

BAB III

METODE PENELITIAN

A. Tempat dan Waktu Penelitian

Penelitian dilakukan pada bulan Maret 2019 sampai dengan Januari 2020. Dalam penelitian tersebut meliputi beberapa kegiatan yaitu identifikasi permasalahan, pengumpulan data, pemrosesan data, penggunaan metode untuk penelitian data serta penulisan laporan. Penelitian bertempat di STMIK Amikom Purwokerto.

B. Metode Pengumpulan Data

Metode pengumpulan data adalah cara yang digunakan untuk mengumpulkan data yang diperlukan dan informasi yang dibutuhkan. Metode pengumpulan data dan informasi yang dibutuhkan adalah sebagai berikut :

1. Wawancara

Wawancara adalah pengambilan data yang dilakukan dengan cara bertatap muka atau menanyakan secara langsung dengan orang-orang yang terlibat didalam objek yang sedang diamati. Wawancara dilakukan dengan bagian kemahasiswaan STMIK Amikom Purwokerto untuk menggali informasi mengenai bagaimana proses dari penerimaan beasiswa. Sebelum melakukan wawancara peneliti terlebih dahulu mempersiapkan pertanyaan

yang akan ditanyakan pada bagian kemahasiswaan STMIK Amikom Purwokerto agar mendapatkan data dan informasi yang dibutuhkan.

2. Studi Pustaka

Studi pustaka yaitu metode yang dilakukan dengan cara menggunakan buku-buku dan literatur sebagai acuan dan referensi untuk mengumpulkan data-data dan informasi yang dibutuhkan. Selain menggunakan buku-buku, literatur juga dapat berupa jurnal, ataupun hasil penelitian terdahulu.

3. Data Primer

Menurut Amrin (2016) data primer adalah data yang diperoleh dari sumber-sumber asli. Sumber asli disini diartikan sebagai sumber pertama data tersebut diperoleh.

C. Alat dan Bahan Penelitian

1. Alat Penelitian

Alat yang digunakan pada penelitian terdiri dari perangkat keras (*hardware*) dan perangkat lunak (*software*):

a. Perangkat Keras (*Hardware*)

Perangkat keras yang digunakan dalam penelitian ini yaitu laptop dengan spesifikasi sebagai berikut:

- 1) Layar 14'' inci
- 2) Prosesor : *Intel Core i5*

3) *Memori RAM : 4,00 GB*

4) *Tipe Sistem : 64-bit Operating System*

b. *Perangkat Lunak (Software)*

Adapun perangkat lunak yang digunakan dalam penelitian yaitu:

1) *Sistem Operasi Windows 7*

Sistem operasi *windows 7* adalah sistem operasi yang berjalan pada laptop peneliti digunakan untuk mendukung penelitian.

2) *Weka*

Weka adalah sebuah paket *toolsmachine* learning praktis. Masalah-masalah dalam *data mining* khususnya tentang klasifikasi yang mendasari pendekatan *machine learning* merupakan tujuan dari *weka*. Tahap *pre-processing* data membuat pengguna berfokus pada algoritme yang digunakan tanpa memperhatikan detail lain yang ada seperti detail pembacaan data, implementasi algoritme *filtering*, dan penyediaan kode untuk evaluasi menggunakan *tools*.

Penerapan suatu sistem pembelajaran pada keseluruhan data dan menganalisis hasilnya untuk mendapatkan suatu informasi dari data atau menerapkan sistem dan melakukan perbandingan performanya adalah contoh penggunaan dari *weka*.

3) *Microsoft Office 2013*

- a. *Microsoft Word 2013* digunakan peneliti untuk membuat proposal dan laporan penelitian akhir.

b. *Microsoft Excel 2013* digunakan peneliti untuk media penginputan serta penyimpanan *dataset*, melakukan pengisian data atau nilai yang hilang (*missing value*) pada *dataset*.

c. *Microsoft Power Point 2013* digunakan peneliti untuk membuat slide presentasi yang digunakan pada saat ujian.

4) *Notepad*

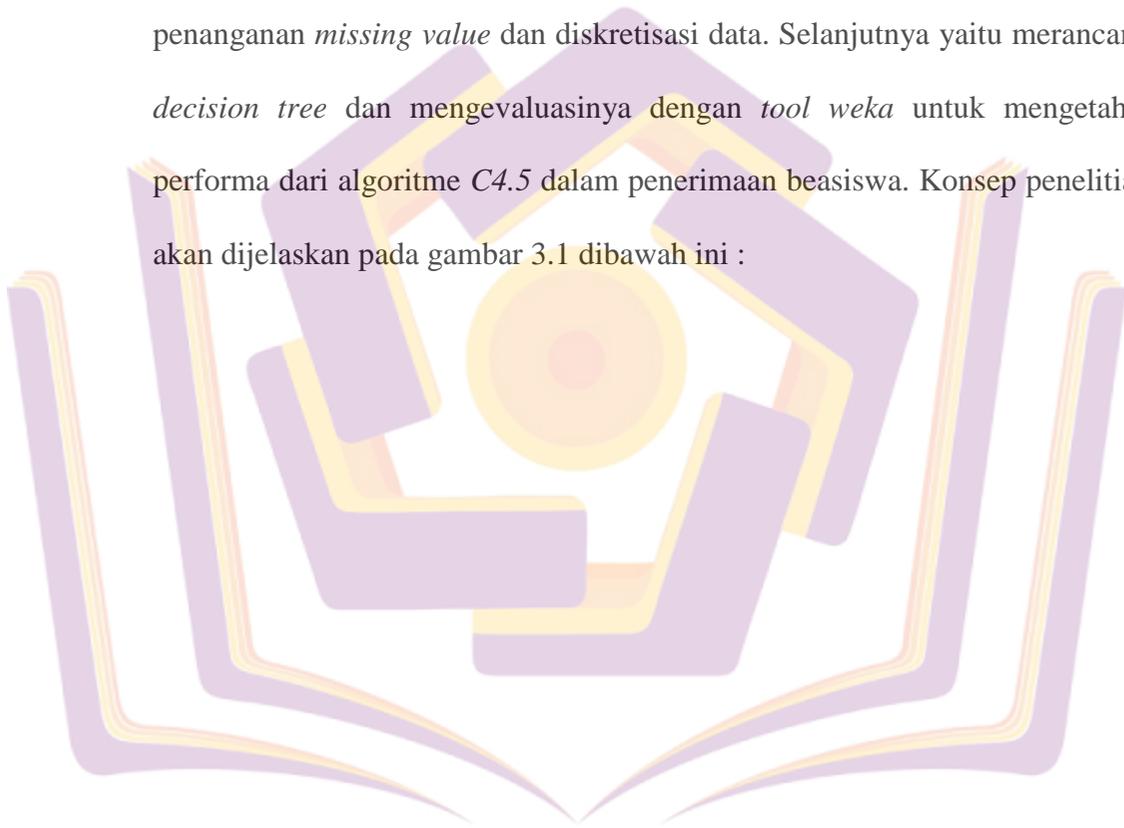
Notepad digunakan peneliti untuk membuat file dengan ekstensi *.arff* (dot *arff*) dari *dataset* yang digunakan.

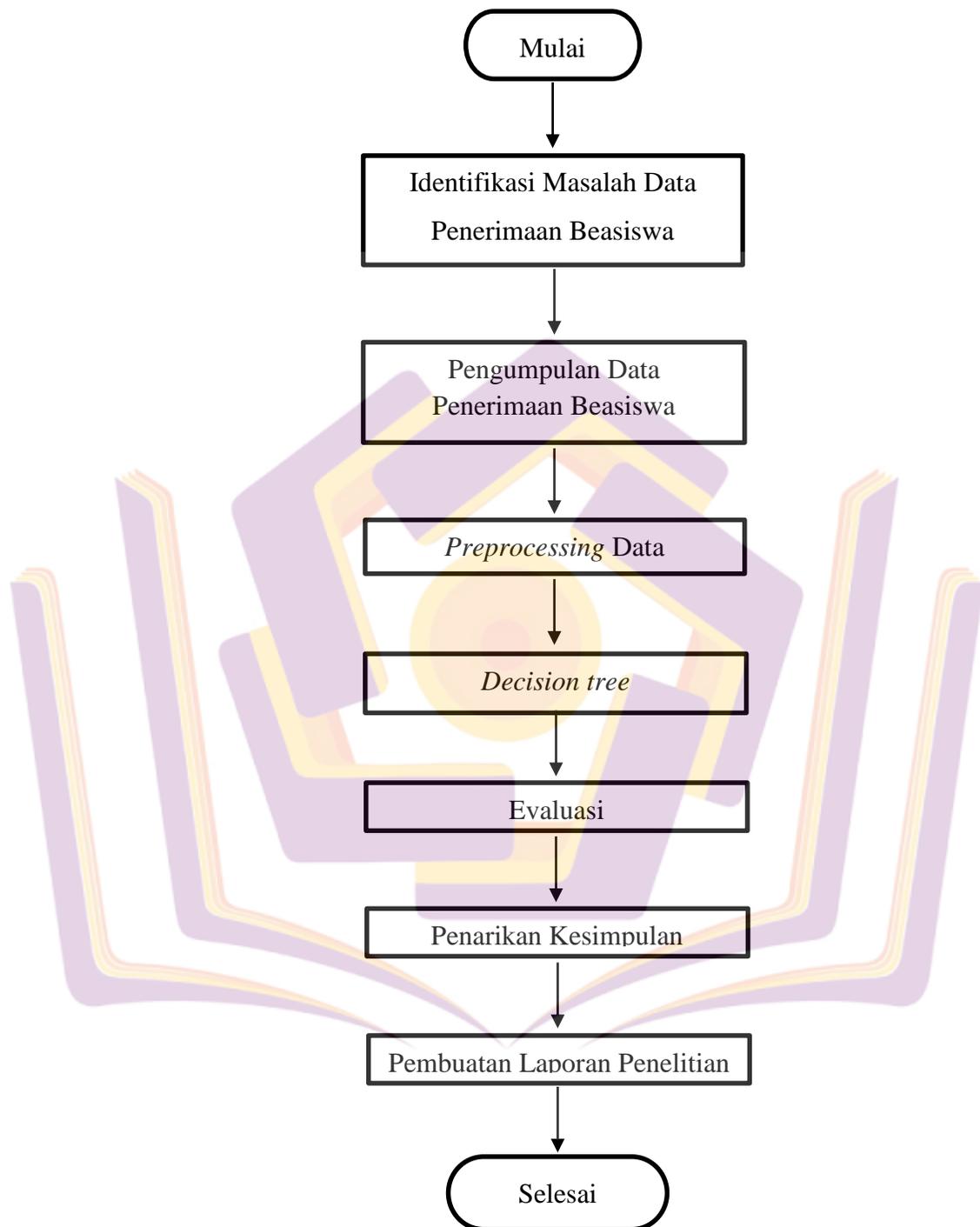
2. Bahan Penelitian

Data yang digunakan untuk penelitian sebanyak 279 *record* yang terdiri dari 120 orang penerima beasiswa dengan status diterima dan 159 orang penerima beasiswa dengan status tidak diterima. Serta terdapat 10 atribut yang akan digunakan pada *dataset* penerimaan beasiswa. Ke 10 atribut ini yang akan digunakan yaitu ada 9 atribut dan 1 atribut yang digunakan untuk menentukan status penerimaan beasiswa diterima atau tidak. Pada penelitian ini atribut yang digunakan ada 10 atribut yaitu IPK, semester, usia, jumlah anggota keluarga, prestasi, keaktifan, penghasilan, mendapat beasiswa dan kelengkapan serta satu atribut yang digunakan sebagai status untuk menentukan penerima beasiswa yang diterima atau tidak.

D. Konsep Penelitian

Konsep penelitian yang digunakan yaitu konsep penelitian berbentuk analisis bagan alur. Penelitian ini menggunakan *dataset* penerimaan beasiswa. Konsep penelitian memiliki beberapa tahapan diantaranya yaitu identifikasi masalah, pengumpulan data, *preprocessing* data dengan melakukan penanganan *missing value* dan diskretisasi data. Selanjutnya yaitu merancang *decision tree* dan mengevaluasinya dengan *tool weka* untuk mengetahui performa dari algoritme *C4.5* dalam penerimaan beasiswa. Konsep penelitian akan dijelaskan pada gambar 3.1 dibawah ini :





Gambar 3.1 Konsep Penelitian

Berdasarkan pada gambar diatas dapat dilihat bahwa penelitian yang akan dilakukan memiliki beberapa tahapan diantaranya yaitu:

1. Identifikasi Masalah

Proses identifikasi masalah yang dilakukan dalam penelitian ini adalah untuk mengetahui permasalahan serta metode yang sesuai digunakan untuk menangani masalah tersebut. Tahap identifikasi masalah dilakukukan dengan mewawancarai bagian kemahasiswaan STMIK Amikom Purwokerto.

Permasalahan yang ada pada proses penerimaan beasiswa di STMIK Amikom Purwokerto yaitu proses penyeleksian data seperti mensortir maupun memfilter data menggunakan bantuan *Microsoft Excel*, setelah data di sortir dan difilter berdasarkan kriteria tertentu, data tersebut masih harus diseleksi kembali secara manual, karena aspek atau kriteria penerimaan yang banyak. Proses penyeleksian manual harus dilakukan secara hati-hati, karena bisa saja mahasiswa memenuhi satu kriteria tapi gagal dikriteria yang lainnya. Selama ini solusi yang dilakukan yaitu membuka pendaftaran lebih dahulu sebelum ada pengumuman dari pemerintah. Sehingga diperlukan suatu metode untuk menganalisis proses penerimaan beasiswa di STMIK Amikom Purwokerto.

2. Pengumpulan Data

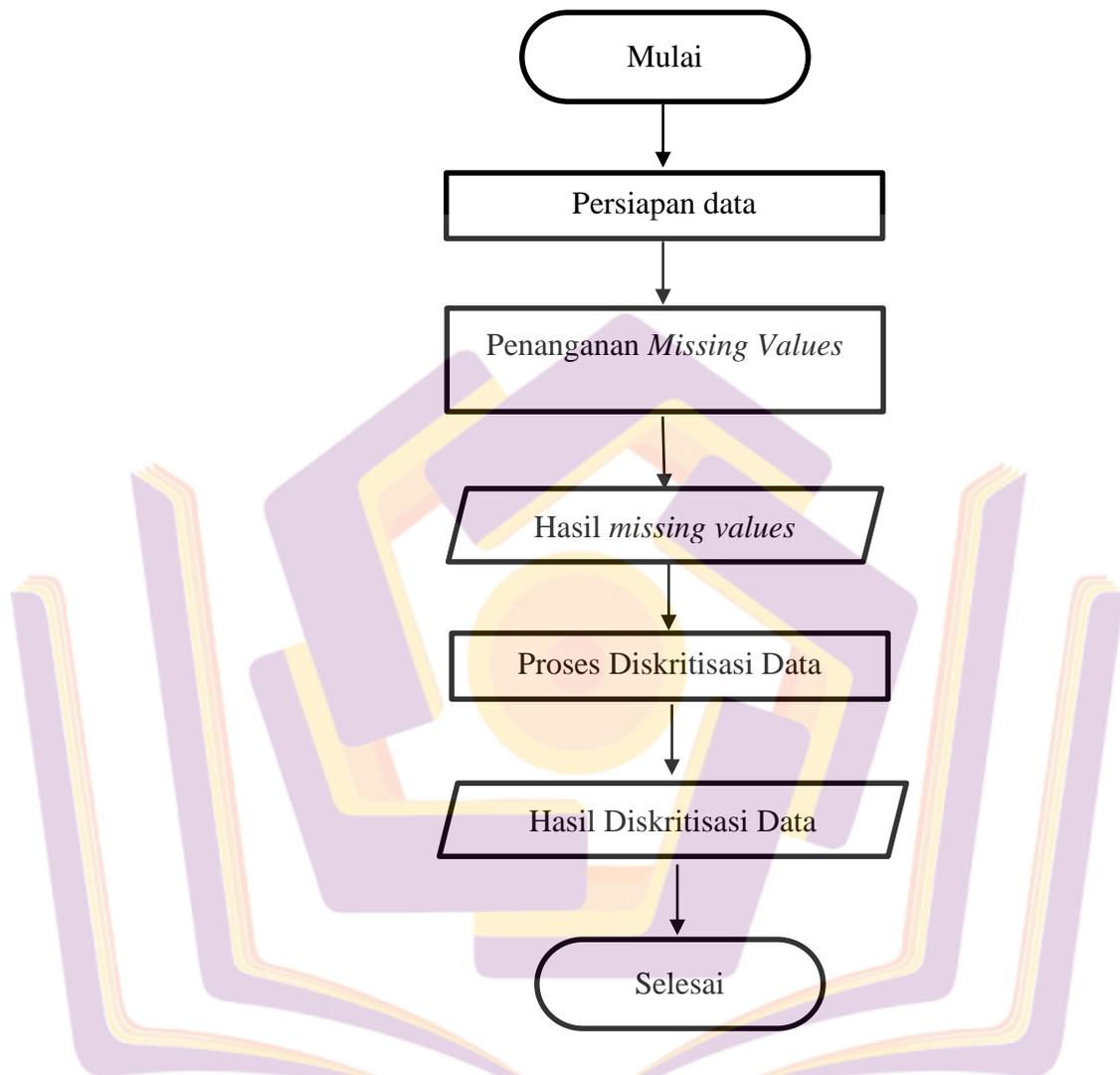
Pengumpulan data yang dilakukan oleh peneliti adalah dengan melakukan proses wawancara dengan bagian kemahasiswaan STMIK Amikom Purwokerto, studi pustaka dan data primer. Wawancara dilakukan oleh peneliti

dengan bagian kemahasiswaan untuk memperoleh informasi tentang proses penerimaan beasiswa dan penyeleksian datanya. Studi pustaka dilakukan oleh peneliti sebagai acuan dan referensi untuk mengumpulkan data-data dan informasi yang dibutuhkan. Seperti menggunakan beberapa buku, jurnal, atau hasil penelitian terdahulu untuk mengumpulkan data yang berkaitan dengan proses penerimaan beasiswa dengan beberapa algoritma yang digunakan. Data primer dari STMIK Amikom Purwokerto berupa data penerimaan beasiswa pada tahun 2015 sampai dengan 2018 yang digunakan untuk proses penelitian.

3. *Preprocessing* Data

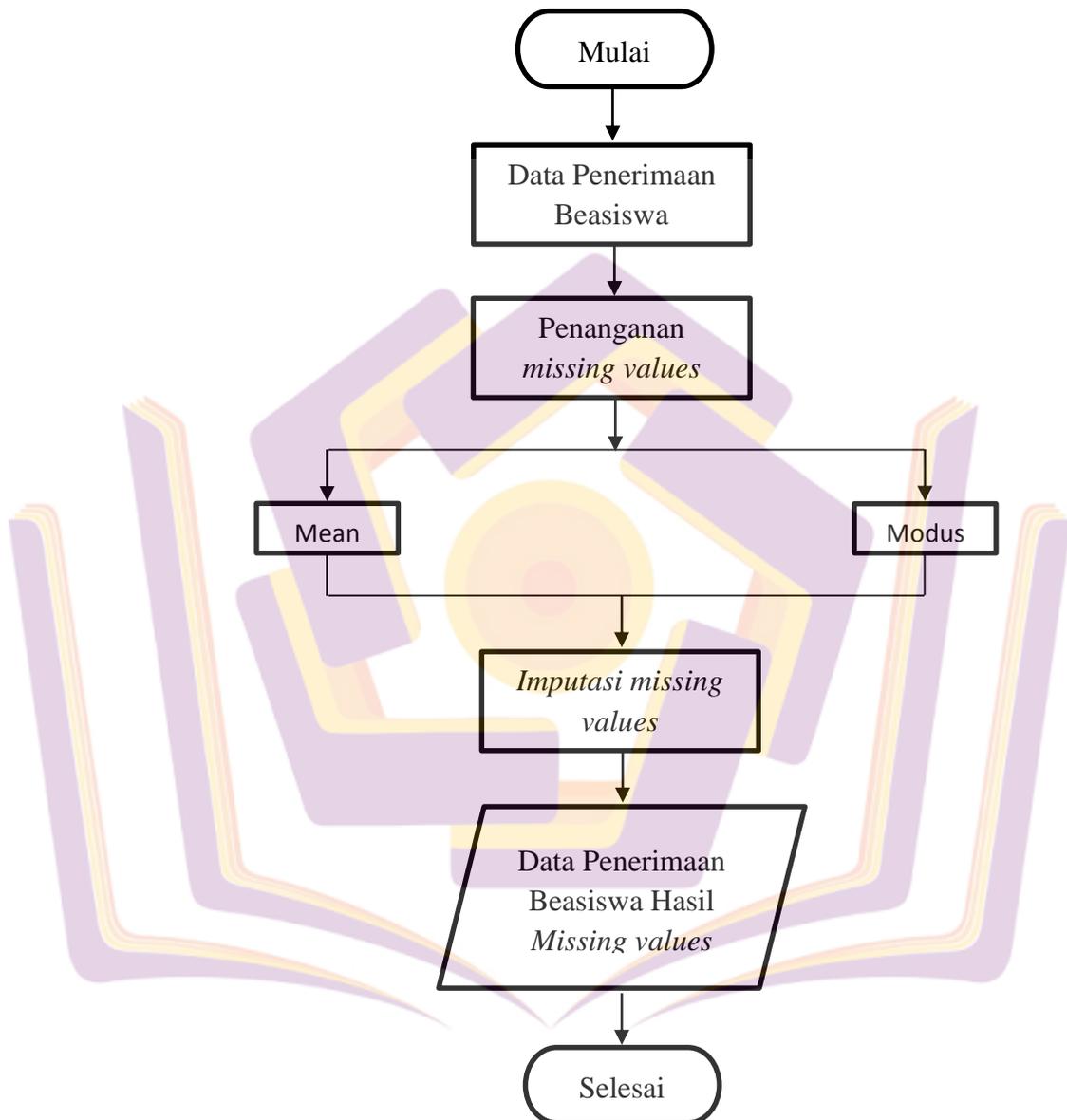
Tahapan dalam *preprocessing* data yang dilakukan peneliti yaitu melakukan persiapan data, penanganan *missing values*, proses diskritisasi data. Setelah itu, proses selanjutnya yaitu mengubah data dari bentuk *Excel* ke bentuk CSV (*Comma-Separated Values*) kemudian dikonversi menjadi *file* dengan ekstensi *.arff* (dot *arff*) dengan menggunakan *software* notepad. Kemudian data dari notepad diubah ke dalam bentuk *.arff* (dot *arff*) agar data yang digunakan bisa diolah untuk menghasilkan sebuah *rule* menggunakan algoritme *C4.5* dengan menggunakan bantuan *software Weka* untuk mengetahui tingkat performa yang diperoleh dari algoritme *C4.5*.

Berikut merupakan gambar dari proses *preprocessing* data



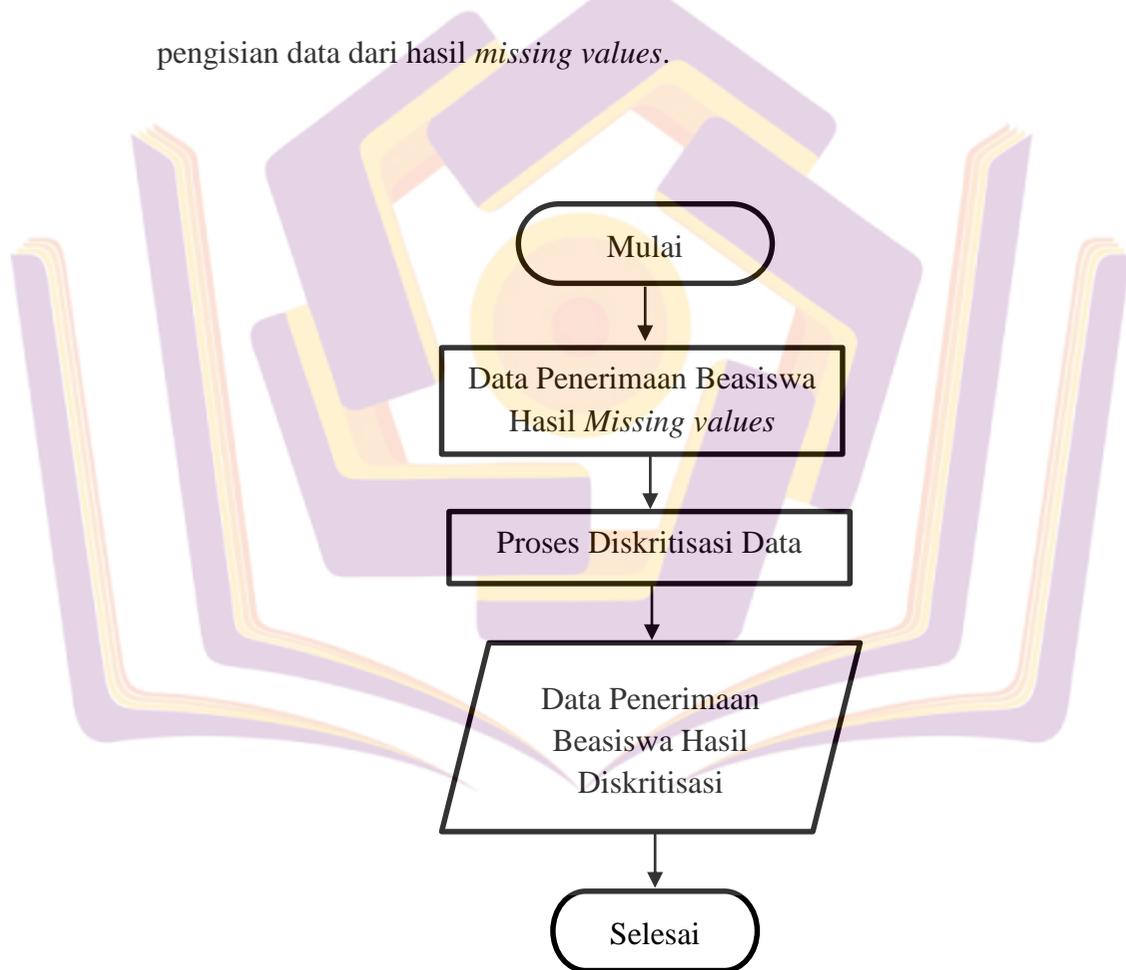
Gambar 3.2 *Preprocessing* Data

Berikut merupakan bagan pada tahap *missing values*



Gambar 3.3 Penanganan *missing values*

Berdasarkan gambar 3.3 menjelaskan tahapan dalam *preprocessing* data yaitu tahap penanganan *missing values*. Tahap penanganan *missing values* dilakukan pada data penerimaan beasiswa dimana penanganan *missing values* ini menggunakan teknik mean dan modus. Mean digunakan untuk data yang bersifat numerik sementara modus digunakan untuk data yang bersifat kategorikal. Langkah selanjutnya yaitu proses *imputasi* data yaitu proses pengisian data dari hasil *missing values*.



Gambar 3.4 Diskritisasi data

Berdasarkan gambar 3.4 tahapan dalam *preprocessing* data yaitu tahap diskritisasi data. Setelah melakukan tahapan penanganan missing values seperti gambar 3.3 maka selanjutnya dilakukan proses diskritisasi. Proses diskritisasi dilakukan setelah data sudah mengalami proses imputasi *missing values*, proses ini dilakukan untuk mengubah data bertipe numerik menjadi data bertipe string, perubahan dilakukan untuk menyesuaikan kategori yang sudah ditentukan sehingga akan menghasilkan *dataset* penerimaan beasiswa yang sudah mengalami tahap *preprocessing* data dari proses diskritisasi.

4. *Decision tree*

Pohon Keputusan (*Decision tree*) dilakukan secara otomatis dengan menggunakan *software weka*. Pada tahap ini, dilakukan proses perhitungan nilai *entropy*, *gain*, *split info* dan *gain ratio* dari masing-masing atribut untuk membentuk suatu pohon keputusan. Proses perhitungan dimulai dengan mencari nilai *entropy* terlebih dahulu, kemudian mencari nilai *gain* dan *split info*. Selanjutnya dicari nilai *gain ratio*, atribut dengan nilai *gain ratio* tertinggi akan menjadi *node* akar atau akar pertama dari suatu pohon keputusan.

5. Evaluasi

Tahapan evaluasi nilai performa didapatkan dari beberapa proses yang telah dilakukan yaitu *preprocessing* data dengan menggunakan *missing values* melalui teknik mean dan modus, dimana nilai mean digunakan untuk data numerik sementara nilai modus digunakan untuk data kategorikal. Setelah hasil

dari proses penanganan *missing values* didapatkan selanjutnya *imputasi dataset* yaitu proses pengisian data atau pengganti nilai-nilai yang hilang (*missing*), setelah didapatkan data dari hasil *imputasi missing values* selanjutnya dilakukan proses diskritisasi data agar data dapat dikategorikan berdasarkan kategori yang sudah ditentukan, dari proses tersebut maka akan diperoleh hasil *decision tree* dari algoritme *C4.5* dengan menggunakan parameter pengukuran yaitu akurasi. Pada tahap evaluasi terdapat tiga pengujian akurasi yaitu pengujian dengan *imputasi missing values*, pengujian tanpa *imputasi missing values* dan pengujian dengan menghilangkan atribut yang tidak berpengaruh.

6. Penarikan Kesimpulan

Penarikan kesimpulan menghasilkan kesimpulan berupa nilai performa dari algoritme *C4.5* yang menghasilkan suatu pohon keputusan

7. Pembuatan Laporan Penelitian

Pembuatan laporan digunakan untuk melaporkan seluruh kegiatan penelitian yang dilakukan saat ini.