

## Abstrak

Pelabelan data membutuhkan *cost* yang mahal dan besar, untuk itulah diperlukan suatu sistem dimana data dapat dilabelkan dengan mudah dan tepat. *Semi-supervised clustering* adalah suatu teknik learning untuk mengelompokkan atau melabelkan data *unsupervised* menggunakan *supervised* data sebagai acuannya. HMRF-KMeans merupakan algoritma *semi-supervised clustering*, dimana algoritma ini menggunakan Hidden Markov Random Field, untuk mengambil dan mengobservasi data secara acak dan menghitung probabilitas alaminya melalui komponen parameter hidden (tersembunyi). HMRF-KMeans menggabungkan *constraint-based* dan *distance-based learning* dalam fungsi objektif HMRF-KMeans. Fungsi objektif HMRF-KMeans yang minimum akan menghasilkan kualitas *cluster* yang baik. Dengan *constraint based*, proses inialisasi centroid menjadi tepat dan *distance learning* membantu untuk meminimumkan fungsi objektif HMRF-KMeans.

**Kata kunci:** *cost, semi-supervised clustering, HMRF-KMeans, algoritma, supervised, unsupervised, constraint, distance.*