

BAB 6

KESIMPULAN DAN SARAN

6.1 Kesimpulan

Pada bagian ini akan diberikan kesimpulan atas penelitian yang telah dilakukan. Berikut adalah beberapa kesimpulan tersebut:

1. Berdasarkan penelitian yang sudah dilakukan, *Order Preserving Encryption* dapat diimplementasikan dengan menggunakan Spark.
2. *Ciphertext* yang dihasilkan dari *Order Preserving Encryption* dapat dimanfaatkan untuk penambahan data jika pengetahuan yang diinginkan memanfaatkan sifat keterurutan.
3. Kunci yang dihasilkan dari algoritma *Order Preserving Encryption* dapat digunakan untuk mengenkripsi data baru selama nilai data baru tersebut berada dalam nilai yang dideskripsikan pada kunci.
4. Nilai median dan modus dapat langsung dihitung *ciphertext* sedangkan nilai nilai mean tidak dapat langsung didapatkan dari *ciphertext*.
5. Hasil *clustering* pada *ciphertext* dan *plaintext* tidak memberikan label yang sama.
6. Dalam membuat model K-means pada Spark, nilai K , jumlah iterasi akan memengaruhi nilai error (WSSSE). Semakin tinggi nilai K dan jumlah iterasi yang digunakan akan membuat nilai WSSSE semakin kecil.

6.2 Saran

Pengujian pada lingkungan *big data* seharusnya menggunakan data yang berukuran besar dan selalu terjadi pertumbuhan terhadap datanya. Sedangkan dalam implementasi yang sudah dilakukan program dikembangkan hanya dengan menggunakan *framework* yang mendukung proses komputasi pada *big data*, belum menggunakan data nyata. Selain itu proses komputasi yang dilakukan belum menggunakan *cluster* sehingga proses komputasi paralel belum bisa diuji.

Oleh sebab itu terdapat beberapa saran dari penulis mengenai kelanjutan dari penelitian ini. Berikut adalah saran-saran tersebut:

1. Melakukan pengujian dengan menggunakan data yang bervolume besar dan memiliki fitur yang lebih beragam agar penambahan data yang dilakukan dapat memberikan gambaran yang lebih nyata.
2. Melakukan pengujian pada *Hadoop Cluster*. Pengujian yang sudah dilakukan belum dapat menggunakan *Hadoop Cluster* dikarenakan pandemi *covid-19* sehingga akses dalam menggunakan *Hadoop Cluster* menjadi sulit.
3. Melakukan pengujian dengan memanfaatkan Spark Streaming, sehingga pengujian dapat dilakukan untuk memproses data yang terus bertambah secara *realtime*.
4. Melakukan modifikasi pada *prune phase* agar dapat menanggulangi masalah enkripsi di luar batas kunci. Hal ini disebabkan karena pada pengujian yang sudah dilakukan, *prune phase* yang diimplementasi menghilangkan beberapa langkah dikarenakan keterbatasan waktu.

5. Melakukan pembulatan *floating point* untuk nilai hasil enkripsi. Hal ini dilakukan agar saat dilakukan dekripsi pada cluster tidak terjadi perbedaan perhitungan *floating point* karena setiap komputer bisa saja melakukan perhitungan nilai *float* dengan ketelitian yang berbeda.
6. Melakukan modifikasi terhadap program pengujian K-means sehingga proses pemodelan yang dilakukan akan menggunakan *centroid* yang memiliki nilai yang sama atau mencari cara lain untuk menilai *cluster* yang dihasilkan. Pada pengujian yang sudah dilakukan hal ini belum dapat diimplementasi karena keterbatasan waktu.

DAFTAR REFERENSI

- [1] Kumari, S. (2017) A research paper on cryptography encryption and compression techniques. *International Journal Of Engineering And Computer Science*, **6**, 20915–20919.
- [2] Han, J., Kamber, M., dan Pei, J. (2010) *Data Mining: Concepts and Techniques*, 3rd edition. Morgan Kaufmann, Waltham.
- [3] Saputra, H. (2019) Pembangunan perangkat lunak enkripsi data untuk penambahan data. Skripsi. Universitas Katolik Parahyangan, Indonesia.
- [4] Ariel, M. (2019) Reduksi big data dengan algoritma clustering agglomerative untuk sistem terdistribusi spark. Skripsi. Universitas Katolik Parahyangan, Indonesia.
- [5] Guller, M. (2015) *Big Data Analytics with Spark: A Practitioner's Guide to Using Spark for Large Scale Data Analysis*, 1st edition. Apress, New York.
- [6] Lam, C. (2010) *Hadoop in Action*, 1st edition. Manning Publications Co., USA.