

entropy textual remarkably euclidean lc phase
rsh clustering
kmeans matrix weights evaluation clustered mining eigenvalues data
document
used problem text
category spectral intraccluster number metrics
similarity speech hybrid algorithm minnewsgroups
distance
purity documents metric sum condition tfidf
relaxed solution different comparable
eg documents metric sum condition tfidf