

COL761 Assignment 3 Report

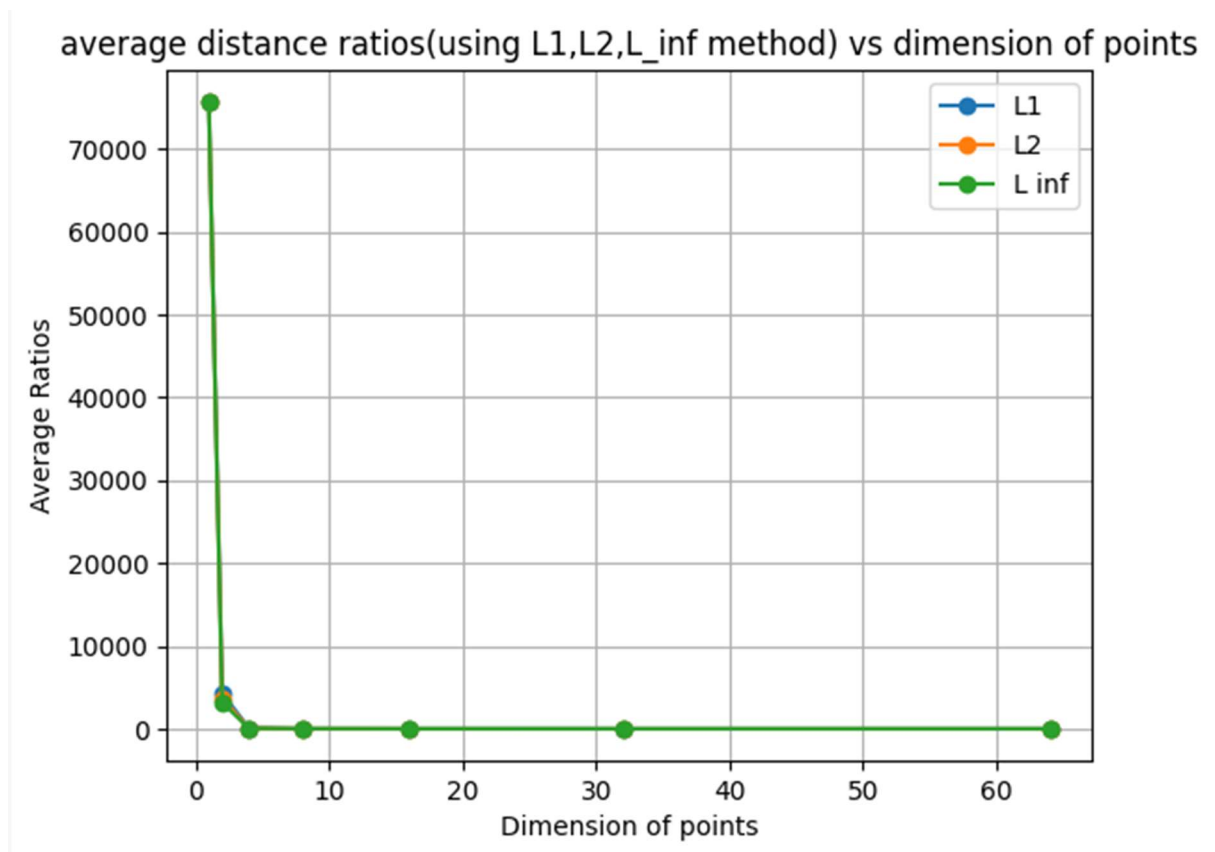
NAME	Entru No.	Contribution
Dharmeswar Basumatary	(2020CS50423)	33.33%
Sahil Sahare	(2020CS50440)	33.33%
Vibhor Sengar	(2020EE30313)	33.33%

Q1.

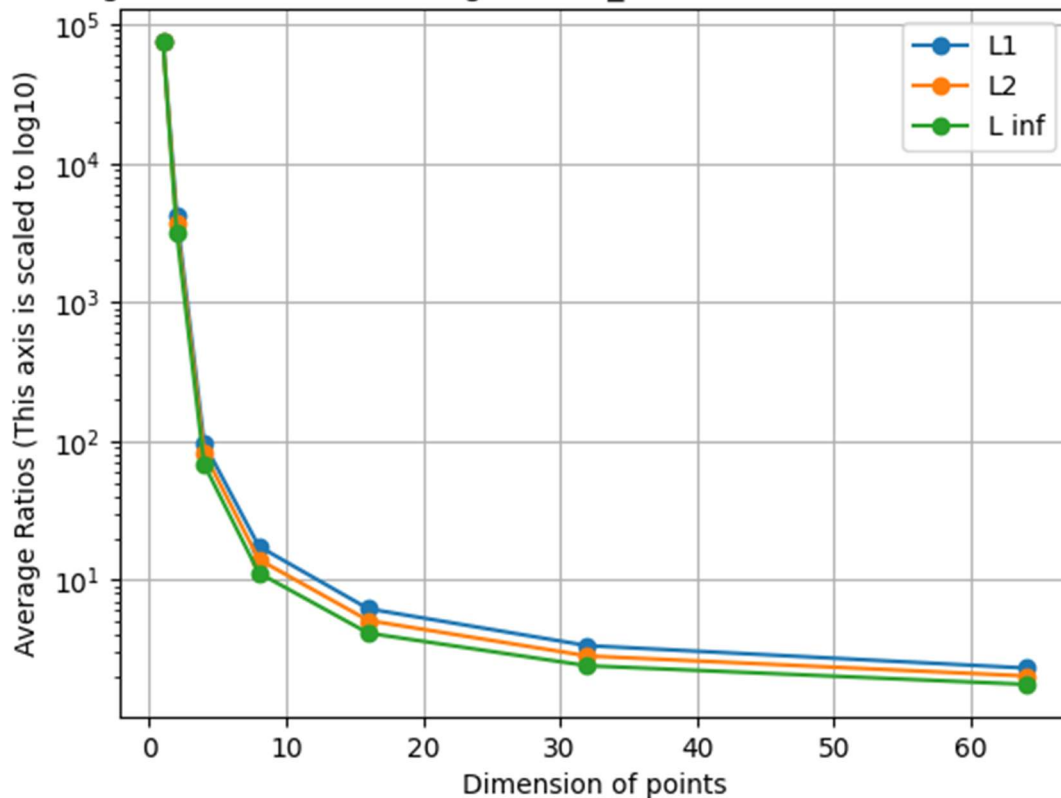
Task:

Plot the average ratio of farthest and the nearest distances versus d for the three distance measures(L1,L2,L_{inf}) .

Plot 1:



average distance ratios(using L1,L2,L_inf method) vs dimension of points



Observation:

The average ratio of farthest and the nearest distances versus d for the three distance measures (L1, L2, L-infinity) decreases with dimensionality.

In high-dimensional spaces, points become increasingly sparse. As the number of dimensions increases, the volume of the space grows exponentially. Consequently, the available data becomes sparser, and points are farther away from each other. Also, the distances between points tend to converge i.e., the difference in distances between pairs of points becomes less pronounced. In other words, the ratio of distances tends to become more uniform. So, the average ratio of farthest distance / nearest distance decreases with increasing dimensions of points.

The curse of dimensionality has practical implications for algorithms that rely on distance measures, such as clustering or nearest neighbour algorithms. In high-dimensional spaces, the concept of proximity becomes less meaningful, as points may appear to be equally distant from each other.