

## BAB III

### METODE PENELITIAN

#### 3.1 Objek Penelitian

Penelitian ini dilaksanakan di Laboratorium Riset Universitas Buana Perjuangan Karawang. Data yang dipakai dalam penelitian ini diambil dari Pusat Data dan Informasi (PUSDATIN) program studi Teknik informatika dari tahun 2016-2020.

#### 3.2 Waktu Penelitian

Penelitian dilakukan dari bulan Juli hingga November tahun 2023. Tabel 3.1 adalah Tabel rincian penelitian.

Tabel 3. 1 Tabel Rincian Penelitian

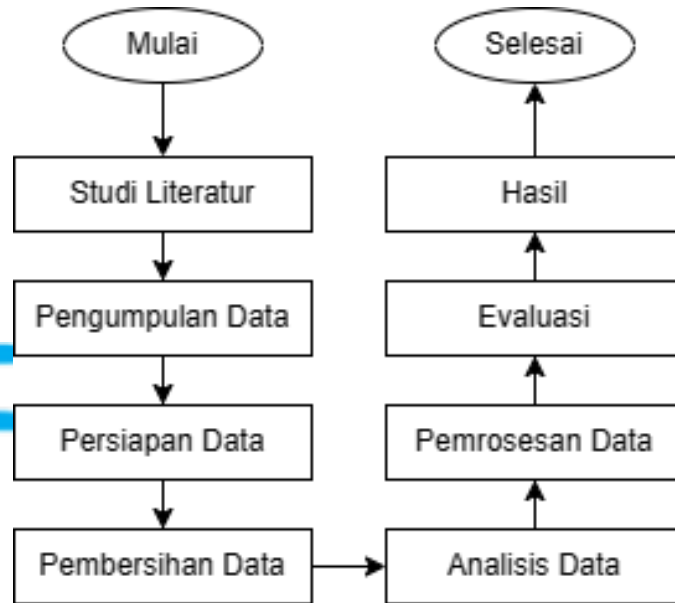
| No | Uraian           | Juli |   |   |   | Agustus |   |   |   | September |   |   |   | Oktober |   |   |
|----|------------------|------|---|---|---|---------|---|---|---|-----------|---|---|---|---------|---|---|
|    |                  | 2023 |   |   |   | 2023    |   |   |   | 2023      |   |   |   | 2023    |   |   |
|    |                  | 1    | 2 | 3 | 4 | 1       | 2 | 3 | 4 | 1         | 2 | 3 | 4 | 1       | 2 | 3 |
| 1  | Studi Literatur  |      |   |   |   |         |   |   |   |           |   |   |   |         |   |   |
| 2  | Pengumpulan Data |      |   |   |   |         |   |   |   |           |   |   |   |         |   |   |
| 3  | Persiapan Data   |      |   |   |   |         |   |   |   |           |   |   |   |         |   |   |
| 4  | Pembersihan Data |      |   |   |   |         |   |   |   |           |   |   |   |         |   |   |
| 5  | Analisis Data    |      |   |   |   |         |   |   |   |           |   |   |   |         |   |   |
| 6  | Pemrosesan Data  |      |   |   |   |         |   |   |   |           |   |   |   |         |   |   |
| 7  | Evaluasi         |      |   |   |   |         |   |   |   |           |   |   |   |         |   |   |
| 8  | Hasil            |      |   |   |   |         |   |   |   |           |   |   |   |         |   |   |

#### 3.3 Peralatan Penelitian

Laptop Asus Vivobook F412DA, Ryzen 3 3250, RAM 8GB, 256ssd Vega3 W10 14.0FHD.

#### 3.4 Prosedur Penelitian

Prosedur penelitian yang akan dijalankan untuk tugas akhir dapat disusun sebagai berikut:



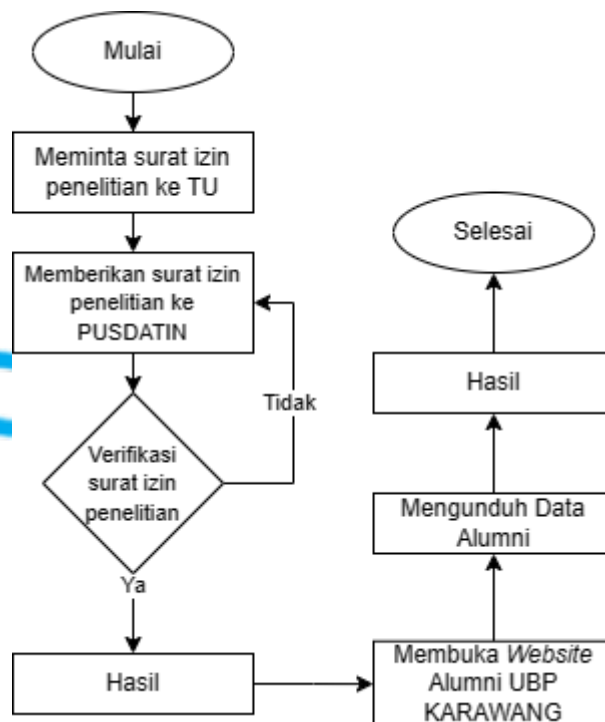
Gambar 3. 1 Alur Prosedur Penelitian

### 3.5 Studi Literatur

Studi literatur dilakukan dengan cara mencari sumber penelitian yang relevan dalam bentuk artikel jurnal. Pendekatan ini bertujuan untuk memahami kerangka kerja konseptual yang sudah ada, mengevaluasi metode-metode yang telah digunakan oleh peneliti terdahulu, serta mengidentifikasi temuan-temuan penting yang dapat mendukung penelitian yang sedang dilakukan.

### 3.6 Pengumpulan Data

Pengumpulan data dilakukan dengan cara mengajukan surat izin penelitian kepada Tata Usaha Fakultas Ilmu Komputer kemudian surat izin diserahkan ke bagian PUSDATIN dan mengunduh data kelulusan di *website* Alumni UBP Karawang yang akan diuraikan pada Gambar 3.2.



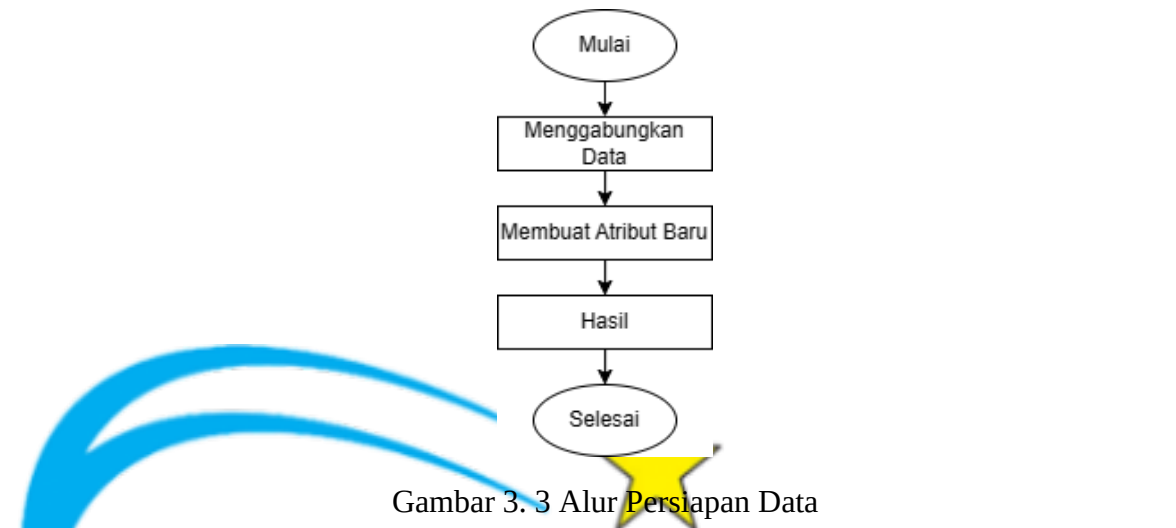
Gambar 3. 2 Alur Pengumpulan Data

Meminta surat izin penelitian ke staf Tata Usaha FIK. Data yang dibutuhkan merupakan data mahasiswa Teknik Informatika Angkatan 2016-2020. Atribut yang terdapat pada data tersebut, yaitu: NIM, jenis kelamin, status pekerjaan, IP Semester 1-8, IPK Semester 1-8 dalam format csv. Setelah mendapatkan surat izin penelitian lalu diserahkan ke bagian PUSDATIN, lalu surat dicek apakah sudah sesuai dengan peraturan dan izin meminta data penelitian. Apabila belum sesuai, mahasiswa diminta membuat surat lagi ke TU, dan apabila sudah sesuai, data akan dikirim melalui email. File akan dikirim setelah dua minggu atau lebih setelah verifikasi surat.

Mengunduh seluruh data lulusan mahasiswa UBP Karawang dari tahun 2016-2020 melalui *website* alumni UBP Karawang (<http://alumni.ubpkarawang.ac.id/>) yang telah disediakan.

### 3.7 Persiapan Data

Persiapan data yang akan dilakukan menggunakan Teknik penggabungan seluruh data menjadi satu data. Berikut adalah gambar 3.3 tentang alur persiapan data:



Gambar 3. 3 Alur Persiapan Data

Pertama, dalam persiapan data adalah Menggabungkan seluruh data yang telah didapatkan. Kedua, membuat atribut baru pada data yang telah digabungkan. Atribut baru tersebut digunakan untuk mengkalkulasi pola kelulusan mahasiswa tepat waktu. Gambar 3.4 adalah hasil dari penggabungan data dan pembuatan atribut baru.

|                        | nim            | angkatan      | jenis_kelamin | status_bekerja | tgl_kelulusan | tahun_kelulusan | status_kelulusan | ipk_semester1 | ipk_semester1 | ipk_semester2 | ... |
|------------------------|----------------|---------------|---------------|----------------|---------------|-----------------|------------------|---------------|---------------|---------------|-----|
| 0                      | 16416255201001 | 2016          | L             | 0.0            | 2021-08-31    | 2021.0          | 0.0              | 3.044445      | 3.044445      | 3.142500      | ... |
| 1                      | 16416255201002 | 2016          | L             | 0.0            | 2020-09-17    | 2020.0          | 1.0              | 3.502778      | 3.502778      | 3.805555      | ... |
| 2                      | 16416255201003 | 2016          | L             | 1.0            | 2023-08-31    | 2023.0          | 0.0              | 3.362778      | 3.362778      | 3.468044      | ... |
| 3                      | 16416255201004 | 2016          | L             | 0.0            | 2021-08-20    | 2021.0          | 0.0              | 3.030556      | 3.030556      | 3.245278      | ... |
| 4                      | 16416255201005 | 2016          | L             | 1.0            | NaT           | NaN             | NaN              | 3.238111      | 3.238111      | 2.362500      | ... |
| ...                    | ...            | ...           | ...           | ...            | ...           | ...             | ...              | ...           | ...           | ...           | ... |
| 1042                   | 20416255201219 | 2020          | L             | 0.0            | NaT           | NaN             | NaN              | 3.292632      | 3.292632      | 2.833500      | ... |
| 1043                   | 20416255201220 | 2020          | L             | 0.0            | NaT           | NaN             | NaN              | 3.435790      | 3.435790      | 3.457000      | ... |
| 1044                   | 20416255201222 | 2020          | P             | 1.0            | NaT           | NaN             | NaN              | 0.500000      | 0.500000      | 0.844500      | ... |
| 1045                   | 20416255201223 | 2020          | L             | 0.0            | NaT           | NaN             | NaN              | 1.315789      | 1.315789      | 1.563250      | ... |
| 1046                   | 20416255201224 | 2020          | L             | 1.0            | NaT           | NaN             | NaN              | NaN           | NaN           | 3.288095      | ... |
| 1047 rows x 23 columns |                |               |               |                |               |                 |                  |               |               |               |     |
|                        | ipk_semester4  | ipk_semester4 | ipk_semester5 | ipk_semester5  | ipk_semester6 | ipk_semester6   | ipk_semester7    | ipk_semester7 | ipk_semester8 | ipk_semester8 |     |
|                        | 3.289000       | 3.321429      | 3.298300      | 3.337500       | 3.301157      | 3.324288        | 3.320211         | 3.430000      | 3.1856081.0   | 0.0           |     |
|                        | 3.644750       | 3.680000      | 3.620800      | 3.525000       | 3.612479      | 3.572857        | 3.627747         | 3.715714      | 3.6428378.0   | 4.0           |     |
|                        | 3.248375       | 3.037143      | 3.199300      | 3.003000       | 3.214545      | 3.287143        | 3.273028         | 3.158750      | NaN           | NaN           |     |
|                        | 3.248625       | 3.322857      | 3.249200      | 3.251500       | 3.256033      | 3.288571        | 3.313239         | 3.642857      | 3.1789188.0   | 0.0           |     |
|                        | NaN            | NaN           | NaN           | NaN            | NaN           | NaN             | NaN              | NaN           | NaN           | NaN           |     |
| ...                    | ...            | ...           | ...           | ...            | ...           | ...             | ...              | ...           | ...           | ...           |     |
|                        | 2.309390       | 2.423913      | 2.335098      | 1.614000       | 2.143740      | 1.062500        | NaN              | NaN           | NaN           | NaN           |     |
|                        | 3.088537       | 2.430000      | 3.163235      | 3.287500       | 3.166585      | 3.182857        | NaN              | NaN           | NaN           | NaN           |     |
|                        | NaN            | NaN           | NaN           | NaN            | NaN           | NaN             | NaN              | NaN           | NaN           | NaN           |     |
|                        | NaN            | NaN           | NaN           | NaN            | NaN           | NaN             | NaN              | NaN           | NaN           | NaN           |     |
|                        | 2.398939       | 1.178572      | 2.649188      | 3.021739       | 2.479720      | 1.562500        | NaN              | NaN           | NaN           | NaN           |     |

Gambar 3. 4 Dataset Hasil Persiapan Data

### 3.8 Pembersihan Data

Pembersihan data dilakukan dengan teknik menghapus atribut dan mengubah nilai yang ada didalam atribut lain. Gambar 3.6 merupakan alur pembersihan



Gambar 3. 5 Alur Pembersihan Data

Langkah pertama adalah mengubah nilai pada kolom jenis kelamin dari alfabetikal menjadi numerikal. Langkah kedua adalah mengubah nilai N/A dan huruf menjadi bernilai 1 atau 0 dengan tujuan mengubah data yang tidak terstruktur menjadi terstruktur.

### 3.9 Analisis Data

Analisis data dilakukan dengan cara menghapus kolom dan baris pada *dataset*. Gambar 3.6 merupakan proses alur analisis data.



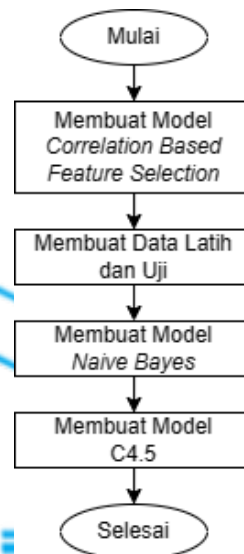
Gambar 3. 6 Alur Analisis Data

Langkah pertama adalah menentukan kolom yang tidak akan dipakai dalam perhitungan kemudian menghapus kolom tersebut. Langkah kedua yaitu menghapus baris yang memiliki nilai 0.

### 3.10 Pemrosesan Data

Pemrosesan data dilakukan dengan cara membuat model *Correlation Feature Selection* (CFS), lalu membuat data latih dan uji dari hasil CFS. Membuat

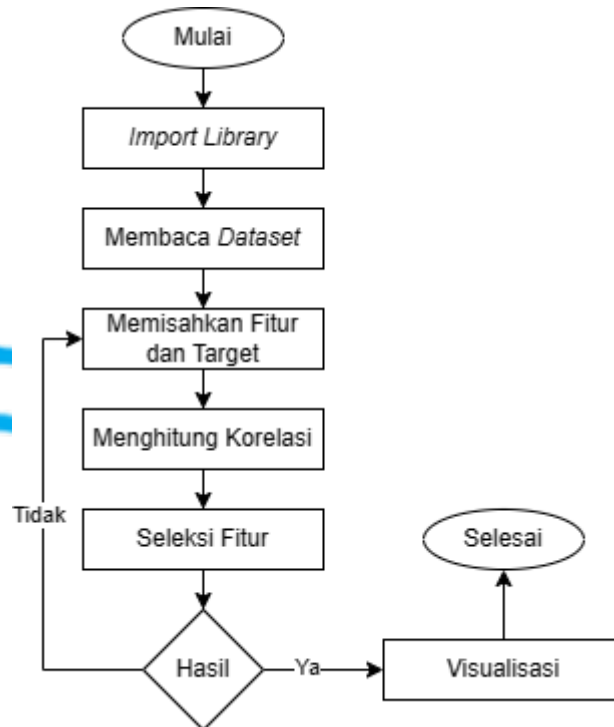
model algoritma Naïve Bayes, serta mengembangkan model algoritma C4.5. Alur pemrosesan data dapat dilihat pada Gambar 3.7.



Gambar 3. 7. Alur Pemrosesan Data

#### 3.10.1 Membuat Model *Correlation Based Feature Selection* (CFS)

Membuat model pemilihan fitur menggunakan teknik *Correlation Based Feature Selection* (CFS). Adapun langkah-langkah dalam pembuatan model CFS dapat dilihat pada Gambar 3.8



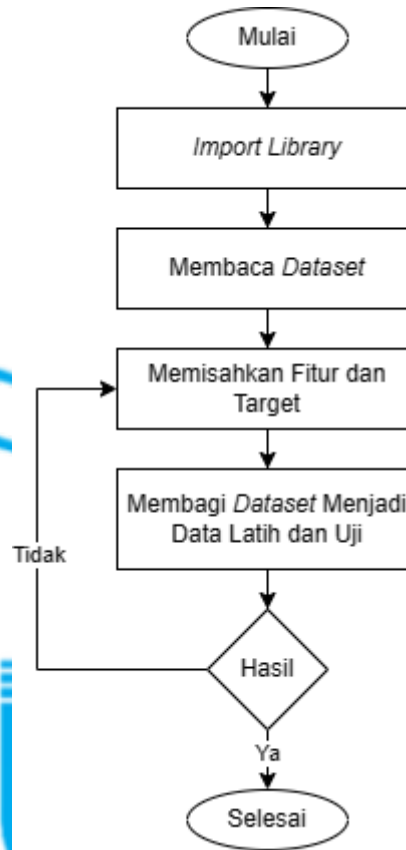
Gambar 3. 8 Alur Pemodelan *Correlation Based Feature Selection*

Langkah pertama dalam pembuatan model pemilihan fitur dengan menggunakan *Correlation Based Feature Selection (CFS)* adalah mengimpor *library* yang dibutuhkan. Langkah kedua yaitu membaca *dataset* yang telah dianalisis. Langkah ketiga adalah memisahkan *dataset* menjadi fitur dan target. Langkah keempat adalah menghitung korelasi antara fitur dan target. Langkah kelima yaitu memvisualisasi hasil yang didapatkan dari proses seleksi fitur menjadi matriks konfusi berbentuk *heatmap*.

### 3.10.2 Membuat Data Latih Dan Uji

Pembuatan pembagian data latih dan uji dilakukan dengan beberapa langkah yang meliputi membaca *dataset* CFS yang telah dibuat sebelumnya, memisahkan fitur serta memisahkan *dataset* menjadi data latih dan data uji. Alur pembuatan data latih dan uji dapat dilihat pada Gambar 3.9.





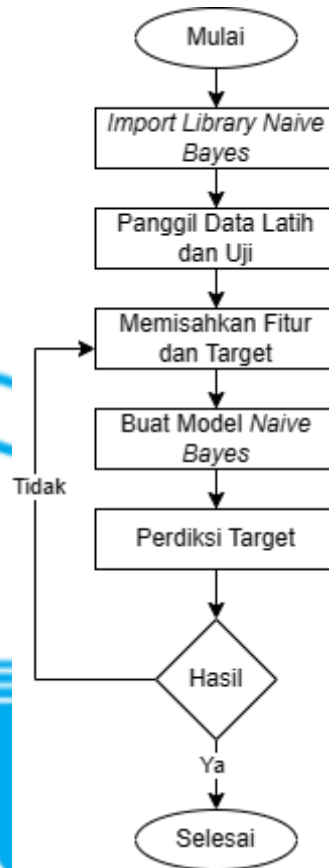
Gambar 3. 9 Alur Pembuatan Data Latih dan Uji

Langkah pertama yaitu mengimpor *library* yang dibutuhkan. Langkah kedua yaitu membaca *dataset* yang telah melewati tahapan *Correlation Based Feature Selection*. Langkah ketiga adalah memisahkan fitur dan target. Langkah keempat yaitu memisahkan *dataset* menjadi data latih dan data uji.

### 3.10.3 Pemodelan Algoritma *Naïve Bayes*

Pemodelan algoritma *Naïve Bayes* yang akan dilakukan dapat dilihat pada Gambar 3.10.



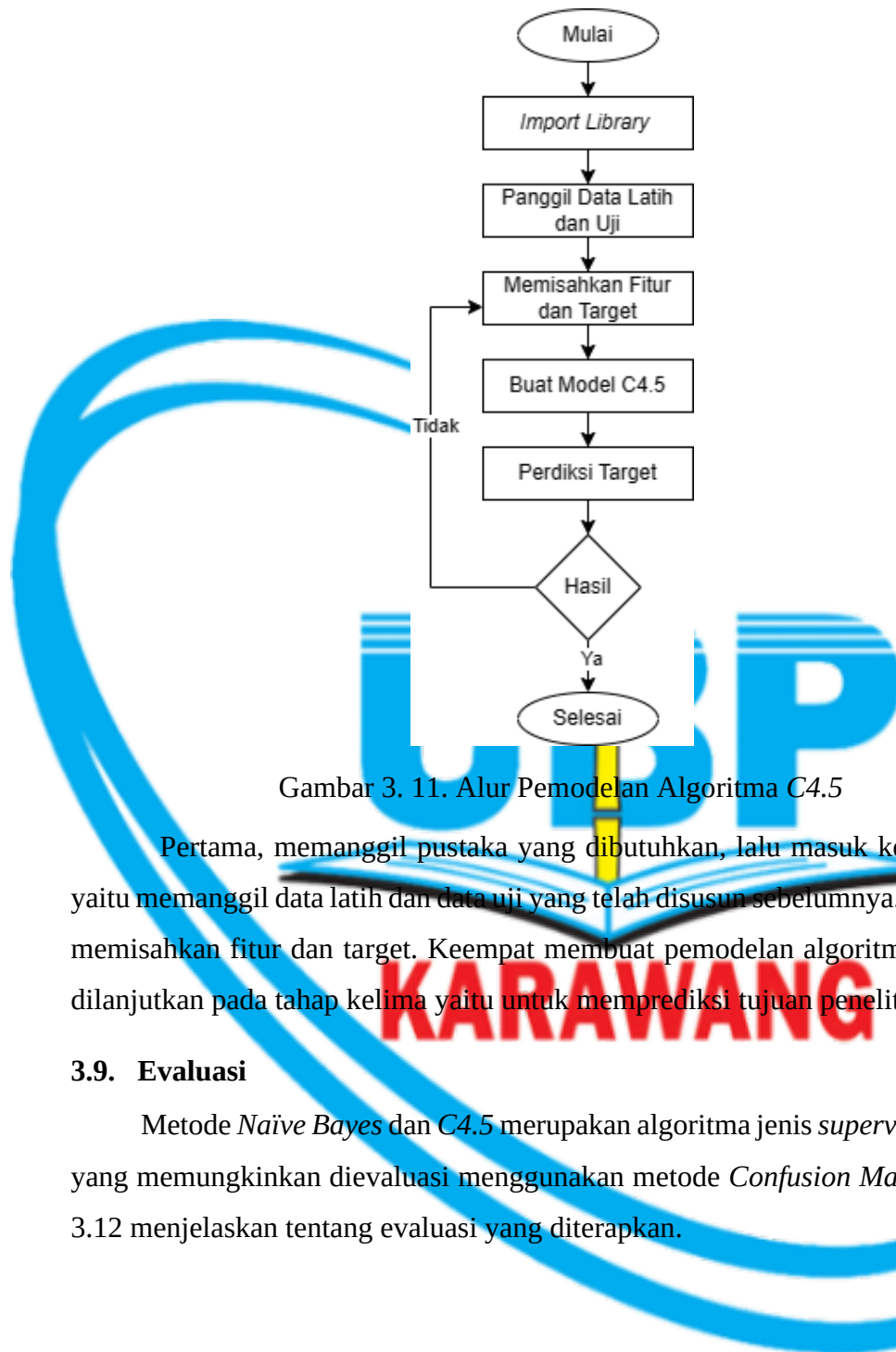


Gambar 3. 10. Alur Pemodelan Algoritma *Naïve Bayes*.

Pertama, impor *library* yang dibutuhkan lalu masuk ketahap kedua yaitu memanggil data latih dan data uji yang telah disiapkan sebelumnya. Ketiga yaitu memisahkan fitur dan target. Keempat membuat pemodelan algoritma *Naïve Bayes* yang dilanjutkan pada kelima untuk memprediksi tujuan penelitian.

#### 3.10.4 Membuat Model C4.5

Langkah-langkah pemodelan algoritma C4.5 yang akan dilaksanakan dapat dilihat pada Gambar 3.11.

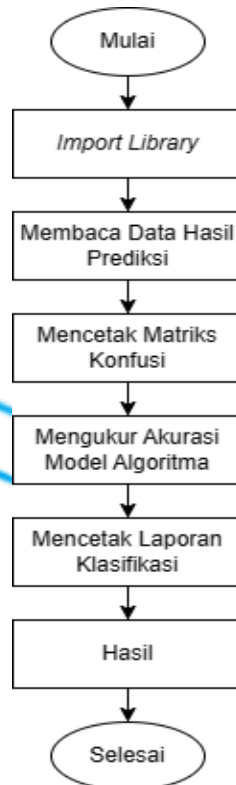


Gambar 3. 11. Alur Pemodelan Algoritma C4.5

Pertama, memanggil pustaka yang dibutuhkan, lalu masuk ketahap kedua yaitu memanggil data latih dan data uji yang telah disusun sebelumnya. Ketiga yaitu memisahkan fitur dan target. Keempat membuat pemodelan algoritma C4.5 yang dilanjutkan pada tahap kelima yaitu untuk memprediksi tujuan penelitian.

### 3.9. Evaluasi

Metode *Naïve Bayes* dan C4.5 merupakan algoritma jenis *supervised learning* yang memungkinkan dievaluasi menggunakan metode *Confusion Matrix*. Gambar 3.12 menjelaskan tentang evaluasi yang diterapkan.



Gambar 3. 12 Evaluasi

Dalam tahap evaluasi, langkah-langkah berikut dilakukan. Pertama, memanggil *library* yang dibutuhkan untuk melakukan metode *confusion matrix*. Kedua, membaca hasil prediksi yang telah diterapkan sebelumnya dari model algoritma *Naïve Bayes* dan model *C4.5*. Ketiga, mencetak matriks konfusi menggunakan fungsi `.confusion_matrix()`. Keempat, mengukur akurasi dari tiap model algoritma menggunakan fungsi `.accuracy_score()`. Kelima, mencetak laporan klasifikasi menggunakan fungsi `.classification_report()`.

### 3.10 Hasil

Dalam tahap menjelaskan hasil adalah dengan cara melakukan perbandingan hasil evaluasi *Confussion Matrix* antara algoritma *Naïve Bayes* dan *C4.5* dengan cara membandingkan hasil akurasi, presisi, dan *recall*. Perbandingan dilakukan untuk menentukan algoritma mana yang lebih efektif dalam memprediksi kelulusan mahasiswa tepat waktu.