# Ensemble divide and conquer approach to solve the rating scores' deviation in recommendation system

## ABSTRACT

The rating matrix of a personalized recommendation system contains a high percentage of unknown rating scores which lowers the quality of the prediction. Besides, during data streaming into memory, some rating scores are misplaced from its appropriate cell in the rating matrix which also decrease the quality of the prediction. The singular value decomposition algorithm predicts the unknown rating scores based on the relation between the implicit feedback of both users and items, but exploiting neither the user similarity nor item similarity which leads to low accuracy predictions. There are several factorization methods used in improving the prediction performance of the collaborative filtering technique such as baseline, matrix factorization, neighbour-base. However, the prediction performance of the collaborative filtering using factorization methods is still low while baseline and neighbours-base have limitations in terms of over fitting. Therefore, this paper proposes Ensemble Divide and Conquer (EDC) approach for solving 2 main problems which are the data sparsity and the rating scores' deviation (misplace). The EDC approach is founded by the Singular Value Decomposition (SVD) algorithm which extracts the relationship between the latent feedback of users and the latent feedback of the items. Furthermore, this paper addresses the scale of rating scores as a sub problem which effect on the rank approximation among the users' features. The latent feedback of the users and items are also SVD factors. The results using the EDC approach are more accurate than collaborative filtering and existing methods of matrix factorization namely SVD, baseline, matrix factorization and neighbours-base. This indicates the significance of the latent feedback of both users and items against the different factorization features in improving the prediction accuracy of the collaborative filtering technique.

**Keyword:** Collaborative filtering; Matrix factorization; K-means; Divide and conquer