

BAB V

KESIMPULAN DAN SARAN

1.1. Kesimpulan

Kesimpulan dari penelitian ini adalah algoritme *cluster* cocok untuk mengoptimasi *heatmap* untuk menemukan pola yang baru, sementara validasi data didapatkan dari keseluruhan *preprocessing* dan seleksi fitur ditambah dengan evaluasi kualitas *cluster* itu sendiri. Akurasi *cluster* adalah dari 51% dari kedua pendekatan seleksi fitur, hal ini dikarenakan *Silhouette Coefficient* dirancang untuk menerima data numerik, sehingga tidak menutup kemungkinan jika *cluster* dievaluasi dengan metode yang dikhususkan untuk mengevaluasi *K-Prototype* menghasilkan skor yang lebih baik atau bahkan lebih buruk. Dengan data yang lebih spesifik, hasil *heatmap* bisa diproyeksikan ke peta dan kemudian dapat diuji kualitas dari *heatmap* itu sendiri. Dengan ini dapat disimpulkan jawaban rumusan masalah sebagai berikut:

1. Dengan *preprocessing* pada *item data* dan seleksi fitur yang benar, keseluruhan validitas data sebagai masukan *heatmap* lebih terjamin. Pengelompokan menggunakan *K-Prototype Clustering* menghasilkan pola *heatmap* yang berbeda dengan jumlah kelompok similaritas data yang lebih banyak dibandingkan tanpa melalui *clustering*. Sehingga polisi dapat menggunakan data setelah *cluster* sebagai data *input Geographic Profiling*. Jika terdapat data untuk diproyeksikan ke peta, maka akurasi dari *heatmap* yang terbentuk dapat dicek berdasarkan kasus yang sebelumnya sudah selesai.

1.2. Saran

Berdasarkan hasil penelitian dari masalah dan metode yang dilakukan penulis, penulis menuliskan beberapa saran untuk pengembangan penelitian pada masalah ini yaitu:

1. Perlu adanya data lokasi yang lebih spesifik agar *heatmap* bisa diproyeksikan ke dalam *heatmap* dan diukur kualitas *heatmap* tersebut.

2. Perlu adanya informasi eksternal mengenai data dari ahli dibidang terkait untuk memperbaiki seleksi fitur secara heuristik.
3. Memodifikasi persamaan *K-Modes* yang digunakan dalam *K-Prototype* karena nilai yang dihasilkan *dissimilarity measure* bersifat diskrit dan kurang mencerminkan jarak antar data.
4. Metode pengelompokan *density based* dapat digunakan untuk menemukan pola data berbentuk *arbitrary*.
5. Diperlukan *evaluation metrics* untuk *mixed-type data*.