Project Proposal

Morgan Baker

2013-09-21

My project idea comes from Kaggle Competitions (Kaggle, [1]). The project is taking a set of handwritten integers and determining the integer from the handwriting. So, when the machine reads a persons handwriting, the machine can easily transpose that information into a document or command line. The motivation for the project is that my handwriting was terrible when I was in school and would get points off multiple times on math tests because a 4 would look like a 9 and a 3 would look like a 5. This also helps with banking, like cashing checks and making transactions, as well as government forms that are on paper. Plus, because the project proposal is from Kaggle, the data is already included in .csv files and ready to be downloaded and manipulated. The images in the dataset have 784 pixels, and are 28 pixels wide and 28 pixels high. The data set has 785 columns, one for the picture identifier and the rest being the 784 pixels. Each entry in all of the pixel columns has a number ranging from 0 to 255, which is the brightness of the pixel. Using python code, I will try to plot and recreate the image and then have the machine look at each pixel to predict the image. Since this is a Kaggle competition, I plan on uploading the project to Git and Kaggle. This also means I have a limited schedule. The way that the competition is being judged is by completion accuracy, or the percentage of accurate images my machine learning algorithm predicts. I hope to have a completion accuracy of .98, meaning only 2 percent of images were identified incorrectly. If this is successful, Ill be happy with myself, knowing that Ill be helping the world change.

References

[1] Kaggle Competitions, *Classify handwritten digits using the famous MNIST data*, `https://www.kaggle.com/c/digit-recognizer`