GNR 652
Assignment 3
Softmax Classifier for Indian Pines Hyperspectral Image data

Amit Lohan
183079033

Objective:- To program a softmax classifier for predicting the classes of pixels of a hyperspectral image given the groundtruth data for image file.

Method:- 50% of the data is used for training and rest 50% is used for testing of the trained model. Following equations have been used for evaluating hypothesis, cost and derivative of cost function with respect to hypothesis for minimizing the cost function with gradient descent:-

$$h_\theta(x^{(i)}) = \begin{bmatrix} p(y^{(i)} = 1|x^{(i)};\theta) \\ p(y^{(i)} = 2|x^{(i)};\theta) \\ \vdots \\ p(y^{(i)} = k|x^{(i)};\theta) \end{bmatrix} = \frac{1}{\sum_{j=1}^{k} e^{\theta_j^T x^{(i)}}} \begin{bmatrix} e^{\theta_1^T x^{(i)}} \\ e^{\theta_2^T x^{(i)}} \\ \vdots \\ e^{\theta_k^T x^{(i)}} \end{bmatrix}$$

$$J(\theta) = -\frac{1}{m}\left[\sum_{i=1}^{m}\sum_{j=1}^{k} 1\left\{y^{(i)} = j\right\} \log \frac{e^{\theta_j^T x^{(i)}}}{\sum_{l=1}^{k} e^{\theta_l^T x^{(i)}}}\right] + \frac{\lambda}{2}\sum_{i=1}^{k}\sum_{j=0}^{n}\theta_{ij}^2$$

$$\nabla_{\theta_j} J(\theta) = -\frac{1}{m}\sum_{i=1}^{m}\left[x^{(i)}(1\{y^{(i)} = j\} - p(y^{(i)} = j|x^{(i)};\theta))\right] + \lambda\theta_j$$

A vectorized form of above equation is used in gradient descent.

Result:- An average accuracy of **82.34%** is obtained for the programed classifier with accuracy varyring in the range of **81.3 to 82.7%** for different runs of program as the 50% of the training data is being sampled randomly.