

## BAB II

### TINJAUAN PUSTAKA

#### 2.1 Karate

##### 1. Definisi Karate

Karate yang terdiri dari dua suku kata yaitu “*kara*” yang artinya kosong dan “*te*” yang artinya tangan merupakan bela diri yang berasal dari Jepang. Dengan demikian karate dapat diartikan sebagai suatu bela diri yang memungkinkan seseorang dalam mempertahankan diri tanpa menggunakan sebuah senjata atau tangan kosong. (Astiawan, 2012).

##### 2. Teknik Dasar Karate

###### a. *Kihon*

Kihon merupakan suatu unsur paling utama yang menjadi dasar pembentuk sebuah teknik. Gerakan dasar yang ada pada kihon seperti kuda-kuda(*dachi*), pukulan(*tsuki*), tendangan(*geri*), dan tangkisan(*uke*) (Ananingdyas, 2016).

###### b. *Kata*

Kata merupakan gabungan dari berbagai gerakan, terdiri dari dasar kuda-kuda, pukulan, tangkisan dan tendangan yang menyatu dalam suatu jurus atau kembangan(P.Victorianus, 2012).

###### c. *Kumite*

Kumite secara harfiah berarti “pertemuan tangan” merupakan pertarungan antara dua *karateka* yang menggunakan teknik menyerang dan bertahan untuk memperbutkan kemenangan. (P.Victorianus, 2012 ).

#### 2.2 Shinkyokushin Karate

*Shinkyokushin* merupakan pecahan aliran *karate kyokushin* yang dibentuk oleh *Shihan* Kenji Midori. *Shinkyokushin* tidak jauh berbeda dengan organisasi *kyokushin* lainnya, yaitu dengan menggunakan *system full body contact*. *Kyokushin*

*karate* bila dilihat dari macam aktifitas latihannya dapat digolongkan sebagai jenis Latihan yang menggabungkan kekuatan, daya tahan dan konsentrasi (Setiabudi, D, 2021).

### 2.3 Data Mining

Data mining adalah proses pembelajaran yang mengumpulkan, membersihkan, dan memperoleh manfaat dari suatu data (Aggarwal, 2015). Data mining juga merupakan proses untuk mencari nilai tambah berupa informasi yang tidak diketahui dari database. KDD (Knowledge Discovery in Database) yang mempunyai arti yaitu menemukan model baru yang efektif dan dapat menjadi pengetahuan dari suatu data (Husain, 2018).

Jadi dapat disimpulkan bahwa *data mining* adalah proses mencari suatu informasi yang baru pada database dengan menggunakan algoritma tertentu, sehingga hasilnya dapat digunakan untuk memecahkan suatu masalah dalam *data mining*.

#### 2.3.1 Tahap-Tahap Data Mining

Data mining sebenarnya merupakan salah satu proses *knowledge discovery in database* (KDD) yang bertugas mengekstrak pola dari data dengan mengfungsikan suatu algoritma. Adapun tahapan dalam proses KDD sebagai berikut (Fatmawati & Windarto, 2018):

1. Data Selection

Merupakan penyeleksian himpunan data yang harus dilakukan sebelum tahap penggalian informasi. Proses *data mining* menggunakan data-data dari hasil seleksi yang dilakukan kemudian disimpan dalam berkas dan terpisah dari basis data.

2. Preprocessing Data (Cleaning)

*Preprocessing data* atau *cleaning data* sangat dibutuhkan dalam pelaksanaan data mining, supaya data yang digunakan tidak terdapat duplikasi data, dan memperbaiki kesalahan pada data.

### 3. Transformasi

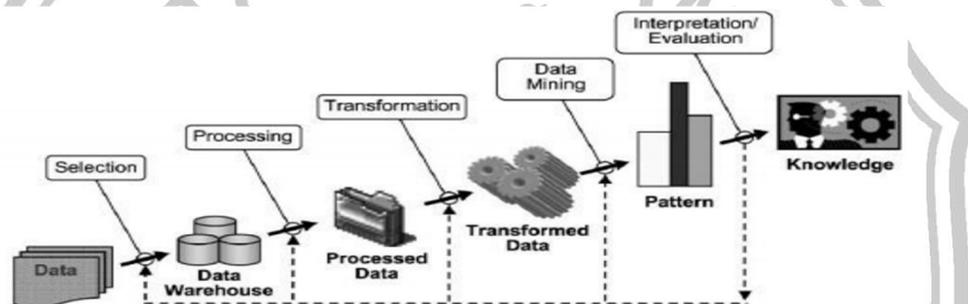
Merubah data yang diproses kedalam bentuk yang sesuai untuk proses data mining.

### 4. Proses Mining

Tahap dimana data diproses menggunakan teknik atau algoritma tertentu yang tujuannya untuk mencari pola atau informasi menarik.

### 5. Interpretasi/Evaluasi

Pengidentifikasian pola informasi yang dihasilkan dari proses mining. pola informasi yang akan dipresentasikan harus dalam format yang mudah dimengerti.



Sumber : (Fatmawati & Windarto, 2018)

### 2.3.2 Pengelompokan Data Mining

Data mining dibagi menjadi sebagian kelompok berdasarkan tugas yang dapat dilakukan yaitu: (Badrul, 2016)

#### 1. Deskripsi

Deskripsi adalah pengidentifikasian pola-pola dalam suatu data yang berulang kemudian pola tersebut diganti menjadi deskripsi aturan yang mudah dipahami. Deskripsi dari pola sering memberikan uraian pada suatu pola.

#### 2. Estimasi

Dalam estimasi mirip dengan klasifikasi hanya saja variabel dan target estimasi lebih tertuju numerik.

### 3. Prediksi

Prediksi hampir sama dengan estimasi dan klasifikasi. Namun dalam prediksi nilai hasilnya akan tetap sama dimasa depan

Contoh prediksi dalam penelitian :

- Prediksi calon pengguna indihome baru dalam bulan berikutnya.
- Prediksi presentase kecelakaan lalu lintas
- Prediksi kelulusan mahasiswa tercepat pada tiap tahunnya .

### 4. Klasifikasi

Klasifikasi yaitu tata cara untuk menemukan model atau fungsi yang mendeskripsikan atau mengelompokan information ke dalam kelas. Dalam klasifikasi ada variabel kategori sasaran.

### 5. Pengklasteran(Clustering )

Merupakan pengumpulan record yang mempunyai kemiripan dan bisa menciptakan kelas objek dari kemiripan tersebut. Klaster juga memiliki ketidakmiripan dengan record-record dalam klaster lain. tidak seperti klasifikasi, clustering tidak mempunyai variabel target.

### 6. Asosiasi

Menemukan atribut asosiasi yang muncul sekali . Asosiasi mempunyai tugas memberikan aturan untuk dapat mengukur hubungan antara dua atribut atau lebih.

## 2.4 Klasifikasi

Klasifikasi adalah proses mengidentifikasi sebuah record data baru yang sudah memiliki *class label* sebelumnya, sehingga proses ini juga disebut sebagai supervised learning. Berbeda dengan *unsupervised learning* yaitu class label dari data latih tidak diketahui dan jumlah kelompok belum diketahui sebelumnya (Han, Kamber & Pei, 2012).



Gambar 2.2 Model Klasifikasi

## 2.5 Rough Set

*Rough set* adalah sebuah metode matematika yang dikembangkan oleh Zdislaw Pawlack yang terutama digunakan untuk menyelesaikan masalah klasifikasi data. Tujuan analisis Rough Set yaitu untuk mendapatkan perkiraan aturan singkat dalam kumpulan data (Beny , 2016 ).

Terdapat dua sistem informasi, yaitu atribut kondisi dan atribut keputusan. Dimana U adalah himpunan tak terbatas yang tidak kosong dari objek dan A adalah set terhingga dari atribut dimana (Nurhayati, 2014):

$$IS = \{U,A\}$$

Dimana : IS = *information Systems*

U = *Object* (sekumpulan example)

A = *Conditional Atribute* (Atribute Kondisi)

Dalam penggunaan *Information Systems* terdapat outcome klasifikasi yaitu, *Decision Systems*, dapat dilihat sebagai

$$IS = (U, \{A,C\})$$

Dimana : IS = *Information Systems*

U = *Object* (sekumpulan example)

A = *Conditional Atribute*( Atribute Kondisi)

C = Atribut Keputusan

*Equivalence class* dan *discernibility matrix* merupakan proses penting dalam teori *Rough Set*, berikut lanjutan tahapan metode *Rough Set* yaitu : (Beny, 2016)

### 1. *Equivalence Class*

*Equivalence Class* merupakan pengelompokan objek-objek yang memiliki nilai atribut kondisi dan keputusan yang sama menjadi satu bagian.

### 2. *Discernibility Matrix*

*Discernibility Matrix* merupakan sekumpulan matrik yang memiliki perbedaan antara objek (i) dengan objek (j)

a. *Discernibility Matrix*

Merupakan kumpulan atribut objek  $X_i$  (Kolom Equivalence Class) dan  $X_j$  (baris Equivalence Class) yang berbeda. Pada matrix discernibility ini semua objek akan dibandingkan. Dalam proses perbandingan ini, tidak memperhatikan atribut keputusannya. Jika atributnya memiliki kesamaan nilai maka tidak akan ada nilai yang dihasilkan, tetapi jika nilainya memiliki perbedaan maka akan diberikan nilai.

b. *Discernibility Matrix Modulo D*.

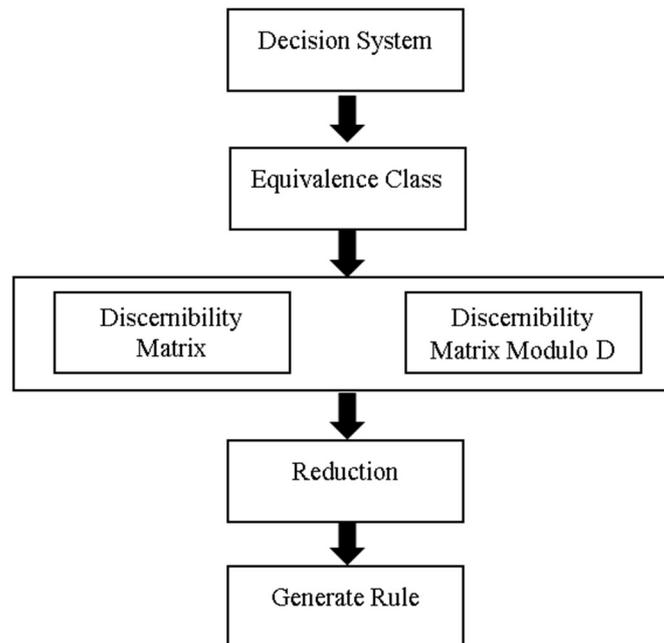
Tahap ini sama seperti *Discernibility Matrix*. Perbedaannya terdapat pada proses membandingkannya. Tidak hanya mempertimbangkan atribut kondisi, tetapi juga mempertimbangkan atribut keputusannya.

3. *Reduct*

*Reduct* merupakan pemilihan atribut minimal (*interesting attribute*) dari suatu himpunan atribut kondisi. Proses ini menggunakan *Prime Implicant Boolean*. (Ardiansyah & Zulfi, 2015).

4. *General Rules*

*General rules* adalah suatu cara untuk membangkitkan *rules* berdasarkan hasil *equivalence class* dan *reduct*. Pada *general rules* terdiri dari *if then* atau *if f then g* yang dapat disebut juga  $f \rightarrow g$ . Bagian *f* disebut sebagai *antecedent* (kondisi) dan bagian *g* disebut *conclusion* (keputusan). *General rules* menghasilkan sebuah pengetahuan atau *rules* dari data yang akan dijadikan sebagai acuan dalam mengambil keputusan. (Nila L. dkk, 2015).



Gambar 2.3 Algoritma Rough Set

## 2.6 Aljabar Boolean

Aljabar boolean merupakan teknik matematika yang digunakan dalam menyelesaikan masalah-masalah logika yang memiliki angka biner, yaitu 1 dan 0 sesuai dengan tipe data boolean yaitu *true* dan *false*. Aljabar boolean mendasari operasi-operasi aritmatika yang dilakukan oleh komputer. Berikut ini adalah hukum-hukum yang berlaku pada aljabar boolean (Arnettha,2016):

Tabel 2.1 Hukum Hukum Aljabar Boolean

No	Hukum	Contoh	
1	Identitas	$A + 1 = A$	$A \cdot 0 = A$
2	Komplemen	$A + A' = 1$	$AA' = 0$
3	Involusi	$(A')' = A$	
4	Komutatif	$A + B = B + A$	$AB = BA$
5	Asosiatif	$(A + B) + C = A + (B + C)$	$(A \cdot B) \cdot C = A \cdot (B \cdot C)$
6	Idempoten	$A + A = A$	$A \cdot A = A$
7	Distribusi	$A \cdot (B + C) = (A \cdot B) + (A \cdot C)$	$A + (B \cdot C) = (A + B) \cdot (A + C)$

8	De Morgan	$(A + B)' = A'B'$	$(AB)' = A' + B'$
9	Absorpsi	$A + AB = A$	$A(A+B) = A$
10	Dominasi	$A + 1 = 1$	$A \cdot 0 = 0$

## 2.7 Rosetta

*Rosetta* merupakan salah satu *software* untuk menganalisis data tabel dalam metode *Rough Set*. *Rosetta* bertujuan untuk menjalankan seluruh proses data mining dan penemuan *knowledge* dari pemrosesan awal data, hingga perhitungan set atribut untuk menghasilkan *rules if-then*. (Nasution, 2016).

## 2.8 Quality Measure

*Quality measure* merupakan ukuran yang dapat diekspresikan dalam jumlah atau ukuran terbatas yang biasanya dihubungkan *decision rule*. *Quality measure* terbagi menjadi 4 yaitu *support*, *strength*, *accuracy*, dan *coverage*. (Nila & dkk, 2011).

Diketahui :

- *Card*  
*Card* Merupakan jumlah dari anggota himpunan
- (U)  
Merupakan jumlah banyak data yang digunakan.
- *Card (f)*  
*Card (f)* merupakan jumlah banyak data pada dataset yang memiliki kesamaan dengan *antecedent*.
- *Card (g)*  
*Card (g)* merupakan jumlah banyak data pada dataset yang memiliki kesamaan dengan *conclusion*.

### 1. *Support*

Nilai *support* dalam *decision rule* merupakan jumlah objek yang sesuai antara atribut dan keputusan.

$$\text{support } (f \rightarrow g) = \text{card } (f \cap g)$$

## 2. *Strength*

Nilai *strength* dalam *decision rules* merupakan seberapa sering objek muncul dalam rule tersebut .

$$Strength = \frac{support}{card(U)} \times 100\%$$

## 3. *Accuracy*

Nilai *accuracy* merupakan rasio perbandingan antara objek yang ada pada atribut dan keputusan terhadap objek yang memenuhi atribut saja.

$$Accuracy = \frac{support}{card(f)} \times 100\%$$

## 4. *Coverage*

Nilai *coverage* merupakan kebalikan dari *accuracy*, yaitu merupakan rasio perbandingan antara objek yang ada pada atribut dan keputusan terhadap objek yang memenuhi keputusan saja.

$$Coverage = \frac{support}{Card(g)} \times 100\%$$

## 2.9 Confusion Matrix

Pengukuran kinerja suatu sistem klasifikasi merupakan suatu hal yang penting untuk memahami seberapa akurat sistem mengklasifikasi suatu data. *Confusion matrix* merupakan salah satu metode yang digunakan untuk melakukan perhitungan akurasi pada *data mining* . *Confusion matrix* berisikan informasi hasil klasifikasi yang sebenarnya dan yang telah diprediksi oleh sistem klasifikasi (Rosandy, 2016).

Tabel 2.2 Confusion Matrix

Prediksi	Kelas Aktual	
	Positif	Negative
Positif	TP	FP
Negative	FN	TN

$$\text{Accuracy} : \frac{TP+}{TP+TN+FN+FP} \times 100 \%$$

$$\text{Precision} : \frac{TP}{TP+FP} \times 100 \%$$

$$\text{Recall} : \frac{TP}{TP+} \times 100 \%$$

#### Keterangan

- TP (*True Positive*) adalah jumlah data positif yang diklasifikasikan positif.
- FP (*False Positif*) adalah jumlah data negatif yang diklasifikasikan positif..
- FN (*False Negatif*) adalah jumlah data positif yang diklasifikasikan negatif.
- TN (*True Negative*) adalah jumlah data negatif yang diklasifikasikan negatif

#### 2.10 F – Measure

*F-measure* adalah salah satu perhitungan evaluasi yang menggabungkan nilai *recall* dan nilai *precision*. Nilai *recall* dan *precision* dalam sebuah kasus dapat memiliki bobot yang berbeda. *F-measure* adalah ukuran yang menampilkan timbal balik antara *recall* dan *precision* yang merupakan bobot harmonic mean. Berikut rumus *f-measure* (Yulianti, 2017) :

$$F1\text{-measure} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$