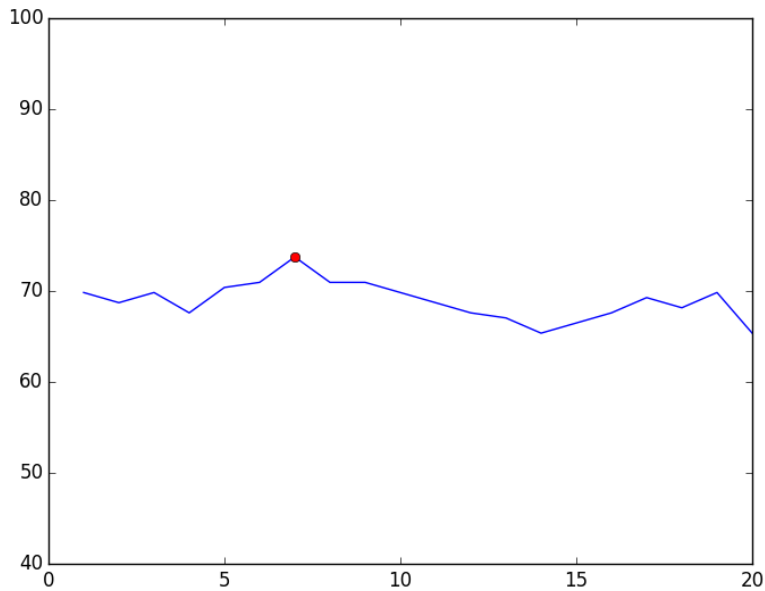


Homework 4 - Adam Crider

CPSC 4383 - Artificial Intelligence

Using the K-Nearest Neighbors algorithm on the dataset provided I got the best K-Value of 7, with an Accuracy of 73.7430%. I generated a plot of K-Values on a range of [1-20], the red dot represents the maximum K-Value found:



The K-Value vs accuracy in this data set is a tough problem because the neighbors are very inconsistent in their values. There is a good K-Value in the [5-8] range that has a good balance of small bias and low variance. If the K-Value is too large then the variance becomes too high very quickly and gives essentially random values. A K-Value too low will be biased towards the immediate neighbors (either good or bad). A well balanced K-Value in this instance is between [5-8] because that gives the data enough points to compare from, while not sampling from a space that will return too much variance.