

Χουλιάρá Μυρτώ  
1115201400227

Παρακάτω παρατίθενται μερικά ενδεικτικά αποτελέσματα της εκτέλεσης του ./cluster. Δεδομένου των πολλών μεταβλητών, οι συγκεκριμένες που χρησιμοποιήθηκαν σε αυτές τις εκτελέσεις είναι:

number\_of\_clusters: 20  
number\_of\_hash\_functions: 5  
number\_of\_hash\_tables: 4  
max\_loops: 20  
input file: twitter\_dataset\_small.csv  
config file: cluster\_config  
metric: euclidean  
w(για το LSH): 0.3

Random X Lloyd X K-Means:  
20 clusters: 0.0683924  
100 clusters: 0.1252675

K-means X Lloyd X K-Means:  
20 clusters: 0.06925745  
100 clusters: 0.12858766666

Random X Lloyd X K-Medoids (PAM):  
20 clusters: 0.04666295  
100 clusters: 0.10915

K-Means X Lloyd X K-Medoids:  
20 clusters: 0.0447155  
100 clusters: 0.110536

Random X LSH X K-Means:  
20 clusters, euclidean: 0.0340281  
20 clusters, cosine: 0.00599693  
100 clusters, euclidean: 0.0529747  
100 clusters, cosine: 0.076563

Random X LSH X K-Medoids:  
20 clusters, euclidean: 0.0160335  
20 clusters, cosine: 0.0334458  
K-Means X LSH X K-Means:  
20 clusters, euclidean: 0.0301091  
100 clusters, euclidean: 0.055799

K-Means X LSH X K-Medoids:  
20 clusters, euclidean: 0.0176527  
100 clusters, euclidean: 0.0179526

Random X HyperCube X K-Means:  
20 clusters, euclidean: 0.0191426  
20 clusters, cosine: 0.0058788  
100 clusters, euclidean: 0.0710123  
100 clusters, cosine: 0.0926861

Random X HyperCube X K-Medoids:  
20 clusters, euclidean: 0.0360712  
100 clusters, cosine: 0.0774142

K-Means X HyperCube X K-Means:  
20 clusters, euclidean: 0.0462884  
100 clusters, cosine: 0.0962552

K-Means X HyperCube X K-Medoids:  
20 clusters, euclidean: 0.0415788  
20 clusters, cosine: 0.0692961

Παρατήρηση: Η cosine μετρική βγάζει καλύτερα αποτελέσματα απο την euclidean, αν και κάνει περισσότερο χρόνο

Πίνακας:

U / IxA	Random X Lloyd	Random X LSH	Random X HC	K-Means X Lloyd	K-Means X LSH	K-Means X HC
K-Means	0.06925745	0.0340281	0.0191426	0.06925745	0.0301091	0.0462884
K-Medoids	0.04666295	0.0160335	0.0360712	0.0447155	0.0176527	0.0415788

Column: InitxAssign  
Row: Update  
Clusters: 20  
Metric: Euclidean