AFRICAN INSTITUTE FOR MATHEMATICAL SCIENCES

(AIMS-GH, ACCRA)

Name: Ijuptil Joseph Kwajighu Assignment Number: 3
Course: Machine Learning and Pattern Recognition Date: May 15, 2021

Support Vector Machine and Iris data

- Using the **ex 9.1.py** file, the following observations were made:
  - With low training data set, the SVM algorithm converges with respect to low SVM regularization parameters, which will strengthen the algorithm.
  - High SVM regularization parameters would results to high training set ratio and low test ratio. This would lead to non-converges.
  - High SVM regularization parameters with low training set ration would results to high test ratio, which implies non-converges.
  - We can apply feature selection method in SVM for the optimal parameter setting in the algorithm.

Deep learning architecture: ResNet-50

Deep networks can hard to train because of vanishing gradient problems. ResNet-50 is a technique that uses skip connection to resolve the problem of vanishing gradient.The ResNet-50 consists of **batch normalization layer, convolutional layer and average pooling layer**. The basic architures for ResNet-50 include the following:

- Shortcut connection is added using skip connection
- Skip connection would mitigate the gradient vanishing problem by allowing the shortcut path for gradient to flow through.
- The connection convert the architecture to residual networks

Manifold learning

1. For the **ex 10.1.py**, the mapping quality is sensitive to the n-neighbors parameter. Setting the parameter too small would results to continuous manifold, which can falsely be divided into disjoint sub-manifolds. In this case the mapping does not reflect any global properties. On the hand, increasing the parameter would results to better and better separation.
2. For the **ex 10.2.py**, the PCA algorithm shows a difference in comparison to the handwritten numbers. The PCA algorithm uses specific rules to choose an arbitrary linear projection. However, the algorithm does not detect the correct dimension. In comparison to the ISOMAP and Locally Linear Embedding (LLE).
    - The main advantage of ISOMAP is that it depend on only one parameter - the neighborhood. Selecting a suitable neighborhood size, determines the locality and globality of the data.
    - The main control parameter of the LLE algorithm is also the number of neighbours of each data point. LLE projects data to a lower space by unfolding the manifold. If n-neigbor is too high, it causes smoothing or elimination of small-scale structures in the manifold, the mapping loses its nonlinear character and the algorithm would tend to behaves like traditional PCA. Unlike the ISOMAP a suitable n-neigbor most be chosen for the algorithm to give better output. The figure generated by the ISOMAP shows similar behavior
    - In conclusion, all three algorithm try to map points on a high-dimensional non-linear manifold to a lower dimensional set of coordinates but one is better than the other.

Prediction in longitudinal studies

One technique for predicting the disease progression is SVM algorithm. To apply the algorithm, we perform the following steps

- Input the data set
- Pre-processed/clean the data
- Split the data into training/test sets
- Create the model using SVM by importing the Support Vector Classifier(SVC).
- Train the model
- Make prediction based on the physiological features
- Evaluate the model and make improvement if possible.

The disease progression depend on the method used. Unlike mixed-effects models, the linear regression models or the generalized estimating equations, the SVM take into account observational dependence within physiological features. VM separates the hyperplane parameters from temporal trend parameters. Hence, to predict whether brain tissue is going to be normal in the future, or will be damaged would depend on the physiological parameters that influence the growth of tumor in the brain.