Group Members:

1. Sri Ramcharan Galla 1212390086
2. Manideep Illendula 1208825003
3. Pujith Sagar Edara 1212331261

**Problem 1**

**Solution for K- Means Clustering:**

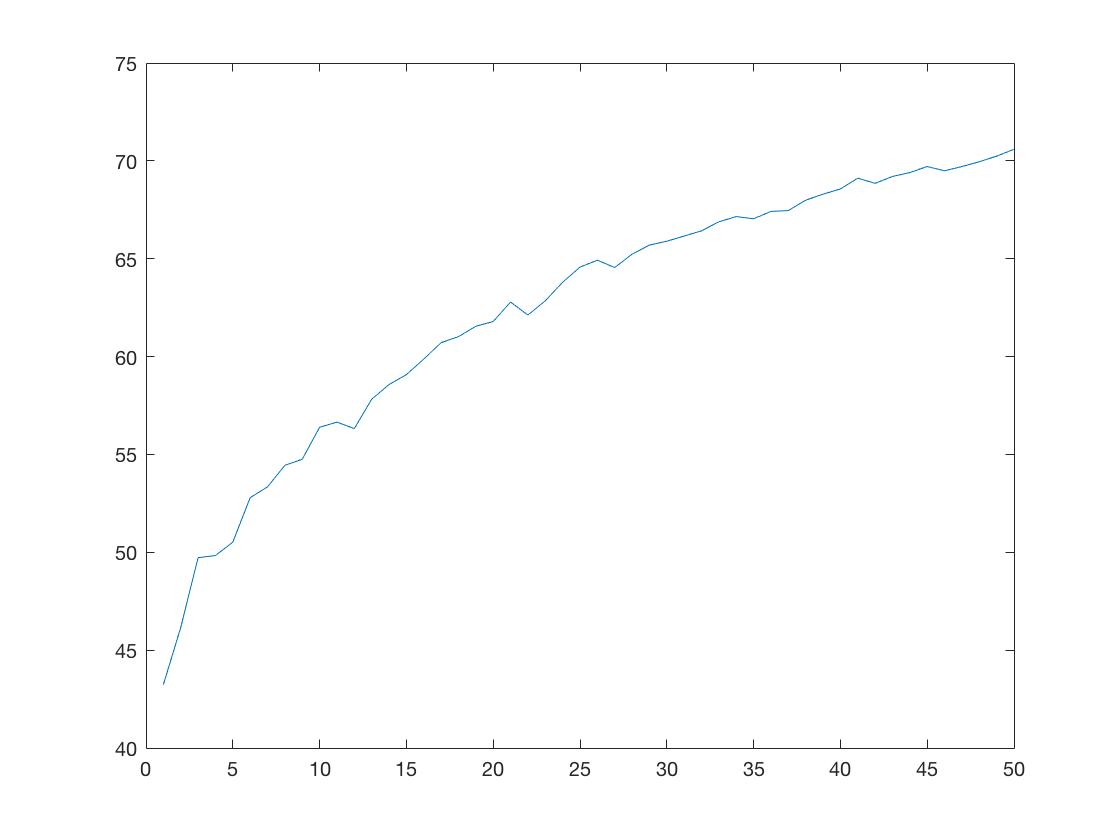
1. For K= 3 the sum of squared errors(sse) varies between 587.3186 to 588.7820. The number of iterations varies between 5 and 13
2. For K= 5 the sum of squared errors(sse) varies between 385.5073 to 449.0513. The number of iterations varies between 6 and 15
3. For K= 7 the sum of squared errors(sse) varies between 277.8805 to 339.4083. The number of iterations varies between 8 and 14

The function used for randomly picking k samples from the dataset to be centroids is datasample.

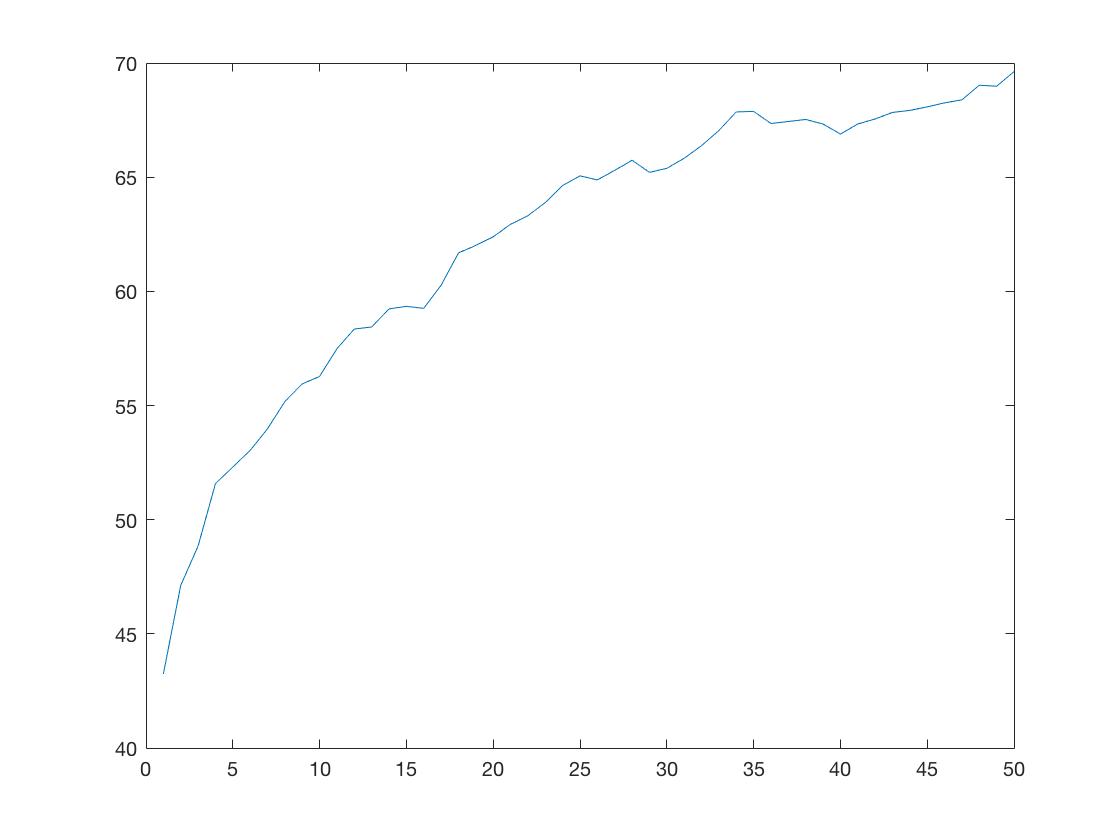
**Problem 2**

Dataset: MindReading

1. Random Sampling

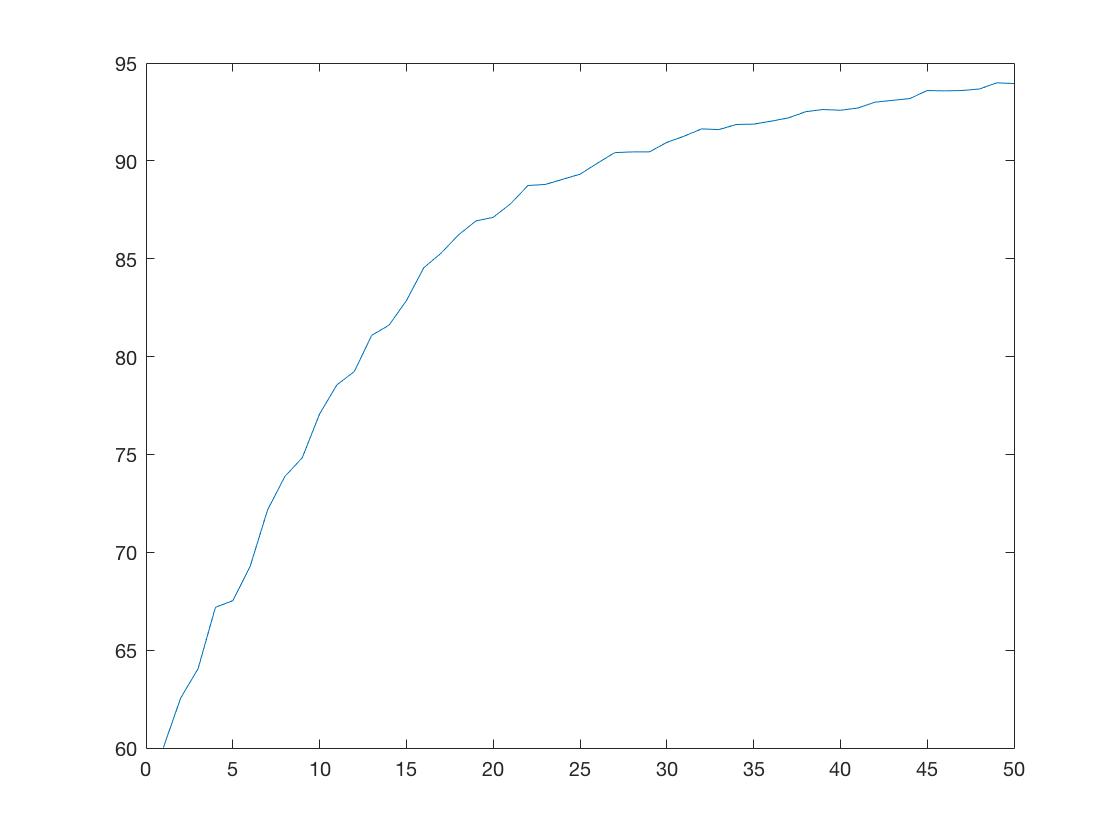


1. Uncertainty based sampling

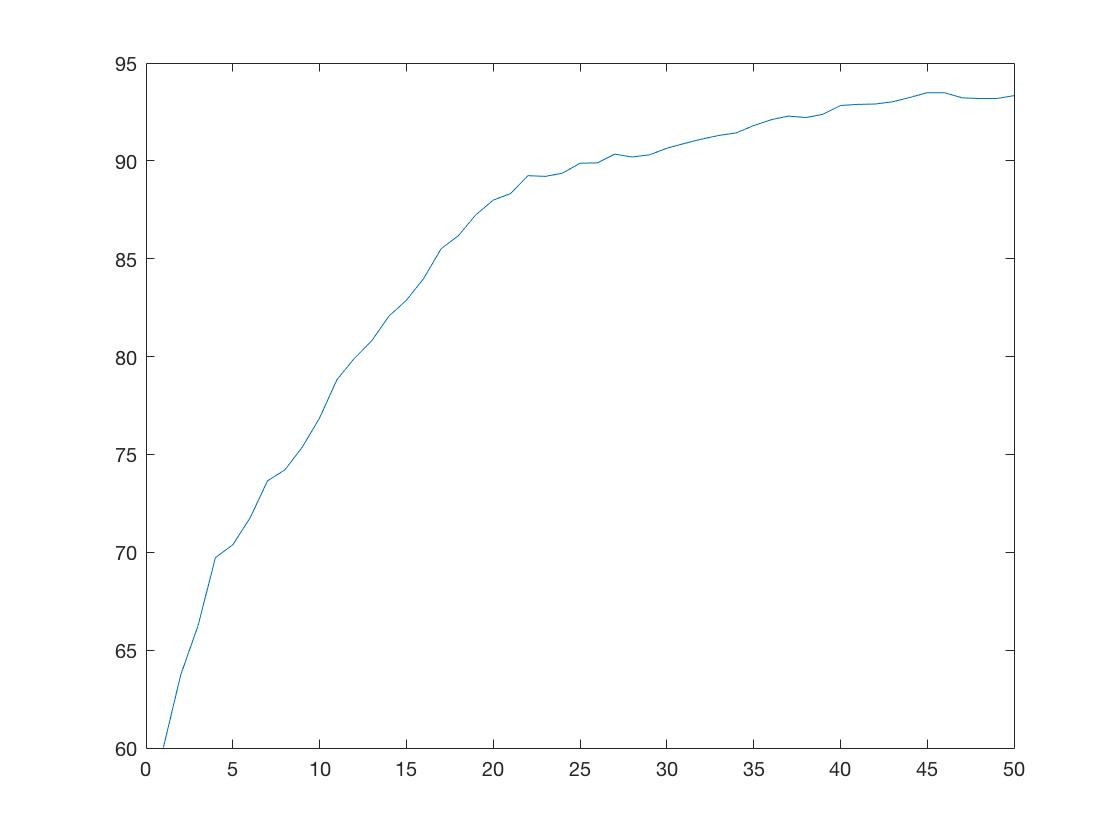


Dataset: MMI

1. Random Sampling



1. Uncertainty based sampling



**Conclusions:**

The accuracy for both the datasets will be increasing with respect to iterations as the model will be trained by the increasing samples with the sampling techniques.