1. Objective of whole project.

Compare MADGE data to other machine learning algorithms like SVM and neural networks. Accuracy and speed will be the two most optimized features, with accuracy without overfitting as the most interesting piece to explore.

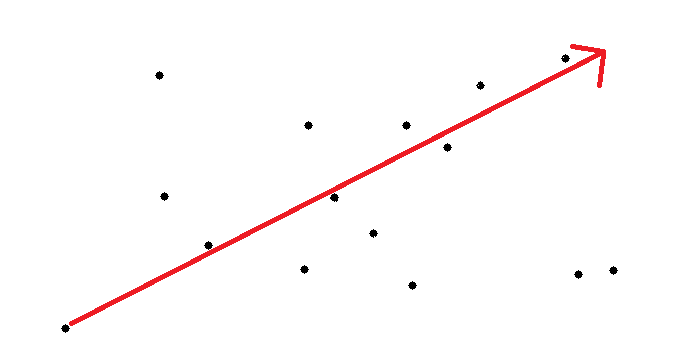
1. What was implemented last week.

- RUML has been updated with the new classes.

<https://docs.google.com/spreadsheets/d/1sktvtEGARj7zPEijXJN4fzFw9FgmM1plFmwsXs3W-pg/edit#gid=1133996158>

An adjustment was made to allow for multi-dimensional sets to be analyzed. The process is done via one main.py function, which I’m not sure how to put on the RUML. We can discuss.

- Implemented a method to import MNIST data and validate the data set. Accuracy for the first 100 points was maximized at 95%. The sigma of the data points was an important factor to the accuracy.



The sigma was determined by taking the magnitude of the vector that spans the range of the points and dividing it by a scalar. The logic is that each point will have an impact on the set with some divided by n value.

1. Plans for upcoming week.

Find an optimal sigma that can be replicated over all ranges.

Start working with other potential data sets.

Last week:

1. Objective of whole project.

Compare MADGE data to other machine learning algorithms like SVM and neural networks. Accuracy and speed will be the two most optimized features, with accuracy without overfitting as the most interesting piece to explore.

1. What was implemented last week.

A few papers were found relating to Gaussian Classification. Most of the methods that I see are related to some probabilistic Gaussian. The ones that use Gaussian for distance measurements seem to be using methods closer to K-means or Regression. (We can discuss in the meeting tomorrow)

Validation data was built for 2-D playground data. Results came back fairly successfully, as it seems as though data classification has high accuracy even with lower training data ratios (50%). This indicates that we are also potentially not overfitting with a sigma of 1. Results were sent via email.

Potential papers for comparing data classification between MADGE and SVM/Neural nets were looked at.

1. Plans for upcoming week.

Update RUML with graphing classification data code

I want to go straight to image classification with this project. It looks like any implementation using neural nets or SVM is with something in higher order, so I’m considering that we might as well bite the bullet and try to raise our dimensionality from 2 to ~65,000. The famous ImageNet paper (ImageNet Classification with Deep Convolutional Neural Networks) uses images with a fixed resolution of 256 × 256 (65,000). If we can generate a MADGE using training data from images, we could potentially try to classify images out of the gate. Runtime would be the most serious issue, but there would be very little image processing.

Regardless, classification comparing to other papers will most likely raise dimensionality significantly if we want to target popular classification methods. We can discuss tomorrow.