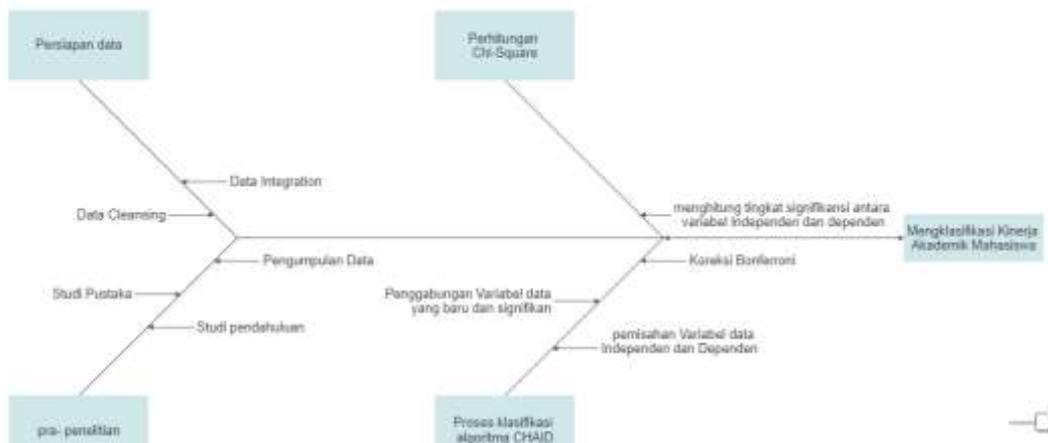


BAB III

METODOLOGI PENELITIAN

3.1 Metodologi Penelitian

Penelitian ini dilakukan dengan menggunakan metode klasifikasi dan Statistik dengan menggunakan algoritma *CHAID* (*Chi Squared Automatic Interaction Detection*). *CHAID* merupakan metode klasifikasi yang akan membagi data menjadi beberapa kelompok yang lebih kecil dari data berdasarkan hubungan variabel dependen dengan variabel independen. *CHAID* adalah sebuah metode untuk mengklasifikasi data yang memiliki tujuan untuk membagi rangkaian data menjadi kelompok-kelompok berdasarkan pada variable dependennya. Metode *CHAID* memiliki tiga tahapan dalam implementasinya yaitu penggabungan (*merging*), pemisahan (*splitting*), dan penghentian (*stopping*). Tahap pertama yaitu *merging*. pada tahap ini data akan di periksa signifikansi dari kategori variable independen terhadap variable dependen. Tahap kedua yaitu *splitting* tahap ini akan memilih variabel independen yang akan digunakan sebagai *split node* (pemisah node) yang terbaik. Pemilihan akan dikerjakan dengan membandingkan *p-value*. Tahap ketiga adalah *stopping* tahap ini dilakukan untuk menghentikan pertumbuhan pohon jika proses pertumbuhan pohon terus berlanjut dan harus sesuai dengan peraturan penghentian.



Gambar 3.1 Diagram *Fish Bone* tahapan penelitian

Penelitian ini di bagi menjadi 4 tahapan utama yaitu :

- 1 Pra penelitian dimulai dengan melakukan studi pendahuluan untuk mengidentifikasi masalah, studi pustaka dan pengumpulan data mahasiswa Universitas Siliwangi. Pada tahapan ini peneliti mencari topik penelitian yang akan di angkat dalam proses pengerjaan penelitian.
- 2 Persiapan data terdiri dari *data cleaning*, dan *data Integration* . persiapan data ini memilih data data yang bisa di gunakan dan bisa di proses dalam penelitian
- 3 Proses klasifikasi algoritma CHAID. Data yang sudah proses akan di klasifikasi menggunakan algoritma CHAID, pada proses klasifikasi ini akan menghasilkan data yang sesuai dengan variabelnya masing masing.
- 4 Perhitungan Chi-Square. Setelah data sesuai dengan variabelnya, data akan di hitung tingkat signifikansi antara variabel satu dengan variabel lainnya, ini bertujuan untuk melihat apakah variabel memiliki saling keterkaitan atau tidak.

3.1.1 Pra Penelitian

a. Studi Pustaka

Studi pustaka dilakukan untuk membuka wawasan dalam mendapatkan peluang penelitian yang akan mengangkat tema yang diteliti, kemudian studi pustaka juga dilakukan untuk mencari referensi baik itu bersumber dari jurnal, buku dan media lain dalam mendapatkan teori-teori tentang permasalahan yang sudah ditetapkan.

b. Pengumpulan Data

Dalam memperoleh data mahasiswa yang akan digunakan pada proses klasifikasi kinerja akademik mahasiswa pada penelitian ini menggunakan data premier yang didapatkan dengan menyebarkan kuesioner kepada mahasiswa universitas Siliwangi yang aktif dari angkatan 2017 sampai dengan 2020 dari semua jurusan dengan mengisi kuesioner pada link berikut <https://forms.gle/qgqgJMraXSZR0Ahh7>.

Jumlah mahasiswa unsil dari angkatan 2017 sampai dengan 2020 adalah sebanyak 12.080 mahasiswa, berdasarkan pengambilan data menggunakan metode slovin dengan tingkat kesalahan 5%. Jumlah responden yang dipilih sudah sesuai dengan standar statistik dengan jumlah sampel minimum adalah lebih besar atau sama dengan 387 mahasiswa ($N \geq 337$). (Puspaningtyas & Dewi, 2020) Berdasarkan metode slovin rumus minimum sampel yang ditetapkan :

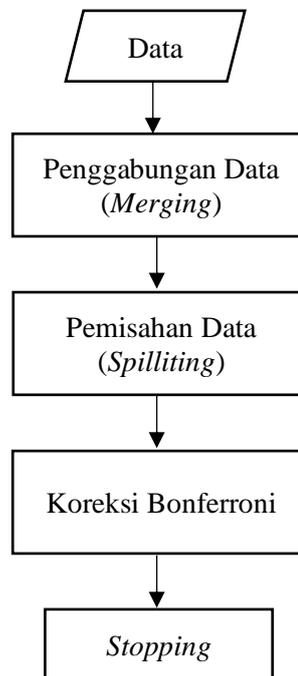
$$n = \frac{N}{1 + N e^2}$$

N = jumlah sampel

n = jumlah populasi 12.080

e = batas kesalan (*error tolerance*) = 0,05 atau 5%

3.1.2 Proses Menggunakan Algoritma CHAID



Gambar 3.2 Flowchart Algoritma CHAID

a. Penggabunga Data (*Merging*)

Tahap pertama dalam algoritma *CHAID* adalah penggabunagn atau (*merging*), pada tahap ini akan diperiksa kesignifikan dari masing masing kategori variabel bebas terhadap variable terkait.

b. Menentukan Variabel Dependen dan Independen (*Splitting*)

Tahap *splitting* memilih variable bebas yang mana akan digunakan sebagai *split mode* (pemisah mode) yang terbaik. Pemilihan dikerjakan dengan membandingkan *p-value* (dari tahap *merging*) pada setiap variabel bebas.

c. Koreksi Bonferroni

Koreksi bonferroni adalah suatu proses koreksi yang digunakan ketika beberapa uji statistik untuk dependen atau independen dan dilakukan secara bersamaan. Koreksi bonferroni biasanya digunakan dalam perbandingan berganda.

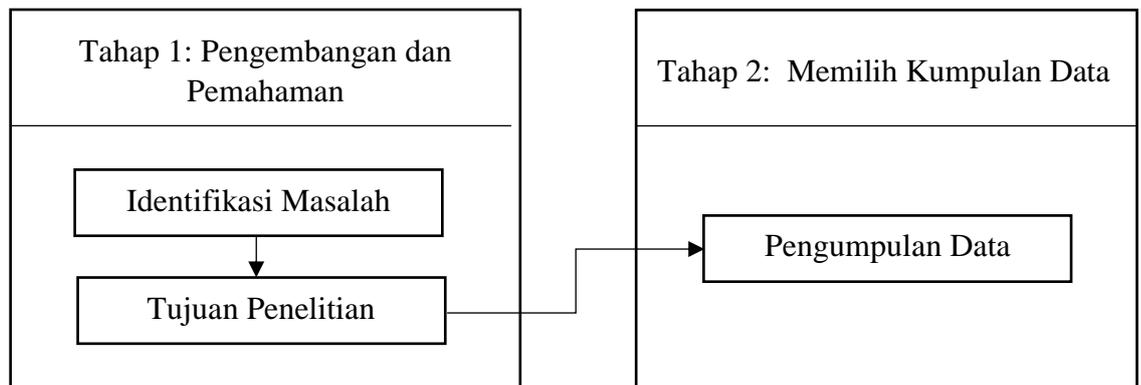
d. *Stopping*

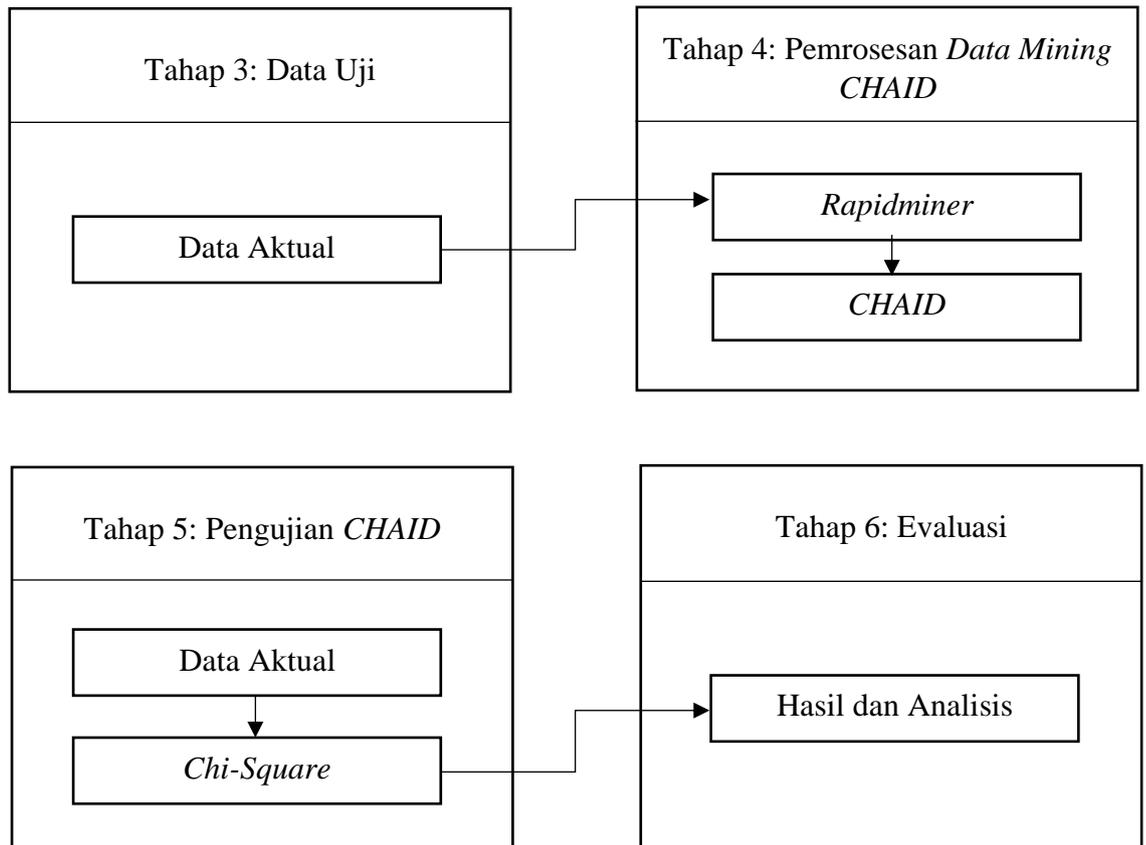
Tahap *stopping* dilakukan jika proses pertumbuhan pohon harus dihentikan jika tidak ada lagi variabel bebas yang signifikan menunjukkan perbedaan terhadap variable terkait.

3.1.3 Hasil

Setelah *dataset* diuji oleh algoritma *CHAID* kemudian *dataset* dikelompokkan dan di hitung menggunakan *Chi-Square* maka akan menghasilkan suatu informasi mengenai klasifikasi data mahasiswa dengan algoritma dan metode ini. Dan akan mendapatkan informasi berupa tingkat signifikansi dari satu variabel yang di uji ke variabel lainnya. Serta menghasilkan berupa pengelompokan data dari hasil klasifikasi.

3.2 Tahapan Penelitian





Gambar 3.3 Tahapan Penelitian

Tahapan alur penelitian ini terdiri dari beberapa proses sampai akhirnya sampai pada hasil dan analisis. Pertama yaitu dengan mengidentifikasi masalah yang akan diteliti, mulai dari informasi terkait, peluang masalah yang bisa diteliti, *dataset* yang akan digunakan, hingga algoritma yang akan digunakan untuk proses *mining*. Selanjutnya untuk menentukan tujuan dalam penelitian ini agar menjadi penelitian yang terstruktur dan terarah untuk bisa menemukan hasil penelitian menjadi hasil akhir. Tahapan selanjutnya yaitu mengumpulkan data, dengan mencari *dataset* mahasiswa Universitas Siliwangi angkatan 2017 – 2020, yang akan digunakan untuk proses penelitian. Selanjutnya yaitu mengelompokkan Data Aktual yang diperoleh dari *dataset*. Data aktual merupakan data yang akan diklasifikasi

dalam *Rapidminer*. Proses *mining* menggunakan *Rapidminer* dan Algoritma klasifikasi *CHAID*. Tahapan selanjutnya penghitungan statistik *Chi-Square*, yaitu untuk memaksimalkan signifikansi pada tiap pembagian data. Tahapan terakhir adalah kesimpulan dan analisis.

