Assignment 3 CSc 486b

4.1 Abstract:

This assignment was incredibly difficult to understand. I skipped part 2 and started off with part 3. I found the train function to be on of the most difficult things to approach. I had to do more personal research into fully understanding what to do and even then I was very unsure as I wouldn't know if it worked until I finished the entirety of it. At first run I continually got both my accuracies to be incorrect. Ranging form 9-20% for both the training and validation. After some difficult reading and assistance from the professor and online sources I was able to find out where I was going wrong and fixed it to the point where my numbers were more adequate within the 35-40% range for Validation accuracy and 55-70% train accuracy especially after my completion with the main method. The main method was where I had trouble with the Cross validation loop. This part confused me at first until I read about how the folding actually works. I then came to the conclusion that I needed to take 1/5 of it out of the list and use that as training to compare it to the other 4/5s. Overall I found this code to be extremely confusing and it really needed a better rundown for how it actually works.

4.2 Derivation of the gradient for cross entropy:

$$
\begin{aligned}
\partial Ex \;/\; \partial oxj &= \partial \;/\; \partial oxj \;\left(-\sum k \;[txk \;\log(oxk)]+(1-txk)\;\log(1-oxk)]\right)\\
&= -\partial \;/\; \partial oxj \;\left(\sum k \;[txk \;\log(oxk)]+(1-txk)\;\log(1-oxk)]\right)\\
&= -\partial \;/\; \partial oxj \;\left([txj \;\log(oxj)] \;+\; (1-txj)\;\log(1-oxj)]\right)\\
&= -(txjoxj \;-1-txj \;/\; 1-oxj)\\
&= -txjoxj \;+\; (1-txj \;/\; 1-oxj)
\end{aligned}
$$

4.3 Training Curve HOG:

*Note: These are as close as I can get as I ran out of time for testing.