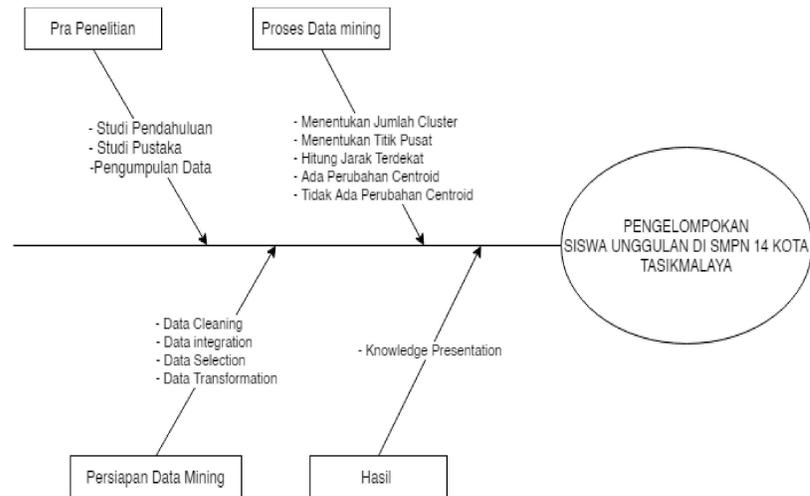


## BAB III METODOLOGI PENELITIAN

### 3.1 Tahap Penelitian



**Gambar 3. 1** Tahap Penelitian

Penelitian ini dibagi menjadi 4 tahapan utama yaitu :

1. Pra penelitian dimulai dengan melakukan studi pendahuluan untuk mengidentifikasi masalah, studi pustaka dan pengumpulan data nilai akademik siswa SMPN 14 Kota Tasikmalaya.
2. Persiapan proses data mining terdiri dari, *Data cleaning*, *Data Authentication*, *Data integration*, *Data Selection* dan *Data Transformation*.
3. Proses data mining dengan mengelompokkan menggunakan algoritma *K-Means*.
4. Melakukan evaluasi hasil pengujian data.

### 3.2 Pra Penelitian

1. Studi Pendahuluan

Studi pendahuluan merupakan tahap awal dalam melakukan penelitian dan pengembangan model. Pada tahap ini, peneliti melakukan observasi dan mengumpulkan data-data terkait pengamatan secara langsung ke pengamatan SMPN 14 Kota Tasikmalaya secara langsung dilakukan dengan wawancara kepada pihak kepala sekolah dan guru-guru.

2. Studi Pustaka

Studi kepustakaan adalah kegiatan untuk menghimpun informasi yang relevan dengan topik atau masalah yang menjadi objek penelitian atau topik cerita yang diusung ke dalam karya tulis non ilmiah. Studi pustaka dilakukan untuk mencari referensi dan jurnal, buku dan media lain dalam mendapatkan teori-teori tentang permasalahan yang sudah ditetapkan.

### 3. Pengumpulan Data

Pengumpulan data merupakan langkah penting suatu penelitian dan pengembangan sistem, dengan teknik pengumpulan data dapat menunjang proses penelitian yang akan diteliti, maka dari itu teknik pengumpulan data harus dilakukan dengan cermat dan benar. Pada analisis kebutuhan sistem teknik pengumpulan data berfokus pada observasi dan wawancara. Data diperoleh di bagian Akademik sehingga didapatkan data nilai rapor siswa dari SMPN 14 Kota Tasikmalaya dari periode tahun 2018/2019 berdasarkan kurikulum 2013.

### 3.3 Proses Data Mining Dengan Algoritma K-Means

#### 1. Menentukan jumlah cluster

Secara acak mendistribusikan data cluster. Pada penelitian ini dibuat 2 cluster yaitu unggulan dan reguler.

#### 2. Menentukan titik pusat

Untuk menentukan titik pusat, titik awal diinisialisasi pada awal data dan untuk titik pusat selanjutnya dilakukan dengan menggunakan rumus persamaan 3

#### 3. Hitung jarak terdekat

Untuk menghitung jarak semua data ke setiap titik pusat cluster menggunakan teori jarak Euclidean.

#### 4. Ada perubahan titik pusat

Jika pusat cluster dan cluster yang baru berubah atau berbeda maka hitung kembali menggunakan langkah yang K .

#### 5. Tidak ada perubahan titik pusat

Jika pusat cluster tidak berubah lagi maka proses clustering selesai.

### 3.4 Persiapan Data Mining

#### 1. Data Cleaning

Tahapan yang perlu dilakukan sebelum analisis data karena biasanya data mentah mengandung informasi yang tidak akurat, tidak tersusun rapi, atau tidak lengkap.

#### 2. Data Integration

Proses ini menggabungkan data dari banyak sumber data. Proses ini juga dapat membantu mengurangi redundansi data dan data yang tidak konsisten yang disebabkan pengambilan data dari banyak sumber.

#### 3. Data Selection

Proses ini dilakukan untuk menyeleksi jumlah data yang akan digunakan pada proses mining dengan tetap merepresentasikan data aslinya.

#### 4. Data Transformation

Setelah dilakukan pada proses data cleaning, data integration, dan data selection maka dilakukan proses Data Transformation. Pada mulanya, data nilai sikap berbentuk non numerik seperti terlihat pada tabel dalam kolom Data Cleaning, kemudian dilakukan inialisasi menjadi numerik untuk nantinya diolah dalam proses data mining.

### 3.5 Hasil

Setelah dilakukan proses data mining, dilanjutkan dengan pengelompokan menggunakan algoritma K-Means dengan menggunakan K yang bernilai 3 dan 5, kemudian dilanjutkan knowledge presentation yaitu proses yang dapat merepresentasikan informasi yang nantinya informasi tersebut dapat digunakan.