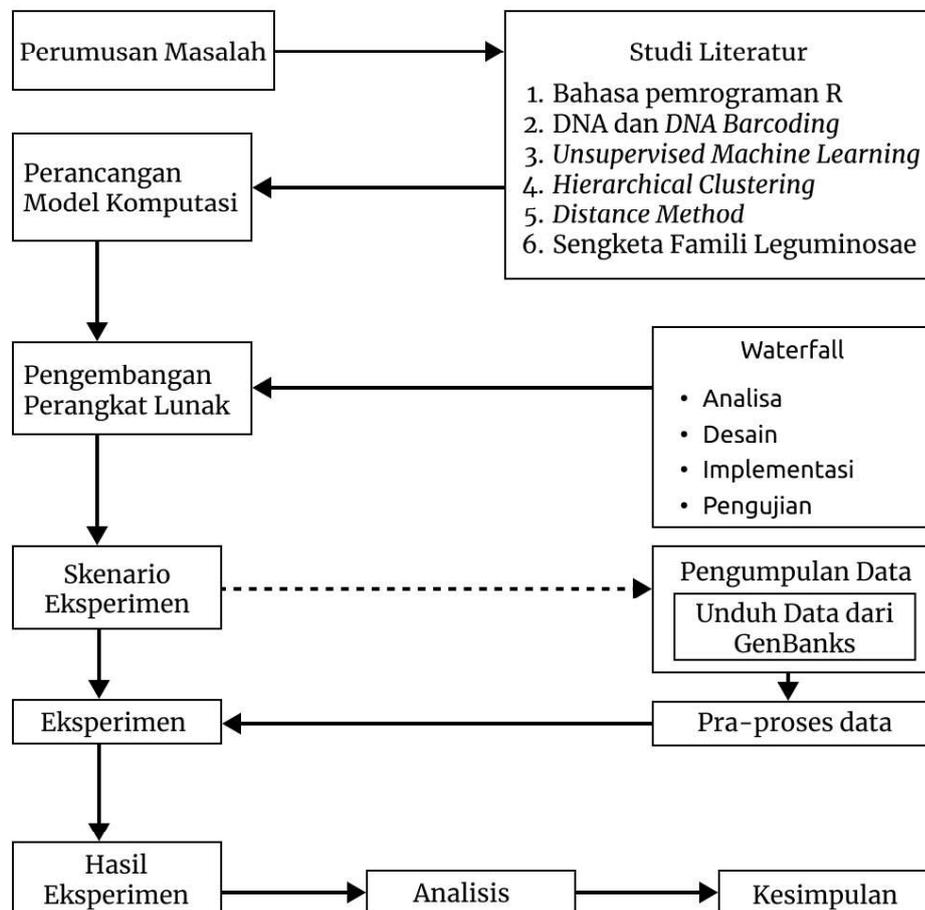


## BAB III METODE PENELITIAN

### 3.1 Desain Penelitian

Desain penelitian merupakan kerangka kerja yang digunakan untuk melakukan penelitian. Pada bagian ini penulis akan memaparkan kerangka kerja terkait penelitian dari memulai penelitian sampai dengan selesai. Desain penelitian dijelaskan lebih spesifik pada Gambar 3.1.



**Gambar 3.1 Desain Penelitian**

Terdapat 7 tahap utama dalam desain penelitian yang dilakukan, yang akan dijelaskan sebagai berikut:

Muhammad Iqbal Zain, 2023  
*IMPLEMENTASI HIERARCHICAL CLUSTERING DALAM DNA BARCODING UNTUK MENENTUKAN TAKSONOMI TUMBUHAN*  
Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

## 1. Perumusan Masalah

Tahap ini merupakan tahap dimana penulis akan mengidentifikasi dan merumuskan masalah yang akan dibahas yang kemudian akan menentukan metode atau algoritma yang akan digunakan untuk menyelesaikan masalah tersebut. Setelah itu, pada tahap ini juga akan ditentukan model penelitian untuk membantu penyelesaian masalah yang telah dirumuskan.

## 2. Studi Literatur

Studi literatur berkaitan dengan proses belajar dan literasi topik yang telah disetujui sebelumnya. Penulis melakukan studi literatur mengenai Bahasa pemrograman R, *DNA dan DNA barcoding*, *Unsupervised Machine Learning*, *Hierarchical Clustering*, *Distance Method*, dan *Sengketa Famili Leguminosae* sebagai landasan penelitian yang dilakukan oleh penulis. Studi literatur berasal dari beberapa sumber seperti buku, jurnal, juga internet, dan bacaan lainnya yang dapat dijadikan bahan literatur.

## 3. Perancangan Model Komputasi

Tahap ini merupakan tahap dimana penulis akan melakukan perancangan model algoritma *hierarchical clustering* untuk kasus *DNA barcoding*. Rancangan model ini dimulai dari pengambilan data, pra proses data, hingga melakukan *clustering*.

## 4. Tahap Pengembangan Perangkat Lunak

Perangkat lunak yang dirancang untuk melakukan *clustering* pada data *barcode* DNA dibuat berdasarkan model komputasi yang dijelaskan pada tahapan sebelumnya. Metode pengembangan perangkat lunak yang digunakan adalah metode *waterfall*.

## 5. Tahap Perancangan Skenario Eksperimen

Setelah dilakukan pengembangan program pada tahap sebelumnya, pada tahap ini akan dilakukan perancangan skenario eksperimen yang akan dilakukan untuk menguji apakah program berjalan sebagaimana mestinya

dan menganalisa hasil keluaran dari program. Skenario yang digunakan pada tahap ini terbagi menjadi dua fase, yaitu fase validasi metode dan fase studi kasus menggunakan metode yang telah teruji.

#### 6. Tahap Pengumpulan Data

Berdasarkan literatur dan tahap skenario eksperimen yang menentukan kebutuhan data yang akan digunakan. Pada tahap ini penulis mengumpulkan data untuk melakukan eksperimen.

#### 7. Pra Proses Data

Sebelum diproses lebih lanjut, data yang dikumpulkan pada tahap sebelumnya terlebih dahulu dilakukan pra proses data. Dalam tahap ini dilakukan verifikasi data yang didapatkan, *sequence alignment*, *sequence trimming*, dan *one hot encoding*.

#### 8. Tahap Eksperimen

Pada tahap ini setelah semua kebutuhan telah terpenuhi, eksperimen dapat dilakukan berdasarkan skenario yang telah dibuat dan menggunakan data yang telah didapat dan di proses.

#### 9. Tahap Hasil dan Analisis

Setelah melakukan perancangan skenario dan melakukan eksperimen sesuai dengan skenario yang telah dirancang pada tahap sebelumnya, hasil dari eksperimen tersebut akan dianalisa. Pada tahap ini diharapkan dapat mendapatkan hasil dari eksperimen yang telah dilakukan sebelumnya.

## 3.2 Metode Penelitian

Metode penelitian yang dilakukan dalam penelitian ini dibagi menjadi dua tahapan, yaitu metode pengumpulan data dan metode pengembangan perangkat lunak.

### 3.2.1 Metode Pengumpulan Data

Pada metode pengumpulan data penulis berusaha mendapatkan data yang valid dan mampu untuk menunjang penelitian yang dilakukan. Adapun metode pengumpulan data pada penelitian ini adalah sebagai berikut:

1. Studi Literatur

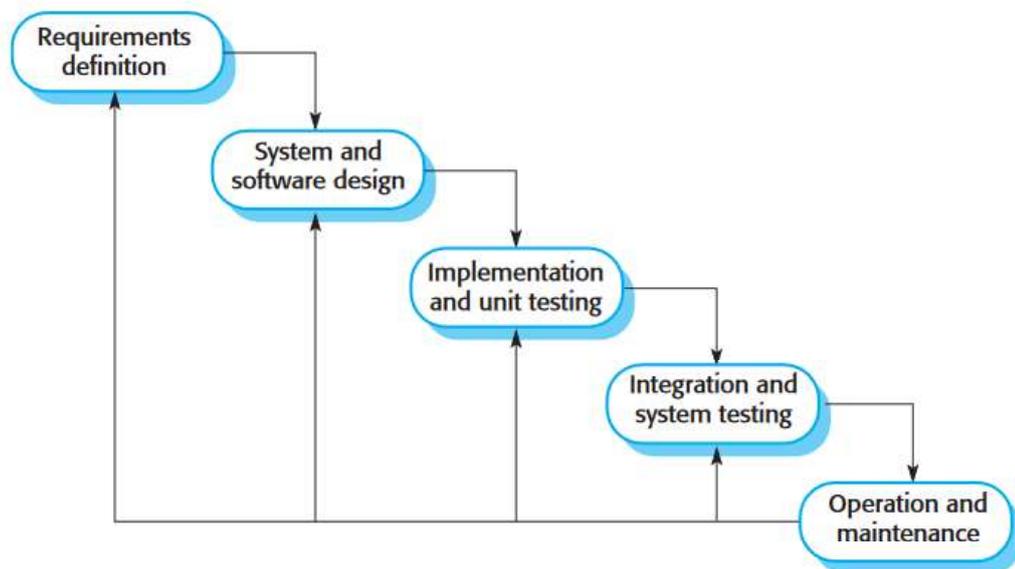
Studi literatur dilakukan dengan cara mempelajari teori dan konsep yang mendukung dalam penelitian ini dengan membaca dari berbagai sumber seperti buku, jurnal, dan sumber-sumber ilmiah lain mengenai topik Bahasa pemrograman R, *DNA dan DNA barcoding*, *Unsupervised Machine Learning*, *Hierarchical Clustering*, *Distance Method*, dan *Sengketa Famili Leguminosae*.

2. Mendapatkan Sekuens Data

Untuk mendukung penelitian yang dilakukan oleh penulis, maka diperlukan data yang akan digunakan dalam penelitian. Data diperoleh dari situs GenBank yang terpercaya seperti NCBI dan hasil sekuensing pribadi. Proses pengumpulan data dilakukan dengan bantuan *library* agar memudahkan dalam pengambilan data yang cukup banyak.

### 3.2.2 Metode Pengembangan Perangkat Lunak

Dalam penelitian ini pengembang menggunakan model waterfall untuk melakukan pembangunan perangkat lunak.



**Gambar 3.2 Model Waterfall dalam Pengembangan Perangkat Lunak**

Model *waterfall* yang digambarkan pada Gambar 3.2 menggambarkan lima tahapan yang dapat membantu dalam proses pengerjaan perangkat lunak (Sommerville, 2016). Pengertian dari setiap tahapan tersebut sebagai berikut:

#### 1. Perumusan Masalah

Analisis dan definisi kebutuhan layanan sistem, kendala, dan tujuan ditetapkan melalui konsultasi dengan pengguna sistem. Pada tahap ini dilakukan analisis terhadap data-data yang didapatkan dan diperlukan seperti data sekuens DNA yang akan digunakan. Tahapan ini harus dilengkapi hingga memenuhi semua kebutuhan.

#### 2. Desain

Proses desain membentuk arsitektur sistem secara keseluruhan. Desain perangkat lunak melibatkan pengidentifikasian dan penggambaran abstraksi sistem perangkat lunak dasar dan hubungannya. Dalam tahapan ini akan dilakukan perancangan desain perangkat lunak yang akan dibuat dari

basis data, arsitektur, dan lain sebagainya. Desain ini dibuat berdasarkan spesifikasi yang telah ditetapkan sebelumnya.

### 3. Implementasi

Pada tahap implementasi, desain perangkat lunak diwujudkan sebagai satu set program atau unit program. Program ini juga nantinya akan diuji agar memenuhi kebutuhan spesifikasi.

### 4. Pengujian

Pada tahap pengujian atau *testing* melibatkan verifikasi bahwa setiap bagian dalam sebuah program memenuhi spesifikasinya. Pada tahapan ini memastikan apakah program sudah sesuai dengan kebutuhan spesifikasi atau belum. Dilakukan juga pengujian dalam hal integrasi untuk memastikan setiap bagian dalam perangkat lunak sudah terintegrasi.

## 3.3 Alat dan Bahan Penelitian

Bagian ini akan menjelaskan secara detail mengenai alat dan bahan penelitian yang digunakan pada penelitian ini.

### 3.3.1 Alat Penelitian

Dalam penelitian ini alat yang digunakan untuk menunjang kebutuhan selama penelitian ini adalah sebagai berikut:

1. Perangkat Keras (*hardware*) yaitu komputer dengan spesifikasi:
  - Processor Intel Core i5-10300H 2.5GHz
  - *Memory* 16GB RAM
  - SSD 1TB
2. Perangkat Lunak (*software*) sebagai berikut:
  - Sistem Operasi Ubuntu 18.04
  - Jupyter Notebook
  - Python 3.8
  - Apache Spark v 2.4
  - *Web Browser* Google Chrome
  - Microsoft Office Excel

### 3.3.2 Bahan Penelitian

Bahan yang diperlukan pada penelitian ini adalah data awal berupa kumpulan DNA dalam format .fasta. Data ini terdiri dari dua jenis data yaitu data fase validasi *hierarchical clustering* dengan membandingkan beberapa *distance method* dan data DNA dari spesies yang sedang dalam sengketa / bermasalah. *DNA barcode* dari spesies yang telah dipilih kemudian diperoleh dari sumber awal yaitu GenBank NCBI (<https://www.ncbi.nlm.nih.gov/genbank/>).