COMP 307 Assignment 1
Basic Machine Learning Algorithms

David Dobbie 300340161

April 12, 2018

Introduction

Part One: Nearest Neighbour Method

The first method of machine learning involves the simplest but also lowest amount of 'learning'. The Nearest-Neighbour classifier. This is due to the algorithm essentially using the entire training data set to guide the classification of each test point.

Each flower instance is composed of four attributes: sepal length, sepal width, petal length, and petal width. This means that each flower instance is a four dimensional vector in a four dimensional space. Therefore, we can find the closeness of an instance to a known training set by measuring its Euclidean distance to already classified instances (the training set). Once normalised, we can change the weighting of each variable, giving a design flexibility for this algorithm as well. However, in these tests the algorithm has been set up to give equal weighting to all variables.

The algorithm to deal with equal amounts of classifiers in a set of k values is to generate a random number so there is 50/50 chance between each classifier type. A note in this implementation is that with a three way tie, there are two 50/50 tests made, biasing the test to the first type iterated through, the Iris-Setosa.

Q1: Basic Nearest Neighbour Effectiveness

With k=1, the Nearest Neighbour method classifies the test instance to be the same as its nearest neighbour in the training set (k=1). This results in a **90.6667%** accuracy of classification. The raw results are provided in greater detail in appendix 1.3.

Q2: k-Nearest Neighbour Effectiveness

When k=3 (i.e. the three closest neighbours are used to help classify them), the classification accuracy is **96%**. This is an increased accuracy of **5.3333%** compared to k=1. This makes intuitive sense since using the nearest neighbour only instead of the modal of the 3 closest will be affect by outliers far more (like a Iris-Setosa in the middle of a cluster of Iris-Virginicas).

Q3: k-NN (k-Nearest Neighbour) Advantages and Disadvantages

Advantages:

- The main advantage with the k-NN algorithm is that it is simple to analyse and to implement. It has a strong link to intuition with the consideration of classifying parts of it into a n-dimensional space.
- Furthermore, it is simple to change the weighting of different variables in this distance modifier.

Disadvantages:

- The computational and memory cost becomes very large when the training set is large. It needs to go through the entire training set for each individual test point in order to give the Euclidean distance metric used. This becomes problematic for impractically large data sets.
- It is not effective if the training set is not labelled and pre-classified. This would need to be addressed first via clustering.
- For many variables there is a lot more dimensional and complexity to go around. This makes it not as feasible or effective in the simple implementation.

Q4: Applying k-fold Cross Validation

1. We would combine all of the data sets given together (training and test) into a large data set (since in our example the test set has classification 'answers' with it as well so it is valid).
2. This set would then be randomised so that any form of division of the set will tend to mirror the same distribution of classifications as the overall data set.
3. The combined data set is then split into K subsets (5 in our case).
4. For each subset it is tested while all of the other subsets would act as the training set. This is done for each subset. This means that each test set is validated with other tests, hence cross-validation.
5. With these test results we can compare the average estimate success of the algorithm with k=5 against other k values.

Q5: Dealing With No Class Labels

We would utilise K-means clustering in order to ascertain our labels for the data for the training and data sets. The methodology used would be to:

1. Place three proposed centroids (centres of mass) into the set's space at random positions.
2. Classify the data set into three sets closest to each centroid

3. Re-calculate the centroid for this data set and place it in a new position
4. Reclassify the data set into three sets of the closest centroid for each data set.
5. Repeat steps 3 and 4 again until the centroids approach asymptotically to their respective positions in the space. This would involve a threshold that would need to be designed; say have a change less than 0.01 units in Euclidean distance to stop.
6. These three sets are then given their respective labels.

Part Two: Decision Tree Learning Method

Q1: Application of the Decision Tree

The decision tree learner implemented utilised the weighted impurity methodology discussed in lectures. The the training set used was from `hepatitis-training.dat`. The testing set used to evaluate the training set was extracting from `hepatitis-test.dat`. The classification accuracy was **92.5926%**. Conversely, the baseline accuracy of classifying everything by the data-set's modal category was **85.1852%**. The resulting learned decision tree is located in appendix 1.1.

Q2: Average Performance

	Baseline Accuracy	**Classifier Accuracy**
Mean	86.49%	85.14%
Median	85.14%	85.41%

Table 1: Average classification accuracy results for the decision tree program

With the results in table 1 we can see that the classifier made from the learned decision tree performs better than the baseline accuracy with simply classifying everything by the mode. as we can see however, the difference is within 1%.

Q3: Pruning Decision Trees

- We can prune the parts of the decision tree that make up a small amount of the classified data. For example, if we have a pure leaf with 1 result and another leaf with 10% impurity but with 20 or so results classified there, we should intuitively prune the small leaf since it has a smaller *potential* effect on the classification accuracy.
- Pruning will reduce the accuracy of the decision tree on the training set since the tree is specially trained to get 100% accuracy on the training set. This is due to each node correlating with the exact composure of the training set.
- Pruning can increase accuracy for the eventual test set as it removes fidelity and detail that may lead to the decision tree to over-fit onto the training set. It allows for the training set's variance to be taken into account and generalises the decision tree for the test set.

Q4: Impurity Measures

When there are three or more classes for the decision tree to distinguish, we can have the mode class in a node be less than the majority (below 50%). Also as we get more types to classify, the size of the impurity measure numerically becomes smaller. This makes it subject more to possible noise or computational rounding off.

Part Three: Perceptron

Q1: Perceptron Performance

The perception achieved a classification accuracy of **97.9798%** after 100 training cycles and a learning rate (η) of 0.2 (results in appendix 1.2). Notably, if the training cycles were increased or the learning rate increased, there was no change in the classification accuracy. This is because it asymptotically approaches that accuracy point after 4 training cycle (fig. 1). When the learning rate was decreased down a couple orders of magnitude, the classification accuracy would decrease since it did not converge quickly enough in 100 tests.

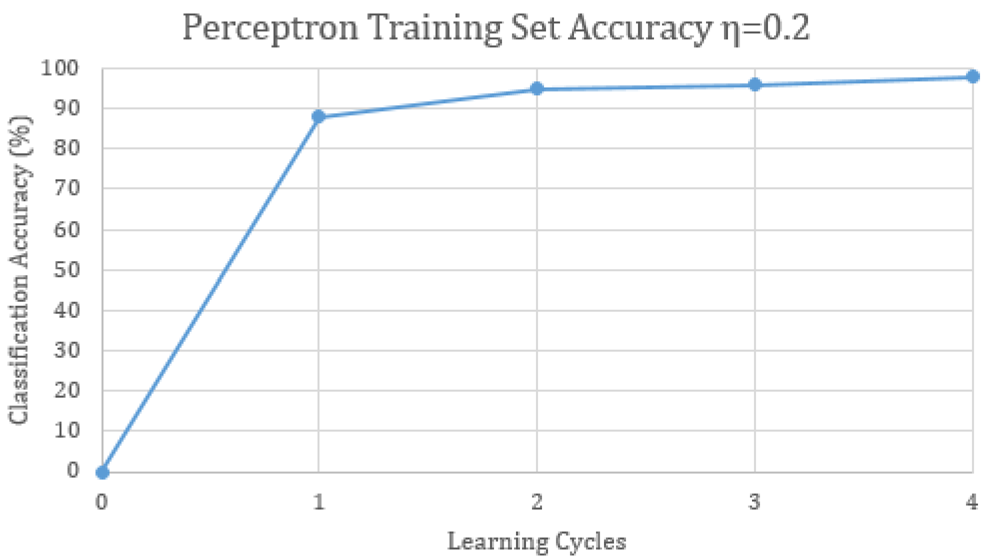

Figure 1: Performance of the Perceptron at η=0.2

The converged accuracy at a point below total classification implies that the image data set is not fully linearly separable. As in, in the multidimensional space of the feature vectors, not all data is able to be completely separated by a hyper plane established by a step threshold function. It is not likely that it is due to computational error inherent with double variables used since it would not lead to this form of linear inseparability.

The use of the feature vector with 3 out of 4 random pixels being chosen leading to a positive result gives a potential imbalance of out readings. This is due to the fact that it takes less to be identified as a white pixel than as a black pixel. This may result in the feature vector not fairly evaluating some pixels over other pixels.

Q2: Performance For Training Data

Evaluating the performance of a perception simply on its performance on the training set is an inherently haphazard methodology. This is due to the fact that it specialises the perception specifically for the training data, removing any useful generality on classifying other sets.

This process of over-fitting can be avoided by maintaining a validation set. The validation set allows for continual testing of the general accuracy of the trained perception. Once the classification inaccuracy rises again, we can stop the learning process as it becomes overly specialised onto the training set. (This assumes that the validation error has a minima at this point.)

Conclusions

These different techniques adaptive techniques of classification form the basis of machine learning techniques. We have found that each method has different aspects to them, from the spatial sorting of nearest neighbour, to logical division with decision trees, to regression of the hyper plane in the perceptron. They also take up different spaces of machine learning frameworks to achieve the same goal of classification.

1 Appendix

1.1 Part 2 Learned Decision Tree

NOTE: The printed decision tree does not suppress the leaf nodes caused by the source of training instances becoming empty resulting in a baseline leaf.

```
HISTOLOGY = true:
   SGOT = true:
      VARICES = true:
         ASCITES = true:
            BIGLIVER = true:
               FATIGUE = true:
                  Class 0, prob = 0.8
               FATIGUE = false:
                  ANTIVIRALS = true:
                     FEMALE = true:
                        Class 0, prob = 1
                     FEMALE = false:
                        BILIRUBIN = true:
                           FIRMLIVER = true:
                              Class 0, prob = 1
                           FIRMLIVER = false:
                              SPIDERS = true:
                                 Class 0, prob = 1
                              SPIDERS = false:
                                 MALAISE = true:
                                    Class 0, prob = 0.8
                                 MALAISE = false:
                                    SPLEENPALPABLE = true:
                                       Class 0, prob = 1
                                    SPLEENPALPABLE = false:
                                       Class 1, prob = 1
                        BILIRUBIN = false:
                           Class 0, prob = 1
                  ANTIVIRALS = false:
                     Class 0, prob = 0.8
            BIGLIVER = false:
               BILIRUBIN = true:
                  SPLEENPALPABLE = true:
                     FIRMLIVER = true:
                        Class 0, prob = 0.8
                     FIRMLIVER = false:
                        STEROID = true:
                           Class 0, prob = 0.8
                        STEROID = false:
                           FEMALE = true:
                              Class 0, prob = 0.8
```

```
                    FEMALE = false:
                       AGE = true:
                          SPIDERS = true:
                             Class 0, prob = 1
                          SPIDERS = false:
                             Class 1, prob = 1
                       AGE = false:
                          Class 0, prob = 0.8
           SPLEENPALPABLE = false:
              Class 0, prob = 0.8
        BILIRUBIN = false:
           Class 0, prob = 0.8
  ASCITES = false:
     BILIRUBIN = true:
        SPIDERS = true:
           Class 0, prob = 0.8
        SPIDERS = false:
           SPLEENPALPABLE = true:
              FIRMLIVER = true:
                 BIGLIVER = true:
                    ANOREXIA = true:
                       Class 0, prob = 0.8
                    ANOREXIA = false:
                       MALAISE = true:
                          Class 0, prob = 0.8
                       MALAISE = false:
                          FATIGUE = true:
                             Class 0, prob = 0.8
                          FATIGUE = false:
                             ANTIVIRALS = true:
                                FEMALE = true:
                                   Class 0, prob = 0.8
                                FEMALE = false:
                                   AGE = true:
                                      Class 0, prob = 0.8
                                   AGE = false:
                                      STEROID = true:
                                         Class 1, prob = 1
                                      STEROID = false:
                                         Class 0, prob = 1
                             ANTIVIRALS = false:
                                Class 0, prob = 0.8
                 BIGLIVER = false:
                    Class 0, prob = 0.8
              FIRMLIVER = false:
                 Class 0, prob = 0.8
           SPLEENPALPABLE = false:
```

```
                Class 0, prob = 0.8
         BILIRUBIN = false:
            Class 0, prob = 0.8
   VARICES = false:
      BILIRUBIN = true:
         ASCITES = true:
            SPIDERS = true:
               Class 0, prob = 0.8
            SPIDERS = false:
               FIRMLIVER = true:
                  Class 0, prob = 0.8
               FIRMLIVER = false:
                  BIGLIVER = true:
                     ANTIVIRALS = true:
                        FEMALE = true:
                           Class 0, prob = 0.8
                        FEMALE = false:
                           SPLEENPALPABLE = true:
                              Class 0, prob = 1
                           SPLEENPALPABLE = false:
                              FATIGUE = true:
                                 Class 0, prob = 0.8
                              FATIGUE = false:
                                 ANOREXIA = true:
                                    Class 1, prob = 1
                                 ANOREXIA = false:
                                    Class 0, prob = 1
                     ANTIVIRALS = false:
                        Class 0, prob = 0.8
                  BIGLIVER = false:
                     Class 0, prob = 0.8
         ASCITES = false:
            SPIDERS = true:
               Class 0, prob = 0.8
            SPIDERS = false:
               FATIGUE = true:
                  Class 0, prob = 0.8
               FATIGUE = false:
                  ANTIVIRALS = true:
                     FEMALE = true:
                        Class 0, prob = 0.8
                     FEMALE = false:
                        AGE = true:
                           Class 0, prob = 0.8
                        AGE = false:
                           FIRMLIVER = true:
                              Class 1, prob = 1
```

```
                        FIRMLIVER = false:
                           ANOREXIA = true:
                              SPLEENPALPABLE = true:
                                 Class 1, prob = 1
                              SPLEENPALPABLE = false:
                                 Class 0, prob = 1
                           ANOREXIA = false:
                              Class 0, prob = 0.8
               ANTIVIRALS = false:
                  Class 0, prob = 0.8
   BILIRUBIN = false:
      Class 0, prob = 0.8
SGOT = false:
   BILIRUBIN = true:
      VARICES = true:
         ANOREXIA = true:
            MALAISE = true:
               BIGLIVER = true:
                  FEMALE = true:
                     Class 0, prob = 0.8
                  FEMALE = false:
                     FIRMLIVER = true:
                        ANTIVIRALS = true:
                           STEROID = true:
                              SPLEENPALPABLE = true:
                                 Class 0, prob = 1
                              SPLEENPALPABLE = false:
                                 AGE = true:
                                    Class 0, prob = 0.8
                                 AGE = false:
                                    ASCITES = true:
                                       Class 0, prob = 1
                                    ASCITES = false:
                                       Class 1, prob = 1
                           STEROID = false:
                              Class 0, prob = 1
                        ANTIVIRALS = false:
                           Class 0, prob = 1
                     FIRMLIVER = false:
                        Class 0, prob = 1
               BIGLIVER = false:
                  Class 0, prob = 0.8
            MALAISE = false:
               ASCITES = true:
                  SPIDERS = true:
                     Class 0, prob = 0.8
                  SPIDERS = false:
```

```
                    FATIGUE = true:
                       Class 0, prob = 0.8
                    FATIGUE = false:
                       ANTIVIRALS = true:
                          STEROID = true:
                             Class 0, prob = 0.8
                          STEROID = false:
                             FEMALE = true:
                                Class 0, prob = 0.8
                             FEMALE = false:
                                AGE = true:
                                   Class 0, prob = 0.8
                                AGE = false:
                                   SPLEENPALPABLE = true:
                                      Class 0, prob = 1
                                   SPLEENPALPABLE = false:
                                      Class 1, prob = 1
                       ANTIVIRALS = false:
                          Class 0, prob = 0.8
              ASCITES = false:
                 Class 0, prob = 0.8
        ANOREXIA = false:
           Class 0, prob = 0.8
     VARICES = false:
        SPLEENPALPABLE = true:
           ANTIVIRALS = true:
              FEMALE = true:
                 Class 0, prob = 0.8
              FEMALE = false:
                 AGE = true:
                    Class 0, prob = 0.8
                 AGE = false:
                    SPIDERS = true:
                       Class 1, prob = 1
                    SPIDERS = false:
                       ASCITES = true:
                          FIRMLIVER = true:
                             Class 1, prob = 1
                          FIRMLIVER = false:
                             Class 0, prob = 1
                       ASCITES = false:
                          Class 1, prob = 1
           ANTIVIRALS = false:
              Class 0, prob = 0.8
        SPLEENPALPABLE = false:
           FIRMLIVER = true:
              ASCITES = true:
```

```
            SPIDERS = true:
               Class 0, prob = 0.8
            SPIDERS = false:
               BIGLIVER = true:
                  FATIGUE = true:
                     Class 0, prob = 0.8
                  FATIGUE = false:
                     ANTIVIRALS = true:
                        STEROID = true:
                           Class 0, prob = 0.8
                        STEROID = false:
                           ANOREXIA = true:
                              Class 0, prob = 1
                           ANOREXIA = false:
                              Class 1, prob = 1
                     ANTIVIRALS = false:
                        Class 0, prob = 0.8
               BIGLIVER = false:
                  Class 0, prob = 0.8
         ASCITES = false:
            Class 0, prob = 0.8
      FIRMLIVER = false:
         BIGLIVER = true:
            ANOREXIA = true:
               MALAISE = true:
                  FATIGUE = true:
                     Class 0, prob = 0.8
                  FATIGUE = false:
                     ANTIVIRALS = true:
                        STEROID = true:
                           FEMALE = true:
                              Class 0, prob = 0.8
                           FEMALE = false:
                              AGE = true:
                                 ASCITES = true:
                                    Class 0, prob = 1
                                 ASCITES = false:
                                    Class 1, prob = 1
                              AGE = false:
                                 Class 0, prob = 0.8
                        STEROID = false:
                           Class 0, prob = 0.8
                     ANTIVIRALS = false:
                        Class 0, prob = 0.8
               MALAISE = false:
                  Class 0, prob = 0.8
            ANOREXIA = false:
```

```
                    Class 0, prob = 0.8
                BIGLIVER = false:
                    Class 0, prob = 0.8
    BILIRUBIN = false:
        SPIDERS = true:
            FIRMLIVER = true:
                BIGLIVER = true:
                    ANOREXIA = true:
                        ANTIVIRALS = true:
                            FEMALE = true:
                                Class 0, prob = 0.8
                            FEMALE = false:
                                AGE = true:
                                    Class 0, prob = 0.8
                                AGE = false:
                                    VARICES = true:
                                        SPLEENPALPABLE = true:
                                            ASCITES = true:
                                                Class 0, prob = 1
                                            ASCITES = false:
                                                Class 1, prob = 1
                                        SPLEENPALPABLE = false:
                                            Class 0, prob = 0.8
                                    VARICES = false:
                                        Class 1, prob = 1
                        ANTIVIRALS = false:
                            Class 0, prob = 0.8
                    ANOREXIA = false:
                        Class 0, prob = 0.8
                BIGLIVER = false:
                    Class 0, prob = 0.8
            FIRMLIVER = false:
                VARICES = true:
                    ASCITES = true:
                        SPLEENPALPABLE = true:
                            STEROID = true:
                                Class 0, prob = 0.8
                            STEROID = false:
                                FEMALE = true:
                                    Class 0, prob = 0.8
                                FEMALE = false:
                                    AGE = true:
                                        Class 0, prob = 0.8
                                    AGE = false:
                                        ANTIVIRALS = true:
                                            BIGLIVER = true:
                                                Class 1, prob = 1
```

```
                                    BIGLIVER = false:
                                       Class 0, prob = 1
                                 ANTIVIRALS = false:
                                    Class 0, prob = 1
                     SPLEENPALPABLE = false:
                        Class 0, prob = 0.8
                  ASCITES = false:
                     Class 0, prob = 0.8
               VARICES = false:
                  Class 0, prob = 0.8
         SPIDERS = false:
            VARICES = true:
               ASCITES = true:
                  SPLEENPALPABLE = true:
                     ANOREXIA = true:
                        FEMALE = true:
                           Class 0, prob = 0.8
                        FEMALE = false:
                           FIRMLIVER = true:
                              Class 1, prob = 1
                           FIRMLIVER = false:
                              Class 0, prob = 1
                     ANOREXIA = false:
                        Class 0, prob = 1
                  SPLEENPALPABLE = false:
                     Class 0, prob = 0.8
               ASCITES = false:
                  Class 0, prob = 0.8
            VARICES = false:
               Class 0, prob = 0.8
HISTOLOGY = false:
   SGOT = true:
      VARICES = true:
         ASCITES = true:
            SPLEENPALPABLE = true:
               ANOREXIA = true:
                  ANTIVIRALS = true:
                     AGE = true:
                        Class 0, prob = 1
                     AGE = false:
                        FEMALE = true:
                           Class 0, prob = 1
                        FEMALE = false:
                           BILIRUBIN = true:
                              SPIDERS = true:
                                 Class 0, prob = 1
                              SPIDERS = false:
```

```
                              Class 1, prob = 1
                        BILIRUBIN = false:
                           Class 0, prob = 1
               ANTIVIRALS = false:
                  Class 0, prob = 1
            ANOREXIA = false:
               Class 0, prob = 1
         SPLEENPALPABLE = false:
            BILIRUBIN = true:
               FATIGUE = true:
                  Class 0, prob = 0.8
               FATIGUE = false:
                  ANTIVIRALS = true:
                     STEROID = true:
                        FEMALE = true:
                           Class 0, prob = 0.8
                        FEMALE = false:
                           SPIDERS = true:
                              Class 0, prob = 1
                           SPIDERS = false:
                              Class 1, prob = 1
                     STEROID = false:
                        Class 0, prob = 0.8
                  ANTIVIRALS = false:
                     Class 0, prob = 0.8
            BILIRUBIN = false:
               Class 0, prob = 0.8
      ASCITES = false:
         Class 0, prob = 0.8
   VARICES = false:
      Class 0, prob = 0.8
SGOT = false:
   BILIRUBIN = true:
      VARICES = true:
         ASCITES = true:
            ANOREXIA = true:
               SPIDERS = true:
                  SPLEENPALPABLE = true:
                     FEMALE = true:
                        Class 0, prob = 1
                     FEMALE = false:
                        BIGLIVER = true:
                           AGE = true:
                              Class 0, prob = 1
                           AGE = false:
                              FATIGUE = true:
                                 Class 0, prob = 1
```

```
                        FATIGUE = false:
                           MALAISE = true:
                              Class 0, prob = 1
                           MALAISE = false:
                              STEROID = true:
                                 Class 0, prob = 0.8
                              STEROID = false:
                                 ANTIVIRALS = true:
                                    Class 0, prob = 1
                                 ANTIVIRALS = false:
                                    FIRMLIVER = true:
                                       Class 0, prob = 1
                                    FIRMLIVER = false:
                                       Class 1, prob = 1
                  BIGLIVER = false:
                     Class 0, prob = 1
            SPLEENPALPABLE = false:
               Class 0, prob = 1
         SPIDERS = false:
            Class 0, prob = 1
      ANOREXIA = false:
         Class 0, prob = 1
   ASCITES = false:
      SPLEENPALPABLE = true:
         FIRMLIVER = true:
            BIGLIVER = true:
               ANOREXIA = true:
                  Class 0, prob = 0.8
               ANOREXIA = false:
                  MALAISE = true:
                     Class 0, prob = 0.8
                  MALAISE = false:
                     FATIGUE = true:
                        Class 0, prob = 0.8
                     FATIGUE = false:
                        ANTIVIRALS = true:
                           FEMALE = true:
                              Class 0, prob = 0.8
                           FEMALE = false:
                              SPIDERS = true:
                                 Class 0, prob = 1
                              SPIDERS = false:
                                 Class 1, prob = 1
                        ANTIVIRALS = false:
                           Class 0, prob = 0.8
            BIGLIVER = false:
               Class 0, prob = 0.8
```

```
            FIRMLIVER = false:
               Class 0, prob = 0.8
         SPLEENPALPABLE = false:
            Class 0, prob = 0.8
   VARICES = false:
      Class 0, prob = 0.8
BILIRUBIN = false:
   VARICES = true:
      ASCITES = true:
         SPLEENPALPABLE = true:
            BIGLIVER = true:
               AGE = true:
                  Class 0, prob = 1
               AGE = false:
                  ANOREXIA = true:
                     MALAISE = true:
                        FEMALE = true:
                           Class 0, prob = 1
                        FEMALE = false:
                           ANTIVIRALS = true:
                              FIRMLIVER = true:
                                 STEROID = true:
                                    FATIGUE = true:
                                       Class 0, prob = 1
                                    FATIGUE = false:
                                       SPIDERS = true:
                                          Class 0, prob = 1
                                       SPIDERS = false:
                                          Class 1, prob = 1
                                 STEROID = false:
                                    Class 0, prob = 1
                              FIRMLIVER = false:
                                 Class 0, prob = 1
                           ANTIVIRALS = false:
                              Class 0, prob = 1
                     MALAISE = false:
                        Class 0, prob = 1
                  ANOREXIA = false:
                     Class 0, prob = 1
            BIGLIVER = false:
               Class 0, prob = 1
         SPLEENPALPABLE = false:
            Class 0, prob = 1
      ASCITES = false:
         Class 0, prob = 0.8
   VARICES = false:
      Class 0, prob = 0.8
```

```
Classification accuracy: 92.5926%
Baseline accuracy: 85.1852%
```

1.2 Part 3 Perceptron Feature and Weight Vectors ($\eta = 0.2$)

```
Feature values {row,col,sgn} (all in order)
(dummy)
{1, 7, 0} {0, 9, 0} {8, 8, 0} {4, 5, 1}
{1, 7, 1} {1, 5, 0} {7, 6, 1} {4, 2, 1}
{2, 2, 1} {6, 8, 1} {7, 6, 1} {8, 9, 0}
{7, 9, 1} {4, 3, 1} {2, 3, 1} {4, 1, 1}
{3, 8, 1} {4, 2, 1} {7, 9, 1} {1, 9, 0}
{6, 5, 0} {2, 8, 0} {0, 2, 0} {8, 6, 1}
{0, 9, 0} {0, 6, 1} {3, 8, 1} {3, 4, 0}
{6, 0, 0} {6, 1, 0} {4, 9, 0} {3, 7, 0}
{8, 2, 1} {1, 3, 1} {9, 8, 0} {0, 7, 0}
{3, 6, 1} {5, 4, 0} {0, 9, 1} {3, 7, 0}
{6, 0, 1} {6, 5, 1} {5, 4, 1} {2, 0, 0}
{1, 4, 0} {0, 7, 1} {7, 7, 1} {7, 3, 1}
{5, 9, 1} {8, 1, 0} {2, 6, 0} {0, 3, 0}
{0, 1, 0} {5, 0, 1} {4, 7, 0} {3, 5, 1}
{2, 0, 1} {6, 4, 0} {6, 1, 0} {9, 8, 0}
{1, 4, 1} {9, 8, 0} {0, 8, 1} {7, 8, 1}
{8, 3, 1} {1, 0, 1} {3, 4, 1} {6, 5, 1}
{0, 9, 1} {6, 8, 1} {4, 8, 0} {9, 9, 0}
{5, 5, 1} {3, 3, 1} {4, 3, 0} {0, 8, 0}
{0, 6, 0} {1, 9, 0} {9, 7, 0} {2, 8, 0}
{8, 9, 0} {7, 8, 1} {5, 8, 0} {1, 2, 0}
{2, 5, 0} {6, 2, 0} {5, 3, 1} {2, 4, 0}
{1, 8, 0} {1, 1, 1} {7, 6, 0} {9, 5, 0}
{0, 0, 1} {9, 1, 0} {1, 9, 1} {3, 2, 1}
{2, 5, 0} {6, 7, 1} {2, 2, 1} {4, 1, 1}
{6, 9, 0} {2, 5, 1} {4, 9, 1} {2, 5, 0}
{8, 3, 1} {3, 9, 0} {7, 9, 1} {7, 6, 1}
{3, 5, 1} {6, 6, 1} {8, 2, 1} {4, 4, 1}
{6, 1, 0} {3, 3, 1} {5, 3, 0} {8, 2, 1}
{3, 6, 1} {8, 6, 0} {1, 4, 0} {2, 9, 1}
{5, 0, 1} {6, 4, 1} {2, 9, 1} {4, 4, 0}
{5, 9, 0} {3, 4, 0} {8, 8, 1} {5, 8, 1}
{2, 5, 1} {4, 4, 0} {4, 4, 0} {9, 3, 0}
{7, 7, 0} {7, 3, 0} {3, 8, 0} {5, 2, 0}
{3, 7, 0} {9, 7, 1} {3, 4, 1} {2, 1, 1}
{3, 2, 1} {3, 5, 1} {2, 2, 0} {6, 4, 0}
{7, 0, 1} {4, 0, 0} {5, 3, 1} {0, 6, 1}
{5, 0, 0} {4, 5, 0} {9, 5, 0} {4, 8, 0}
{1, 6, 0} {3, 1, 0} {8, 4, 1} {2, 0, 1}
{0, 6, 0} {7, 0, 1} {6, 7, 1} {8, 4, 1}
```

```
{2, 6, 1} {7, 7, 0} {5, 7, 1} {8, 3, 0}
{2, 2, 1} {6, 7, 1} {5, 5, 0} {0, 2, 0}
{0, 6, 1} {1, 1, 1} {3, 0, 0} {4, 8, 1}
{0, 1, 0} {9, 0, 1} {8, 8, 1} {7, 7, 0}
{8, 9, 1} {0, 7, 0} {7, 8, 1} {7, 1, 0}
{0, 8, 1} {4, 5, 0} {1, 4, 1} {6, 2, 0}
{3, 5, 0} {6, 6, 1} {6, 7, 1} {1, 9, 1}
{6, 2, 1} {2, 7, 0} {5, 1, 1} {9, 9, 1}
{9, 2, 0} {5, 3, 0} {4, 4, 0} {2, 7, 1}

The weights vector acquired w:
-1.2
0.2
0.2
0.4
0.4
0.8
-0.2
0.2
0.4
0.8
0.4
5.55112e-017
-0.4
5.55112e-017
-0.6
-0.6
0.4
5.55112e-017
1
0.6
0.4
0.2
5.55112e-017
0.6
-0.4
1
1
0.2
5.55112e-017
0.6
0.4
-0.2
0.2
-0.2
0.4
1
```

```
0.8
0.2
0.6
-0.4
0.2
5.55112e-017
0.4
1
-1
-0.2
5.55112e-017
-0.2
1
-0.4
```

1.3 Part 1 Nearest Neighbour k=1 Results

ID	Test Type	Answers Type	Matching
0	Iris-Setosa	Iris-Setosa	Yes
1	Iris-Setosa	Iris-Setosa	Yes
2	Iris-Setosa	Iris-Setosa	Yes
3	Iris-Setosa	Iris-Setosa	Yes
4	Iris-Setosa	Iris-Setosa	Yes
5	Iris-Setosa	Iris-Setosa	Yes
6	Iris-Setosa	Iris-Setosa	Yes
7	Iris-Setosa	Iris-Setosa	Yes
8	Iris-Setosa	Iris-Setosa	Yes
9	Iris-Setosa	Iris-Setosa	Yes
10	Iris-Setosa	Iris-Setosa	Yes
11	Iris-Setosa	Iris-Setosa	Yes
12	Iris-Setosa	Iris-Setosa	Yes
13	Iris-Setosa	Iris-Setosa	Yes
14	Iris-Setosa	Iris-Setosa	Yes
15	Iris-Setosa	Iris-Setosa	Yes
16	Iris-Setosa	Iris-Setosa	Yes
17	Iris-Setosa	Iris-Setosa	Yes
18	Iris-Setosa	Iris-Setosa	Yes
19	Iris-Setosa	Iris-Setosa	Yes
20	Iris-Setosa	Iris-Setosa	Yes
21	Iris-Setosa	Iris-Setosa	Yes
22	Iris-Setosa	Iris-Setosa	Yes
23	Iris-Setosa	Iris-Setosa	Yes
24	Iris-Setosa	Iris-Setosa	Yes
25	Iris-versicolor	Iris-versicolor	Yes
26	Iris-versicolor	Iris-versicolor	Yes
27	Iris-Virginica	Iris-versicolor	No
28	Iris-versicolor	Iris-versicolor	Yes
29	Iris-versicolor	Iris-versicolor	Yes
30	Iris-versicolor	Iris-versicolor	Yes
31	Iris-versicolor	Iris-versicolor	Yes
32	Iris-versicolor	Iris-versicolor	Yes
33	Iris-Virginica	Iris-versicolor	No
34	Iris-versicolor	Iris-versicolor	Yes
35	Iris-versicolor	Iris-versicolor	Yes
36	Iris-versicolor	Iris-versicolor	Yes
37	Iris-versicolor	Iris-versicolor	Yes
38	Iris-versicolor	Iris-versicolor	Yes
39	Iris-versicolor	Iris-versicolor	Yes
40	Iris-versicolor	Iris-versicolor	Yes
41	Iris-versicolor	Iris-versicolor	Yes
42	Iris-versicolor	Iris-versicolor	Yes
43	Iris-versicolor	Iris-versicolor	Yes
44	Iris-versicolor	Iris-versicolor	Yes
45	Iris-versicolor	Iris-versicolor	Yes
46	Iris-versicolor	Iris-versicolor	Yes
47	Iris-versicolor	Iris-versicolor	Yes
48	Iris-versicolor	Iris-versicolor	Yes
49	Iris-versicolor	Iris-versicolor	Yes
50	Iris-Virginica	Iris-Virginica	Yes
51	Iris-Virginica	Iris-Virginica	Yes
52	Iris-versicolor	Iris-Virginica	No
53	Iris-Virginica	Iris-Virginica	Yes
54	Iris-Virginica	Iris-Virginica	Yes
55	Iris-Virginica	Iris-Virginica	Yes
56	Iris-Virginica	Iris-Virginica	Yes
57	Iris-Virginica	Iris-Virginica	Yes
58	Iris-versicolor	Iris-Virginica	No
59	Iris-versicolor	Iris-Virginica	No
60	Iris-Virginica	Iris-Virginica	Yes
61	Iris-Virginica	Iris-Virginica	Yes
62	Iris-Virginica	Iris-Virginica	Yes
63	Iris-versicolor	Iris-Virginica	No
64	Iris-Virginica	Iris-Virginica	Yes
65	Iris-Virginica	Iris-Virginica	Yes
66	Iris-Virginica	Iris-Virginica	Yes
67	Iris-Virginica	Iris-Virginica	Yes
68	Iris-Virginica	Iris-Virginica	Yes
69	Iris-Virginica	Iris-Virginica	Yes
70	Iris-Virginica	Iris-Virginica	Yes
71	Iris-Virginica	Iris-Virginica	Yes
72	Iris-Virginica	Iris-Virginica	Yes
73	Iris-Virginica	Iris-Virginica	Yes
74	Iris-versicolor	Iris-Virginica	No