Lab_9_Bucketized_Features_Using_Quantiles_and_Feature_Crosses

April 24, 2019

Copyright 2017 Google LLC.

```
In [0]: # Licensed under the Apache License, Version 2.0 (the "License");
        # you may not use this file except in compliance with the License.
        # You may obtain a copy of the License at
        #
        # https://www.apache.org/licenses/LICENSE-2.0
        #
        # Unless required by applicable law or agreed to in writing, software
        # distributed under the License is distributed on an "AS IS" BASIS,
        # WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
        # See the License for the specific language governing permissions and
        # limitations under the License.
```

1 Lab 9: Bucketized Features Using Quantiles and Feature Crosses

Learning Objectives: * Learn to use quantiles to create bucketized features. * Learn how to introduce feature crosses. * Starting from just having the data loaded, train a linear classifier to predict if an individual's income is at least 50k using numerical features, categorical features, bucketized features, and feature crosses.

1.0.1 Standard Set-up

We begin with the standard set-up as seen in the last lab again using the census data set.

```
In [0]: import math

        from IPython import display
        from matplotlib import cm
        from matplotlib import gridspec
        from matplotlib import pyplot as plt
        import numpy as np
        import pandas as pd
        from mpl_toolkits.mplot3d import Axes3D
        from sklearn import metrics
        import tempfile
```