

BAB III METODE PENELITIAN

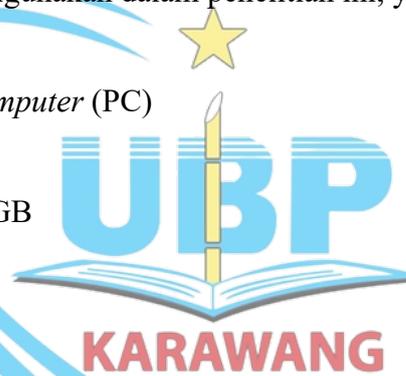
3.1. Bahan Penelitian

Dalam penelitian ini digunakan data nilai rapor siswa Sekolah Menengah Kejuruan Negeri (SMKN) jurusan Teknik Komputer Jaringan (TKJ). Data didapat langsung dari pihak SMKN bagian kurikulum dalam bentuk format .xslm (*excel*). Data yang dibutuhkan yaitu nilai mata pelajaran produktif seperti Rancang Bangun Jaringan, Administrasi Server, Jaringan Nirkabel dan mata pelajaran umum yaitu Matematika.

3.2. Peralatan Penelitian

Ada beberapa alat yang digunakan dalam penelitian ini, yaitu:

- a. Perangkat keras
 - 1) *Personal Computer (PC)*
 - 2) *Laptop*
 - 3) *Flashdisk 8 GB*
 - 4) *Printer*
 - 5) *Smartphone*
- b. Perangkat lunak
 - 1) *Sistem Operasi Windows 7 Ultimate*
 - 2) *Microsoft Office Word*
 - 3) *Microsoft Office Excel*
 - 4) *Sistem Operasi Linux Mint*
 - 5) *IDLE Python*
 - 6) *Google Chrome*



3.3. Lokasi Penelitian dan Waktu Penelitian

Dalam penulisan proposal tugas akhir ini penulis melakukan penelitian di Sekolah Menengah Kejuruan Negeri (SMKN) 1 Karawang yang terletak di Jalan Pangkal Perjuangan, RT.05/RW.04, Kelurahan Tanjungpura, Kecamatan Karawang Barat, Kabupaten Karawang. SMK Negeri ini merupakan salah satu

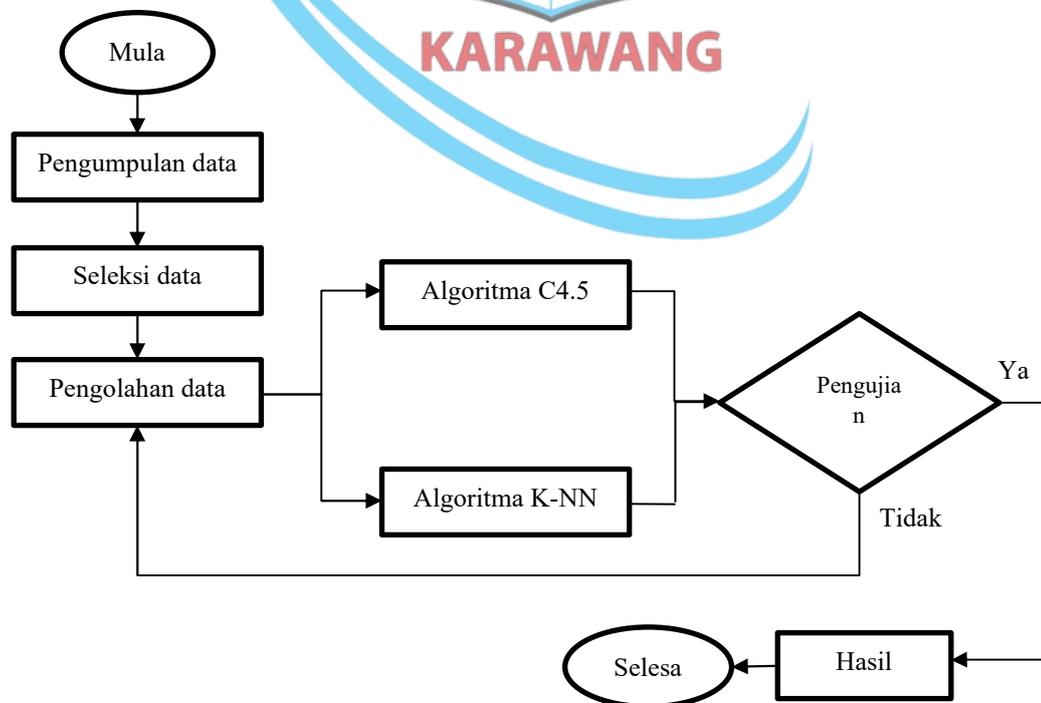
sekolah yang memiliki program keahlian bidang teknologi di Kabupaten Karawang. Adapun waktu penelitian terdapat pada tabel dibawah ini:

Tabel 3.1 Waktu penelitian

No	Kegiatan	Bulan																
		Bulan Ke-1					Bulan Ke-2				Bulan Ke-3							
		1	2	3	4	5	1	2	3	4	1	2	3	4				
1	Pengajuan judul	■																
2	Studi literatur	■		■			■											
3	Analisis kebutuhan	■																
4	Pengumpulan data													■				
5	Seleksi data													■				
6	Implementasi algoritma													■				
7	Pengujian menggunakan tool data <i>mining</i>													■				
No	Kegiatan	Bulan																
		Bulan Ke-4					Bulan Ke-5				Bulan Ke-6							
		1	2	3	4	5	1	2	3	4	1	2	3	4				
8	Studi literatur	■																
9	Penulisan hasil dan pembahasan	■																
10	Pembuatan aplikasi pengujian	■																
11	Pengujian menggunakan aplikasi	■																
12	Evaluasi	■																

3.4 Prosedur Percobaan

Berikut adalah alur prosedur percobaan dalam penelitian ini:



Gambar 3.1 Alur prosedur percobaan

3.4.1. Pengumpulan data

Pada tahap pengumpulan data penulis mendatangi SMKN 1 Karawang bertujuan meminta izin melakukan penelitian untuk mengumpulkan data-data yang diperlukan. Data tersebut diantaranya nama siswa, kelas, nilai mata pelajaran produktif dan mata pelajaran umum. Berikut dataset nilai siswa sebelum dilakukan penyeleksian.

Tabel 3.2 Dataset Sebelum Tahap Seleksi

No	Kelas	Nama	L/P	Mata Pelajaran				
				B.Indo	MTK	AS	RBJ	JN
1	TEI 1	Abdul Rizal	L	81	78	79	83	80
2	TEI 1	Ade Karmila	P	83	79	85	84	83
3	TEI 1	Agil Ghalib	L	80	78	79	83	85
4	TEI 1	Ahmad Aziz	L	81	78	87	88	87
5	TEI 1	Amalia Sandra	P	83	78	85	86	85
6	TEI 1	Amelia	P	81	78	85	86	83
...
...
...
353	RPL 3	Wiwin	P	86	79	79	80	81

3.4.2. Seleksi data

Seleksi data bertujuan untuk membersihkan data dari atribut-atribut yang tidak diperlukan. Setelah tahap penyeleksian dilakukan, maka akan menghasilkan dataset yang akan diolah pada tahap selanjutnya, seperti nama siswa, kelas, nilai mata pelajaran produktif dan nilai mata pelajaran umum. Penulis menggunakan data tersebut karena merupakan salah satu dasar atau materi prasyarat dari beberapa mata kuliah di program studi Teknik Informatika dan Sistem Informasi.

Tabel 3.3 Dataset Setelah Tahap Seleksi

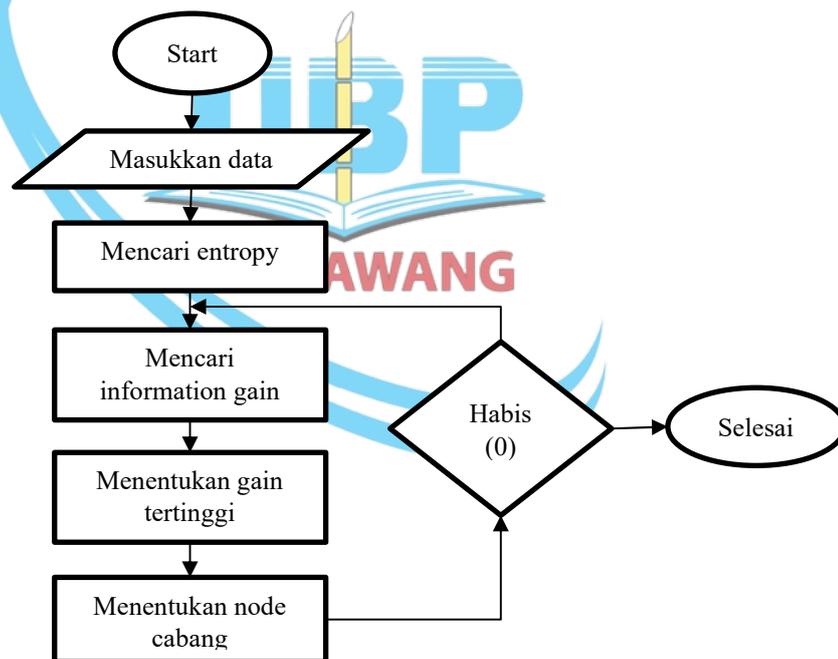
No.	Nama	Jurusan	Nilai Rata-Rata Mata pelajaran			
			MTK	AS	RBJ	JN
1	Ade	TEI	79,5	85	85	83
2	Agil	TEI	79	79	81,5	85
3	Ahmad	TEI	78,5	87	88	87
4	Amalia	TEI	78,5	85	85,5	85
5	Amelia	TEI	79	85	85,5	83
6	Arsyi	TEI	79	85	85,5	80
7	Atih	TEI	78,5	80	81,5	80

No.	Nama	Jurusan	Nilai Rata-Rata Mata pelajaran			
			MTK	AS	RBJ	JN
...
...
353	Wiwin	RPL	79,5	79,5	81,5	81

3.4.3. Pengolahan Data

Pada tahap ini proses klasifikasi dilakukan dengan menggunakan algoritma C4.5 dan algoritma K-NN. Nilai yang dihasilkan dalam penelitian ini berupa tingkat akurasi dari kedua algoritma tersebut, serta pohon keputusan dari algoritma C4.5 yang menentukan nilai siswa memenuhi kriteria program studi Teknik Informatika atau Sistem Informasi.

A. Tahapan Algoritma C4.5



Gambar 3.2 Tahapan Algoritma C4.5

Tahapan dalam membuat pohon keputusan dengan algoritma C4.5 (Haryati *et al.*, 2015) yaitu:

1. Memasukkan *data training*. *Data training* biasanya dari data histori yang pernah terjadi sebelumnya dan sudah dikelompokkan ke dalam kelas-kelas tertentu.

2. Mencari nilai entropy menggunakan rumus sebagai berikut:

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

Keterangan:

S : himpunan kasus

A : atribut

n : jumlah partisi S

p_i : proporsi dari S_i terhadap S

3. Kemudian hitung nilai *Gain* dengan metode *information gain*:

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} \times Entropy(S_i) \quad (2)$$

Keterangan:

S : himpunan kasus

A : atribut

n : jumlah partisi atribut A

$|S_i|$: jumlah kasus pada partisi ke-i

$|S|$: jumlah kasus dalam S

4. Menentukan nilai *Gain* tertinggi dari setiap atribut.
5. Menentukan akar dari pohon. Akar akan diambil dari atribut yang terpilih dengan nilai *Gain* paling tinggi yang akan menjadi akar pertama.
6. Ulangi langkah kedua hingga semua tupel terpartisi.

B. Tahapan Algoritma K-NN

1. Menentukan parameter K (jumlah tetangga paling dekat)
2. Hitung jarak antara data yang akan dievaluasi (data *testing*) dengan semua pelatihan (data *training*)

Rumus *Euclidean Distance* untuk menghitung jarak :

$$d_{ij} = \sqrt{[(x_i - x_j)^2 + (y_i - y_j)^2]} \quad (3)$$

Keterangan :

X_1 : Sampel data

X_2 : Data uji atau *testing*

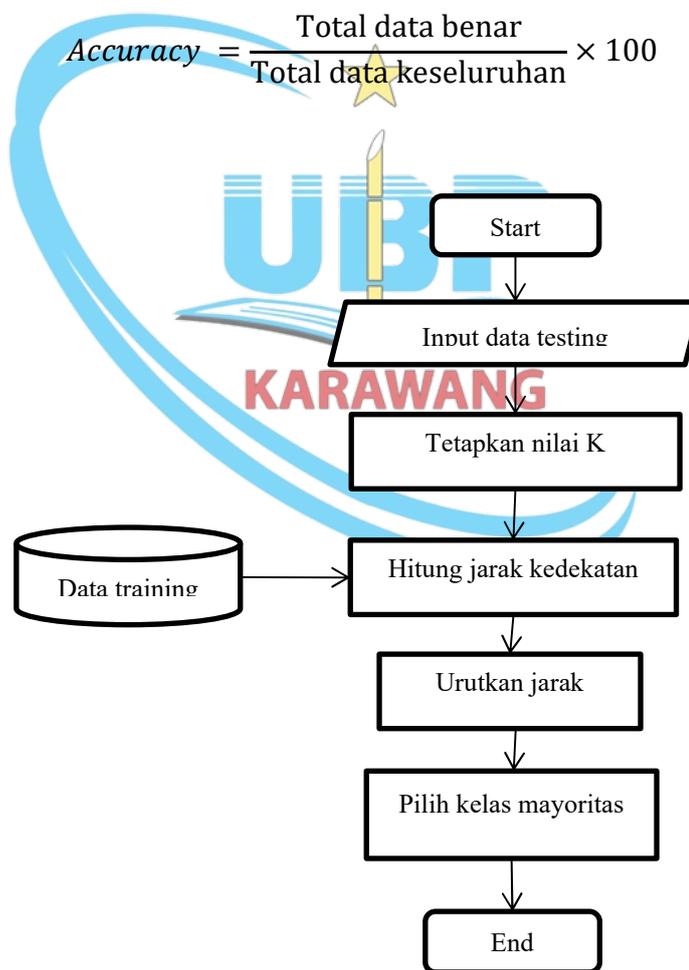
I : Variabel data

d : Jarak

p : Dimensi data

3. Urutkan jarak yang terbentuk dari jarak yang terkecil hingga jarak terbesar
4. Tentukan tetangga terdekat berdasarkan jarak minimum ke-K
5. Menentukan kelas mayoritas sebagai klasifikasi objek atau data baru
6. Cari jumlah kelas dari tetangga yang terdekat dan tetapkan kelas tersebut sebagai kelas data yang akan dievaluasi dan hitung nilai akurasi menggunakan rumus berikut:

$$Accuracy = \frac{\text{Total data benar}}{\text{Total data keseluruhan}} \times 100$$



Gambar 3.3 Tahapan Algoritma K-NN

3.4.4. Pengujian

Pada tahap pengujian bertujuan untuk memastikan nilai tingkat akurasi yang dihasilkan. Dalam proses pengujian penulis melakukan beberapa tahapan yaitu:

1. Perhitungan secara manual
2. Pengujian menggunakan *tool* RapidMiner

Implementasi algoritma pada sistem menggunakan bahasa Python

