Semester Project Update #2 – The Fantastic Four

Portia Allen, Connor Caldwell, Arnaud Filliat, Johnny Murphy

The main algorithm we want to use is a machine learning technique using support vector machines. These will allow us to be able to predict where new data points will be using previous data points, which is perfect for tracking a 3D particle. Since the particles we are using are in a multidimensional space, the best way to use the SVM software is to utilize the kernel function, allowing for nonlinear models to be produced.

In this case the nonlinear models would be most useful in the X and Y position of the particles, allowing us to block off different groups and relate them to other features, such as the X and Y velocity. This will hopefully allow for more accurate predictions while maintaining data integrity.

Sklearn has multiple functions that can be utilized for SVM machine learning, including the kernel function as mentioned prior, as well as a C variable that allows for flexibility. This prevents overfitting or underfitting allowing a unique fit for each model. The leniency in our function will have to be tested, since the particles are semi-random and are hard to get a locked down number on. It will likely be in the range of C = .1-1 area since there are so many data points to be adjusted for.

Due to the random nature of particles and how they move another possible algorithm to use would be an unsupervised learning algorithm. We were thinking of using k-means as well to split the clustering and possibly look at reducing dimension to only the most important features. These features could include the particles' initial velocities and their charge. This would enable a more "learned" approach which could be significant but due to the large amount of data may be impractical because of the time to run.