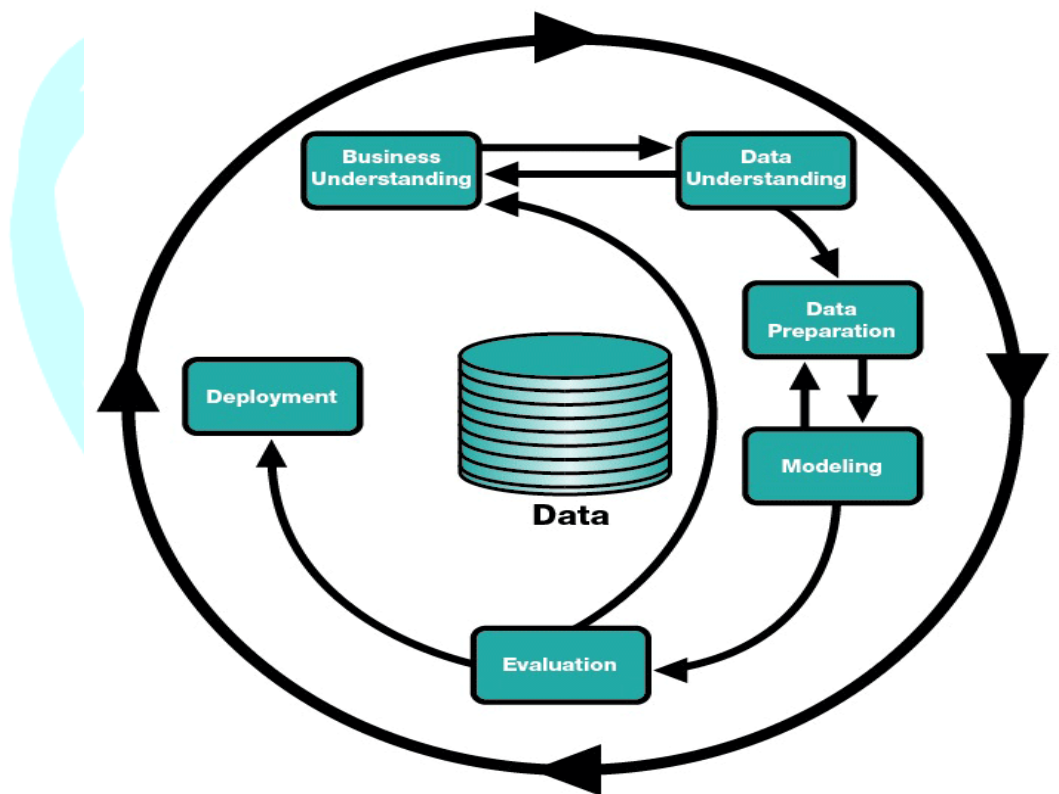


BAB III

METODE PENELITIAN

Metode Penelitian

Penelitian Tugas Akhir menggunakan metode CRISP-DM (Cross Industry Standard Process for Data Mining). Gambar 2.3 menunjukkan prosedur yang akan dilakukan.



Gambar 2.3 Metode CRISP-DM

3.1.1 Objek Penelitian

Pada penelitian ini hal penting yang akan penelitian lakukan yaitu mempersiapkan bahan penelitian, dalam hal ini peneliti akan mengkaji dataset yang meliputi data Huruf Mutu (HM), Satuan Kredit Mahasiswa (SKS), Angka Mutu (AM), dan Kumulatif Nilai (KM) per semester yang dihimpun dari Pusdatin Universitas Buana Perjuangan Karawang dan mahasiswa, sehingga data tersebut dapat peneliti klasifikasikan dengan menerapkan algoritma C4.5.

3.1.2 Lokasi Penelitian

Program Studi Teknik Informatika di Universitas Buana Perjuangan Karawang melakukan penelitian ini di Universitas Buana Perjuangan Karawang (UBP Karawang) adalah PTS swasta yang terletak di Karawang, Jawa Barat, Indonesia. Berlokasi di Jl. H.S. Ronggowaluyo, Teluk Jambe Timur, Karawang, Jawa Barat, 41361.

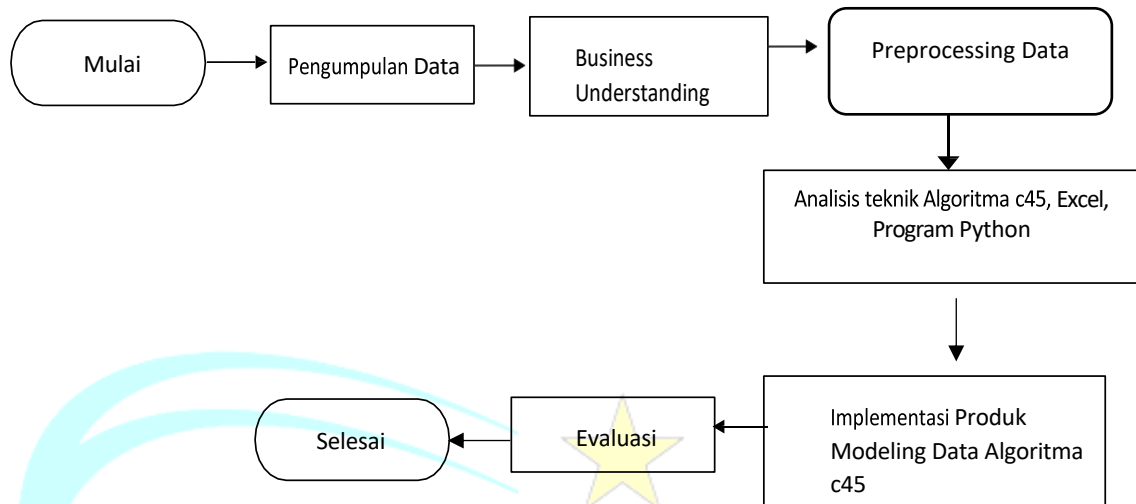
3.1.3 Alur Penelitian

Untuk menilai penelitian ini, peneliti menggunakan pendekatan studi kasus. Tujuan penelitian ini terdiri dari dua tujuan. Yang pertama adalah untuk memprediksi lulus tepat waktu mahasiswa Ubp Karawang berdasarkan predikat kelulusan mereka. Tujuan kedua adalah untuk mengembangkan prototype sistem informasi prediksi IPK dengan pemodelan Algoritma C4.5.

Dua model konseptual dan procedural digunakan untuk mencapai tujuan tersebut. Untuk memecahkan masalah secara teoritis, penelitian ini menggunakan pendekatan konseptual. Algoritma C4.5, yang telah lama digunakan untuk mengklasifikasikan data dengan atribut numerik dan kategorial, adalah model konseptual yang digunakan. Nilai atribut bertipe diskret dari rekaman baru dapat diprediksi dengan menggunakan aturan yang dihasilkan dari proses klasifikasi. (Irmayansyah & Kastrilia, 2020).

Dalam proses pengembangan sistem, metode prosedural digunakan. Untuk sistem informasi prediksi IPK mahasiswa, metode prototipe digunakan (Anestiviya, 2021). Metode ini dimulai dengan mengumpulkan kebutuhan pengguna; dalam hal ini, mahasiswa akan menjadi pengguna perangkat yang dikembangkan. Setelah itu, membuat rancangan kilat yang akan dievaluasi kembali sebelum diproduksi secara tepat.

Tujuan penelitian menentukan variabel yang digunakan. Penelitian ini menggunakan data primer dan sekunder. Data primer diperoleh melalui wawancara pengguna untuk mengetahui tanggapan mereka tentang kualitas produk sistem berdasarkan fitur dan fungsionalitas sistem atau perangkat lunak secara keseluruhan. Data sekunder termasuk statistik dan IPK siswa. Adapun alur penelitian ini dapat dilihat pada Gambar 3.1



Gambar 3.1 Alur Penelitian

3.2.4 Pengumpulan Data

Tahapan ini menjelaskan sumber data penelitian dan informasi yang dapat digunakan untuk penelitian. Dalam penelitian ini, pendekatan kualitatif digunakan untuk mengumpulkan data, perolehan informasi detail relatif jangka panjang (Sugiyono, 2019).

Dalam pengumpulan data, peneliti mendatangi pihak Pusdatin UBP Karawang dan data yang didapat yaitu Nilai IPK Mahasiswa UBP Karawang Program Studi Teknik informatika Angkatan 2015, 2018, 2020. Data keseluruhan belum di dapatkan, namun berikut kolom dan data yang nantinya di dapatkan dari UBP Karawang.

Tabel 3.2 Sample Data IPK Mahasiswa

No	NIM	Nama Mahasiswa	IPK	Program Studi	Semester
1	19416255201042	Abigiel Jen	3,10	Teknik Informatika	1/2019/ganjil
2	19416255201043	Bernard Ima	3,20	Teknik Informatika	3/2019/ganjil
3	19416255201044	Chika Hani	3,30	Teknik Informatika	5/2019/ganjil
4	19416255201045	Dimas Guard	3,40	Teknik Informatika	7/2019/ganjil

No	NIM	Nama Mahasiswa	IPK	Program Studi	Semester
5	19416255201046	Erlangga Fin	3,50	Teknik Informatika	9/2019/ganjil
6	19416255201047	Fina Dhea	3,60	Teknik Informatika	11/2019/ganjil
7	19416255201048	Gissel	3,70	Teknik Informatika	6/2019/genap
8	19416255201049	Hindustan	3,80	Teknik Informatika	8/2019/genap
9	19416255201111	Zaenal Abidin	4,10	Teknik Informatika	10/2019/genap
10	19416255201042	Abigiel Jen	4,12	Teknik Informatika	12/2019/genap
11	19416255201043	Bernard Ima	4,15	Teknik Informatika	14/2019/genap
12	19416255201044	Nadia	4,15	Teknik Informatika	1/2019/ganjil
13	19416255201045	Rafli	4,15	Teknik Informatika	13/2019/ganjil
14	19416255201046	Riska	4,15	Teknik Informatika	15/2019/ganjil
15	19416255201047	Desty	4,15	Teknik Informatika	17/2019/ganjil
16	19416255201048	Amelia	4,15	Teknik Informatika	19/2019/ganjil
17	19416255201049	Rizky	4,15	Teknik Informatika	21/2019/ganjil
18	19416255201050	Ridho	4,15	Teknik Informatika	16/2019/genap
19	19416255201051	Tiara	4,15	Teknik Informatika	18/2019/genap
20	19416255201052	Siti	4,15	Teknik Informatika	20/2019/genap

No	NIM	Nama Mahasiswa	IPK	Program Studi	Semester
21	19416255201053	Ajeng	4,15	Teknik Informatika	22/2019/genap
22	19416255201054	Sintya	4,15	Teknik Informatika	24/2019/genap
23	19416255201055	Reza	4,15	Teknik Informatika	26/2019/genap
24	19416255201056	Aditsti	4,15	Teknik Informatika	23/2019/ganjil
25	19416255201057	Putri	4,15	Teknik Informatika	23/2019/ganjil
26	19416255201058	Putra	4,15	Teknik Informatika	25/2019/ganjil
27	19416255201059	Noval	4,15	Teknik Informatika	27/2019/ganjil
28	19416255201060	Devan	4,15	Teknik Informatika	29/2019/ganjil
29	19416255201061	Alya	4,15	Teknik Informatika	28/2019/genap
30	19416255201062	Syawali	4,15	Teknik Informatika	30/2019/genap

Tabel 3.3 Data Mahasiswa

No	NIM	Nama Mahasiswa	IPK	IPS	Sttus bekerja	Tahun lahir	Program Studi	Semester
1	19416255201042	Abigiel Jen	3,1	3,1	0	1992	Teknik Informatika	1/2019/ganjil
2	19416255201043	Bernard Ima	3,2	3,2	1	1999	Teknik Informatika	3/2019/ganjil

No	NIM	Nama Mahasiswa	IPK	IPS	Sttus bekerja	Tahun lahir	Program Studi	Semester
3	19416255201044	Chika Hani	3,3	3,3	1	1995	Teknik Informatika	5/2019/ganjil
4	19416255201045	Dimas Guard	3,4	3,4	1	1999	Teknik Informatika	7/2019/ganjil
5	19416255201046	Erlangga Fin	3,5	3,5	1	1999	Teknik Informatika	9/2019/ganjil
6	19416255201047	Fina Dhea	3,6	3,6	1	1999	Teknik Informatika	11/2019/ganjil
7	19416255201048	Gissel	3,7	3,7	1	1997	Teknik Informatika	6/2019/genap
8	19416255201049	Hindustan	3,8	3,8	1	1998	Teknik Informatika	8/2019/genap
9	19416255201042	Zaenal Abidin	3,1	3,1	0	1992	Teknik Informatika	10/2019/genap
10	19416255201043	Abigiel Jen	3,2	3,2	1	1999	Teknik Informatika	12/2019/genap
11	19416255201044	Bernard Ima	3,3	3,3	1	1992	Teknik Informatika	14/2019/genap
12	19416255201045	Nadia	3,4	3,4	1	1999	Teknik Informatika	1/2019/ganjil
13	19416255201046	Rafli	3,5	3,5	1	1999	Teknik Informatika	13/2019/ganjil
14	19416255201047	Rizky	3,6	3,6	1	1999	Teknik Informatika	15/2019/ganjil
15	19416255201049	Ridho	3,3	3,3	1	1995	Teknik Informatika	17/2019/ganjil
16	19416255201050	Tiara	3,4	3,4	1	1999	Teknik Informatika	19/2019/ganjil

No	NIM	Nama Mahasiswa	IPK	IPS	Sttus bekerja	Tahun lahir	Program Studi	Semester
17	19416255201051	Siti	3,3	3,3	1	1999	Teknik Informatika	21/2019/ga njil
18	19416255201052	Ajeng	3,4	3,4	1	1999	Teknik Informatika	16/2019/ge nap
19	19416255201053	Sintya	3,5	3,5	1	1997	Teknik Informatika	18/2019/ge nap
20	19416255201054	Reza	3,6	3,6	1	1998	Teknik Informatika	20/2019/ge nap
21	19416255201055	Aditsti	3,7	3,7	1	1992	Teknik Informatika	22/2019/ge nap
22	19416255201056	Putri	3,8	3,8	1	1999	Teknik Informatika	24/2019/ge nap
23	19416255201057	Putra	3,1	3,1	0	1992	Teknik Informatika	26/2019/ge nap
24	19416255201058	Noval	3,2	3,2	1	1999	Teknik Informatika	23/2019/ga njil
25	19416255201059	Devan	3,3	3,3	1	1999	Teknik Informatika	23/2019/ga njil
26	19416255201060	Alya	3,4	3,4	1	1999	Teknik Informatika	25/2019/ga njil
27	19416255201061	Ridho	3,5	3,5	1	1995	Teknik Informatika	27/2019/ga njil
28	19416255201062	Tiara	3,6	3,6	1	1999	Teknik Informatika	29/2019/ga njil
29	19416255201049	Siti	3,3	3,3	1	1999	Teknik Informatika	28/2019/ge nap

No	NIM	Nama Mahasiswa	IPK	IPS	Sttus bekerja	Tahun lahir	Program Studi	Semester
30	19416255201111	Zaenal Abidin	4,1	4,1	1	1999	Teknik Informatika	30/2019/ge nap

Tabel 3.4 Sample Data IPK Mahasiswa

NIM	Nama	lulus tepat waktu
19416255201042	Abigiel Jen	Y
19416255201043	Bernard Ima	T
19416255201044	Chika Hani	Y
19416255201045	Dimas Guard	T
19416255201046	Erlangga Fin	Y
19416255201047	Fina Dhea	T
19416255201048	Gissel	Y
19416255201049	Hindustan	Y

Data yang di dapat mempunyai 2 data tabel. Tabel 3.2 Dan 3.3 adalah sample data mahasiswa per semester Dan tabel berikutnya adalah data Tabel untuk mengetahui mahasiswa ini lulus tepat waktu atau tidak.

Data yang ada pada database seringkali tidak digunakan sepenuhnya, jadi hanya data yang sesuai untuk dianalisis yang dapat diambil dan diproses dari Database. Hasil seleksi ini kemudian disimpan dalam berkas yang berbeda dari basis data operasional. (Elis Lisnawati, 2019).

3.2.5 Seleksi Data

Seleksi data adalah proses pemilihan data dari atribut yang tidak diperlukan (Permana & Patwari, 2019). Seleksi data pada penelitian ini berguna untuk mengidentifikasi dan memilih atribut yang paling relevan dan signifikan dalam membangun pohon keputusan (Ryanwar, 2020).

Algoritma C4.5 digunakan untuk membangun model prediksi berdasarkan data yang diberikan. Proses seleksi data dalam konteks ini proses ini adalah langkah penting dalam pembentukan pohon keputusan yang efektif dan efisien (Widiastuti., 2023). Tujuan dari seleksi data dalam algoritma C4.5 adalah;

Meningkatkan akurasi dimana dengan menghilangkan atribut yang kurang relevan atau memiliki kontribusi yang rendah terhadap pembentukan pohon

keputusan, model yang dihasilkan dapat memiliki akurasi yang lebih tinggi karena fokus pada informasi yang lebih penting.

Mengurangi overfitting yang terjadi ketika model terlalu rumit dan mengadaptasi noise dalam data pelatihan. Dengan memilih hanya atribut yang paling informatif, maka dapat membantu mencegah overfitting dan membuat model yang lebih umum.

- 1 Meningkatkan keterbacaan model pohon keputusan yang lebih sederhana dan lebih mudah dibaca akan mempermudah interpretasi hasil prediksi dan pengambilan keputusan.
- 2 Mengurangi biaya komputasional yang di mana dengan mengurangi jumlah atribut yang dipertimbangkan dalam pembentukan pohon keputusan, maka dapat mengurangi biaya komputasional yang diperlukan untuk melatih atau mengevaluasi model.

Setelah langkah ini selesai, selanjutnya adalah membuat kumpulan data yang digunakan untuk komputasi manual dan penerapan algoritma pengelompokan C4.5 dalam aplikasi *Python*.

Tabel 3.5 Seleksi Data

No	NIM	Nama Mahasiswa	IPK
1	19416255201042	Abigiel Jen	3,10
2	19416255201043	Bernard Ima	3,20
3	19416255201044	Chika Hani	3,30
4	19416255201045	Dimas Guard	3,40
5	19416255201046	Erlangga Fin	3,50
6	19416255201047	Fina Dhea	3,60
7	19416255201048	Gissel	3,70
8	19416255201049	Hindustan	3,80
9	19416255201111	Zaenal Abidin	4,10
10	19416255201042	Abigiel Jen	4,12
11	19416255201043	Bernard Ima	4,15
12	19416255201044	Nadia	4,16

No	NIM	Nama Mahasiswa	IPK
13	19416255201045	Rafli	4,18
14	19416255201046	Riska	4,20
15	19416255201047	Desty	4,23
16	19416255201048	Amelia	4,26
17	19416255201049	Rizky	4,28
18	19416255201050	Ridho	4,30
19	19416255201051	Tiara	4,32
20	19416255201052	Siti	4,7
21	19416255201053	Ajeng	4,8
22	19416255201054	Sintya	4,9
23	19416255201055	Reza	4,6
24	19416255201056	Aditsti	4,5
25	19416255201057	Putri	4,4
26	19416255201058	Putra	4,3
27	19416255201059	Noval	4,2
29	19416255201060	Devan	4,1
29	19416255201061	Alya	4,0
30	19416255201062	Syawali	4,33

3.2.6 Analisis Teknik

Analisis teknik yang dilakukan pada penelitian ini adalah sebagai berikut: mahasiswa Ubp Karawang di Program Studi Teknik Informatika yang akan lulus tepat waktu.

1. Algoritma C4.5: Untuk membangun pohon keputusan, algoritma C4.5

biasanya digunakan dengan langkah-langkah berikut:

- a. pilih atribut sebagai akar;
- b. buat cabang untuk tiap-tiap nilai;
- c. bagi kasus dalam cabang; dan d. ulangi prosedur untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Nilai gain tertinggi dari atribut yang ada menentukan pemilihan atribut sebagai akar. Nilai gain dihitung dengan hasil perhitungan nilai entropy. Untuk melakukan ini, rumus yang ditunjukkan dalam persamaan digunakan.

berikut:

$$Gain(S, A) = (S) - \sum_{i=1}^n |S_i| \times Entropy(S_i) \quad (4)$$

Keterangan:

S : himpunan kasus

A : atribut

n : jumlah partisi atribut A

|S_i|: jumlah kasus pada partisi ke-i

|S| : jumlah kasus dalam S

Setelah mendapatkan nilai gain, terdapat satu hal lagi yang perlu dilakukan dalam perhitungan yaitu mencari nilai entropy.

Entropy digunakan untuk menentukan seberapa informatif sebuah input atribut untuk menghasilkan output atribut. Rumus dasar dari entropy tersebut adalah sebagai berikut:

$$Entropy(S) = \sum_{i=1}^n -p_i \log_2 p_i \quad (5)$$

Keterangan:

S : himpunan kasus n : jumlah partisi S

p_i : proporsi dari S_i terhadap S

1. Program Python

Peneliti menggunakan library Google Colab untuk membuat program python dengan tujuan untuk meningkatkan perhitungan manual dan penggunaan RapidMiner.

- a. Siapkan *dataset* yang akan digunakan.
- b. Lakukan *preprocessing* pada *dataset*.
- c. Buat fungsi untuk menghitung *entropy* dan *gain ratio*.
- d. Buat fungsi untuk membangun pohon keputusan.
- e. Buat fungsi untuk melakukan *pruning* pada pohon keputusan.

3.2.7 Implementasi produk

Pada tahap implementasi ini adalah bagaimana pengolahan datanya diterapkan dengan bantuan sebuah tools. Tools yang akan digunakan dalam implementasi penelitian ini adalah dengan menggunakan program Python. Selain itu hasil produk akan diimplementasikan pada platform yang di mana pada tahap ini juga akan dilakukan perhitungan secara manual dengan menggunakan rumus algoritma C4.5.

3.2.8 Evaluasi

Evaluasi pada penelitian ini adalah Akurasi yaitu ukuran yang digunakan untuk mengukur seberapa baik sebuah model dalam melakukan klasifikasi atau prediksi pada dataset yang diberikan.

Perangkat sistem yang sudah disampaikan kepada pengguna pasti akan mengalami perubahan. Perubahan tersebut dapat karena mengalami kesalahan karena perangkat sistem harus menyesuaikan dengan lingkungan yang baru, seperti sistem operasi baru atau karena pengguna membutuhkan perkembangan fungsional atau untuk kerja (Purwaningsih & Nurelasari, 2023). Tahapan ini adalah tahap terpanjang karena akan dilakukan uji coba terhadap perangkat.

Akurasi dapat dihitung dengan membandingkan jumlah prediksi yang benar dengan jumlah total data yang diuji. Rumus akurasi adalah:

$$\text{Akurasi} = (\text{Jumlah prediksi benar} / \text{Total jumlah data}) \times 100\%.$$

Akurasi sangat penting dalam bidang kesehatan, seperti dalam tes pap smear untuk deteksi kanker serviks, dan dalam pemberitaan, seperti dalam laporan berita yang benar, tepat, dan cermat.

Dengan menggunakan metode Davies Bouldin Index (DBI). Metode ini merupakan cara untuk menyempurnakan kinerja sebuah metode atau algoritma optimasi demi mendapatkan jumlah cluster yang terbaik (Cesar et al., 2023). matrik kohesi dalam sebuah cluster ke-i yang dirumuskan sebagai berikut (ZulfaNabila,2021):

$$SSW = 1 \sum_{i=1}^m d(x_i, c_i) \quad (6)$$

Keterangan:

m : jumlah data dalam *cluster* ke- i

c : *centroid cluster* ke- i

d : jarak

Untuk menghitung separasi antar *cluster* dapat menggunakan *Sum of Square Between Cluster* (SSB) dengan penjabaran sebagai berikut (Dekaprasetya et al., 2022):

$$SSB_{ij} = d(c_i c_j) \quad (7)$$

Keterangan:

c : *centroid cluster* ke- i

d : jarak

Setelah memperoleh nilai kohesi dan separasi langkah selanjutnya melakukan pengukuran rasio (R_{ij}) untuk mengetahui nilai perbandingan antar *cluster* ke- i dan ke- j . *Cluster* yang memiliki nilai kohesi terkecil dan separasi yang terbesar adalah *cluster* yang baik. Berikut persamaan dalam menghitung nilai rasio (Dekaprasetya et al., 2022):

$$DBI = 1 \sum_{i \neq j}^k \max_{(ij)} (R) \quad (8)$$

Jumlah kelompok yang digunakan adalah k . Pendapatan DBI yang diperoleh (tidak negatif kurang dari 0), lebih baik. Pada titik ini, pengujian akan dilakukan untuk memastikan apakah penelitian telah mencapai tujuan yang diharapkan. Pengujian ini mencakup penyesuaian hasil sistem dengan perhitungan manual dan pengujian validasi algoritma.

3.2.9 Kesimpulan

Kesimpulan yang diperoleh dari prediksi ketepatan mahasiswa untuk lulus tepat waktu. Dengan mengukur akurasi dari penerapan algoritma C4.5 pada penelitian ini.

2.1 Populasi dan Sampel Penelitian

2.1.1 Populasi Penelitian

Menurut Sugiyono (2019), populasi adalah area umum yang terdiri dari objek atau subjek yang memiliki kualitas dan karakteristik tertentu yang dipilih

oleh peneliti untuk diteliti dan kemudian membuat kesimpulan. Untuk populasi penelitian ini, populasinya adalah mahasiswa Program Studi Informatika angkatan 2018 dan 2019 Universitas Buana Perjuangan Karawang.

2.1.2 Sampel Penelitian

Dalam penelitian ini, pengambilan sampel dilakukan secara random, tanpa mempertimbangkan kelas populasi atau Program Studinya. Karena ukuran populasi sudah diketahui, rumus slovin dapat digunakan untuk menghitung sampel minimum yang diperlukan. Peneliti menggunakan kelonggaran atau margin kesalahan sebesar 95% dalam kegiatan penelitian ini.

$$n = \frac{N}{1 + Ne^2} \quad (9)$$

Keterangan:

n : hasil perhitungan sampel,

N : adalah populasi keseluruhan,

e : angka kelonggaran/margin kesalahan populasi keseluruhan.

3.1.3 Deskripsi Penelitian

Secara umum, alur sistem yang akan dibangun akan mengumpulkan data siswa dalam bentuk tabel dengan beberapa atribut, termasuk NIM., Huruf Mutu (HM), Satuan Kredit Mahasiswa (SKS), Angka Mutu (AM), dan Kumulatif Nilai (KM) per semester Program Studi Informatika Universitas Buana Perjuangan Karawang.

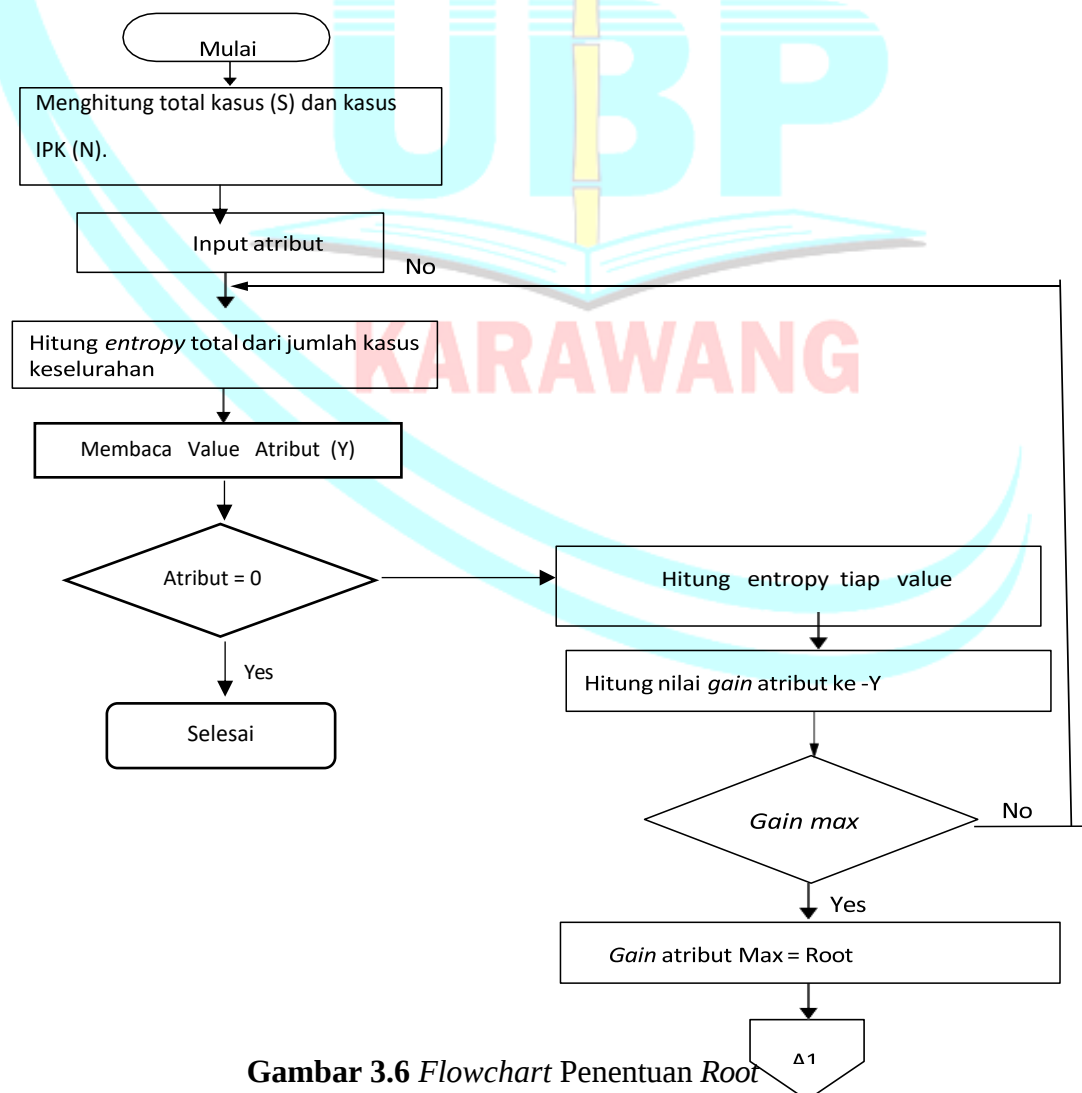
Setelah data siswa diklasifikasikan berdasarkan tujuan yang ditentukan, metode Decision Tree, algoritma C4.5, akan digunakan untuk menghitung nilai entropy dan pendapatan informasi. Setelah proses perhitungan selesai, akan dihasilkan aturan atau kondisi yang digunakan dalam penentuan keputusan pada proses prediksi. Keterangan IPK siswa adalah hasil dari sistem ini. Kemampuan sistem termasuk sebagai berikut :

1. Mengubah format data menjadi tabel data;

2. Menghitung nilai gain dan entropy;
 3. Memberikan update dengan melatih sampel data baru;
 4. Memberikan informasi baru tentang penentuan target yang akan diprediksi
- Dalam hal ini, untuk memprediksi nilai IPK mahasiswa

Gambar 3.6 menunjukkan flowchart pembentuk pohon keputusan yang berjalan pada sistem prediksi IPK mahasiswa dengan Algoritma C4.5 saat penentuan akar.

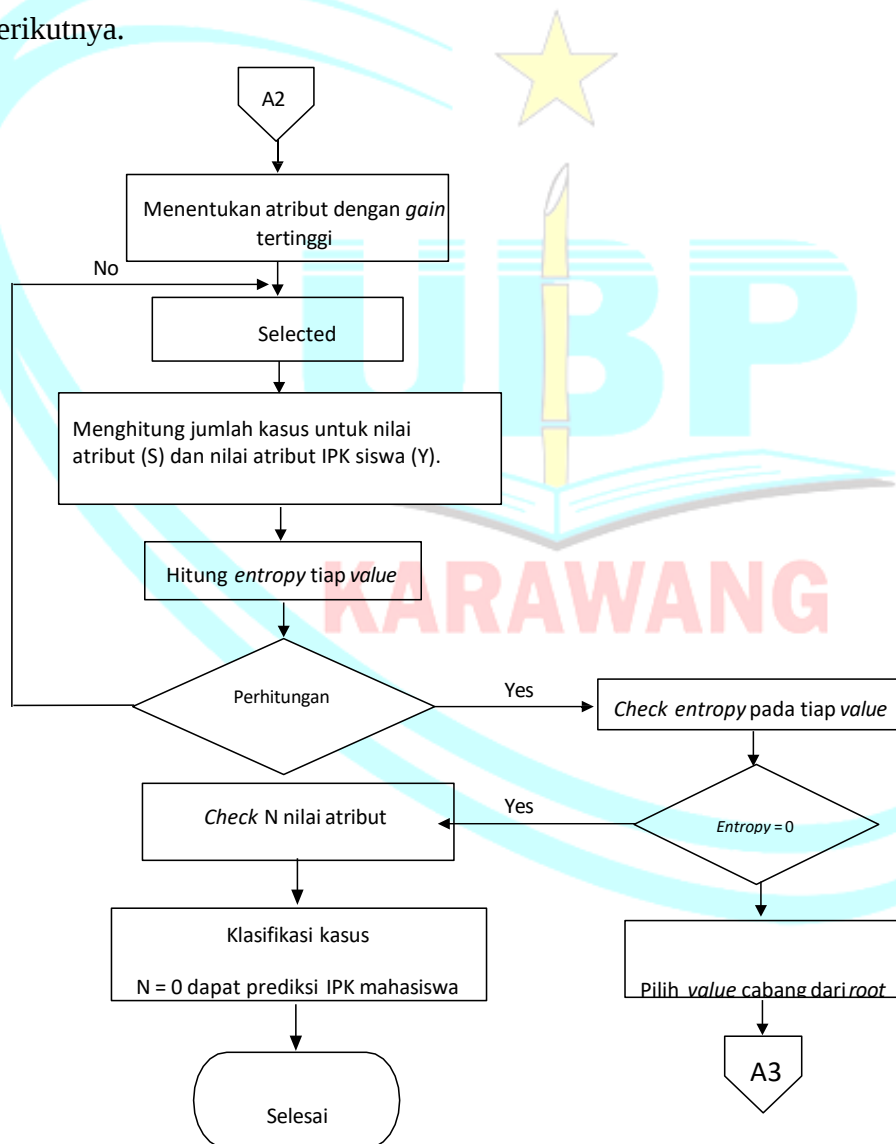
1. Sistem akan menghitung nilai informasi total dari data pelatihannya
2. Mendaftarkan atribut A
3. Menghitung nilai entropy dan nilai gain tiap atribut
4. Membandingkan nilai gain terbesar dari tiap atribut
5. Kemudian memilih atribut dengan nilai gain terbesar untuk dijadikan akar.



Gambar 3.6 Flowchart Penentuan Root

Selanjutnya adalah menentukan cabang pada *tree* yang dapat dilihat pada Gambar 3.7 di bawah ini.

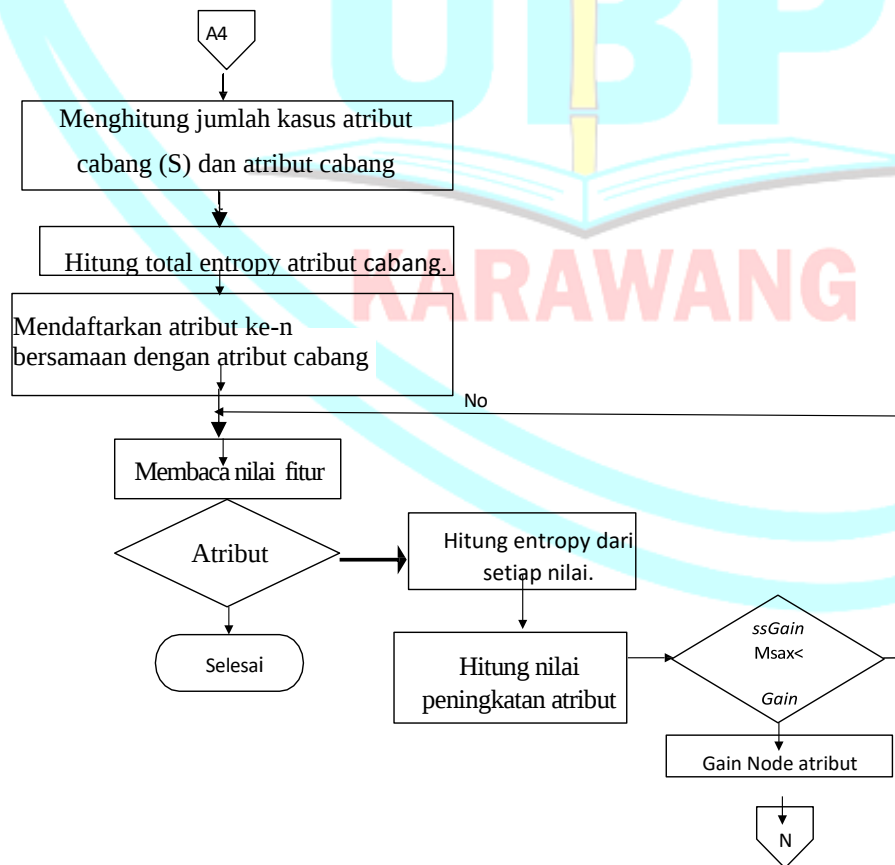
1. Atribut dengan gain tertinggi akan dipilih oleh sistem.
2. Nilai yang ada pada atribut tertinggi akan diklasifikasikan berdasarkan prediksi nilai IPK mahasiswa (N).
3. Entropi akan dihitung oleh setiap atribut sampai semua atribut selesai.
4. Nilai entropy yang nol (0) akan dikoreksi untuk menentukan klasifikasi kasus.
5. Nilai entropy yang lebih besar dari nol (0) akan dijadikan cabang pada node berikutnya.



Gambar 3.7 Flowchart Penentuan Cabang

Kemudian menentukan *node* pada pembentukan *tree* yang dapat dilihat pada Gambar 3.8 di bawah ini.

1. Sistem akan menghitung jumlah kasus pada atribut dengan *gain* tertinggi.
2. Menghitung nilai *entropy* total dari atribut dengan *gain* tertinggi.
3. Mendaftar atribut A selain atribut dari *gain* tertinggi.
4. Menghitung *entropy* dari masing-masing nilai atribut ke-n hingga habis terhitung semua.
5. Menghitung nilai *gain* dari masing-masing atribut.
6. Menentukan *gain* maksimal sebagai penentuan *node* selanjutnya.
7. Setelah *node* selanjutnya terpilih maka proses perhitungan akan berulang lagi mulai dari penentuan cabang dan penentuan *node* hingga semua atribut habis dieksekusi dan mencapai end of tree.
8. habis dieksekusi dan mencapai *end of tree*.



Gambar 3.8 Flowchart Penentuan Node