**Reflection on Image Processing w/ PySpark Final Project Assignment**

For the final project, I decided to create an assignment that walked a student through creating a simple PySpark setup for image classification. I originally started with a much more traditional assignment structure, where the student was given a small amount of starter code and they were asked to complete it, with the help of some given instructions. However, I didn't really like how the assignment flowed. My goal was to create an assignment for people that had never used PySpark before, and I felt that this wasn't a great way to introduce it. I then had the idea of using Jupyter notebooks for the assignment, which I really like. That allowed me to create an assignment that easily included questions and directions while allowing students to run small code snippets, and edit the code themself. The assignment follows the basic structure of having students read the given code, and then answer questions to get them thinking about the functionality of the code. They are also asked to extend the code in several cases, opening up an avenue for them to show their understanding of the code they have already seen. At the end, it becomes more open ended, asking the students to edit the code themself and to see if they can increase the accuracy of the classification system through either feature engineering or changing classification model. The students aren't expected to make significant breakthroughs, but it should give a good perspective on thinking about changes that could be made to the very simple pipeline.

I've included three different documents with the assignment: assignment.ipynb which is the jupyter notebook that will be given to the students, completed_assignment.ipynb which is the jupyter notebook with the question filled out, and rubric.pdf which contains basic instructions on how to grade the jupyter notebook to try and ensure consistency. The CIFAR batches are also included within the assignment to be used by pyspark. The assignment is overall designed to be easy to complete, if a student is willing to put in the effort to read through the documentation.

Overall I'm happy with the way the assignment worked out, it does a good job walking a potential student through a pyspark assignment. I think the choice to change to a jupyter style notebook worked out really well :).