Project Summary:

###Data Set Summary & Exploration

####1. Provide a basic summary of the data set. In the code, the analysis should be done using python, numpy and/or pandas methods rather than hardcoding results manually.

I used the numpy library to calculate summary statistics of the traffic
signs data set:

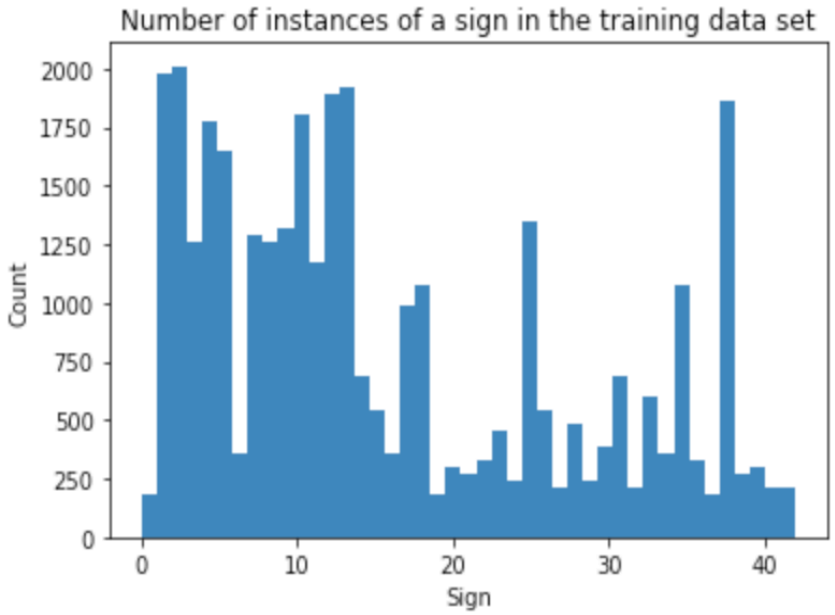
* The size of training set is 34799
* The size of the validation set is 4410
* The size of test set is 12630
* The shape of a traffic sign image is 32 x 32 x 3
* The number of unique classes/labels in the data set is 43

####2. Include an exploratory visualization of the dataset.

Here is an exploratory visualization of the data set. It is a bar chart showing how the data …

A few images displayed from the training set:

Histogram of the images with unique labels:



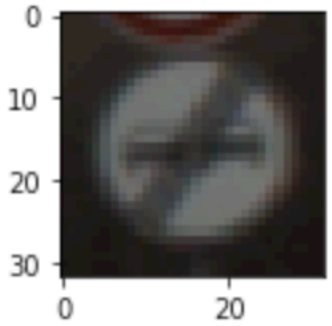
###Design and Test a Model Architecture

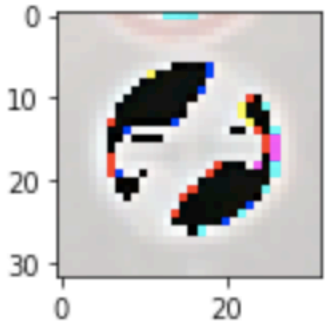
####1. Describe how you preprocessed the image data. What techniques were chosen and why did you choose these techniques? Consider including images showing the output of each preprocessing technique. Pre-processing refers to techniques such as converting to grayscale, normalization, etc. (

As a first step, I did not convert the images to grayscale because with the assumption that color data is important to recognize traffic signs.

I normalized the image data because so that the data has mean zero and equal variance.

The difference between the original data set and the augmented data set is the following





####2. Describe what your final model architecture looks like including model type, layers, layer sizes, connectivity, etc.) Consider including a diagram and/or table describing the final model.

My final model consisted of the following layers:

|

Layer	**Description**
Convolutional Layer 1	Input = 32x32x3. Output = 28x28x6.
RELU	
Pooling 1	Input = 28x28x6. Output = 14x14x6.
Convolutional Layer 2	Convolutional. Output = 10x10x16.
RELU	
Pooling 2	Input = 10x10x16. Output = 5x5x16.
Fully Connected layer 1	Input = 400. Output = 120.
Fully Connected layer 2	Input = 120. Output = 84.
Fully Connected layer 3	Fully Connected. Input = 84. Output = 43.

####3. Describe how you trained your model. The discussion can include the type of optimizer, the batch size, number of epochs and any hyperparameters such as learning rate.

I used the following:
EPOCHS = 40
BATCH_SIZE = 128
rate = 0.001 (learning rate kept low as suggested in the course lectures)
used AdamOptimizer as shown in the course lectures

To train the model, I used a shuffled training set to avoid training the images in order of the labels.

####4. Describe the approach taken for finding a solution and getting the validation set accuracy to be at least 0.93. Include in the discussion the results on the training, validation and test sets and where in the code these were calculated. Your approach may have been an iterative process, in which case, outline the steps you took to get to the final solution and why you chose those steps. Perhaps your solution involved an already well known implementation or architecture. In this case, discuss why you think the architecture is suitable for the current problem.

I uses the Lenet-5 architecture as was described in the course lecture which without modification is plug and play and gives an accuracy of 89 percent. I tweaked the number of Epochs to 10 (learning rate of 0.001) which capped the accuracy to about 91 percent.

I then tweaked the number of epochs to 50 but the accuracy dipped after about 45epochs to 0.91. I therefore capped the number of epochs to 40 to get an accuracy of **93.1%**

My final model results were:

- **validation set accuracy of 93.1%**

Test set accuracy of 92.6%

Lenet-5 architecture was plug and play for the traffic signs classification and therefore I chose this with the following modifications:

- The images need not be padded as they were already 32 x 32
- Validation set was already provided so it was easy to plug in to Lenet 5
- Lenet-5 used grayscale images but the traffic signs images are colored, this information was used for visualization and to modify the input depth to 3 for the three color channels.
- Rest of the architecture was compatible with the traffic classier image inputs
- On the output layer, MNIST only had 10 labels whereas for the traffic sign classification there are 43 labels.

###Test a Model on New Images

####1. Choose five German traffic signs found on the web and provide them in the report. For each image, discuss what quality or qualities might be difficult to classify.

Here are five German traffic signs that I found on the web:

These were chosen specifically because of the following reasons:

- The No Entry sign is blurred therefore the model might have trouble accurately classifiying it.
- The priority road sign image has portions of other road signs on the edges and also the background is dark and therefore the model might have trouble accurately classifiying it.
- The roundabout sign has an additional square section inside the circle which the training set will not have so the model might have trouble accurately classifiying it.
- The upper left edge of the Stop sign image is washed out and the other background is dark and therefore the model might have trouble with the boundaries of the Stop Sign to accurately classify it.
- The yield sign image is entirely washed out and the interior if the image has almost the same color information as the background and therefor the model might have trouble accurately classifying it.

####2. Discuss the model's predictions on these new traffic signs and compare the results to predicting on the test set. At a minimum, discuss what the predictions were, the accuracy on these new predictions, and compare the accuracy to the accuracy on the test set (OPTIONAL: Discuss the results in more detail as described in the "Stand Out Suggestions" part of the rubric).

Here are the results of the prediction:

Sign	**Prediction**
No Entry	Turn Left Ahead
Priority road	Right-of-way at the next intersection
Roundabout	Speed limit (30km/h)
Stop	Keep right
Yield	Yield

The model was able to correctly guess 1 of the 5 traffic signs, which gives an accuracy of **20%**. This accuracy is pretty low as compared to the validation set accuracy of **93.1 %** and the test set accuracy of 92.6 %. This is because the model might be overfitting for the training data sets. This can be improved by generating additional data points by preprocessing the existing images in the training set by random scaling, random translate and randomly changing the brightness.

####3. Describe how certain the model is when predicting on each of the five new images by looking at the softmax probabilities for each prediction. Provide the top 5 softmax probabilities for each image along with the sign type of each probability.

F**or the No Entry Sign, the model assigned a softmax probability of 5.43 Percent**

Image: No Entry:	**Label- SoftMax Probabilities * 100**
Turn left ahead	34 95.6162452698
Turn right ahead	33 4.38376143575
No Entry	17 5.4353214179e-15
Go straight or left	37 1.33065164041e-16
Roundabout mandatory	40 7.91777443894e-20

For the priority road sign the model assigned a softmax probability of 1.91579784263e-11 which is very very small.

Image: Priority road	**Label- SoftMax Probabilities * 100**
Right-of-way at the next intersection	11 100
Beware of ice/snow	30 1.97120564316e-07
Double curve	21 7.52803097637e-08
Slippery road	23 2.07471865209e-10
Priority road	12 1.91579784263e-11

For the roundabout sign the model could not classify accurately and the first five labels were all different than the roundabout sign

Image: Roundabout:	**Label- SoftMax Probabilities * 100**
Speed limit (30km/h)	1 - 99.8874723911
General caution	18 -0.105301092844
Speed limit (20km/h)	0 -0.00669502042001
Dangerous curve to the left	19 - 0.000433421519119
Right-of-way at the next intersection	11 - 8.4803542677e-05

Similarly , For the stop sign the model could not classify accurately and the first five labels were all different than the roundabout sign

Image: Stop	**Label - SoftMax Probabilities * 100**
Keep right	38 - 95.9829688072
Go straight or right	34 - 4.01638671756
Turn left ahead	9 - 0.000439411587649
Traffic signals	35 - 0.00019238036657
Turn right ahead	33 - 7.92266448002e-06

Model had 100% certainty for the Yield sign and correctly classified the sign

Image: Yield	**Label - SoftMax Probabilities**
Yield	13 100.0
Speed limit (20km/h)	0 0.0
Speed limit (30km/h)	1 0.0
Speed limit (50km/h)	2 0.0
Speed limit (60km/h)	3 0.0