Hand-Written Digits Identification Using SVM and CNN

Sahil Khanna
Department of Telecommunication
University of Maryland College Park
skhanna9@umd.edu

I. INTRODUCTION

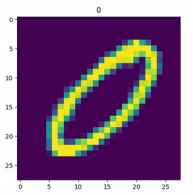
Deep Networks play significant role in many emerging applications such as face identification, speech recognition because of its high accuracy results. In this project, I am comparing two conventional approaches Support Vector Machines and the Convolution Neural Networks on Mnist data. The purpose of this project is to gain insights on the implementation of both the methods.

II. ALGORITHM

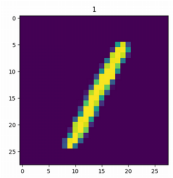
The algorithm used for both the models is written in python. I have used python's Tensor flow library to implement deep network. Some other modules like Matplotlib, Pyplot, Numpy, Scikit Learn are used to support the implementation of the models. I have used svm() function from scikit learn library to implement support vector machines.

III. DATA SET

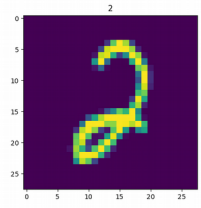
The Mnist data consists of 60000 training images of the hand-written digits with 10 classes and 10000 testing images. Each image has a dimension of 28*28. The all ten classes are shown in the following figures.



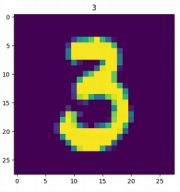
0



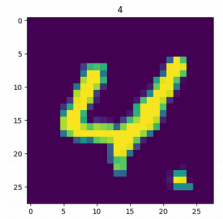
1



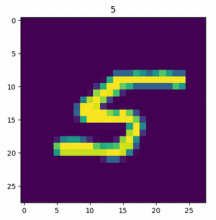
2



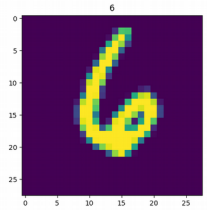
3



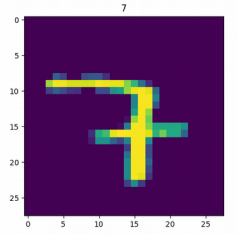
4



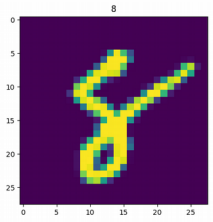
5



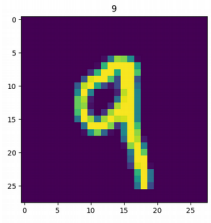
6



7



8



9

IV. RESULTS

The efficiency of the models is calculated based on the accuracy of the predicted labels. The observations are taken by varying the different parameters. First results of the SVM model will be discussed, later for CNN model.

The SVM models is analyzed based on the regularization parameter and different kernels.

- The first result is taken by using different kernels. The variation in kernels give different test accuracy. This difference is due to the shape of data in the multidimensional space. It can be seen from the table 1 that the RBF

kernel is the best fit for this problem, thus the data is non linearly distributed.

Table 1 Accuracy of the SVM

Support Vector Machine			
	RBF	Linear	Polynomial (3 degree)
Accuracy	93.2	91.96	83.78

- The second result is taken by varying the regularization factor in linear as well as gaussian support vector machine. The regularization is very important to make our model generalized. There are two major problems to consider while designing a model: over-fitting and under-fitting. Over-fitted model performs very good on the training data but performs poor on the test data, contrary under-fitted model performs poor in training data and sometimes performs good in test data. This problems are regulated by using the C parameter in SVM method. It can be seen from the figures 1 that the increase in C value leads to more generalized model (increasing the accuracy on test data) and avoid over-fitting problem. However, in figure 2 it can be seen that if we keep on increasing the C parameter after a certain value the model starts under-fitting the data and results in poor performance. Hence, it is very important to select appropriate values for the C. Many techniques are being used to calculate the C value, I used the grid search method. It calculates the accuracy with step increment of the C value and selects the value with high accuracy on validation data.

- The third experiment is to check the incorrect predictions by Linear SVM model. The figures 3, 4, 5, 6 and 7 shows the incorrectly predicted digits. It can be concluded from the pictures that the noisy picture can fall onto other digit tangential region thus results in wrong prediction.

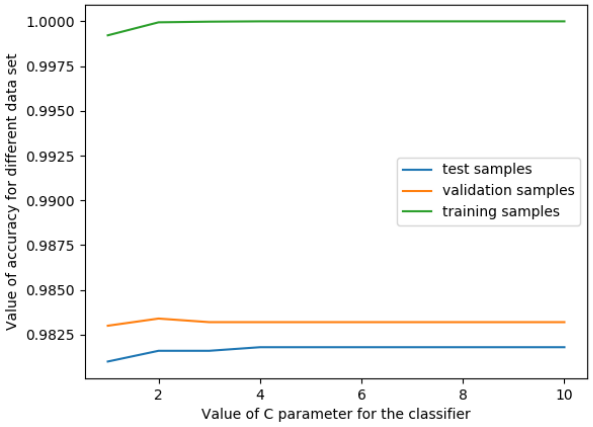


Figure 1 Relationship between C and accuracy in RBF SVM

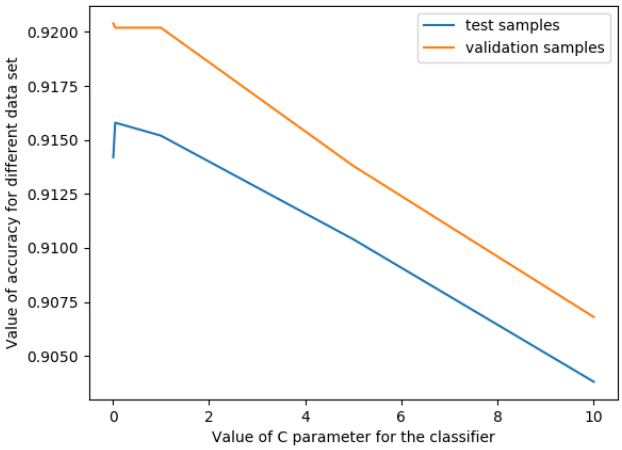


Figure 2 Relationship between C and accuracy in Linear SVM

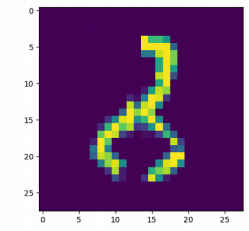


Figure 3 Linear SVM predicts the above 8 digit as 1

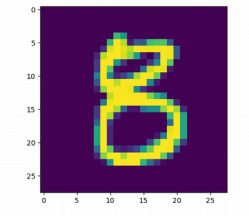


Figure 4 Linear SVM predicts the above 8 digit as 5

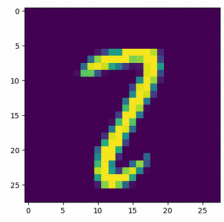


Figure 5 Linear SVM predicts the above 2 digit as 7

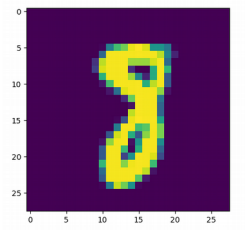


Figure 6 Linear SVM predicts the above 8 digit as 7

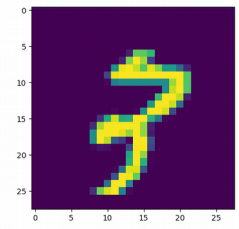


Figure 7 Linear SVM predicts the above 3 digit as 9

The experiments with convolution neural network is taken by composing a deep network model using tensor flow library. I used the LeNet – 5 designed to compose CNN network. The network uses Adam Optimizer as a convergence algorithm, Soft-max Cost function, Gaussian initialization, Relu as a non linear function for fully connected layers and drop out method to generalize my model. The learning curve is seen in figure 8. It can be concluded from the figure that the model is converges after 300 epochs. The figure 9 shows the accuracy of the network on test data.

- The convolution neural network convergence is hugely depended on learning rate. It can be seen from the figure 10 that if the learning rate is very high the model do not converges, thus results into low accuracy. On contrary, if the learning is very small, the model will converge but it takes large time. The figure 10 shows the relationship between the accuracy and log of the learning rate given number of epochs 500.

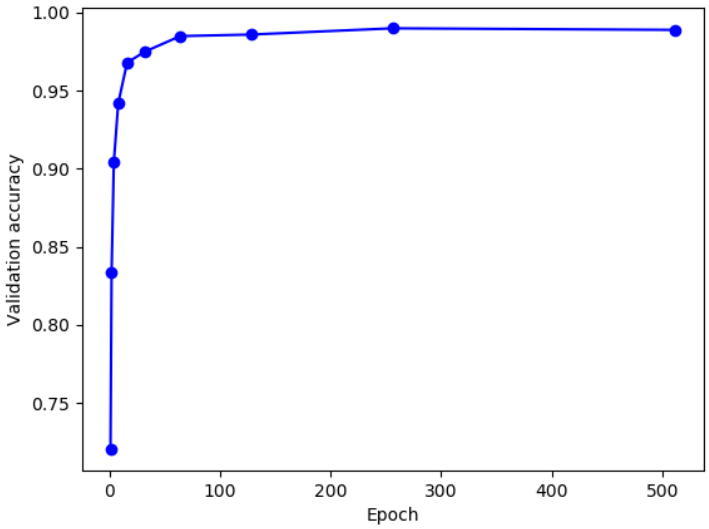


Figure 8 Learning curve for the CNN

Figure 9 Test accuracy of the convolution neural network

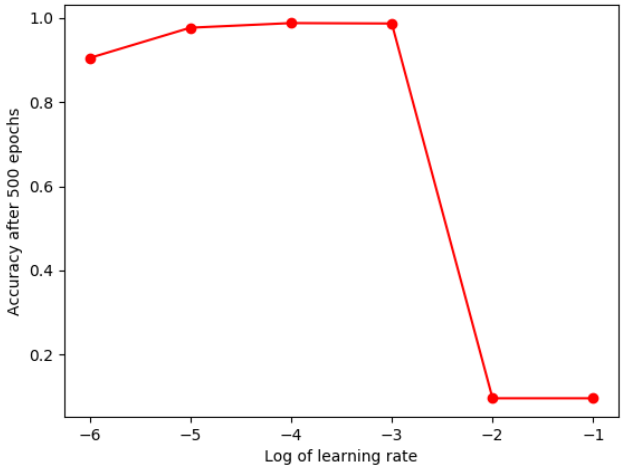
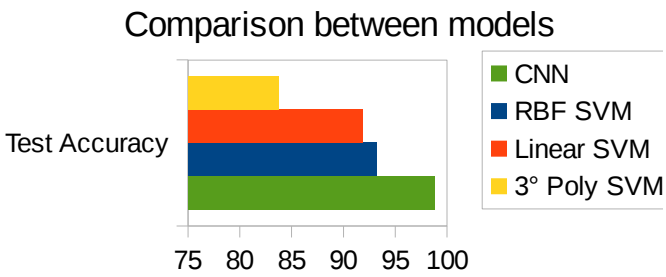


Figure 10 Relationship between the log10 of learning rate with the validation accuracy given 500 epochs

V. CONCLUSION

It can be seen from the results that the accuracy of prediction depends upon various factors such as value of C, kernel selection and model structure. There relations can be seen through these results. It can be

concluded from the experiment that the deep network is the better choice to identify the hand-written digits as compared to support vector machine model because it provides better accuracy.



VI. Acknowledgment

This project required in-depth research, a huge amount of work and dedication. Implementation would not have been possible if I did not have the support of many individuals.

I would like to extend mine sincere gratitude to all of them. First of all, I am thankful to professor Ramalingam Chellappa and professor Behtash Babadi for the provision of expertise and technical support in the completion of this project. Without their superior knowledge and experience, it would have been impossible for me to complete the project in the given timeline.

I would also like to express mine sincere thanks to Huai Jen Liang who devoted his time and knowledge in the implementation of this project and constant encouragement which helped me in completion of this project.