COSC 480 – Miniproject #1

Task: Using keras, TensorFlow, and Python, create a sentiment analysis for Amazon product reviews.

For this project you will create a python program that will do the following things:

1.) Import the data set posted to the course website. The data is relatively clean, but it is in a pseudo-XML format and will need to be parsed in order to be useful. You should split your data into 50/50 training and test sets. Make sure that it's an even split across all product categories.

2.) Develop a convolutional neural network model to predict the polar sentiment value based on the text of the review.

3.) Evaluate your CNN on accuracy. Make tweaks to the CNN parameters as necessary to maximize your accuracy.

4.) Submit your group's sufficiently commented notebook file via Blackboard by **11:59pm on September 28th**. This project will be worked on in groups of up to 3. We will choose groups in class on September 13th. The header comment for your notebook should include:

File name
Names of all members of the group
Project name and description
Any special execution instruction (extra libraries you had to install, etc.)
Date

Credit: Mark Dreze, JHU for the dataset
(http://www.cs.jhu.edu/~mdredze/datasets/sentiment/index2.html)

And Kaggle for the inspiration.