

## BAB V

### KESIMPULAN DAN SARAN

#### 5.1. Kesimpulan

Setelah melakukan penelitian mengenai implementasi Apache Spark dalam melakukan *String Matching* pada Algoritma Boyer-Moore Horspool untuk Bioinformatika, maka penulis mendapatkan beberapa kesimpulan yang berhubungan dengan tujuan penelitian. Berikut kesimpulan yang dapat penulis jabarkan.

1. Berhasil merancang dan mengimplementasikan *cluster* Apache Spark untuk pencocokkan *pattern* pada *string* menggunakan algoritma Boyer-Moore Horspool dengan jumlah *node* 1, 3, 5, 11, 16 dengan menerapkan beberapa fungsi yang berada di Apache Spark sehingga algoritma Boyer-Moore Horspool dapat berjalan secara paralel sesuai dengan pada tujuan penelitian ini.
2. Telah melakukan total 102 eksperimen yang terbagi ke dalam skenario *stand alone* dan skenario *cluster* Apache Spark dan pembagian ke dalam beberapa skenario data.
3. Melakukan analisa hasil eksperimen dengan kesimpulan bahwa pemangkasan waktu tidak beriringan dengan jumlah *node*. Dapat diambil kesimpulan, bahwa pada penelitian ini penggunaan *platform* Big Data belum maksimal. Sebab kurva yang dihasilkan pada diagram di seluruh eksperimen tidak menurun. Bahkan mengalami kenaikan dan penurunan yang tidak stabil. Hal itu menjadi indikasi bahwa pemanfaatan Big Data untuk kasus pada Penelitian ini kurang tepat sasaran.

#### 5.2. Saran

Dalam pelaksanaan penelitian, penulis menyadari bahwa masih banyak kekurangan yang dilakukan oleh penulis dalam penelitian ini. Oleh karena itu, penulis menyampaikan beberapa saran yang dapat dilakukan di kemudian hari. Penelitian selanjutnya dapat menghasilkan sebuah program yang jauh lebih baik dalam hal kecepatan. Berikut beberapa saran yang dapat penulis anjurkan.

1. Dalam melakukan input data masih dilakukan di dalam kode program sehingga harus dilakukan oleh yang paham dengan pemograman.
2. Terkait tema penelitian *big data* harus terlebih dahulu ditelaah secara matang apakah data dan studi kasus cukup berat dan besar sehingga *stand alone* akan memakan waktu yang lama sehingga penggunaan *platform* Big Data akan lebih maksimal.
3. Membuat beberapa fungsi praproses yang lebih beragam dan dinamis sehingga dapat digunakan pada data yang universal, tidak berpatok pada data dengan standar format NCBI/Ensembl dan juga penamaan header ketika data dipanggil di dalam satu direktori sekaligus sehingga sekuens kromosom satu dan lainnya tidak akan bersatu.
4. Membuat sebuah fungsi untuk mengatasi perpotongan sekuens DNA yang terjadi sehingga tidak ada pola yang terpotong di dalam sekuens dan menghasilkan nilai akurasi lebih baik.
5. Penulis berharap program ini dapat digunakan untuk menunjang perkembangan ilmu pengetahuan khususnya dalam bidang ilmu komputer dan biologi.
6. Penulis berharap program ini dapat dikembangkan jauh lebih baik lagi