In The Name Of God

University of Tehran

Department of Engineering Science

Algorithms of Machine Learning

## Spring 2022

Project 3

Professor: Dr. Ali Kamandi
Students: Abbas Mohamadiyan 810898059
Mohamad Esfandiyarpoor 810898038

Kmeans Algorithm

well, firstly we import required libraries, then reading data and drawing a scatter plot to have a broad overview of how data looks like. as we can see there are five crowds(just like we we've been asked) and then we wish to split these clusters in the end of this process.
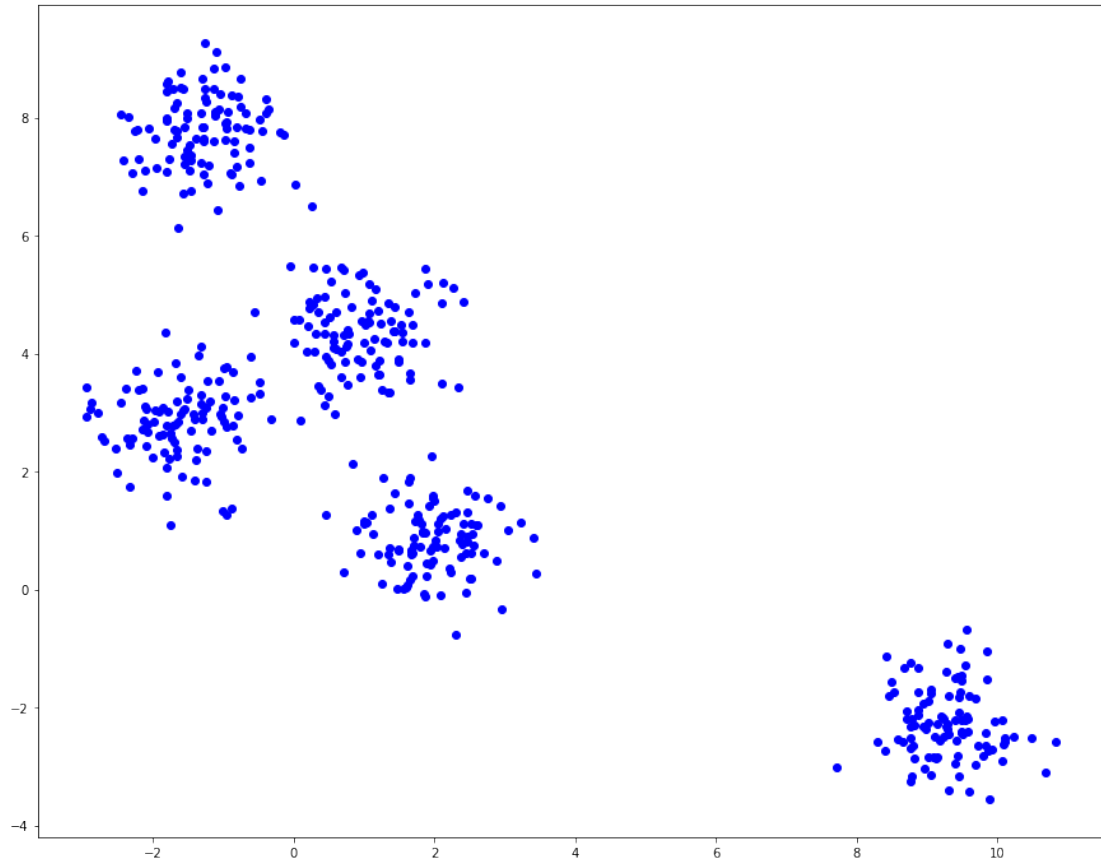
it seems like best distance we could work with is Euclidean(L2 norm)and so we do!

finally we are so good to start our algorithm

    define our parameters

    making our starting point through the way commented in the code

   and starting the main "for" loop

Distance matrix (Dist) has diatance of points from means in columns 1:5 and from zero in column 0 and "which" is list which says which mean point is the closest for each datapoint.

assigning datapoints to their corresponding clusters with the help of "which" and we have our first cluster!

finally meanpoints would get updated (they're in the middle of their cluster) and doing the loop over and over to see better resuls
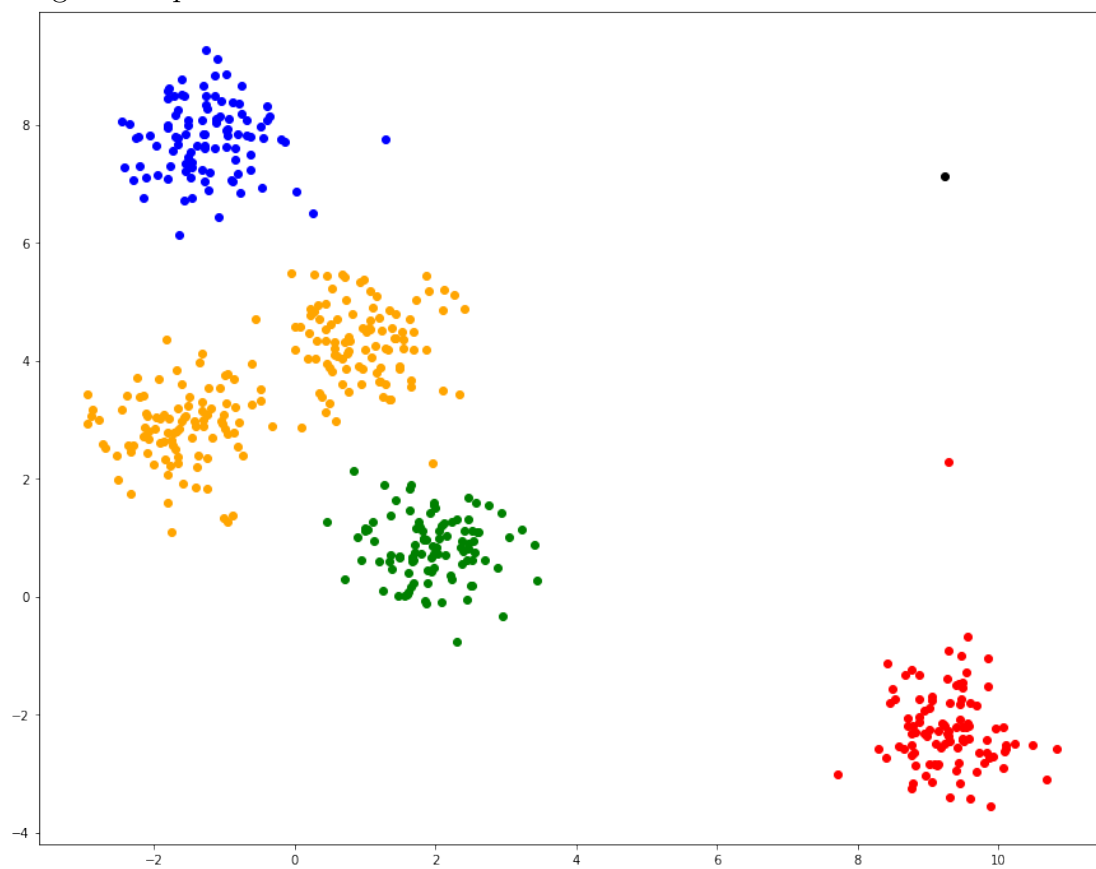
it's not the desired clustering we had in mind, but thats all we could do.
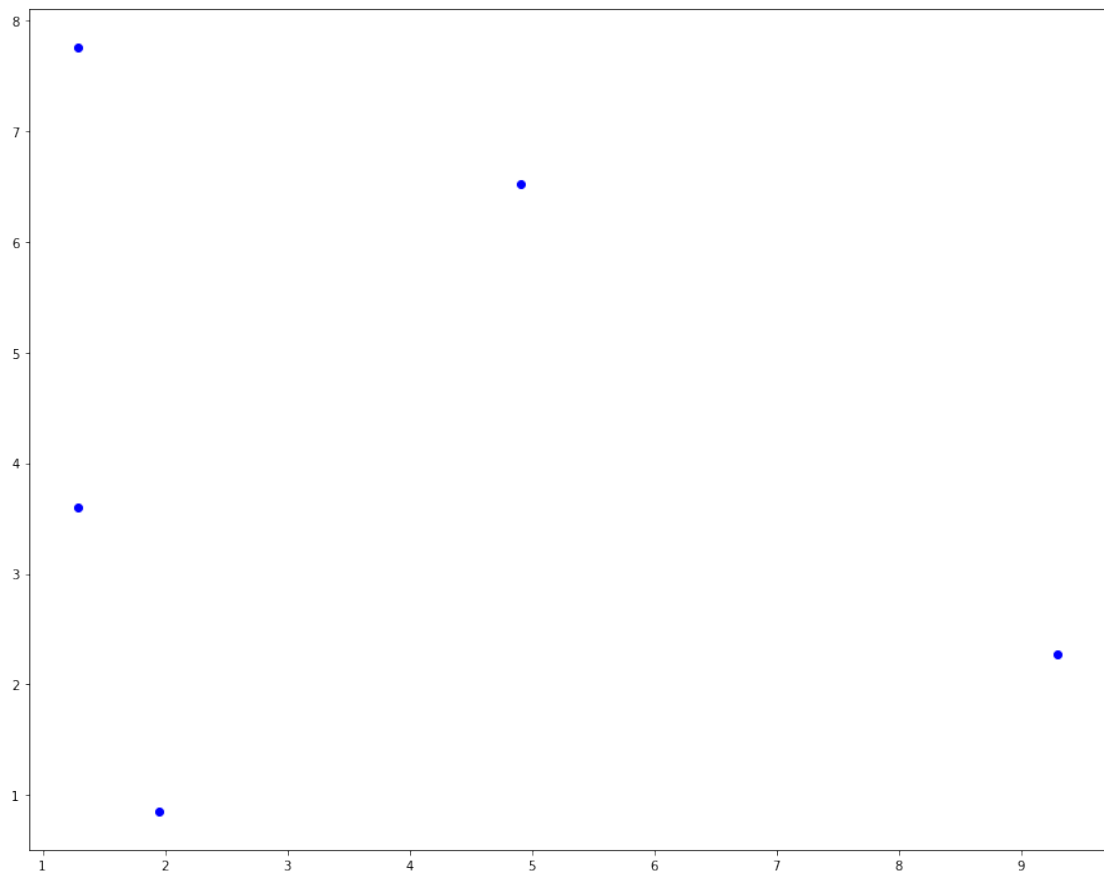
and these are the mean points which clusters are formed around them

Thanke you