****

Институт информационных и вычислительных технологий

Кафедра управления и интеллектуальных технологий

**Отчет по лабораторной работе №2**

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

**«Многослойный персептрон»**

Выполнили студенты: Михайловский Михаил, Озеров Сергей

Группа: А-03-21

Проверил: Колпинский Сергей Викторович

**Москва 2024**

Оглавление

[Введение 2](#_Toc159764672)

[Многоклассовая классификация прямой 2](#_Toc159764673)

[Многоклассовая классификация softmax 4](#_Toc159764674)

[Скрытый линейный слой 4](#_Toc159764675)

[Обучение по методу распространения обратной ошибки 5](#_Toc159764676)

[Обучение по методу сопряжённых градиентов 6](#_Toc159764677)

[Скрытый логистический слой 6](#_Toc159764678)

[Обучение по методу распространения обратной ошибки 7](#_Toc159764679)

[Обучение по методу сопряжённых градиентов 8](#_Toc159764680)

[Обучение по выборке без одного класса 9](#_Toc159764681)

Введение

Исследуется датасет с минеральными водами. В нём присутствует 5 классов минеральной воды. И каждый объект описывается 23 признаками, исключая сам класс. Будем проводить многоклассовую классификацию по типу минеральной воды.

Для построения модели выделим значимые признаки, по которым проведение классификации должно быть наиболее простым. Для этого найдём пары признаков, для которых на диаграмме рассеянности классы будут хорошо разделимы (рис. 1, 2).

|  |  |
| --- | --- |
| Рис. 1. Пример значимых признаком | Рис. 2. Пример незначимых признаков. |

В результате было выделено 6 значимых признаков, по которым модель будет проводить многоклассовую классификацию.

Для каждого класса имеется по 8 представителей в используемом датасете. Случайным образом для каждого класса выберем одно наблюдение в валидационную выборку и одно в тестовую.

Многоклассовая классификация прямой

Построим линейную регрессию, которая будет давать предсказания классов. Для этого настраиваем один выходной линейный нейрон и тренируем модель (рис. 3, 4). Как мы видим, в результате обучения модель приходит к ошибке порядка 0.2 по валидационной выборке.

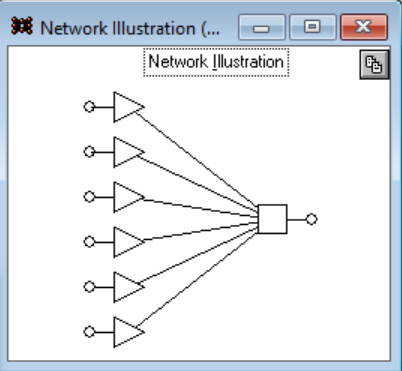
****

Рис. 3. Модель для построения линейной регрессии.

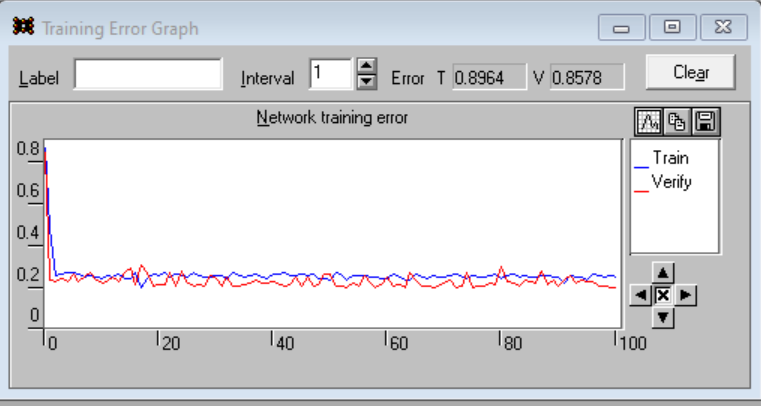
****

Рис. 4. График ошибки в течение обучения модели.

Анализируя результаты предсказания модели для датасета (рис. 5), можно сказать, что даже, если её предсказания округлять, для получения метки класса, предсказания дают неверный класс.

Использование линейной регрессии для многоклассовой классификации является плохой практикой, поскольку таким образом подразумевается некое отношение порядка между выбранными классами (2-ой класс больше 1-го, но меньше 3-го), что может влиять на получаемые модели, а также результаты такой модели сложно интерпретировать. Например, значение 1.5. Такое значение может значить и то, что наблюдение похоже как на 1, так и на 5 класс, или то, что оно очень похоже на 1 и 5 классы и немного похоже на 2 класс.

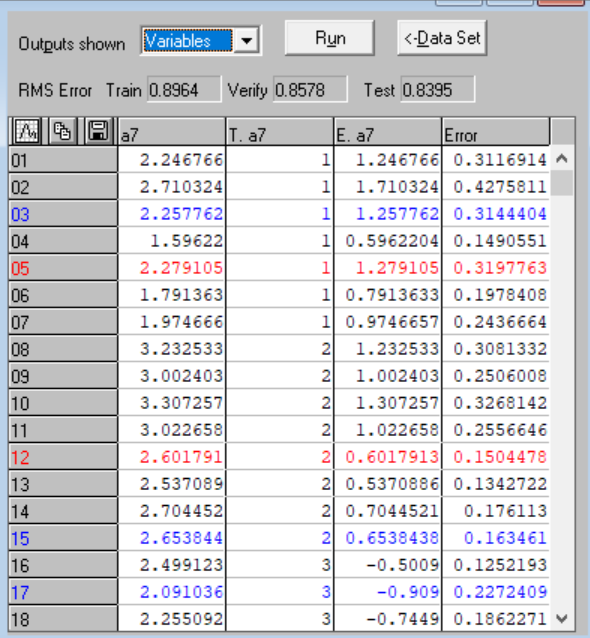
****

Рис. 5. Предсказания модели для датасета.

Многоклассовая классификация softmax

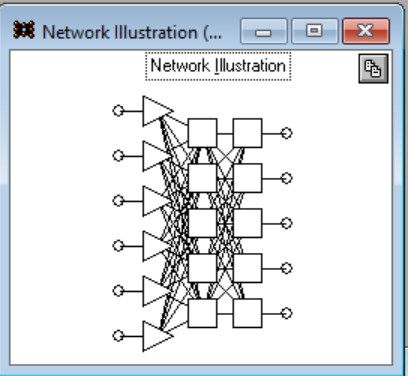
Скрытый линейный слой

Для проведения многоклассовой классификации построим следующую модель (рис. 6). Скрытый слой является линейным, а выходные нейроны используют функцию активации softmax:

*–* функция активации softmax

Такая функция активации оценивает уверенность модели в данном классе (результат – вероятность). Таким образом, 5 нейронов с функцией активации softmax оценивают уверенность модели в том, что данное наблюдение принадлежит тому или иному классу из 5.

Для обучения такой модели, необходимо преобразовать и исходную таблицу, добавив 5 переменных C­i, где переменная Ci равна 1, если данное наблюдение принадлежит i-ому классу, иначе равна 0.

****

softmax

Рис. 6. Модель для многоклассовой классификации.

Обучение по методу распространения обратной ошибки

Обучим модель по методу распространения обратной ошибки. Как видим, обучение проходит очень быстро (рис. 7).

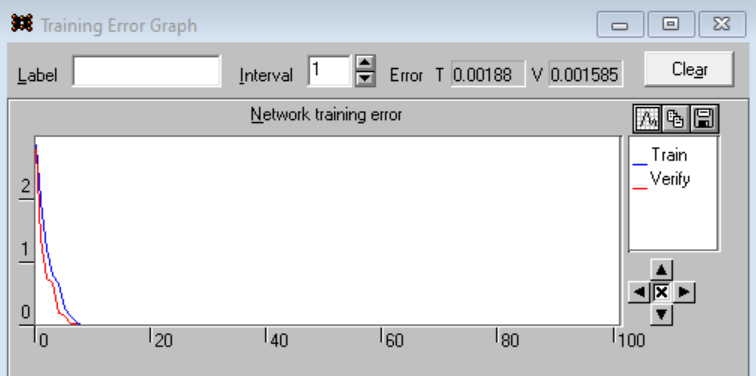
****

Рис. 7. График обучения модели.

В результате модель очень хорошо предсказывает классы минеральной воды, показывая хорошие результаты по тестовой выборке.

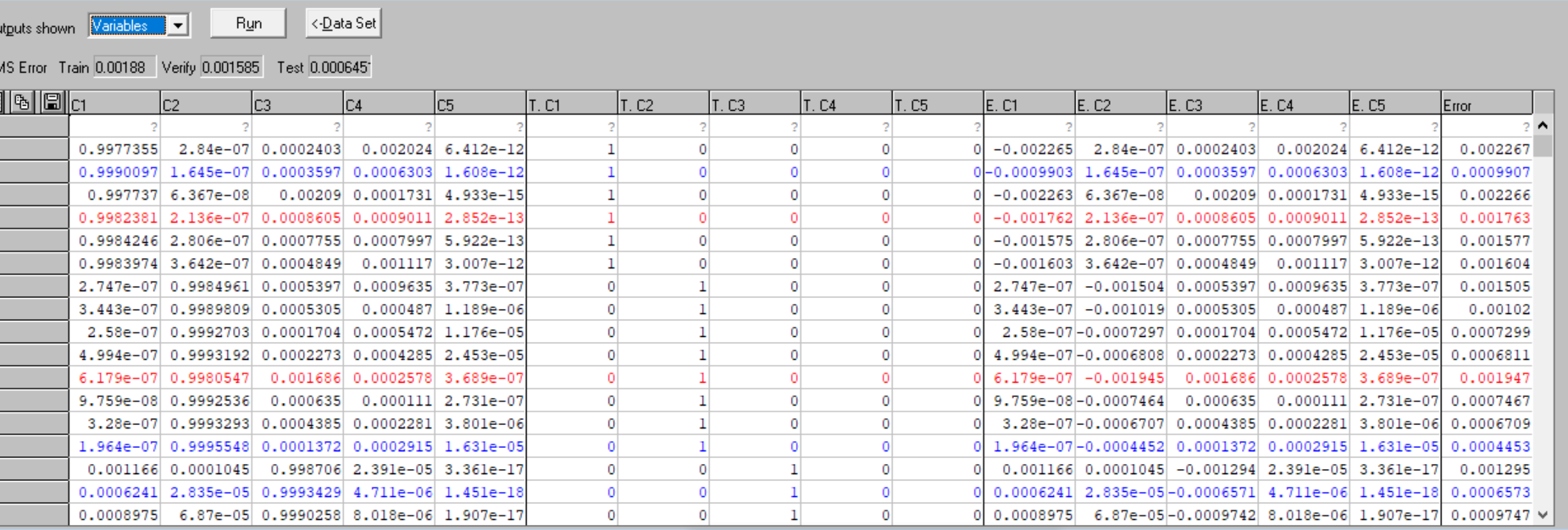
****

Рис. 8. Предсказания модели для датасета.

Обучение по методу сопряжённых градиентов

Обучим модель по другому методу: методу сопряжённых градиентов. В отличие от предыдущего метода, в данном методе предусмотрена автоматическая остановка по достижении нулевой ошибки или установившегося значения ошибки (рис. 9).

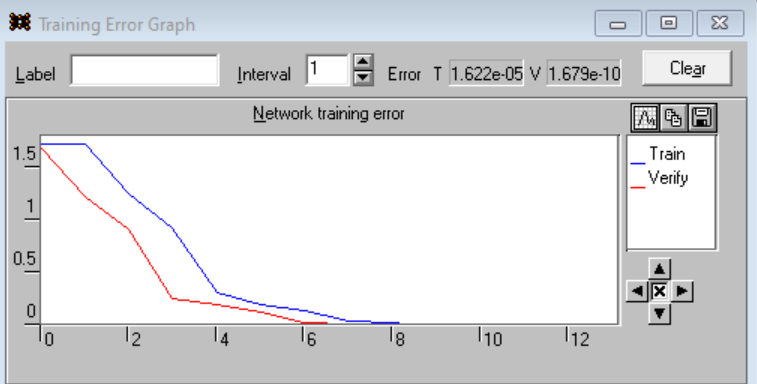
****

Рис. 9. График обучения модели.

Как видим по полученным ошибкам, даная модель получилась более качественной, показав гораздо меньшую ошибку на тестовой выборке.

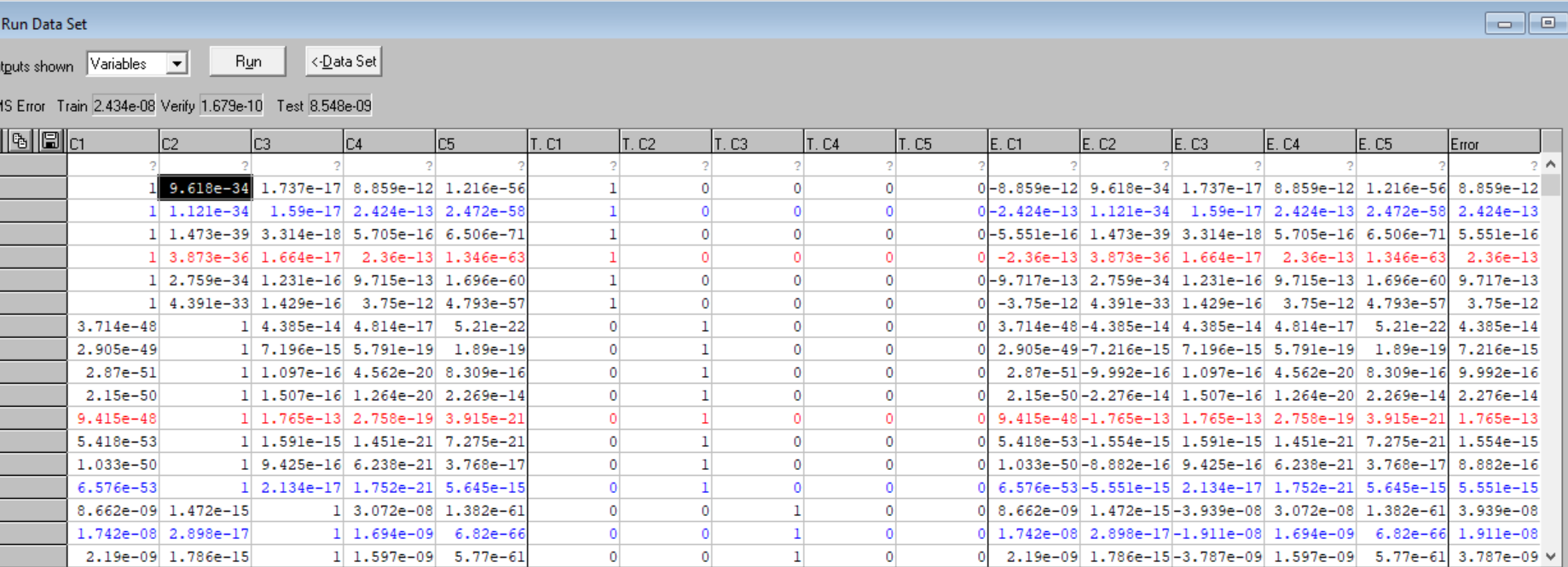
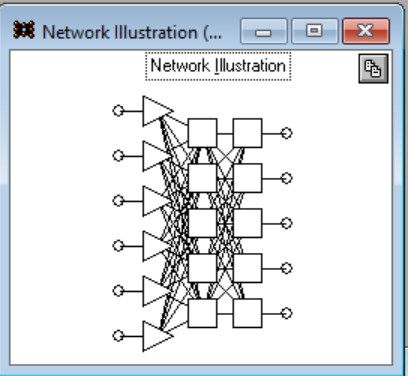
****

Рис. 10. Предсказания модели для датасета.

Скрытый логистический слой

Изменим структуру нейросети. Теперь функция активации скрытого слоя будет логистической (рис. 11). Проведём обучение по тем же методам.

****

softmax

Рис. 11. Структура нейросети.

Обучение по методу распространения обратной ошибки

При наличии сигмоиды в структуре нейросети обучение по методу распространения обратной ошибки проходит заметно медленее. Однако, этот процесс можно ускорить, подобрав другие параметры обучения.

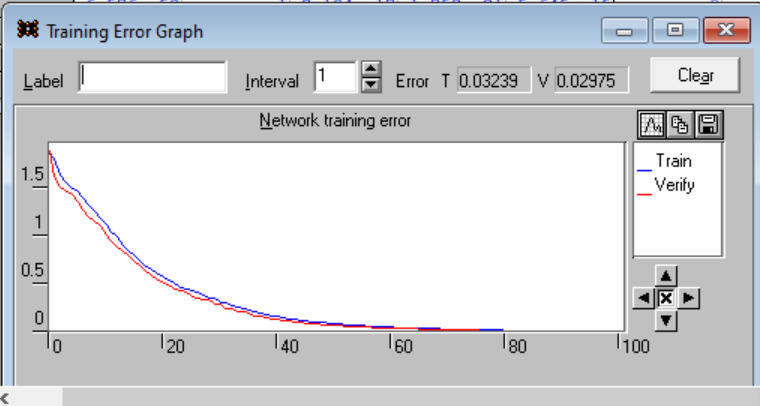
****

Рис. 12. График обучения модели.

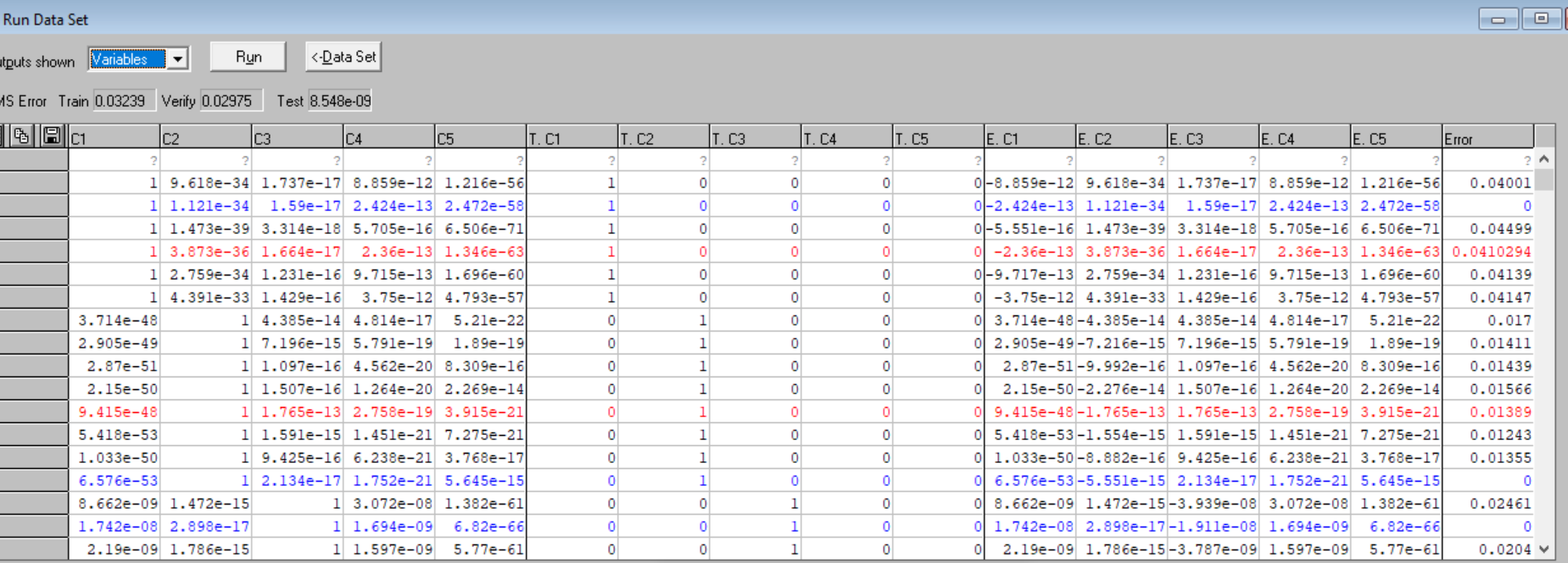
****

Рис. 13. Предсказания моделия для датасета

Обучение по методу сопряжённых градиентов

По сравнению с предыдущим методом, обучение по методу сопряжённых градиентов проходит гораздо быстрее. В данном случае это плюс, поскольку этот метод имеет меньше гиперпараметров, и более прост в использовании (рис. 14). Однако, значения ошибки для тестовой выборки получились большими, чем для предыдущего метода (рис. 15).

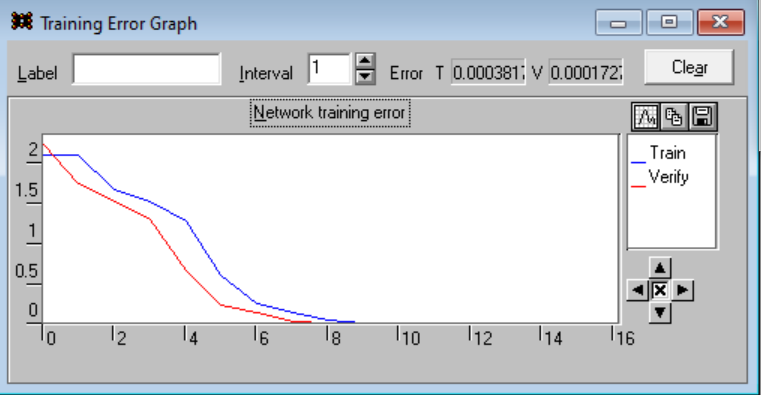
****

Рис. 14. График обучения модели.

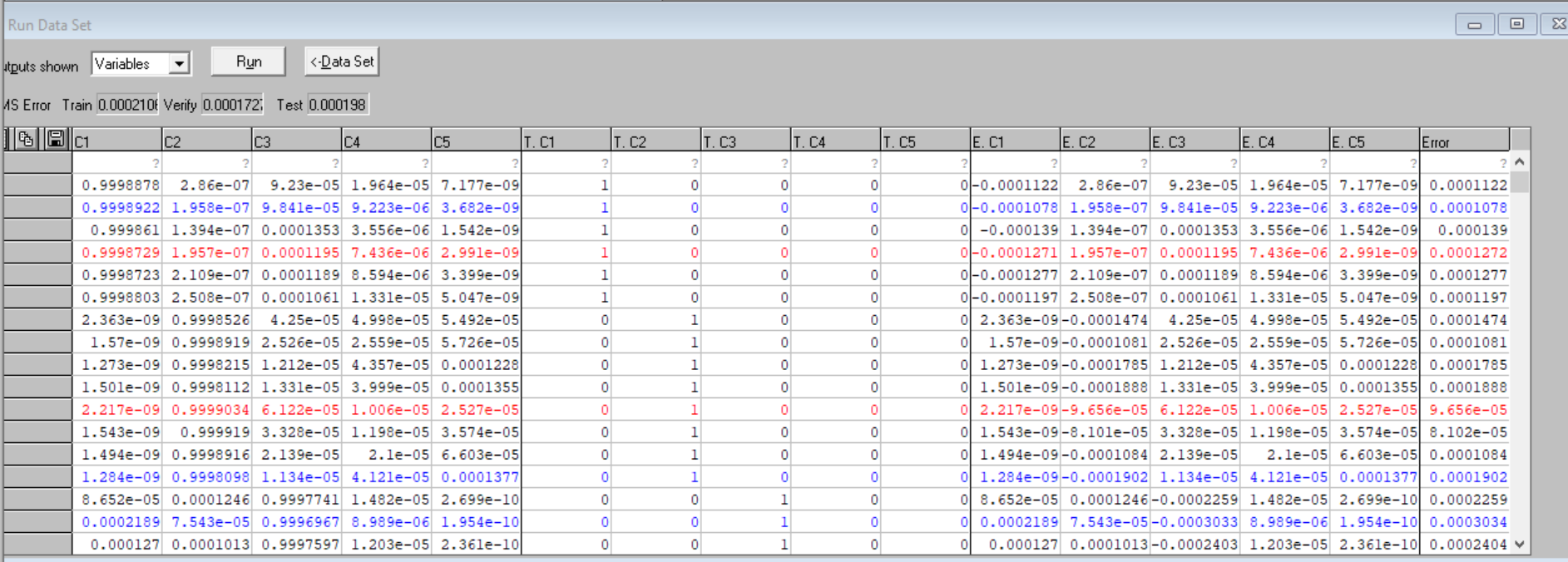
****

Рис. 15. Предсказания модели для датасета.

Обучение по выборке без одного класса

Попробуем исключить из обучающего процесса один из классов (в нашем случае 5-ый класс).

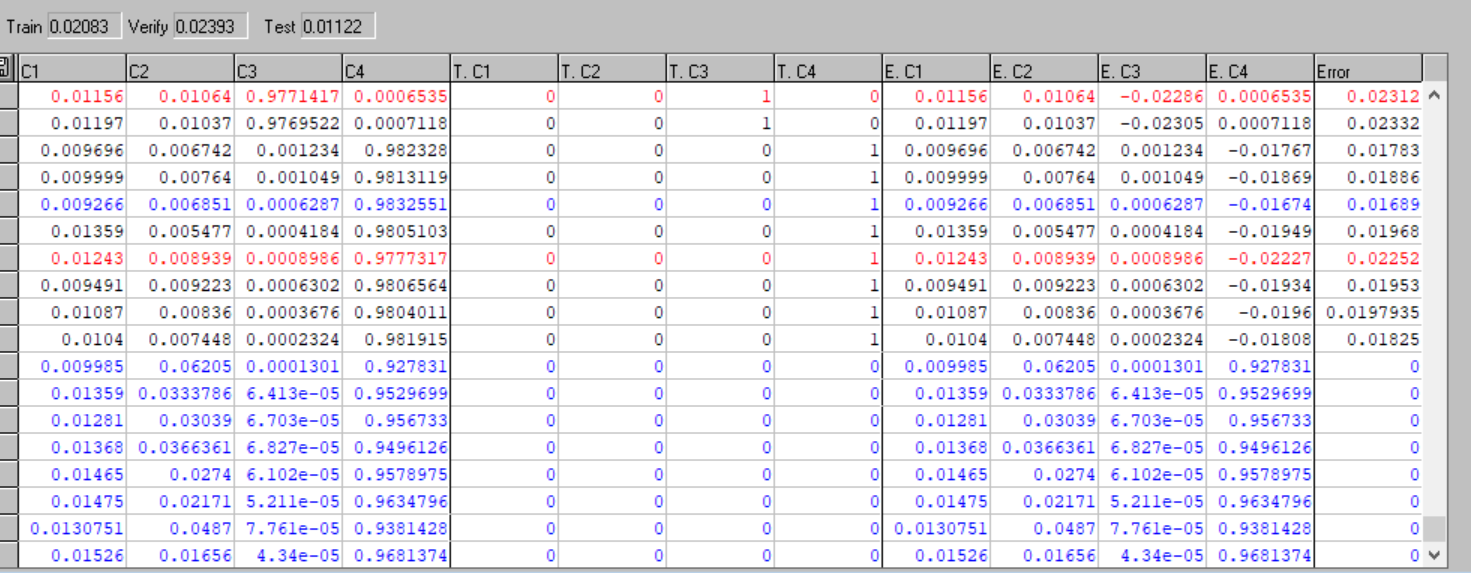
****

Рис. 16. Предсказания модели для датасета.

Новая модель содержала только 4 класса и обучалась только по их представителям. 5-й класс был исключен из процесса обучения. При тестировании модели каждому экземпляру 5-го класса она присуждает метку 4-го.

Как мы видим, для всех объектов 5-го класса была присуждена метка 4-го класса. Учитывая полученные результаты во время выбора значимых признаков, это связано с тем, что 4 и 5 классы на диаграммах рассеянности находятся довольно рядом, и порой перемешаны. За счёт этого, когда модель не знала о существовании данного класса, она не научилась отличать его от 4 класса, и всем его представителям присудила метку 4 класса.

Однако, возможен и другой исход. При подаче модели наблюдений неизвестного ей класса, она может отказаться его классифицировать, присудив всем меткам равные вероятности.