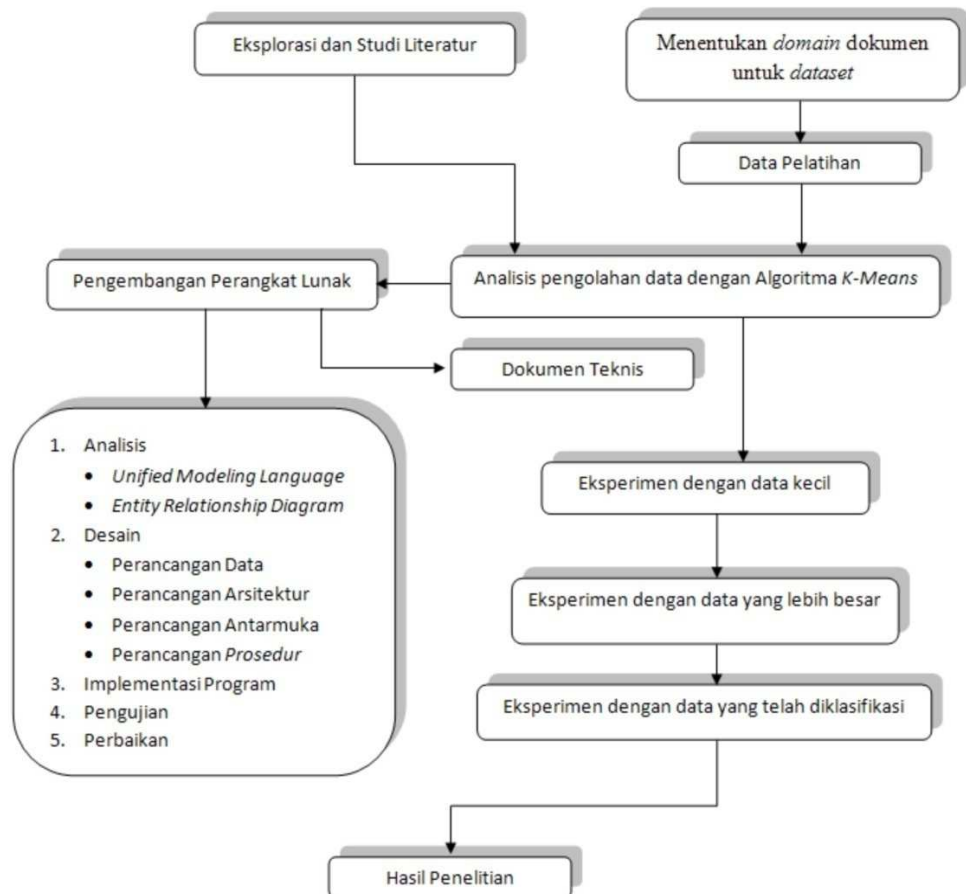


## BAB III

### METODOLOGI PENELITIAN

#### 3.1 Desain Penelitian

Gambar 3.1 merupakan desain penelitian yang akan digunakan dalam proses penelitian penerapan algoritma K-Means pada clustering berita berbahasa Indonesia.



Gambar 3.1 Desain Penelitian

## **3.2 Metode Penelitian**

### **3.2.1 Eksplorasi dan studi literatur**

Eksplorasi dan studi literatur dilakukan dengan mempelajari konsep-konsep yang berkaitan dengan skripsi ini, seperti *clustering* pada aliran berita yang terdiri dari karakteristik berita dan transformasi berita menjadi dataset, algoritma *Clustering*, melalui literatur-literatur seperti buku (*textbook*), paper, dan sumber ilmiah lain seperti situs internet ataupun artikel dokumen teks yang berhubungan.

### **3.2.2 Metode Pengembangan Perangkat Lunak**

Dalam melakukan pengembangan perangkat lunak ini, penulis menggunakan pendekatan terstruktur yang merupakan suatu pendekatan berorientasi objek terfokus pada pengolahan data dan proses dari suatu perangkat lunak.

Pada tahap awal pengembangan perangkat lunak, dibutuhkan suatu model untuk mengacu dalam pengembangan perangkat lunak. Model itu disebut model proses. Pada penelitian ini penulis menggunakan model sekuensial linier.

Model sekuensial linier untuk software engineering, sering disebut juga dengan siklus kehidupan klasik atau model air terjun. Model ini mengusulkan sebuah pendekatan kepada perkembangan software yang sistematis dan sekuensial yang mulai pada tingkat dan kemajuan sistem pada seluruh analisis, desain, kode, pengujian, dan pemeliharaan. Dimodelkan setelah siklus rekayasa konvensional, model sekuensial linier melingkupi aktivitas – aktivitas sebagai berikut (Yandra, 2010):

## 1. Rekayasa dan permodelan sistem

Karena sistem merupakan bagian dari sebuah sistem yang lebih besar, kerja dimulai dengan membangun syarat dari semua elemen sistem dan mengalokasikan beberapa subset dari kebutuhan ke software tersebut. Pandangan sistem ini penting ketika software harus berhubungan dengan elemen-elemen yang lain seperti software, manusia, dan database.

## 2. Analisis kebutuhan software

Proses pengumpulan kebutuhan diintensifkan dan difokuskan, khususnya pada software. Untuk memahami sifat program yang dibangun, analis harus memahami domain informasi, tingkah laku, unjuk kerja, dan interface yang diperlukan dalam penelitian penggunaan algoritma *K-Means* untuk *clustering* dokumen berita berbahasa Indonesia ini. Kebutuhan baik untuk sistem maupun software didokumentasikan.

Analisis terhadap kebutuhan dilakukan dengan menggunakan beberapa alat (*tools*) seperti :

- Kamus Data (*Data Dictionary*)

Digunakan untuk menyimpan deskripsi untuk semua objek data yang digunakan pada Aplikasi *clustering engine*. (Merujuk ke dokumen teknis).

- Diagram Keterhubungan Entitas (*ERD/Entity Relationship Diagram*)

ERD ini menggambarkan hubungan antar objek data (entitas) *clustering engine*

- *Unified Modeling Language*

UML adalah bahasa pemodelan standar atau kumpulan teknik-teknik pemodelan untuk menspesifikasi, mem-visualisasi, mengkonstruksi dan mendemonstrasikan hasil kerja dalam pengembangan perangkat lunak (Fowler, 2004). UML lahir dari penggabungan banyak bahasa pemodela grafis berorientasi obyek yang berkembang pesat pada akhir tahun 1980an dan awal 1990an.

Secara sederhana UML digunakan untuk menggambar sketsa sistem. Pengembang menggunakan UML untuk menyampaikan beberapa aspek dari sebuah perangkat lunak melalui notasi grafis. UML mendefinisikan notasi dan semantik. Notasi merupakan sekumpulan bentuk khusus yang memiliki makna tertentu untuk menggambarkan berbagai diagram piranti lunak dan semantik mendefinisikan bagaimana bentuk-bentuk tersebut dapat dikombinasikan. Ada beberapa jenis diagram yang disediakan dalam UML, antara lain (Tn, 2010):

1. Use-case diagram

Diagram ini berguna untuk menggambarkan interaksi antara pengguna dengan sebuah perangkat lunak.

2. Activity diagram

Diagram ini berguna untuk menggambarkan prosedur-prosedur perilaku perangkat lunak.

3. Class diagram

Diagram ini berguna untuk menggambarkan class, fitur, dan hubungan-hubungan yang terjadi. Pada diagram ini pendekatan berorientasi obyek memegang peranan yang sangat penting.

#### 4. Sequence diagram

Diagram ini berguna untuk menggambarkan interaksi antar obyek dengan penekanan pada urutan proses atau kejadian.

#### 5. State machine diagram

Diagram ini digunakan untuk menggambarkan bagaimana suatu kejadian mengubah obyek selama masa hidup obyek tersebut.

### 3. Desain

Desain software sebenarnya adalah proses multi langkah yang berfokus pada empat atribut sebuah program yang berbeda, struktur data, arsitektur software, representasi interface, dan detail algoritma. Proses desain menterjemahkan syarat/kebutuhan ke dalam sebuah representasi software yang dapat diperkirakan demi kualitas sebelum dimulai pemunculan kode. Sebagaimana persyaratan, desain didokumentasikan dan menjadi bagian dari konfigurasi software.

#### 4. Generasi kode

Desain harus diterjemahkan kedalam bentuk mesin yang bisa dibaca. Langkah pembuatan kode melakukan tugas ini. Jika desain dilakukan dengan cara yang lengkap, pembuatan kode dapat diselesaikan secara mekanis.

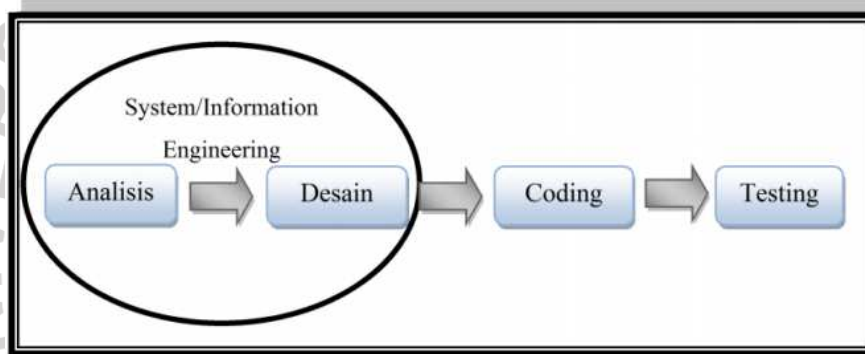
#### 5. Pengujian

Sekali program dibuat, pengujian program dimulai. Proses pengujian berfokus pada logika internal software, memastikan bahwa semua pernyataan sudah diuji, dan pada eksternal fungsional, yaitu mengarahkan pengujian untuk menemukan kesalahan – kesalahan dan memastikan bahwa input yang dibatasi

akan memberikan hasil aktual yang sesuai dengan hasil yang dibutuhkan. Proses pengujian berfokus kepada keakuratan *clustering engine* dalam mengelompokkan data. Sebarapa besar tingkat keakuratan dari Algoritma *K-Means*. (Merujuk ke Dokumen Teknis).

### 3.2.3 Eksperimen terhadap Algoritma

Eksperimen terhadap algoritma berfokus kepada algoritma yang digunakan pada penelitian ini, yakni memberikan berbagai perlakuan terhadap masukan untuk algoritma K-Means.



**Gambar 3.2** Model Sekuensial Linier

## 3.3 Alat dan Bahan Penelitian

### 3.3.1 Alat Penelitian

Alat-alat yang digunakan dalam penelitian ini terbagi kedalam dua kelompok, yaitu perangkat lunak dan perangkat keras. Untuk perangkat keras yang digunakan dalam pengembangan *clustering engine* ini, penulis menggunakan perangkat keras dengan spesifikasi sebagai berikut :

1. Processor AMD Phenom 9650 Quad-Core 2.3 GHz
2. Memori 2048 MB
3. Harddisk 160 GB

4. VGA 700MB *Shared*
5. Monitor 17 Inch dengan Resolution 1152 x 864 px
6. Mouse
7. Keyboard
8. Perangkat keras penyimpan data berupa flashdisk, CD.

Sedangkan kebutuhan perangkat lunak yang digunakan untuk pengembangan aplikasi ini, penulis menggunakan perangkat lunak dengan spesifikasi sebagai berikut :

1. Sistem Operasi Microsoft Windows XP service Pack 2
2. JDK 1.6 Update 18
3. Netbeans 6.8 yang telah terintegasi dengan apache Tomcat
4. XAMPP versi 1.7.3

### **3.3.2 Bahan Penelitian**

Bahan penelitian yang digunakan dalam penelitian ini adalah *content* berita yang telah dikumpulkan dari situs <http://www.detik.com> .