The MNIST is an extremely large dataset consisting of 60000 training and 10000 testing data, which are handwritten digits. In this part of the project, I will use the decision forest model of machine learning to train models with different depths of trees and numbers of trees and compare their accuracies. Also, I will compare the performance of models that are trained by the original untouched data and a particular transformation of the original data, called stretched bounding box. The stretched bounding box is obtained by cutting the horizontal and vertical dark range out of the original image and resizing it to a square bounding box that the horizontal range of dark pixels is centered in the box.

Accuracies of untouched raw pixels:

	Depths = 4	Depths = 8	Depths = 16
Trees = 10	0.7329	0.9042	0.9478
Trees = 20	0.7775	0.9156	0.959
Trees = 30	0.8133	0.9247	0.963

Accuracies of stretched bounding box:

	Depths = 4	Depths = 8	Depths = 16
Trees = 10	0.7611	0.9144	0.951
Trees = 20	0.7848	0.9255	0.9641
Trees = 30	0.7885	0.9274	0.9674

With the increase in the depth and number of trees, the accuracies improve from around 70% to 96%. I also found that most of the accuracies of the stretched bounding box are higher than that of the untouched raw pixels. By the data, the stretched bounding box is better than untouched raw pixels. The model of a stretched bounding box can enlarge the number in the image increasing the proportion of the pixels that create the number in the image, allowing the computer to identify the number easily.