**Machine Learning Assessment Assignment**

**Performed by – Kaushik Gnanasekar, Shreyas Ramanathan**

This documentation contains the following –

1. Architecture diagram
2. Approach of the project
3. Technology used
4. Deployment of the code

**Architecture Diagram**

Image Processing

Convolutional Neural Network

Prediction

Fig: Architecture of the project

**Approach of the solution**

Since captchas are images, this project requires some image processing .OpenCV was used for image processing.For the recognition of the captcha, Deep Learning was used. For training the deep learning model we segmented the captcha, i.e, if the captcha is ABCDE, we divide the image as { A, B, C, D }. We do this by finding the continuous contours and dividing it accordingly ( *findContours()* function in OpenCV ). We trained the model using clean captcha images downloaded online and also the dataset given by your team. So we had around 10,000 captchas for training the model.

Deep learning is a concept under Machine learning where the working of a human brain is mimicked.  
Data is processed through a network known as neural network. A neural network consists of nodes and edges. Each node and edge carries some value known as weights. These weights are acquired from the input that is given while training the model. In training of the model, these weights are optimized so as to create a knowledge base and can be used for prediction and recognition of captcha.  
  
After the image processing part, we used Convolutional Neural Network consisting of 2 layers, 1 hidden layer and 1 output layer. Hidden layer consists of 512 nodes, and the output layer consists of 24 nodes (I and O were omitted because of the lack of dataset and the ambiguity). The segmented captchas are used for training this model. The trained model can now be used to solve captchas.

The captcha image is segmented and fed as an input, and the prediction is represented as a 2D matrix containing the probabilities of the 24, the max probability is fetched and the corresponding alphabet is displayed.

**Technology Used**

1. Python 3.6.5
2. Tensorflow
3. Keras
4. OpenCV

**Deployment of the Project**

The basic modules of the project are

1. OpenCV
2. Keras using TensorFlow Backend

First we start with segmenting the captchas for generating the training and test data. We’ve written a script for the purpose of segmenting the captchas.

* Run captcha\_vision\_dataset.py

The images are segmented and ready to be used for training the model. Before that we have manually partitioned the training and test data so we have written the following scripts:

* segment\_data.py
* Run train\_model.py

Now the model is trained and we can use the .hdf5 (knowledge base) file created.

* Run captcha\_prediction.py

This returns a 2D matrix with the probabilities, that is, it’s the result of the output nodes. The maximum value of the matrix is found out and the corresponding alphabet is displayed.