датасету.

Пороговая функция активации

Обучим модель с пороговой выходной функцией активации. По графику ошибки в течение обучения можно сказать, что обучение пороговой функции активации проходит более хаотично и менее гладко. Это связано с тем, что метод обратного распространения ошибки основывается на методах градиентного спуска. Однако, при использовании пороговой функции активации, вычисляемый градиент почти всегда получается равным нулю.

Это значит, что градиентный спуск в данном случае бесполезен, и приходится применять методы нулевого порядка для оптимизации заданной целевой функции.

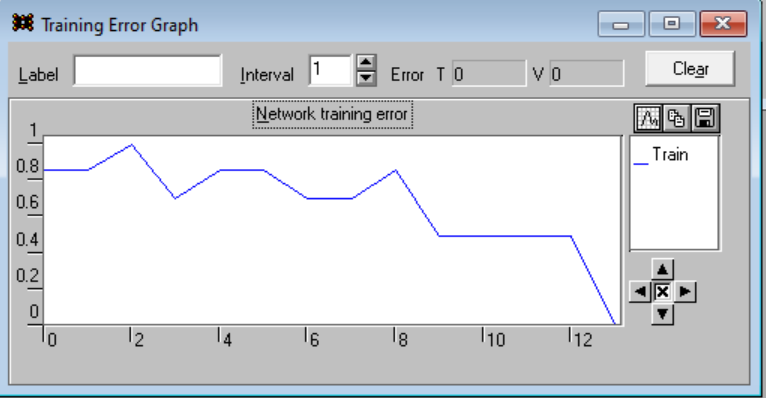
****

Рис. 7. График ошибки в течение обучения модели с пороговой функцией активации.

Как результат, данная модель полностью обучается данной выборке, и может предсказывать все её значения (рис. 8).

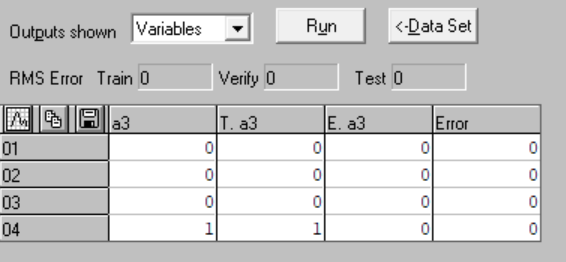
****

Рис. 8. Предсказания модели значений AND по датасету.

Логистическая функция активации

Обучим модель с логистической выходной функцией активации. По графику ошибки в течения обучения, можно сказать, что такая модель ведёт себя похоже на модель с линейной функцией активации. Однако, в данном случае, обучение модели проходило дольше, и более гладко. Связано это с тем, что для обучения данной модели выставленные параметры обучения меньше подходят.

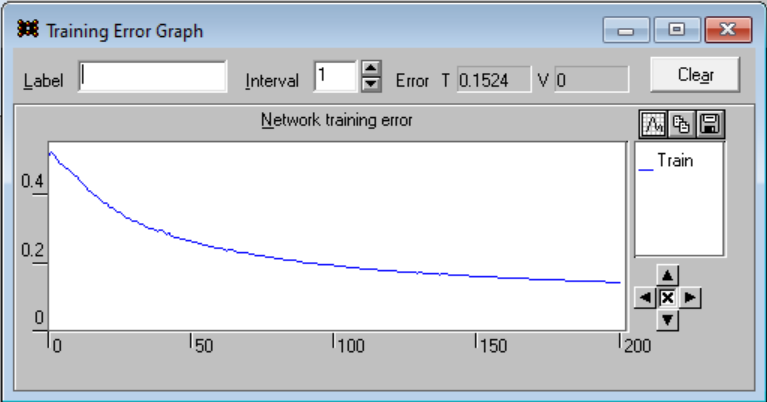
****

Рис. 9. График ошибки в течение обучения модели с логистической функцией активации.

В результате прогона таблицы истинности получаем аналогичные модели с линейной функцией активации результаты.

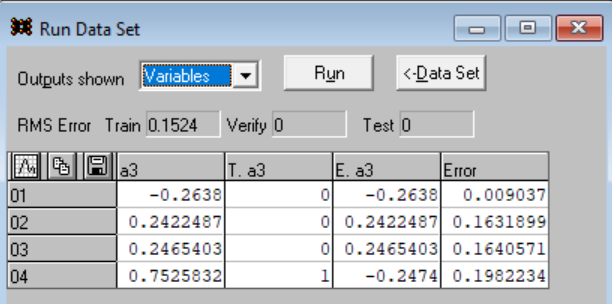
****

Рис. 10. Предсказания модели значений AND по датасету.

Логическая **функция** OR

Проведём аналогичные исследования для таблицы истинности функции OR. Стоит отметить, что так же, как и функция AND, её значения являются линейно разделимыми.

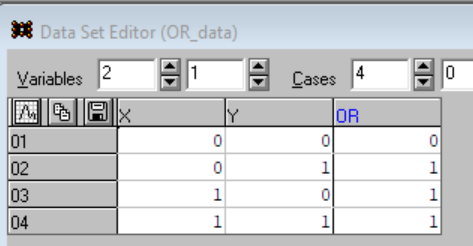
****

Рис. 11. Таблица истинности логической функции OR.

Линейная функция активации

Обучение с линейной функцией активации проходит достаточно быстро.

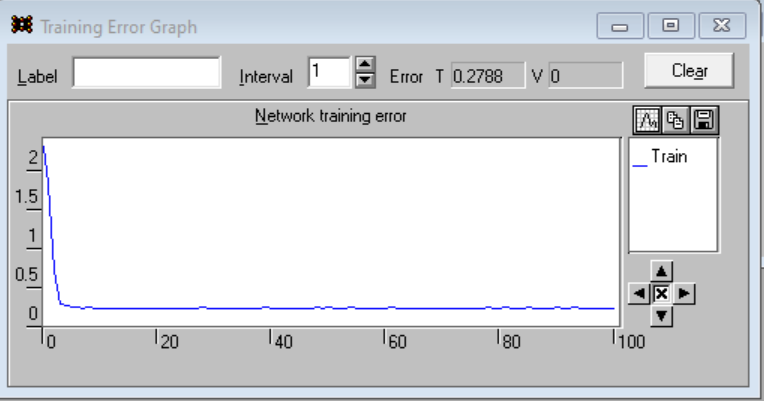
****

Рис. 12. График ошибки в течение обучения модели с линейной функцией активации.

Получаем результаты аналогичные полученным результатам для линейной выходной функции активации для таблицы истинности AND.

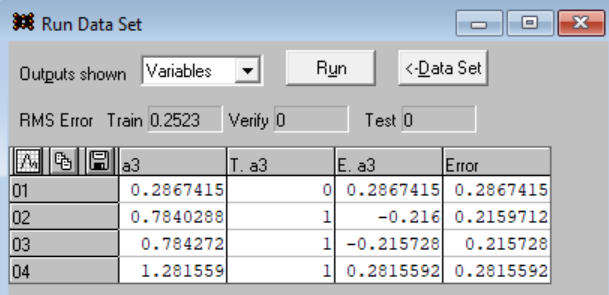
****

Рис. 13. Предсказания модели значений OR по датасету.

Пороговая функция активации

В результате обучения модель приходит к ненулевой ошибке.

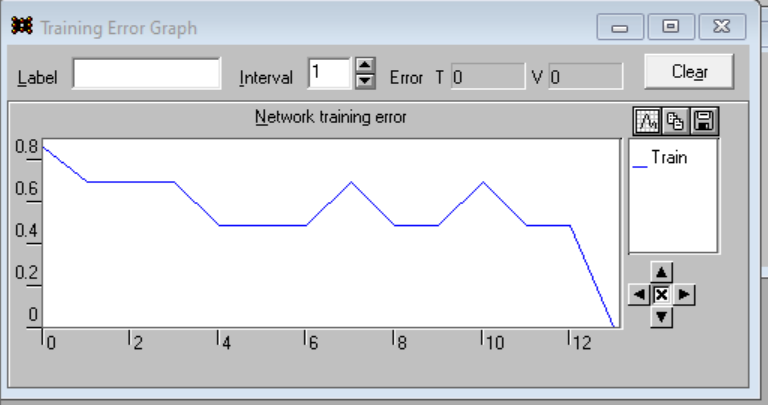
****

Рис. 14. График ошибки в течение обучения модели с пороговой функцией активации.

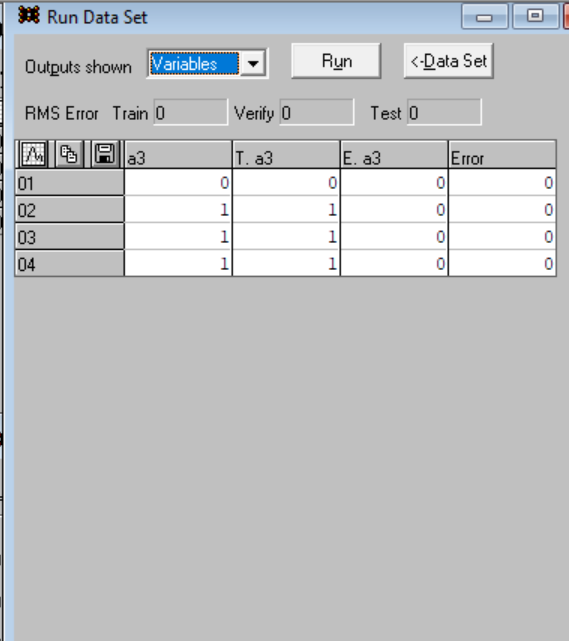
****

Рис. 15. Предсказания модели значений OR по датасету.

Логистическая функция активации

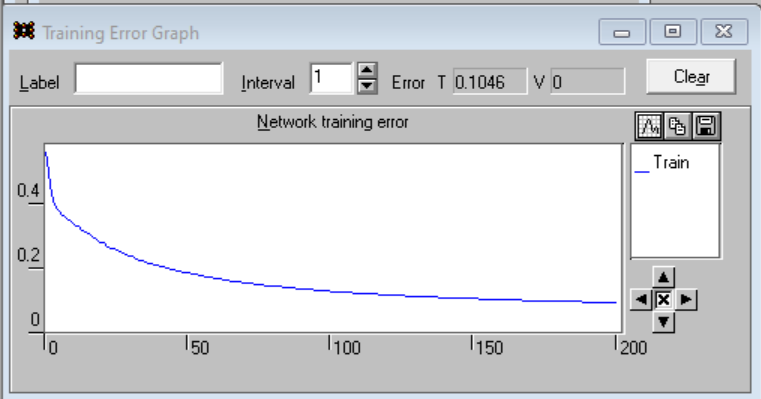
****

Рис. 16. График ошибки в течение обучения модели с логистической функцией активации.

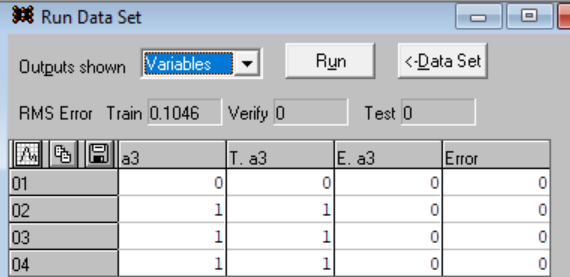
****

Рис. 17. Предсказания модели значений OR по датасету.

Логическая **функция** XOR

Для функции исключающего ИЛИ картина будет иной, поскольку при графическом представлении можно увидеть, что её значения не являются линейно разделимыми. По крайней мере одной прямой.

Используемая нами структура нейросети может построить только линейное разделение, поэтому теоретически от неё можно ожидать некоторых ошибочных предсказаний значений функции.

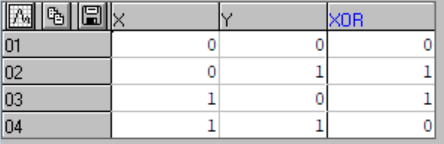
****

Рис. 18. Таблица истинности логической функции XOR.

Линейная функция активации

Как мы видим по графику ошибки в течение обучения, в результате модель смогла достичь ошибки предсказания порядка 0.5. Этот результат хуже, чем для двух предыдущих функций.

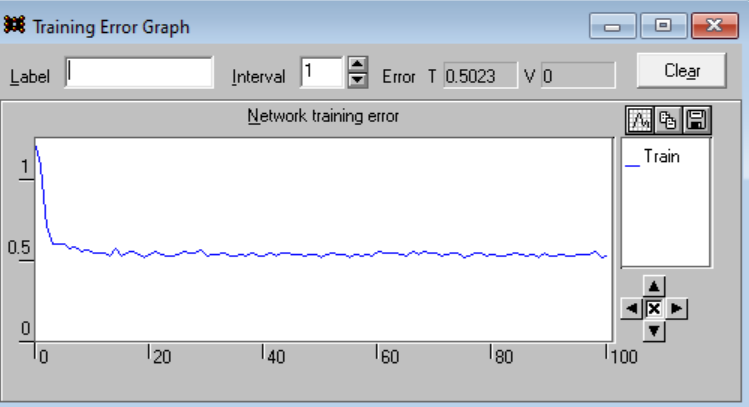
****

Рис. 19. График ошибки в течение обучения модели с линейной функцией активации.

В результате прогона таблицы значений XOR через нашу модель видим, что модель не смогла научиться предсказывать значения функции. Она для любого входного значения предсказывает приблизительно среднее значение между возможными.

Таким образом, такая модель никак не предсказывает ни одного значения, отказываясь давать метки объектам.

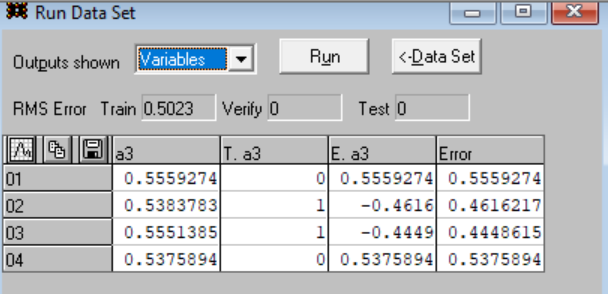
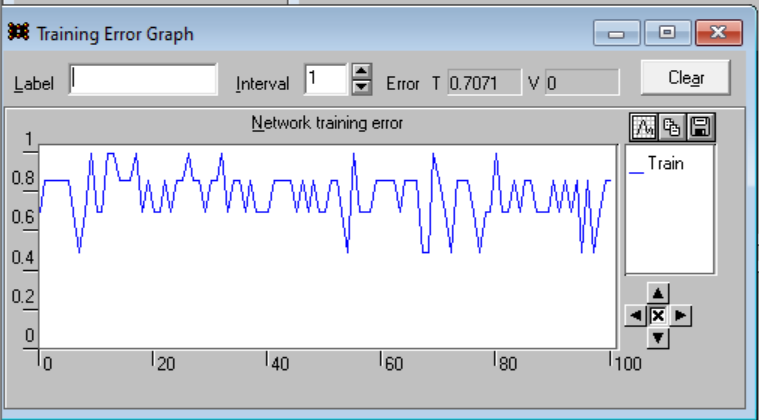
****

Рис. 20. Предсказания модели значений XOR по датасету.

Пороговая функция активации

Для пороговой функции активации обучение имеет колибательный характер. Оно не приходит к некоторому значению, а ошибка остаётся почти на каждом шаге больше 0.5.

****Рис. 21. График ошибки в течение обучения модели с пороговой функцией активации.

Как мы видим, модель для каждого объекта предсказывает одно и то же значение. Таким образом, у неё получается в половине случаев «угадать» правильный ответ, но тем не менее, такая модель не несёт какой-то пользы.

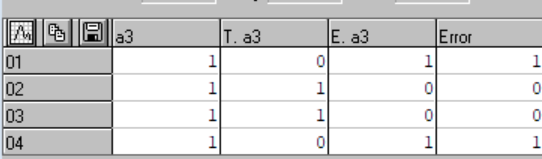
****

Рис. 22. Предсказания модели значений XOR по датасету.

Логистическая функция активации

Так же как и предыдущие функции активации, в течение обучения ошибка остаётся высокой.

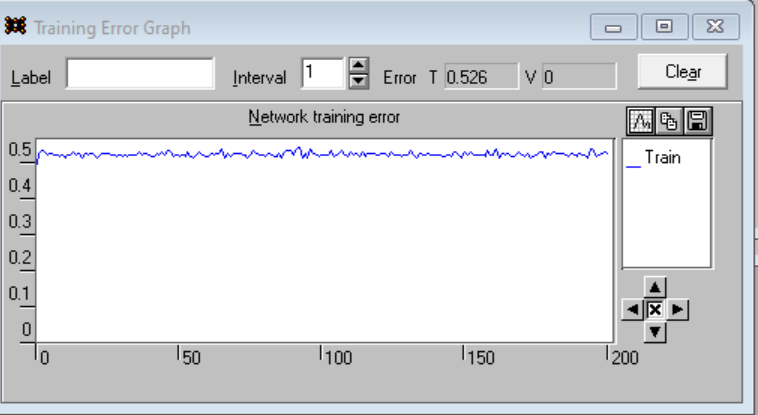


Рис. 23. График ошибки в течение обучения модели с логистической функцией активации.

По таблице истинности, так же как и модель с линейной функцией активации, данная модель предсказывает просто средние значения между «1» и «0».

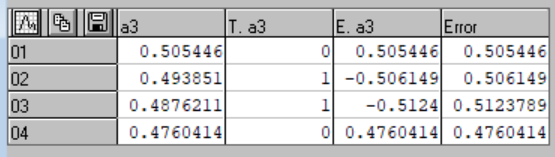


Рис. 24. Предсказания модели значений XOR по датасету.

Модель достаточная для предсказания XOR

Попробуем построить модель, которая сможет предсказывать значения функции XOR. Для этого добавим 1 скрытый слой состоящий из 3 нейронов с логистической функцией активации. Входной и выходной слои будут иметь линейную функцию активации (рис. 25).

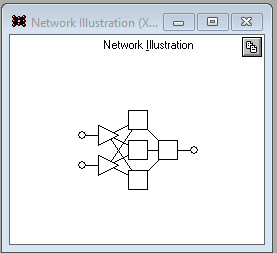
****

Рис. 25. Структура созданной нейросети для предсказывания XOR.

При попытках обучить данную модель, получаем различные результаты. При обучении по методу обратного распространения ошибки модель не даёт должных результатов и выдаёт средние одинаковые значения для всех входных данных.

При использовании обучения по методу сопряжённых градиентов при удачных начальных весах модель обучается до практически нулевой ошибки.

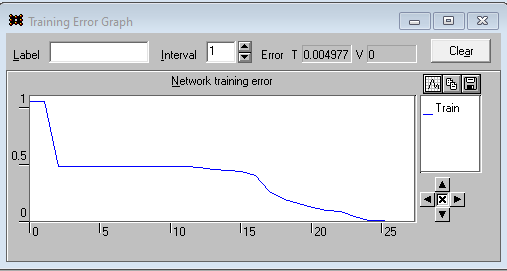
****

Рис. 26. График ошибки для обучения данной модели

Как мы видим по результатам прогона таблицы истинности, данная модель очень точно предсказывает значения исключающего ИЛИ.

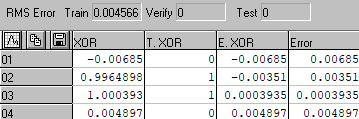
****

Рис. 27. Предсказания модели для XOR.