Project Milestone Report

Wentao Xu(wx225) Xingyue Dai(xd86) Ning Xu(nx43)

1 Motivation

Face recognition has been a hot application and research topic over the years. Recognizing faces with technical tools is important not only because it is useful to verify personal identity but also uncover criminals and deter crimes.In this project, we would like to build an application to identify and verify a person and his/her gender from a digital image and test it on a large dataset of faces. After realizing these basic functions, ideally we would like to apply our application for a broader use such as classifying ages, automatically completing the missing parts of faces on images and so on.

2 Method

Techniques: Logistic Regression, One-vs-All(multi-class classification), One-layer Convolutional Neural Network

Up until this point, we used logistic regression and Convolutional neural network for data training and modeling. However,as logistic regression is mainly used for binary classification, the classification problem having only two class labels, it gets complex to apply when we have more than two class labels in input train data. Thus, we used one-vs-all techniques(logistic regression) for categorization. If we have N-class instances dataset, we can generate N binary classifier models each trains and models one class of training dataset with logistic regression.

For Convolutional Neural Network, we applied basic one-layered convolutional neural network to the training dataset. The one-layered convolutional neural network has 64 filters with size 3x3, a max-pooling layer and a dense layer with size 100 and an output layer. And, we used stochastic gradient descent as our gradient method.

3 Preliminary experiments

Describe the experiments that you've run, the outcomes, and any error analysis that you've done. You should have tried at least one baseline.

Brief Introduction

To build the application for face recognition, we referred to a machine learning project at the university of Toronto, and worked with a subset of the 'FaceScrub' dataset which includes over 100000 face images of 530 People(265 males and 265 females). We selected out 620 face images of 6 people, specifically Fran Drescher, America Ferrera, Kristin Chenoweth, Alec Baldwin, Bill Hader and Steve Carell, as our train data for milestone purposes. Below are some sample face images.

After some pre-processing of the image data, we separated the dataset into three sub datasets for training, validation and test purposes respectively. Our baseline then was to train and model on the training set and make accurate predictions on the validation set and test set.

For logistic regression model, We did so by splitting this multi classification problem into multiple binary classification problems by building multi-class classier models with One-vs-all and logistic regression in order to accurately categorize images and make predictions, which were made using

the model that is the most confident. For convolutional neural network model, we just input the pre-processed data into the neural network and get the prediction result.

Part 1 Preprocessing image data before modeling

We created arrays to store the data of each image and their corresponding labels 0,1,2,3,4,5 which represent Steve Carell, Michael Vartan, Bill Hader, Alec Baldwin, Daniel Radcliffe and Gerard Butler in order. The original face images were retrieved from the internet so they were of different sizes and qualities. For the training purposes, we converted the 620 images with RGB to gray-scale images and resized them each to 32x32. Each of image data was saved as a numpy array with (32,32,1) originally. For logistic regression, we flattened each picture such that the shape of array became (620, 1024). While for Convolutional neural network, we kept each picture with its original shape. We then separated the dataset into three non-overlapping sets: train set(496 images) for training purposes, validation set(62 images) and test set(62 images).

Part 2 Modeling

One-vs-All with logistic regression model Since this is a multi-classification problem, as we stated earlier, we applied multi-class classification technique one-vs-all to break down a multi-classification problem into multiple binary classification problems.

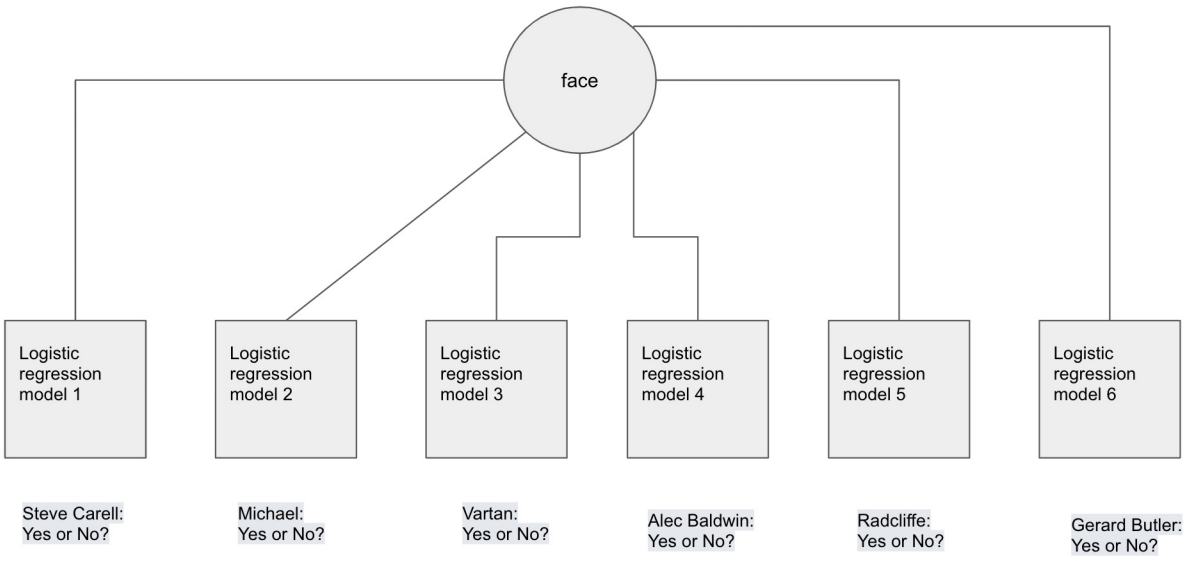
With 6 class labels(6 actors) in the dataset, we created 6 binary classifier models such that each of them is a logistic regression model that uses the Sigmoid function $g(z) = \frac{1}{1+e^{-z}}$ as its activation function and the cost(loss) function

$$J(\theta) = \sum_{i=1}^{m} -y^{(i)} \times log(sigmoid_\theta(x^{(i)})) - (1 - y^{(i)}) \times log(1 - sigmoid_\theta(x^{(i)}))$$

where m = # of training examples in this case.

Since for each of the labels 0,1,2,3,4,5, we have a corresponding binary classifier model. We modified the input training dataset for each binary classifier model accordingly. Specifically, for the binary

classifier model corresponding to label i, we set the label of each image in the input training set to 1 if its original label equals to i or 0 otherwise. This way we guaranteed that the number of label classes for each binary classifier model is indeed 2. With the training set and its corresponding modified label set as input training data, by applying gradient decent,we obtained 6 most optimized θs for the cost functions of the 6 binary classifier models respectively. We then made predictions by using the model that was most confident, i.e. returned the label with highest probability. Below is a simple illustration of one-vs-all.



It turns out that our model has about 88% training accuracy of the predictions on the training set, about 76% accuracy of the predictions on the validation set and about 74% accuracy of the predictions on the test set.However, due to relatively small training dataset(only 500 images), the percentages of accuracy on predictions fluctuate a little.

Convolutional Neural Network For Convolutional neural network, we input the data with shape (500 x (32, 32, 1)) into the one-layered Convolutional Neural Network, then set the number of training epoch to be 100. At then end we got 99% accuracy on the training set and The test accuracy of the model reached 88%.

Error analysis of low validation accuracy in Convolutional Neural Network One problem we encountered while modeling with CNN was that with 0.1 as the validation split parameter,the validation accuracy was only 70%. We analyzed possible causes and tried several other validation split.Then, we realized that the validation accuracy changed as our choices on validation split parameter changed. For example, when we set the parameter to be 0.01, the validation accuracy reached 0.88 which was relatively high. We then concluded that we should choose a proper size of validation set for it to be beneficial to the model.

4 Future work

Using multi classification models to accurately recognize and predict face images is just our initial step to build the application for face recognization. We'll improve our model with new machine learning techniques such as VGG model and train and test it with much larger dataset, ideally with the whole 'FaceScrub' dataset. As we stated earlier in our project proposal, we aim to do more structured deep neural networks to improve the accuracy of facial recognition, and will try to do some experiments of face completion algorithms on the 'FaceScrub' dataset. Ideally, such algorithms will have some observed data as input and make predictions about the missing part of the image.

References

[1] vision & interaction group, National university of Singapore. (n.d.).

FaceScrub Dataset.http://vintage.winklerbros.net/facescrub.html

[2] CSC411 Proejct 1: Face Recognition and Gender Classification with Regression. (n.d). http://www.cs.toronto.edu/%7Eguerzhoy/411/proj1/

[3] Band, A. (2020, May 14). Multi-class Classification — One-vs-All & One-vs-One. Medium. https://towardsdatascience.com/multi-class-classification-one-vs-all-one-vs-one-94daed32a87b