Data

Divide-and-conquer based Landmark Selection
- Divide the landmark selection into small sub-optimization problems
  - Use light-k-means for large subset
  - Use k-means for large subset
- Set the subset centers as landmarks

Approximate Similarity Matrix Construction
- Find the K-nearest landmarks
- Construct cross-similarity matrix between data points and K-nearest landmarks

Bipartite Graph Partition
- Interpret the cross-similarity matrix as a bipartite graph
- Find a low dimensional embedding via a small eigen-problem
- Apply k-means on the embedding to obtain clustering results.