Machine Learning Advanced Nanodegree

Capstone Proposal

Veluvolu Akhil

(akhilchowdary78@gmail.com)

August 27th, 2018

Proposal

Abstract: -

This project is going to be designed and developed to perform facial expression recognition using Convolutional Neural Networks (CNN). The goal is to classify each facial image into one of the seven facial emotions: anger, disgust, fear, happiness, sadness, surprise and neutrality. To reduce the overfitting of the models, techniques like dropout and batch normalization will be used.

Domain background: -

I enrolled for a course on 'Deep Learning and Reinforcement Learning through machine learning' by Udacity. I'm very much fascinated in the areas of image and speech recognition with deep learning. The accuracy we can achieve is better compared to other ML models in these areas. Deep Learning is a branch of machine learning that allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction. Deep learning techniques have dramatically improved the state of art in areas such as image recognition, speech recognition, natural language processing.

Problem statement: -

Learning facial expressions of a person from an image. Facial behavior is one of the most important cues for sensing human emotion and intentions among people. In today's business world retail-stores are investing more to satisfy the customer with there products, the reactions of customers could be used as feedback to improve the offers on products or to remove certain products not liked by many. Similarly, analysis of facial expressions of drivers would help in determining their stress level and such information could be used to alert drivers if they are stressed and, in a state, unsafe for driving (or) should transfer control from manual to self-drive mode. With these applications in mind, this proposal describes an attempt to learn facial expressions from an image using deep learning.

Datasets and inputs: -

ICML 2013 hosted a facial expression recognition challenge based on a new dataset assembled using images crawled from the internet. In this project I will be working on the same problem, to classify an image of a human face into one of 7 expressions. The sample data set is available in kaggle. The data consists of 48x48 pixel grayscale images of faces. The whole dataset contains 35887 pixels of images in one column & it's corresponding labels in another. The data is split into train, valid and

test dataset. The training set consists of 22,967 images, Validation set contains 5,742 images and test set contains 7178 images.

Solution statement: -

This project will be done using various methods (Convolutional Neural Networks, Feed Forward Neural Networks) to identify the seven key human expressions. Convolutional Neural Network (CNN), a Deep Learning model is more effective at determining the important features in a given image than any other model. The major reason for CNN to be more effective is it's implementation of finding patterns within images, which uses filters to find specific pixel groups. This project idea will also be implemented using other technique, Feed Forward Neural Network, but CNN is expected to give a better result.

Benchmark model: -

As this is a Kaggle competition, we can compare our solution against the top kaggle solutions.

Evaluation metrics: -

Accuracy is a common metric for binary classifiers, it takes both true positives & true negatives with equal weight

	Predicted: Yes	Predicted: NO
Actual: Yes	TP	FN
Actual: NO	FP	TN

Accuracy = (TP+TN) / (TP + FP + TN + FN)

Project design: -

Step 0: Download & Extract the dataset.

Step 1: Explore & Review the dataset.

Step 2: Read pixels of Images & its labels from the dataset into variables.

Step 3: Preprocess the image pixels into workable format and apply Normalization. Split the dataset into 3 parts as training, Validation and testing sets.

Step 4: Train and Validate the model based on training & validation dataset, Test the data on test set to obtain accuracy (algorithms considered are Feed Forward Neural Networks and Convolutional Neural Networks).

Step 5: Compile and display the performance metrics & results (Confusion Matrix of TP/TN/FP/FN as well as precision and recall).