
Contents

The Star Clustering Algorithm for Information Organization <i>J.A. Aslam, E. Pelekhev, and D. Rus</i>	1
A Survey of Clustering Data Mining Techniques <i>P. Berkhin</i>	25
Similarity-Based Text Clustering: A Comparative Study <i>J. Ghosh and A. Strehl</i>	73
Clustering Very Large Data Sets with Principal Direction Divisive Partitioning <i>D. Littau and D. Boley</i>	99
Clustering with Entropy-Like k-Means Algorithms <i>M. Teboulle, P. Berkhin, I. Dhillon, Y. Guan, and J. Kogan</i>	127
Sampling Methods for Building Initial Partitions <i>Z. Volkovich, J. Kogan, and C. Nicholas</i>	161
TMG: A MATLAB Toolbox for Generating Term-Document Matrices from Text Collections <i>D. Zeimpekis and E. Gallopoulos</i>	187
Criterion Functions for Clustering on High-Dimensional Data <i>Y. Zhao and G. Karypis</i>	211
References	239
Index	265