**EX5 – K-Means Clustering**

%% =================== Part 1: K-Means Clustering ======================

load('ex7data2.mat');

K = 3; max\_iters = 10;

initial\_centroids = [3 3; 6 2; 8 5];

[centroids, idx] = runkMeans(X, initial\_centroids, max\_iters);

%% ============= Part 2: K-Means Clustering on Pixels ===============

A = double(imread('bird\_small.png'));

A = A / 255; % Divide by 255 so that all values are in the range 0 - 1

img\_size = size(A);

X = reshape(A, img\_size(1) \* img\_size(2), 3);

K = 16; max\_iters = 10;

initial\_centroids = kMeansInitCentroids(X, K);

[centroids, idx] = runkMeans(X, initial\_centroids, max\_iters);

%% ================= Part 3: Image Compression ======================

idx = findClosestCentroids(X, centroids);

X\_recovered = centroids(idx,:);

X\_recovered = reshape(X\_recovered, img\_size(1), img\_size(2), 3);

subplot(1, 2, 1);

imagesc(A);

title('Original');

subplot(1, 2, 2);

imagesc(X\_recovered)

function idx = **findClosestCentroids**(X, centroids)

%FINDCLOSESTCENTROIDS computes the centroid memberships for every example

K = size(centroids, 1);

idx = zeros(size(X,1), 1);

dist\_matrix = zeros(size(X,1),K);

for i=1:K

dist\_matrix(:,i) = sqrt(sum(bsxfun(@minus,X,centroids(i,:)).^2,2));

end

[dummy,idx] = min(dist\_matrix,[],2);

end

function centroids = **computeCentroids**(X, idx, K)

%COMPUTECENTROIDS returs the new centroids by computing the means of the

%data points assigned to each centroid.

[m n] = size(X);

centroids = zeros(K, n);

for i=1:K

centroids(i,:) = mean(X(idx==i,:));

end

end

function [centroids, idx] = **runkMeans**(X, initial\_centroids, ...

max\_iters)

%RUNKMEANS runs the K-Means algorithm on data matrix X, where each row of X

%is a single example

[m n] = size(X);

K = size(initial\_centroids, 1);

centroids = initial\_centroids;

idx = zeros(m, 1);

for i=1:max\_iters

idx = findClosestCentroids(X, centroids);

centroids = computeCentroids(X, idx, K);

end

function centroids = **kMeansInitCentroids**(X, K)

%KMEANSINITCENTROIDS This function initializes K centroids that are to be

%used in K-Means on the dataset X

rand\_list = randperm(size(X,1),K);

centroids = X(rand\_list',:);

end