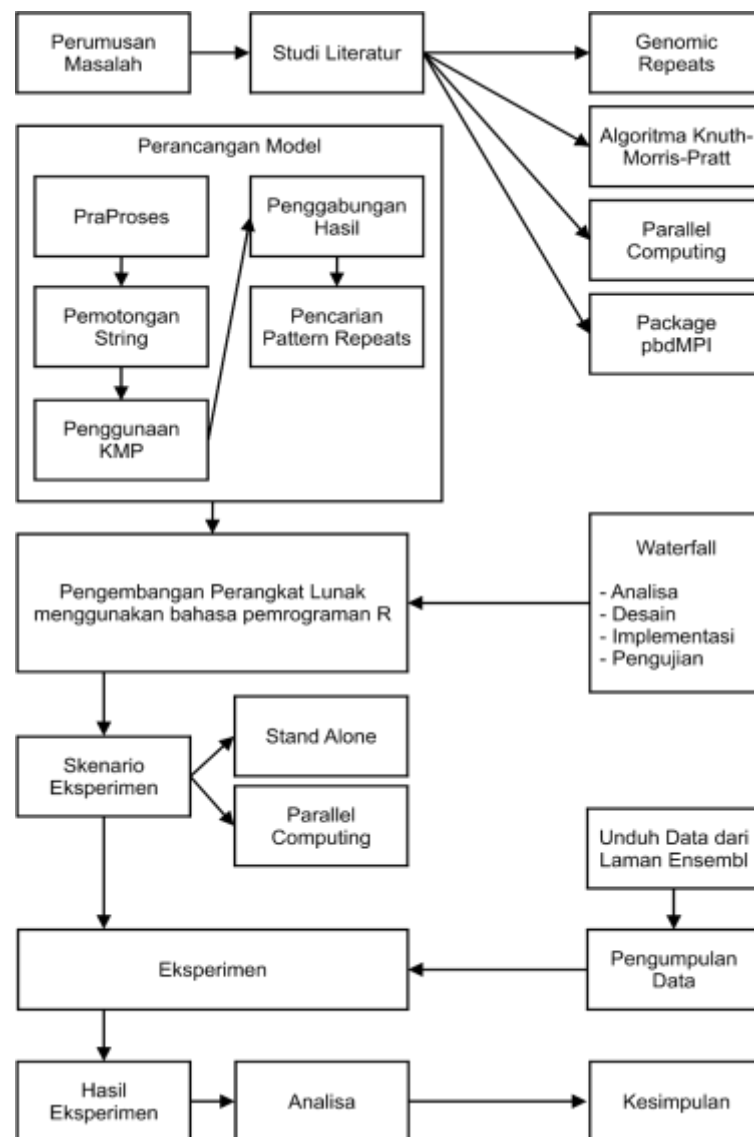


## BAB III METODOLOGI PENELITIAN

### 3.1 Desain Penelitian

Desain penelitian adalah kerangka kerja yang digunakan untuk melakukan penelitian. Pada bagian ini penulis akan memaparkan kerangka kerja dari mulai penelitian sampai selesai. Desain penelitian digambarkan pada Gambar 3.1.



**Gambar 3. 1 Desain proses penelitian deteksi *genomic repeats* menggunakan algoritma Knuth-Morris-Pratt pada R *high-performance computing package***

1. Pada tahap pertama penulis melakukan diskusi untuk menemukan dan menentukan rumusan masalah yang dapat diangkat menjadi topik pada penelitian skripsi ini. Pada tahap ini pula penulis merumuskan tujuan dari penelitian yang berkaitan dengan rumusan masalah yang juga telah tertulis pada bab pertama skripsi ini.
2. Selanjutnya penulis melakukan studi literatur berkaitan dengan topik yang telah disetujui pada tahap pertama. Pada tahap ini dilakukan studi literature tentang *genomic repeats* yang mengandung pengetahuan umum tentang DNA serta *trinucleotide repeat disorders*, penyakit genetik yang disebabkan perulangan sejumlah 3 asam pasang basa. Terdapat pula pengetahuan tentang algoritma *string matching* Knuth-Morris-Pratt, pemrograman bahasa R, *parallel computing* dan tentang *package High-Performance Computing* pada bahasa pemrograman R yang akan digunakan pada penelitian ini, pbdMPI.
3. Tahap berikutnya penulis melakukan pengumpulan data yang dapat diunduh dari *genbank* yang tersedia: FTP NCBI atau FTP Ensembl. Pada penelitian kali ini penulis mengunduh 24 file kromosom (kromosom 1-22, kromosom X dan Y) manusia dari publikasi nomor 88 di laman FTP Ensembl yang dapat diunduh pada [ftp://ftp.ensembl.org/pub/release-88/fasta/homo\\_sapiens/dna/](ftp://ftp.ensembl.org/pub/release-88/fasta/homo_sapiens/dna/)
4. Selanjutnya penulis merancang model *parallel computing* untuk pencarian *pattern* pada *string* menggunakan algoritma Knuth-Morris-Pratt. Rancangan model yang dibuat dimulai dari praproses, pemotorngan *string*, penggunaan algoritma Knuth-Morris-Pratt pada setiap iterasi, penggabungan hasil dari setiap iterasi dan pencarian *pattern* yang berulang.
5. Dilanjutkan dengan membuat program pencarian *pattern* pada *string* dengan menggunakan algoritma Knuth-Morris-Pratt yang diimplementasikan secara paralel dengan menggunakan *package* pbdMPI pada bahasa pemrograman R dari model yang telah dirancang pada proses sebelumnya.
6. Setelah melewati pembuatan program dari analisa hingga pengujian, penulis membuat skenario eksperimen yang akan dilakukan selanjutnya.

7. Penulis melakukan eksperimen komputasi *stand alone* dan *parallel computing* dengan berbagai *core* dan nilai iterasi untuk didapatkannya perbandingan dari setiap satu kali percobaan.
8. Langkah berikutnya penulis melakukan analisa dari hasil eksperimen yang telah dilakukan pada tahap sebelumnya. Pada tahap ini penulis diharapkan mendapatkan grafik dari hasil eksperimen berbagai penggunaan jumlah *core* dan nilai iterasi pada setiap satu kali percobaan.
9. Terakhir penulis akan mendokumentasikan hasil penelitian ini dan menyusun laporan penelitian dalam bentuk skripsi.

### 3.2 Alat dan Bahan Penelitian

Bagian ini menjelaskan secara detail alat dan bahan yang digunakan untuk melakukan penelitian.

#### 3.2.1 Alat Penelitian

1. Perangkat Keras (*Hardware*) yaitu komputer dengan spesifikasi:
  - Processor Intel® Core™ i7-5820K CPU @ 3.30Ghz (12 CPUs), ~3.3GHz
  - Memory 32768MB RAM
  - NVIDIA GeForce GTX 980
  - HDD 730 GB
2. Perangkat lunak (*software*) sebagai berikut:
  - Sistem Operasi Windows 10 Pro 64-bit (10.0, Build 14393)
  - R 3.4.0
  - Rstudio 1.0.143
  - Microsoft MPI 8.1.12438.1084
  - Microsoft MPI SDK 8.1.12438.1084
  - Microsoft Office Excel
  - ConEmu 161206.x64

#### 3.2.2 Bahan Penelitian

Bahan yang diperlukan untuk melakukan penelitian yaitu data sekuens DNA manusia yang diunduh dari Ensembl yang dapat diunduh pada

Ahmad Bayu Rachman, 2017

**DETEKSI GENOMIC REPEATS MENGGUNAKAN ALGORITMA KNUTH-MORRIS-PRATT PADA R HIGH-PERFORMANCE COMPUTING PACKAGE**

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

[ftp://ftp.ensembl.org/pub/release-88/fasta/homo\\_sapiens/dna/](ftp://ftp.ensembl.org/pub/release-88/fasta/homo_sapiens/dna/) serta beberapa artikel untuk menunjang pengetahuan guna terlaksanakannya penelitian ini.

### 3.3 Metode Penelitian

Adapun metode yang dilakukan dalam penelitian ini dibagi kedalam dua bagian, yaitu metode pengumpulan data dan metode pengembangan perangkat lunak.

#### 3.3.1 Metode Pengumpulan Data

Penulis berusaha mendapatkan data yang valid dan mampu menunjang penelitian. Ada pun metode pengumpulan data pada penelitian ini adalah sebagai berikut:

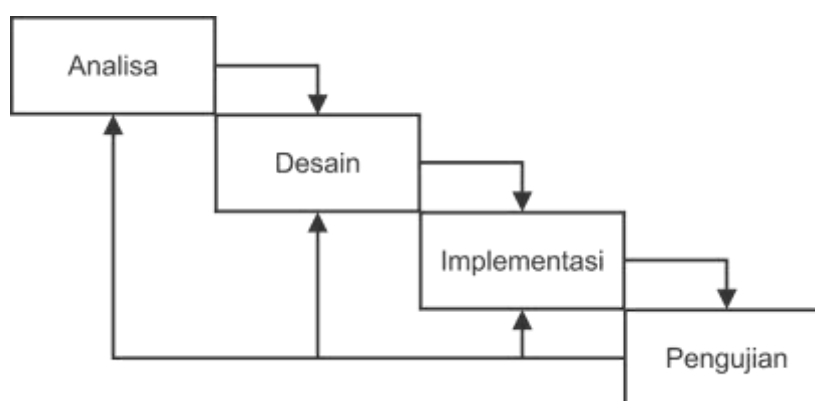
1. Studi Literatur

Studi literatur dilakukan dengan mempelajari teori dan konsep yang menjadi pendukung dalam penelitian ini, yaitu tentang *genomic repeats*, algoritma Knuth-Morris-Pratt, *parallel computing* dan pbdMPI.

2. Mengunduh Data Sekuens

Pengumpulan data sekuens yang akan digunakan pada penelitian ini diunduh dari *genbank* terpercaya, Ensembl. Data yang digunakan dapat diunduh pada sumber yang telah disebutkan pada poin bahan penelitian.

#### 3.3.2 Metode Pengembangan Perangkat Lunak



**Gambar 3. 2 Model Waterfall (Sommerville, 2011)**

Metode pengembangan perangkat lunak dilakukan dengan metode *waterfall*. Model SDLC air terjun (*waterfall*) sering juga disebut model sekuensial

Ahmad Bayu Rachman, 2017

**DETEKSI GENOMIC REPEATS MENGGUNAKAN ALGORITMA KNUTH-MORRIS-PRATT PADA R  
HIGH-PERFORMANCE COMPUTING PACKAGE**

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

linier (*sequential linier*). Model *waterfall* menyediakan pendekatan alur hidup perangkat lunak secara sekuensial atau urut dimulai dari analisis, desain, pengkodean, pengujian dan tahap *support* (Sukamto dan Shalahuddin, 2011). Penulis menggunakan metode *modern waterfall* seperti pada Gambar 3.2 agar jika suatu saat ada kesalahan pada salah satu tahap, bisa dikembalikan ke tahap sebelumnya. Berikut pengertian dari tahap-tahap pada model *waterfall* pada Gambar 3.2 menurut Ian Sommerville (2011) :

1. *Requirments Analysis and Definition* (Analisa)

Analisis adalah tahap menentukan aplikasi atau *software* seperti apakah yang akan dibuat. Analisis merupakan tahapan penetapan fitur, kendala dan tujuan sistem melalui konsultasi dengan pengguna sistem. Semua hal tersebut akan ditetapkan secara rinci dan berfungsi sebagai spesifikasi sistem. Analisis ini terdiri dari analisis kebutuhan dan analisis pembuatan sistem.

2. *System and Software Design* (Desain)

Dalam tahapan ini akan dibentuk suatu arsitektur sistem berdasarkan persyaratan yang telah ditetapkan. Dan juga mengidentifikasi dan menggambarkan abstraksi dasar sistem perangkat lunak dan hubungan-hubungannya. Desain terdiri dari desain database, desain arsitektur system, dan desain antarmuka (*user interface*)

3. *Implementation and Unit Testing* (Implementasi)

*Coding* adalah tahap proses implementasi dari desain, dalam tahapan ini, hasil dari desain perangkat lunak akan direalisasikan sebagai satu set program atau unit program. Setiap unit akan diuji apakah sudah memenuhi spesifikasinya.

4. *Integration and System Testing* (Pengujian)

Proses testing atau pengujian dilakukan pada logika internal untuk memastikan semua pernyataan sudah diuji. Dalam tahapan ini, setiap unit program akan diintegrasikan satu sama lain dan diuji sebagai satu sistem yang utuh untuk memastikan sistem sudah memenuhi persyaratan yang ada. Setelah itu sistem akan dikirim ke pengguna sistem.