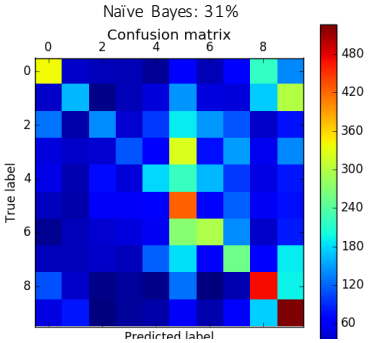
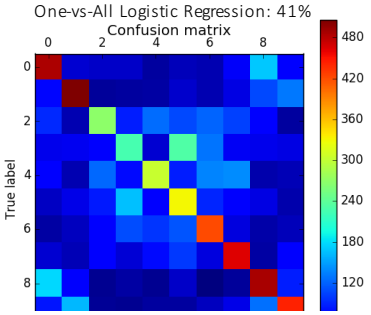
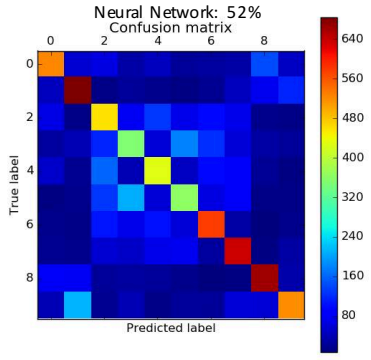
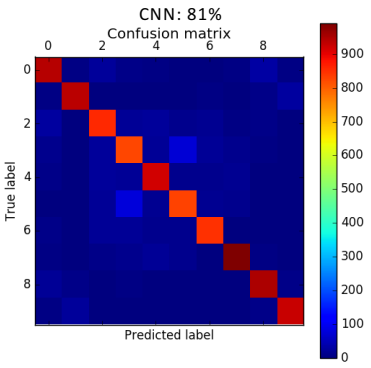
CIFAR 10 Image Recognition:
Ensembling with Support Vector Machines









0: plane
1: car
2: bird
3: cat
4: deer
5: dog
6: frog
7: horse
8: ship
9: truck

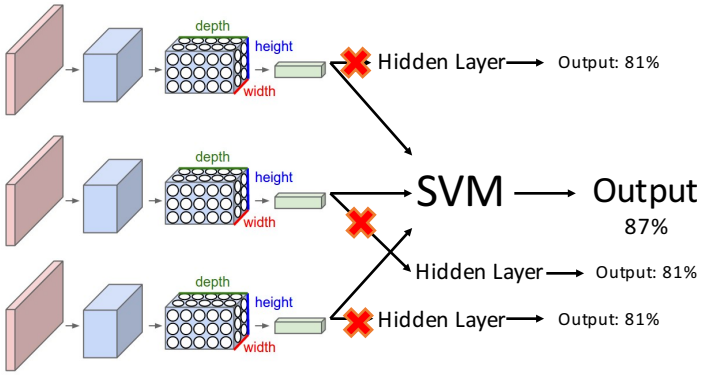
To the left, you will see various confusion matrices of different methods we tried on the CIFAR 10 Problem. We found that Naïve Bayes performed the worst (31%), likely since it does not account for the interdependence of the image pixels. Next was Logistic Regression (41%) which observes inter-pixel relationships, but only to a degree. Then was the Neural Network (52%) which heavily learns inter-pixel relationships. Finally was the Convolutional Neural Network (81%) which heavily learns inter-pixel relationships with a specific emphasis on 2D locality, a crucial aspect of images.

To the right, you will see the CNN structure we used to train the various models that were underlying our SVM. This is a powerful ResNet architecture given by He, Zhang, Ren, and Sun at Microsoft Research for the CIFAR-10 Problem. For the conv layers, only 3x3 filters are used.

CNN Architecture

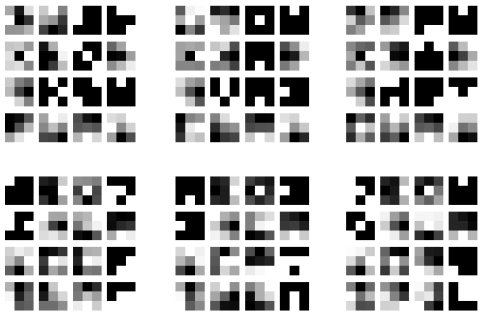
#	name	size
0	Input	3x32x32
1	Convolution	16x32x32
2	BatchNormal	16x32x32
3	Convolution	16x32x32
4	BatchNormal	16x32x32
5	Convolution	16x32x32
6	BatchNormal	16x32x32
7	Convolution	32x16x16
8	BatchNormal	32x16x16
9	Convolution	32x16x16
10	BatchNormal	32x16x16
11	Convolution	64x8x8
12	BatchNormal	64x8x8
13	Convolution	64x8x8
14	BatchNormal	64x8x8
15	globalpool	64
16	output	10

Using the features learned by CNN's as input into a linear SVM worked very well (87%), close to human accuracy (94%). Each CNN summarizes each image into 64 key values, and the SVM processes all of the key values of each image from each CNN into a reasonably accurate classification



Below are the visuals of the CNN's first layer weights. The two rows are from different CNN's. Each column is a representation of the weights for Red, Green and Blue, respectively

Weight Visualizations for CNN's



We built a computer with a high performance GPU to build our models 20-30x more quickly

Building a Machine Learning Computer