

## BAB V

### KESIMPULAN DAN SARAN

#### 1.1 Kesimpulan

Setelah melakukan penelitian mengenai implementasi algoritma Boyer-Moore untuk mendeteksi *genomic repeats* menggunakan Apache Spark Streaming, maka dapat diambil beberapa kesimpulan yang selaras dengan tujuan penelitian. Berikut kesimpulan yang didapat oleh penulis.

1. Penelitian ini berhasil merancang model komputasi *Big Data* untuk pencarian *pattern* pada *string* menggunakan algoritma Boyer-Moore dengan 6 tahap utama yaitu praproses, data input, menjalankan Apache Spark Streaming, memasukkan data kedalam folder streaming, sistem algoritma Boyer-Moore, dan *download output*.
2. Penelitian ini berhasil mengimplementasikan model komputasi *Big Data* menggunakan algoritma Boyer-Moore pada Apache Spark Streaming dengan bahasa pemrograman Python. Indikator keberhasilan merupakan program berhasil mengeluarkan hasil *genomic repeats* dengan kecepatan kurang dari 3 menit atau 180 detik dan mencapai akurasi rata-rata 80%.
3. Penelitian ini berhasil melakukan analisa terkait kecepatan dari hasil eksperimen yang dilakukan yaitu dengan menggunakan 4 *core* dan beberapa *worker nodes* dan 2 *worker nodes* dan beberapa *cores* yang dipakai pada setiap *node* nya.
4. Telah melakukan total 576 eksperimen yang terbagi dalam 2 skenario dengan menggunakan beberapa *worker nodes* dan 4 *core* CPU sebanyak 288 kali eksperimen dan skenario dengan menggunakan beberapa *core* CPU dan 2 *worker nodes* sebanyak 288 kali eksperimen. Dari hasil eksperimen dapat diambil kesimpulan bahwa semakin banyak *worker nodes* atau semakin banyak *core* yang digunakan tidak akan menjamin kecepatan dari proses komputasi maupun pertambahan efisiensi dari penggunaan *worker nodes* dan *core* dalam komputasi.

## 1.2 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. Mengefisiensikan kode program pada fungsi *BM\_algorithm* agar dapat mempercepat proses komputasi.
2. *Output* dari hasil program sebaiknya dapat lebih dirapihkan sehingga pengguna yang memerlukan data *output* dapat membaca *file* dengan mudah dan jelas tanpa bantuan program lain untuk membaca *file*.
3. Dapat menghasilkan akurasi yang lebih tinggi dengan mengubah *block size* pada Apache Spark dalam mempartisi data input agar dapat mencapai rata-rata akurasi hingga 100%.
4. Penulis berharap program ini dapat menjadi modal awal dalam melakukan penelitian - penelitian selanjutnya dengan menggunakan studi kasus yang memerlukan *streaming processing* pada Apache Spark Streaming.