# Machine Learning Engineer Nanodegree

## Capstone Proposal

Marwan Morsy
August 29, 2018

## Landmarks Recognition

### Domain Background

I was interested in Computer vision field which deals with how computers can be made for gaining high-level understanding from digital images or videos. It automates tasks that the human visual system can do. object recognition has gained tremendous interest by engineers and scientists in artificial intelligence and computer vision.
So I chose the Landmark Recognition problem to solve using Convolution Neural Networks which outperforms many traditional machine learning approaches computer vision and pattern recognition.
Some related academic research relevant to this domain :

1. Tour the World: building a web-scale landmark recognition engine
2. Landmark recognition with compact BoW histogram and ensemble ELM

My motivation for using this problem that I'm in senior year in college and my graduation project will be related to computer vision topic and I thought this problem would be a good practice on it especially on a quite large dataset.

### Problem Statement

Did you ever go through your vacation photos and ask yourself: What is the name of this temple I visited in China? Who created this monument I saw in France? Landmark recognition can help! This technology can predict landmark labels directly from image pixels, to help people better understand and organize their photo collections. This problem is inspired by Google Landmark Recognition Challenge on kaggle.
The problem is a multi classification problem. I will predicting the Landmark class which the input image belongs to.

## Datasets and Inputs

The dataset was originally in this link **landmark3d** of 25 classes with over 55,000 images, It had some problems :

- The dataset was a bunch of lists that have links to download the images from.
- The dataset was quite large and slow to download and upload it on google drive.
- The data was divided into training and validation sets only.

I will use a smaller version of this dataset and divide it into training (16099 images), validation (973 images) and testing (2486 images) sets,There is no imbalance in the dataset the images don't have the same dimensions size (height, width) but all of them are quite big images of size approximately 700px* 1000px or 1000px* 700px
This dataset will be included in this link Modified Dataset

## Solution Statement

I plan to use a simple sequential model with Keras library. This model will train on the bottleneck features from the pre-trained models like VGG16, VGG19 by transfer learning. Accuracy will be used as the model performance evaluation metric.

## Benchmark Model

The Benchmark model will be the basic CNN model. I plan to compare the results of the my Benchmark model to the solution model . I will compare the accuracy of each model to see which is more effective, as well as compare the speed of the two models.
My basic CNN model will be using a convolutional neural network (CNN) with Keras library. The CNN will consists of three main types of neural layers: (1) convolutional layers, (2) Max pooling layers, and (3) fully connected layers.
The convolutional layer use filters to generate feature maps. The pooling layer reduces the dimensions of the input volume and helps to reduce overfitting problems. The fully connected layer converts 2D feature maps into 1D feature vector for classification.
I will use techniques such as dropout in this model also.

## Evaluation Metrics

The overall evaluation metric that will be used for the benchmark model and the solution model is accuracy.

$$Accuracy = number\ of\ correctly\ identified\ instances \div all\ instances$$

## Project Design

First step in the project will take  a subset of **landmark3d** dataset which has over 55,000 images. I will be using 10 classes instead of 25 classes and divide the dataset which is originally divided into  training and validation sets only into training, validation and testing which will reduce the total images to 19558 images. This is done due to computational.
After preparing the data I will do image processing like image scaling and data augmentation to help the model in training and get better results.
After that I will build the basic CNN network which is the benchmark model , begin training on the training data and choosing the best model every epoch by using validation set and will evaluate the best saved model on the testing data and get the testing accuracy.
After that I will apply transfer learning by obtaining the bottleneck features from the pre-trained features such as VGG16 and VGG19 for all the three sets, build a simpler model to train on the bottleneck features and evaluate on the testing set. I will use the best pretrained model that has the best testing accuracy as the solution model and compare it with the basic CNN model testing accuracy.
I will show some visualizations like showing what is the effect of  data augmentation on a sample image and also two graphs through the training one to compare the training accuracy and validation accuracy the other to compare the training loss and validation loss for each model.