## Name: He Zhiming S-ID: 2018227154 Date: 2019-10-14

### The Objective of my Project

I will use 3 ML techniques in *MNIST Handwritten Data*: **1).** Random Forest (Supervisory Algorithms), **2).** Support Vector Machine (Supervisory Algorithms), **3).** K-means (Unsupervised Algorithms). Depending on my completion time and understanding, maybe I will use more ML technology.

### Data of my choice

**1).** I don't have a foundation for ML. **2).** My research direction is Fog/Edge Computing. I try to find data sets related to my current research, but unfortunately, I haven't found them.

Based on the above two points, I use the ***MNIST DATABASE of handwritten digits*** to ensure that I can complete the project and get visible improvements in big data & machine learning.

**MNIST consists of 4 parts:** training set images, training set labels, test set images and test set labels. It has a training set of 60,000 examples, and a test set of 10,000 examples.

**Data Description:** Each data entry is text data that represents a handwritten digit. The value of each label is a number between 0 and 9. Size of each image is 28x28 pixels and are in greyscale. The size of these data sets has been normalized and fixed size has been formed.

#### Expectation of results

**1).** Through continuous practice and parameter improvement, I hope to improve the **accuracy** of the training model and reduce **cross-entropy**. If time and energy are available, I hope I can reduce the training **time** and test **time** to improve performance.

**2).** Compare the **advantages** and **disadvantages** of the 3 or more ML algorithms.

**3).** **Understanding** big data and ML, **familiar with** and **master** simple ML technology.