

## BAB III METODE PENELITIAN

### 3.1 Bahan Penelitian

Bahan penelitian yang akan diteliti adalah data presensi karyawan minimarket Alfamart Hos Cokroaminoto yang berjumlah 4 karyawan dan diambil data presensi 5 bulan setiap karyawannya dengan setiap bulan terdapat rata-rata 25 hari kerja yang diatur jam kerjanya dengan *shift* sehingga berjumlah 405 data. Data diambil 5 bulan dikarenakan data hanya dapat diakses 5 bulan terakhir. Data berupa hasil cetakan presensi karyawan perbulan yang kemudian diketik ulang kedalam *softfile* berformat (xlsx).

### 3.2 Peralatan Penelitian

Peralatan yang digunakan untuk penelitian ini yaitu :

1. *Hardware* : Laptop HP Notebook - 14-am015tx dengan spesifikasi Intel® Core™ i5-6200U (2.3 GHz, up to 2.8 GHz, 3 MB cache, 2 cores), 4 GB DDR4-2133 SDRAM, AMD Radeon™ R5 M430 Graphics, 500 GB 5400 rpm SATA
2. *Software* : 64-bit Windows 10 Operating System, Rapidminer, Python 3.6.8

### 3.3 Lokasi dan Jadwal Penelitian

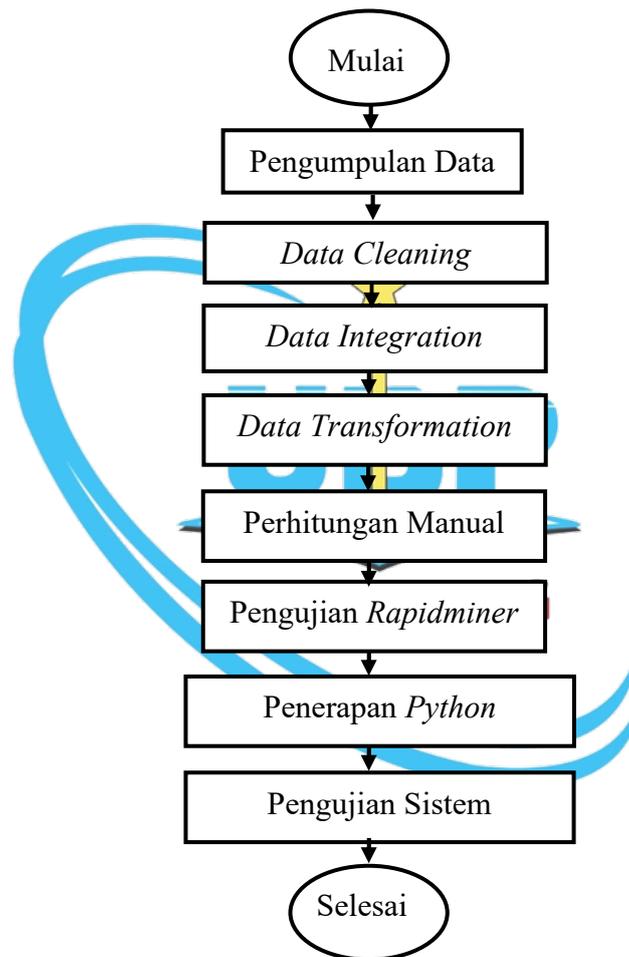
Lokasi penelitian dilakukan di gerai Alfamart. Rincian penelitian ditunjukkan pada Tabel 3.1.

Tabel 3.1 Jadwal Penelitian

No.	Kegiatan	Bulan					
		1	2	3	4	5	6
1.	Studi literatur						
2.	Analisis kebutuhan						
3.	Perancangan sistem						
4.	Implementasi						
5.	Pengujian						

### 3.4 Prosedur Penelitian

Penelitian ini dimulai dengan mengumpulkan data presensi karyawan Alfamart dan menganalisis data tersebut karena tidak semua atribut digunakan dalam penelitian ini. Kemudian data tersebut diolah menggunakan algoritma C4.5 sehingga menghasilkan hasil yang diinginkan yang selanjutnya bisa diterapkan kedalam sistem. Sistem akan menghasilkan beberapa kategori tingkat kedisiplinan karyawan yaitu baik, cukup dan buruk.



Gambar 3.1 *Flowchart* Prosedur Penelitian

#### 3.4.1 Pengumpulan Data

Pengumpulan data dilakukan di minimarket Alfamart Hos Cokroaminoto. Data yang diteliti merupakan data presensi karyawan dengan jumlah 4 orang. Masing-masing karyawan diambil data presensi selama 5 bulan dari Agustus 2019 sampai dengan Desember 2019. Setiap bulan terdapat sekitar 25 hari kerja sehingga total

terdapat 405 data. Data yang digunakan untuk penelitian adalah data presensi karyawan perharinya. Beberapa atribut yang tercantum diantaranya periode, tanggal, jam masuk, jam pulang, total jam dan total hari. Ketentuan jam masuk *shift* pagi adalah jam 07.00 dan *shift* siang adalah jam 15.00.

### 3.4.2 Data Cleaning

Tahapan ini melakukan *cleaning dataset* dengan mencari hubungan atribut-atribut yang dibutuhkan. *Data cleaning* dilakukan untuk mengolah atribut yang bisa dijadikan perhitungan dan benar-benar relevan.

### 3.4.3 Data Integration

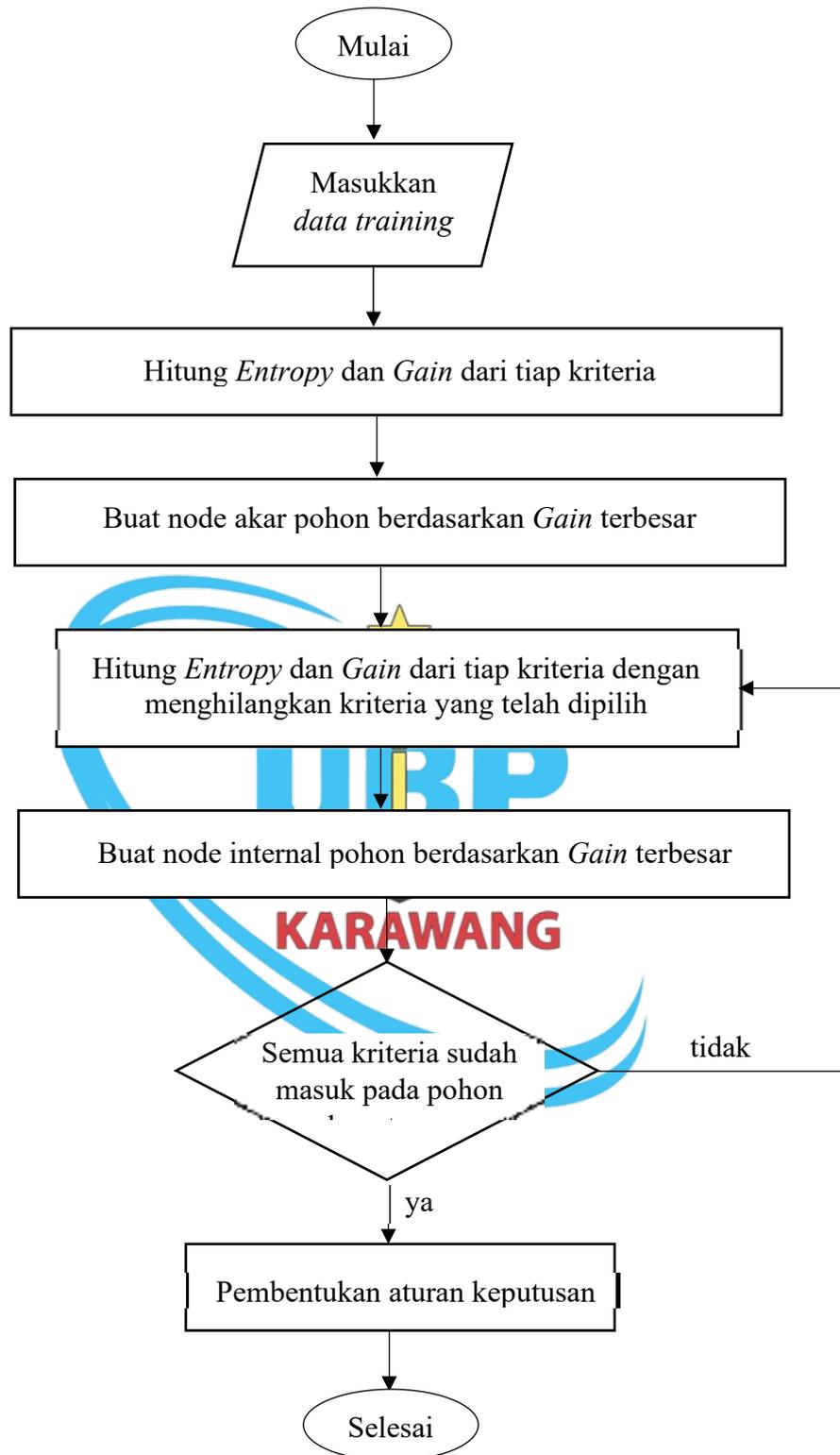
*Dataset* dari karyawan serta bulan yang berbeda tentunya masih terpisah. Pada tahap ini setelah pembersihan data dilakukan penggabungan data dari sumber yang didapat. Data diolah menjadi tabel yang nantinya akan diproses pada tahap selanjutnya.

### 3.4.4 Data Transformation

Tahapan ini dilakukan perubahan bentuk ke *dataset* agar sesuai untuk dilanjutkan ke proses *data mining*. Data numerik harus dirubah menjadi kategorikal dengan cara membagi menjadi beberapa interval.

### 3.4.5 Perhitungan Manual

Perhitungan algoritma C4.5 dimulai dengan memasukkan *data training*, kemudian menghitung *Entropy* dan *Gain* dari masing-masing kriteria data training yang ada untuk menentukan kriteria mana yang akan menjadi akar pada pohon keputusan. Kemudian membuat *node* akar dari pemilihan kriteria dengan memilih *Gain* terbesar. Lalu, menghitung kembali *Entropy* dan *Gain* dari masing-masing kriteria dengan menghilangkan kriteria yang telah dipilih sebelumnya. Lalu cek pada semua kriteria apakah kategori pada kriteria yang memiliki nilai *Gain* tertinggi sudah masuk pada kelas yang sama, jika belum maka ulangi dengan menghitung *Entropy* dan *Gain* kriteria yang belum masuk kedalam pohon keputusan.



Gambar 3.2 *Flowchart* Perhitungan Algoritma C4.5

### 3.4.6 Pengujian Rapidminer

Pengujian dilakukan menggunakan hasil implementasi dari algoritma C4.5 menggunakan *tools Rapidminer*. Pengujian ini bertujuan untuk mengetahui pohon keputusan yang optimal serta mencocokkan kesesuaian hasil dari perhitungan implementasi yang dilakukan.

### 3.4.7 Penerapan Python

Penerapan algoritma pada sistem menggunakan bahasa pemrograman *Python* dengan *web open source Jupyter Notebook*. Algoritma C4.5 pada *Python* akan menggunakan *library* agar lebih mudah lalu selanjutnya membuat program sederhana untuk mengklasifikasikan tingkat kedisiplinan presensi karyawan.

### 3.4.8 Pengujian Sistem

Pengujian dilakukan untuk mengetahui hasil perhitungan manual dan penerapan *python* apakah memiliki kecocokan dan menghasilkan kesesuaian dengan menerapkan algoritma C4.5.

