INCREASING COMPARISON PERFORMANCE USING K-HARMONIC MEAN

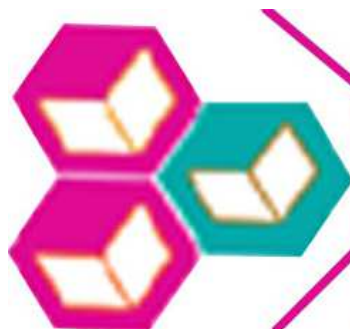## SUCHITRA REYYA[1], M. PUSHPA[2], B. PRIYANKA[3], K. DEVI[4] & G. RUTHROSY[5]

[1]Assistant Professor, Department of Computer Science & Engineering, LENDI Institute of Engineering & Technology,

Andhra Pradesh, India

[2,3,4,5]Department of Computer Science & Engineering, LENDI Institute of Engineering & Technology,

Andhra Pradesh, India

ABSTRACT

Data clustering is one of the common techniques used in data mining. A popular performance function for measuring goodness of data clustering is the total within-cluster variance. The K-Means (KM) algorithm is a popular algorithm which attempts to find a K-clustering. The K-Means [2] algorithm is a centre based clustering algorithm. The dependency of the K-Means performance on the initialization of the centres is a major problem; a similar issue exists for an alternative algorithm, Expectation Maximization (EM) [6]. In this paper, we propose a new clustering method called the K-Harmonic Means algorithm (KHM). KHM [3] is a centre-based clustering algorithm which uses the Harmonic Averages of the distances from each data point to the centres as components to its performance function. It is demonstrated that K-Harmonic Means is essentially insensitive to the initialization of the centres. In certain cases, K-Harmonic Means significantly improves the quality of clustering results comparing with both K-Means and EM, A unified view of the three performance functions, K-Means', K-Harmonic Means 'and EM's, are given for comparison. Experimental results of KHM comparing with KM on Iris [4] data.

KEYWORDS: Clustering, K-Means, K-Harmonic Means, EM, Iris