

PREDIKSI POLA KECELAKAAN KERJA PADA PERUSAHAAN NON EKSTRAKTIF MENGGUNAKAN ALGORITMA DECISION TREE: C4.5 DAN C5.0

***PREDICTION PATTERNS OF OCCUPATIONAL
ACCIDENT IN NON EXTRACTIVE COMPANY
USING DECISION TREE ALGORITHM : C4.5***

NAMA : YOLANDA RIZKITA PUTRI

NRP : 1212100040

DOSEN PEMBIMBING : Dr. IMAM MUKHLASH, S.Si, MT

Drs. NURUL HIDAYAT, M. Kom



JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA



M

PENDAH
ULUAN

TINJA
UAN
PUSTA
KA

METOD
E
PENELI
TIAN

PERANC
ANGAN
DAN
IMPLEME
NTASI
SISTEM

PENGUJ
IAN DAN
PEMBAH
ASAN

KESIMP
ULAN

DAFT
AR
PUST
AKA

Latar Belakang

Rumusan Masalah

Batasan Masalah

Tujuan Penelitian

Manfaat Penelitian



Latar Belakang

Rumusan Masalah

Batasan Masalah

Tujuan Penelitian

Manfaat Penelitian



Latar Belakang

Rumusan Masalah

Batasan Masalah

Tujuan Penelitian

Manfaat Penelitian



KEJADIAN TAK TERENCANA DAN TIDAK TERKONTROL YANG DISEBABKAN OLEH MANUSIA, FAKTOR SITUASI ATAU LINGKUNGAN ATAU ERUPAKAN KOMBINASI DARI FAKTOR-FAKTOR TERSEBUT YANG MENGGANGGU PROSES KERJA YANG MUNGKIN BERAKIBAT ATAU TIDAK BERAKIBAT CEDERA, PENYAKIT, KEMATIAN, KERUSAKAN HARTA BENDA ATAU KEJADIAN LAIN YANG TIDAK DIHARAPKAN.

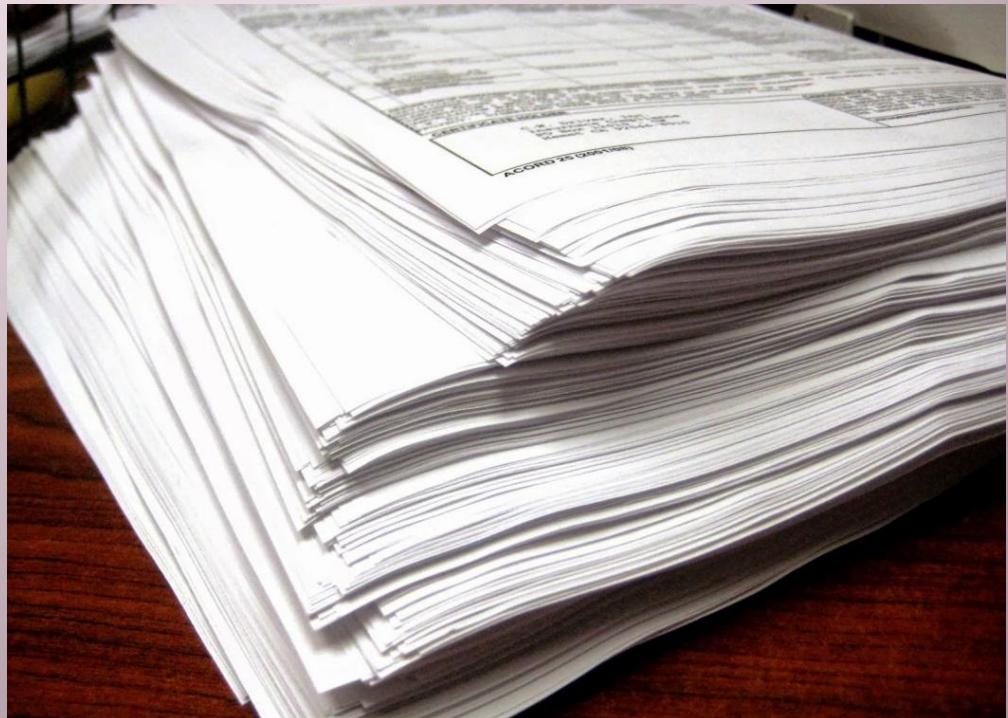
Latar Belakang

Rumusan Masalah

Batasan Masalah

Tujuan Penelitian

Manfaat Penelitian



Latar Belakang

Rumusan Masalah

Batasan Masalah

Tujuan Penelitian

Manfaat Penelitian



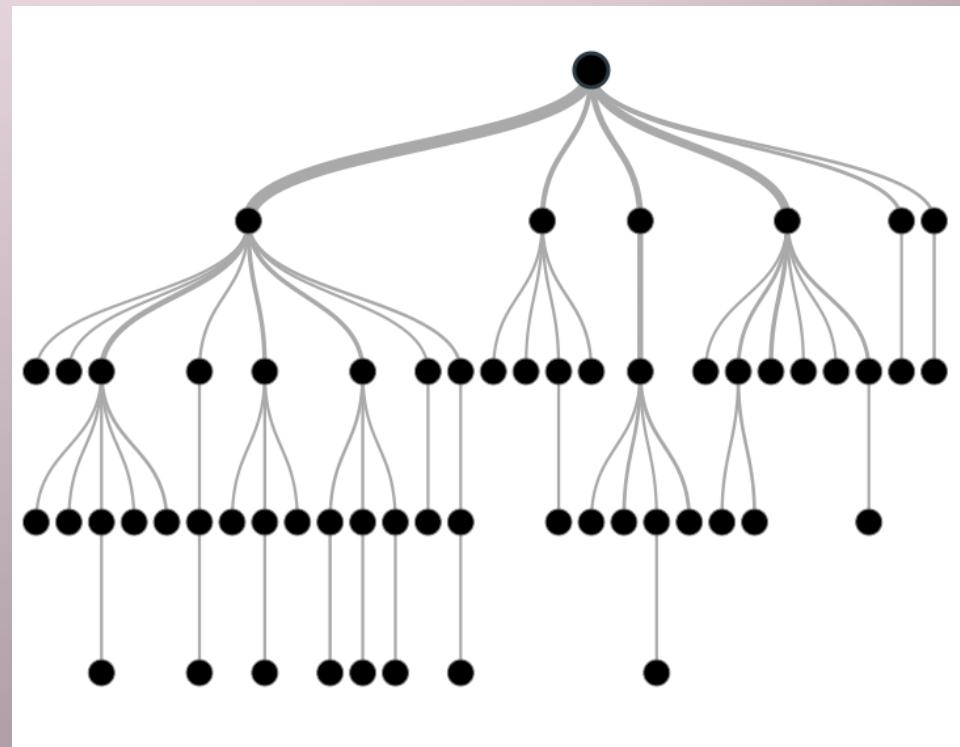
Latar Belakang

Rumusan Masalah

Batasan Masalah

Tujuan Penelitian

Manfaat Penelitian



Latar Belakang

Rumusan Masalah

Batasan Masalah

Tujuan Penelitian

Manfaat Penelitian

C4.5 ALGORITHM

VS

C5.0 ALGORITHM

Latar Belakang

Rumusan Masalah

Batasan Masalah

Tujuan Penelitian

Manfaat Penelitian

1. Bagaimana menggali pola kecelakaan kerja pada suatu perusahaan menggunakan *data mining* dengan metode *classification*?

Latar Belakang

Rumusan Masalah

Batasan Masalah

Tujuan Penelitian

Manfaat Penelitian

2. Berapa banyak model prediksi yang dihasilkan dari *classification* menggunakan *Decision Tree Algorithm* yaitu Algoritma C4.5 dan C5.0

Latar Belakang

Rumusan Masalah

Batasan Masalah

Tujuan Penelitian

Manfaat Penelitian

3. Bagaimana perbandingan kevalidan antara hasil yang diberikan oleh Algoritma C4.5 dan Algoritma C5.0?

Latar Belakang

Rumusan Masalah

Batasan Masalah

Tujuan Penelitian

Manfaat Penelitian

- 1. Objek yang digunakan pada penelitian ini adalah 11 atribut yang berasal dari 200 data kecelakaan kerja.**

Latar Belakang

Rumusan Masalah

Batasan Masalah

Tujuan Penelitian

Manfaat Penelitian

2. **Dalam penelitian ini, terdapat 3 kategori objek, yaitu Event, Karyawan, dan Sistem Perusahaan.**

Latar Belakang

Rumusan Masalah

Batasan Masalah

Tujuan Penelitian

Manfaat Penelitian

3. Data yang digunakan pada penelitian ini adalah data dari kantor produksi PT. Ajinomoto Indonesia cabang Mojokerto.

Latar Belakang

Rumusan Masalah

Batasan Masalah

Tujuan Penelitian

Manfaat Penelitian

UMUM

Implementasi algoritma C4.5 dan C5.0 dalam perangkat lunak untuk mendapatkan pola kecelakaan kerja sehingga dapat diambil keputusan dalam tindakan preventifnya

KHUSUS

- Membuat sebuah perangkat lunak yang memuat algoritma C4.5 dan C5.0 sehingga dapat membuat pola dari data training
- Mencari perbandingan tingkat keakuratan hasil pengklasifikasian dari algoritma C4.5 dan C5.0 dalam perangkat lunak

Latar Belakang

Rumusan Masalah

Batasan Masalah

Tujuan Penelitian

Manfaat Penelitian

1. Diperoleh informasi dan perbandingan mengenai pola kecelakaan kerja menggunakan metode *classification* yaitu C4.5 dan C5.0.

Latar Belakang

Rumusan Masalah

Batasan Masalah

Tujuan Penelitian

Manfaat Penelitian

2. Diperoleh suatu metode untuk mengambil tindakan preventif untuk terjadinya kecelakaan kerja pada perusahaan yang menjadi studi kasus.

Latar Belakang

Rumusan Masalah

Batasan Masalah

Tujuan Penelitian

Manfaat Penelitian

3. Sebagai salah satu referensi penggunaan *classification* dalam penggalian pola dengan Algoritma C4.5 dan Algoritma C5.0.

Penelitian
Terdahulu

**Explaining And Predicting
Workplace Accidents Using
Data Mining Techniques
Dari
T. Rivas, M. Pas, J.M. Matias,
JF Garcia, J.Taboada**

Penelitian
Terdahulu

**Are Decision Tree Always
Greener On The Open (Source)
Side Of The Fence?**
Dari
**Samuel A. Moore, Daniel M D
Addario, James Kurinskas dan
Gary M. Weiss**

Penelitian
Terdahulu

**A Method For Classification Of
Network Traffic Based On C5.0
Machine Learning Algorithm
Dari
Bujlow, Tomasz Riaz, Tahir
Pederson, dan Jens Myrup**

Penelitian Terdahulu

*A Novel Hybrid Intelligent
Method Based On C4.5 Decision
Tree Classifier And One-against-
all Approach For Multi-class
Classification Problem*

DATA MINING

Data mining merupakan proses ekstraksi pola yang penting dari data dalam jumlah besar . *Data mining* merupakan salah satu langkah dalam proses *Knowledge Discovery in Database* (KDD) untuk menemukan pola yang bermanfaat.

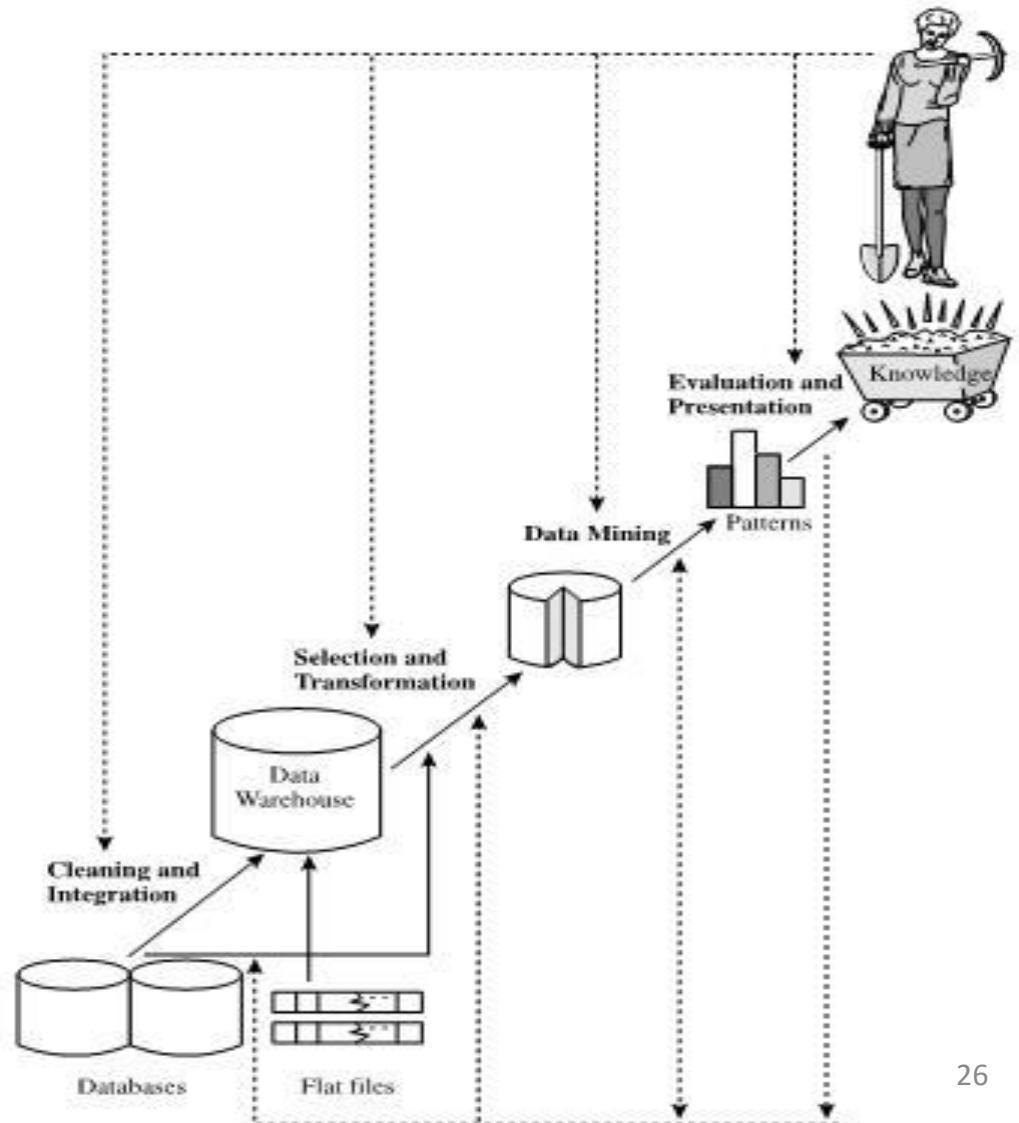
Beberapa karakteristik umum yang dimiliki data yang akan dianalisis adalah sebagai berikut [4]:

- Ukuran Data Besar .
- Data tidak lengkap, sehingga perlu proses cleaning.
- Struktur data kompleks.
- Merupakan kumpulan data yang heterogen.

KNOWLEDGE DISCOVERY IN DATABASE

Proses menemukan informasi yang bermanfaat serta pola-pola yang ada dalam data. KDD merupakan sebuah proses yang terdiri dari serangkaian proses iteratif yang terurut dan *data mining* merupakan salah satu langkah dalam KDD

KNOWLEDGE DISCOVERY IN DATABASE



CLASSIFICATION



Classification adalah proses penemuan model (fungsi) yang menggambarkan dan membedakan kelas data atau konsep yang bertujuan agar bisa digunakan untuk memprediksi kelas dari objek yang label kelasnya tidak diketahui .

Classification data terdiri dari 2 langkah proses, yaitu :

- *Learning (Fase Training)*
- *Classification*

DECISION TREE ALGORITHM

Decision tree adalah salah satu metode *classification* yang paling popular karena mudah untuk diinterpretasi oleh manusia . Metode ini digunakan untuk memprediksi nilai diskrit dari fungsi target.

DECISION TREE ALGORITHM

Terdapat 3 jenis node yang terdapat pada *decision tree* yaitu:

- *Root node*, merupakan node paling atas, pada node ini tidak ada input dan bisa tidak mempunyai output atau mempunyai output lebih dari satu.
- *Internal Node*, merupakan node percabangan. Pada node ini terdapat percabangan. Pada node ini terdapat satu input dan memiliki output minimal dua.
- *Leaf node* atau *terminal node*, merupakan node akhir. Pada node ini terdapat satu input dan tidak mempunyai output

ALGORITMA
C4.5

- Hitung setiap kemunculan kejadian
- Hitung *entropy total* dan *entropy* untuk setiap *value* atribut
- Hitung *information gain* untuk tiap atribut. *Information gain* inilah yang akan digunakan untuk membuat *tree*.
- Bentuk simpul pertama yang berisi atribut yang memiliki *information gain* tertinggi.

ALGORITMA C5.0

C.5.0 merupakan penyempurnaan algoritma sebelumnya yang dibentuk oleh Ross Quinlan pada tahun 1987, yaitu ID3 dan C4.5. Dalam algoritma ini pemilihan atribut diproses menggunakan *gain ratio*.

Langkah kerja pembangunan tree pada algoritma C5.0 mirip dengan pembangunan algoritma C 4.5.

INFORMATION
GAIN
GAIN RATIO

Ukuran *information gain* dan *gain ratio* digunakan untuk memilih atribut uji pada setiap *node* di dalam *tree*. Ukuran ini digunakan untuk memilih atribut atau *node* pada *tree*.

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

(1)

INFORMATION
GAIN
GAIN RATIO

dengan :

S = himpunan kasus

n = jumlah partisi S

p_i = proporsi dari S_i terhadap S

INFORMATION
GAIN
GAIN RATIO

$$\begin{aligned} \text{InformationGain}(S, A) = \\ \text{Entropy}(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * \text{Entropy}(S_i) \end{aligned} \quad (2)$$

Dimana :

- S = himpunan kasus
- A = Atribut
- n = jumlah artisi atribut A
- $|S_i|$ = jumlah kasus pada partisi ke-i
- $|S|$ = jumlah kasus dalam S

INFORMATION
GAIN
GAIN RATIO

Perhitungan *gain ratio* untuk algoritma C5.0 akan berjalan setelah perhitungan *information gain* diatas dilakukan. Perhitungan *gain ratio* selanjutnya menggunakan rumus dibawah ini :

$$GainRatio = \frac{Information\ Gain\ (S, A)}{\sum_{i=1}^n Entropy\ (Si)} \quad (3)$$

Studi
Literatur

7/25/2016

INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA

Pengumpulan
Data

Pengolahan
Data

Perancangan
Interface
dan
Implementasi
Sistem

Evaluasi
dan
Penarikan
Kesimpulan

Studi Literatur ini dilakukan untuk identifikasi permasalahan dengan mencari referensi yang menunjang penelitian yang berupa tugas akhir, jurnal, buku, maupun artikel. Dalam tugas akhir ini, studi literatur yang dilakukan mengenai *data mining*, *classification*, Algoritma C4.5, dan Algoritma C5.0

PENGUMPU
LAN DATA

- | |
|---|
| <ul style="list-style-type: none">• On Work : Kecelakaan kerja yang terjadi saat karyawan sedang dalam proses pekerjaan dan berada dalam area perusahaan |
| <ul style="list-style-type: none">• Traffic Inside Fact : Kecelakaan yang terjadi selama di area perusahaan yang tidak termasuk On Work |
| <ul style="list-style-type: none">• Traffic On Commute : Kecelakaan kerja yang terjadi saat karyawan berada di luar area perusahaan, contoh : waktu berangkat kerja, waktu pulang kerja, dinas luar. |

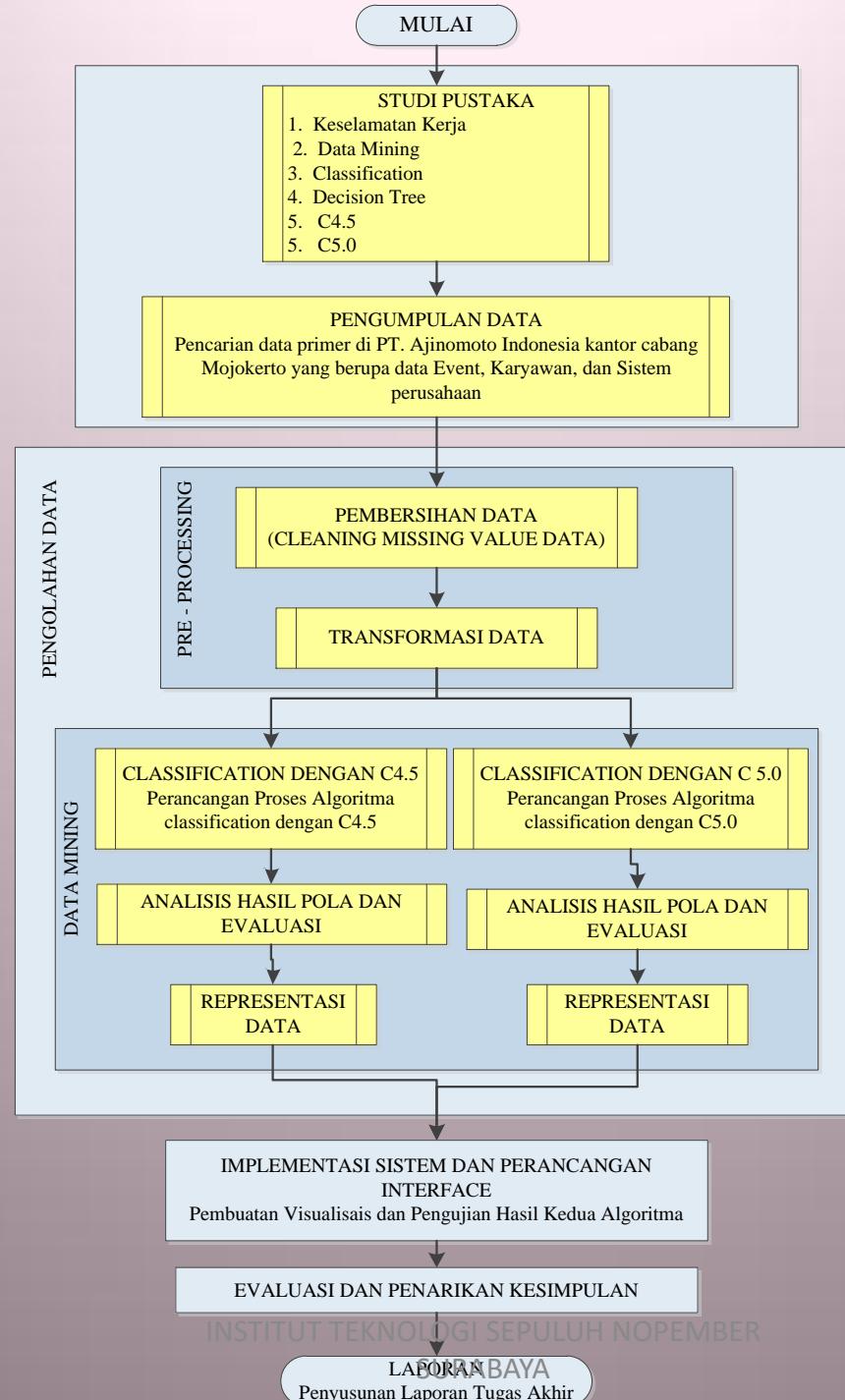
PENGOLAH AN DATA

Pre Processing

- Data Cleaning
- Transformasi data

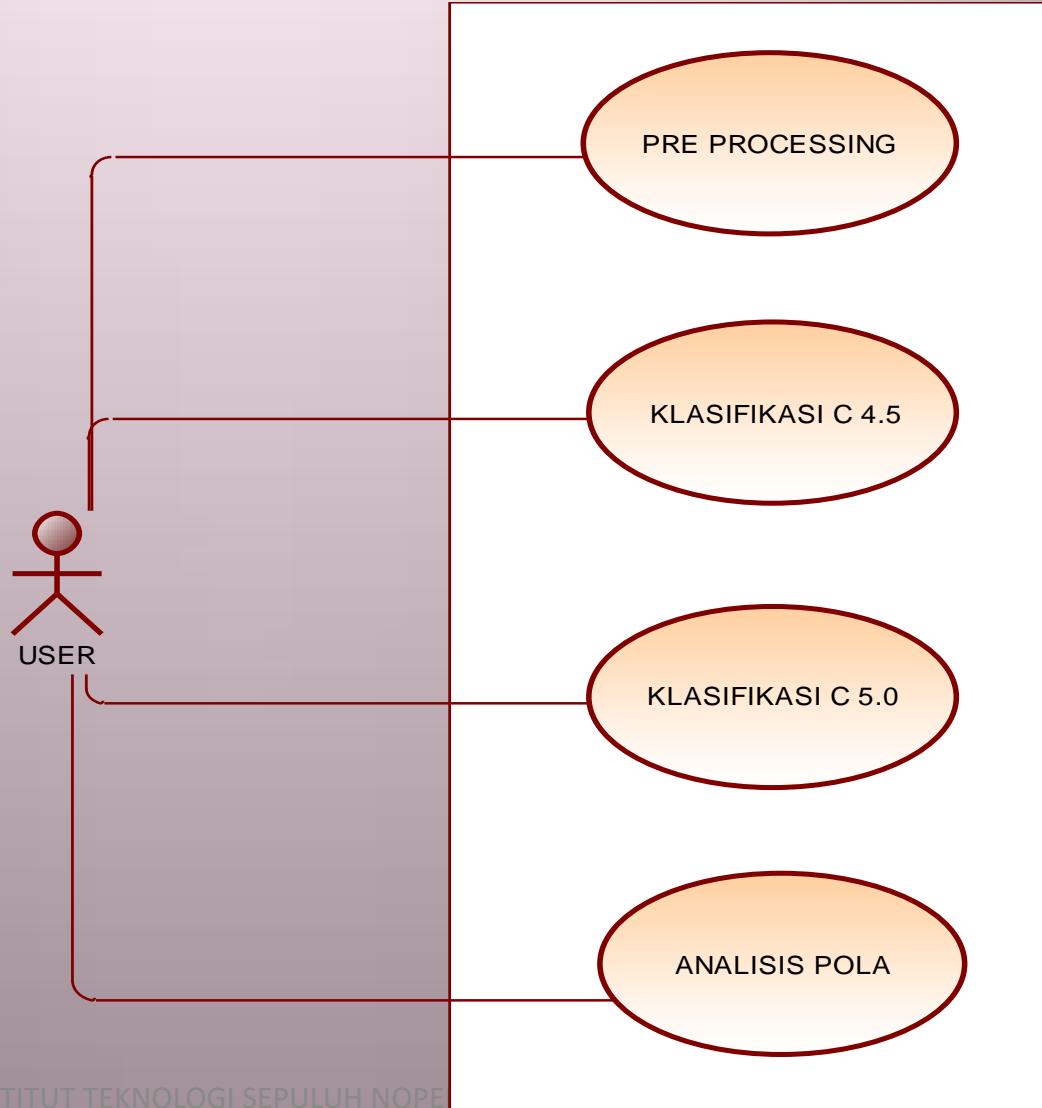
Data Mining

- Perancangan Algoritma
- Pengujian Algoritma
- Analisis Hasil Pola



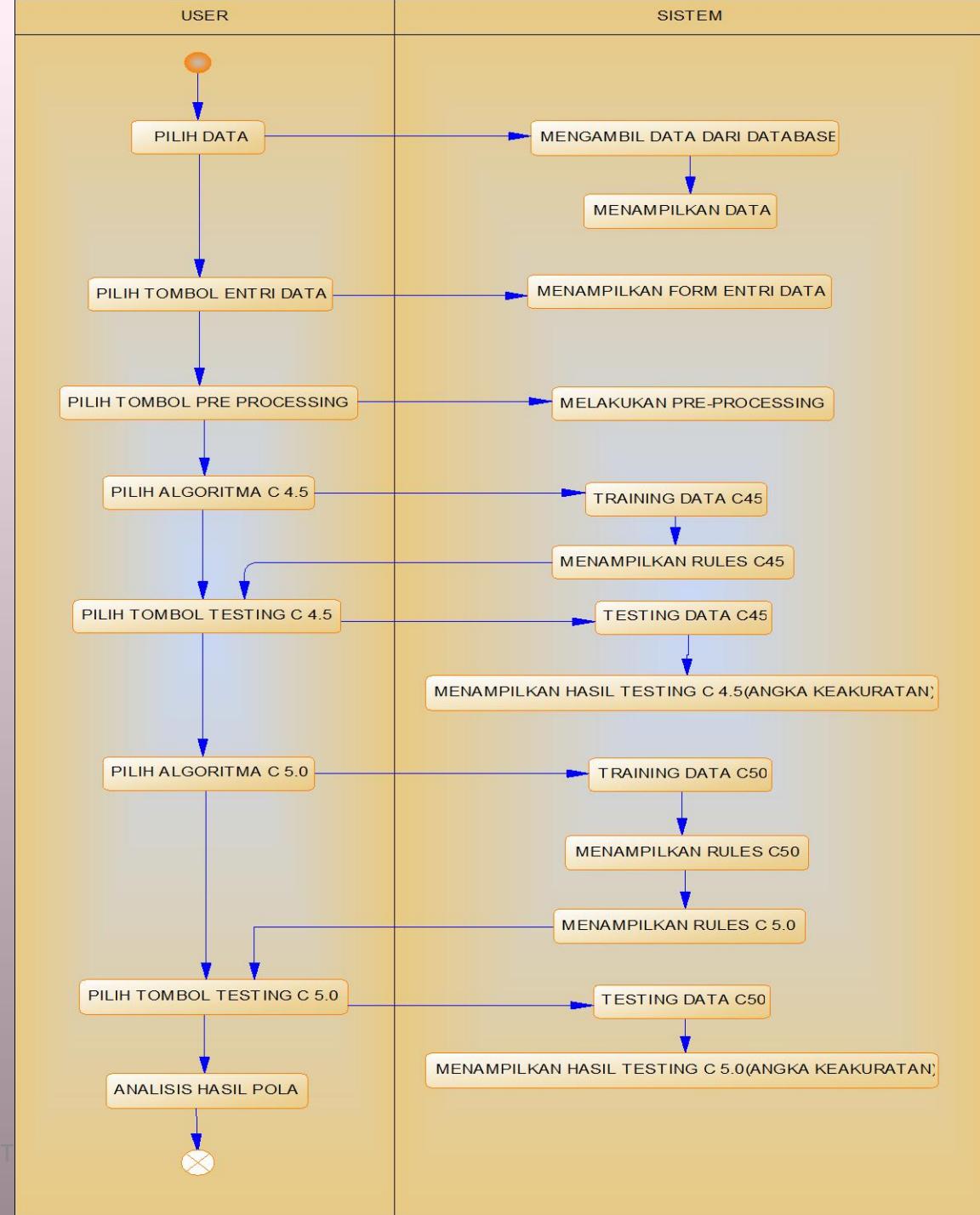
PERANCANGAN DAN IMPLEMENTASI SISTEM

ANALISI SISTEM



PERANCANGAN DAN IMPLEMENTASI SISTEM

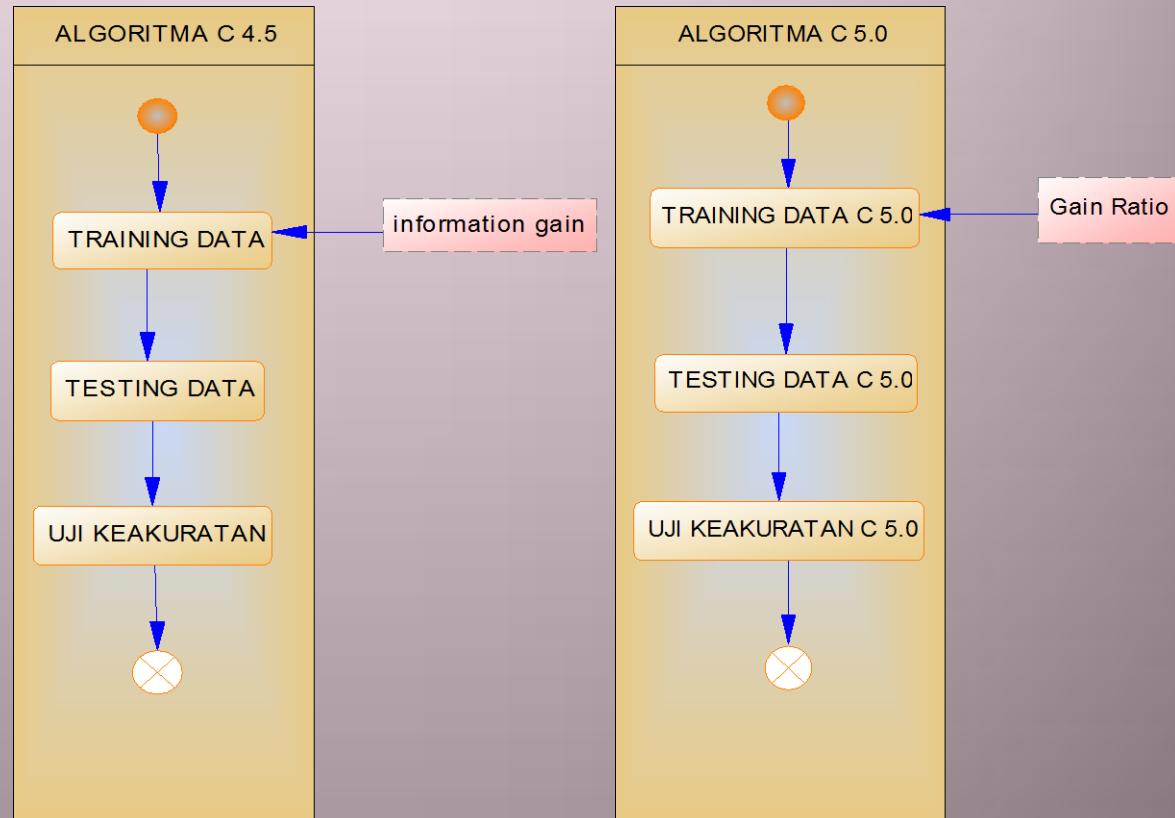
ANALISIS SISTEM



PERANCANGAN DAN IMPLEMENTASI SISTEM

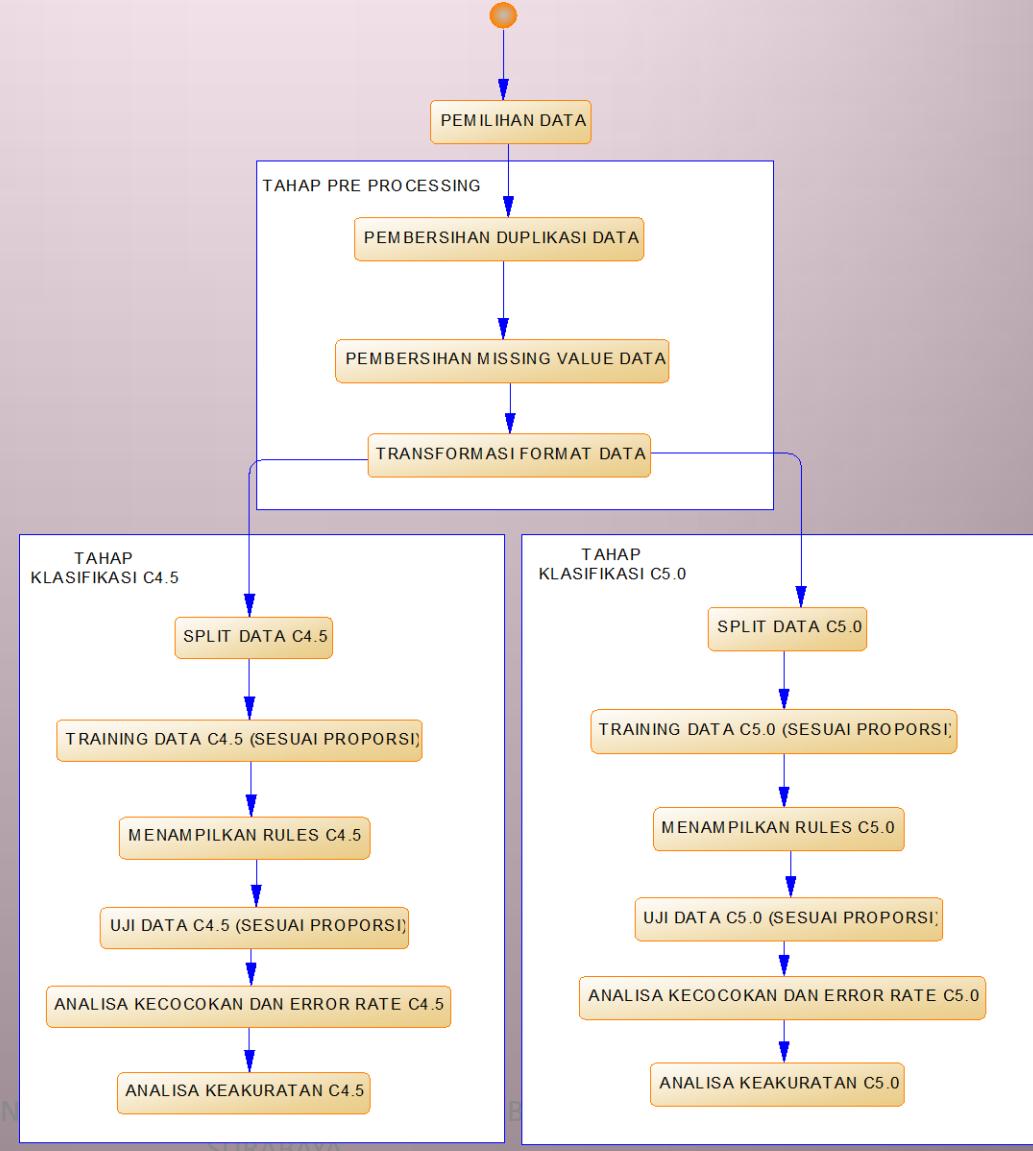
ANALISI SISTEM

M



PERANCANGAN DAN IMPLEMENTASI SISTEM

ANALISIS SISTEM



PERANCANGAN DAN IMPLEMENTASI SISTEM

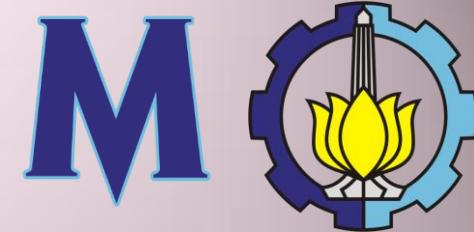
PERANCANGAN ALGORITM A



	A	B	C	D	E	F	G	H	I	J	K	L
1	2031613145	MM	L	40	INDONESI	GRD C	00:00.0	20	ISS	NO	TRAFFIC INSIDE FACT	
2	2051412261	R	L	24	INDONESI	GRD B	00:00.0	2	SWABINA	YES	TRAFFIC ON COMMUTE	
3	2101412431	Y	P	25	INDONESI	GRD B	00:00.0	3	SWABINA	NO	TRAFFIC INSIDE FACT	
4	3011311211	H	L	50	INDONESI	GRD A	00:00.0	26	KONTRAKTOR	YES	TRAFFIC INSIDE FACT	
5	3021612401	LL	L	24	INDONESI	GRD B	00:00.0	2	FUJI	YES	TRAFFIC ON COMMUTE	
6	4051312049	KL	L	41	INDONESI	GRD B	00:00.0	19	REGULER	YES	TRAFFIC INSIDE FACT	
7	4081113170	DG	L	27	INDONESI	GRD C	00:00.0	3	SWABINA	YES	TRAFFIC INSIDE FACT	
8	5011413803	GD	L	31	INDONESI	GRD C	00:00.0	7	SWABINA	NO	ON WORK	
9	7031311139	J	L	39	INDONESI	GRD A	00:00.0	17	SWABINA	NO	ON WORK	
10	8021512313	DD	L	24	INDONESI	GRD B	00:00.0	2	FUJI	NO	TRAFFIC ON COMMUTE	
11	8031311212	I	L	23	INDONESI	GRD A	00:00.0	3	FUJI	NO	ON WORK	
12	8041413320	GT	L	29	INDONESI	GRD C	00:00.0	5	KONTAKTOR	NO	TRAFFIC INSIDE FACT	
13	8061511091	HH	L	23	INDONESI	GRD A	00:00.0	3	FUJI	NO	TRAFFIC ON COMMUTE	
14	8081213079	E	L	35	INDONESI	GRD C	00:00.0	13	REGULER	NO	ON WORK	
15	9031412097	IY	L	49	INDONESI	GRD B	00:00.0	29	REGULER	YES	TRAFFIC INSIDE FACT	
16	9041412930	EB	L	34	INDONESI	GRD B	00:00.0	12	FUJI	YES	TRAFFIC ON COMMUTE	
17	9061512054	II	L	31	INDONESI	GRD B	00:00.0	10	REGULER	NO	ON WORK	
18	9121512205	KK	L	30	INDONESI	GRD B	00:00.0	9	REGULER	YES	TRAFFIC ON COMMUTE	
19	10041313341	IK	L	31	INDONESI	GRD C	00:00.0	10	REGULER	YES	TRAFFIC ON COMMUTE	
20	10081511139	JJ	L	24	INDONESI	GRD A	00:00.0	2	REGULER	NO	TRAFFIC ON COMMUTE	
21	10101412210	Z	L	24	INDONESI	GRD B	00:00.0	4	FUJI	NO	TRAFFIC ON COMMUTE	
22	11091311012	M	L	30	INDONESI	GRD A	00:00.0	7	SWABINA	YES	TRAFFIC ON COMMUTE	

PERANCAN GAN DAN IMPLEME NTASI SISTEM

PERANCAN GAN ALGORITM A



phpMyAdmin

Server: 127.0.0.1 > Database: tugas-akhir

Structure SQL Search Query Export Import Operations Privileges Routines Events

Table	Action	Rows	Type	Collation	Size	Overhead
data_karyawan	Browse Structure Search Insert Empty Drop	76	InnoDB	latin1_swedish_ci	16 KiB	-
data_karyawan1	Browse Structure Search Insert Empty Drop	76	InnoDB	latin1_swedish_ci	16 KiB	-
jenis_data	Browse Structure Search Insert Empty Drop	18	InnoDB	latin1_swedish_ci	16 KiB	-
3 tables	Sum	170	InnoDB	latin1_swedish_ci	48 KiB	0 B

Check All With selected:

Print view Data Dictionary

Create table

Name: Number of columns: 4

Server: 127.0.0.1 > Database: tugas-akhir > Table: data_karyawan

Browse Structure SQL Search Insert Export Import Privileges Operations Tracking Triggers

#	Name	Type	Collation	Attributes	Null	Default	Extra	Action
1	id	varchar(15)	latin1_swedish_ci		No			Change Drop Primary Unique Index Spatial Fulltext More
2	Nama	varchar(30)	latin1_swedish_ci		No	None		Change Drop Primary Unique Index Spatial Fulltext More
3	jenis_kelamin	char(1)	latin1_swedish_ci		No	None		Change Drop Primary Unique Index Spatial Fulltext More
4	Umur	int(2)			No	None		Change Drop Primary Unique Index Spatial Fulltext More
5	kewarganegaraan	varchar(10)	latin1_swedish_ci		Yes	NULL		Change Drop Primary Unique Index Spatial Fulltext More
6	tipe_pekerjaan	varchar(5)	latin1_swedish_ci		No	None		Change Drop Primary Unique Index Spatial Fulltext More
7	waktu_per_hari	time(2)			Yes	NULL		Change Drop Primary Unique Index Spatial Fulltext More
8	lama_bekerja	int(4)			Yes	NULL		Change Drop Primary Unique Index Spatial Fulltext More
9	status	text	latin1_swedish_ci		No	None		Change Drop Primary Unique Index Spatial Fulltext More
10	pelatihan	char(5)	latin1_swedish_ci		Yes	NULL		Change Drop Primary Unique Index Spatial Fulltext More
11	kecelakaan_kerja	varchar(20)	latin1_swedish_ci		No	None		Change Drop Primary Unique Index Spatial Fulltext More

PERANCANGAN DAN IMPLEMENTASI SISTEM

PERANCANGAN ALGORITM A



Sebelum Transformasi	Sesudah Transformasi
<ol style="list-style-type: none">1. Id2. Nama3. Jenis kelamin4. Umur5. Kewarganegaraan6. Tipe pekerjaan7. Waktu per hari8. Lama bekerja9. Status10. Pelatihan11. Kecelakaan Kerja	<ol style="list-style-type: none">1. Id2. Jenis Kelamin3. Umur4. Tipe pekerjaan5. Lama bekerja6. Status7. Pelatihan8. Kecelakaan Kerja

PERANCANGAN DAN IMPLEMENTASI SISTEM

PERANCANGAN ALGORITMA

7/25/2016



Node			Jml Kasus	On Work	Traffic on Commute	Traffic inside fact
	Total		?	?	?	?
	Jenis Kelamin	L	?	?	?	?
		P	?	?	?	?
	Umur	Umur1	?	?	?	?
		Umur2	?	?	?	?
		Umur3	?	?	?	?
		Umur4	?	?	?	?
	Tipe Pekerjaan	GRD A	?	?	?	?
		GRD B	?	?	?	?
		GRD C	?	?	?	?
	Lama Bekerja	Junior	?	?	?	?
		Senior	?	?	?	?
	Status	ISS	?	?	?	?
		Reguler	?	?	?	?
		Fuji	?	?	?	?
		Swabina	?	?	?	?
		Kontraktor	?	?	?	?
	Pelatihan	Yes	?	?	?	?
		No	?	?	?	?

PERANCANGAN DAN IMPLEMENTASI SISTEM

PERANCANGAN ALGORITM A

Node			Entropy	Information Gain
Total			?	
Jenis Kelamin	L		?	?
	P		?	
Umur	Umur1		?	?
	Umur2		?	
	Umur3		?	
	Umur4		?	
Tipe Pekerjaan	GRD A		?	?
	GRD B		?	
	GRD C		?	
Lama Bekerja	Junior		?	?
	Senior		?	
Status	ISS		?	?
	Reguler		?	
	Fuji		?	
	Swabina		?	
Pelatihan	Kontraktor		?	?
	Yes		?	
	No		?	



PERANCANGAN DAN IMPLEMENTASI SISTEM

PERANCANGAN ALGORITM A

7/25/2016



Node			Entropy	Information Gain	Gain Ratio
	Total		?		
Jenis Kelamin	L		?	?	?
	P		?		
Umur	Umur1		?	?	?
	Umur2		?		
	Umur3		?		
	Umur4		?		
Tipe Pekerjaan	GRD A		?	?	?
	GRD B		?		
	GRD C		?		
Lama Bekerja	Junior		?	?	?
	Senior		?		
Status	ISS		?	?	?
	Reguler		?		
	Fuji		?		
	Swabina		?		
	Kontraktor		?		
Pelatihan	Yes		?	?	?
	No		?		

SURABAYA

PEMBAGIAN DATA UNTUK TESTING DAN TRAINING

No	Data Training	Data Testing
1	95%	5%
2	90%	10%
3	85%	15%
4	80%	20%
5	75%	25%

PENGUJIAN DAN PEMBAHASAN



TRAINING C4.5 DENGAN 95% DATA TRAINING

ALGORITMA C 4.5

95

TRAINING

RULES

UJI

```
<----Perhitungan pada cabang ISS----->
Gain jenis_kelamin1 0.0
Gain Umur 0.0
Gain tipe_pekerjaan 0.6500224216483541
Gain lama_bekerja 0.0
Gain pelatihan 0.0
Cabang Selanjutnya adalah tipe_pekerjaan

<----Pengambilan keputusan pada cabang GRD C adalah TRAFFIC INSIDE FACT
<----Pengambilan keputusan pada cabang GRD A adalah TRAFFIC ON COMMUTE
<----Perhitungan pada cabang REGULER----->
Gain jenis_kelamin1 0.0
Gain Umur 0.1878810667094395
Gain tipe_pekerjaan 0.15105426046442028
Gain lama_bekerja 0.13039400326884087
```

RULES	DECISION
ISS --> GRD C	TRAFFIC INSIDE FACT
ISS --> GRD A	TRAFFIC ON COMMUTE
REGULER --> YES --> umur2--> GRD B	TRAFFIC ON COMMUTE
REGULER --> YES --> umur4--> GRD B	TRAFFIC INSIDE FACT
REGULER --> YES --> umur4--> GRD C	TRAFFIC INSIDE FACT
REGULER --> YES --> umur4--> GRD A	TRAFFIC INSIDE FACT
REGULER --> NO --> umur1	TRAFFIC ON COMMUTE
REGULER --> NO --> umur2 --> GRD C	ON WORK
REGULER --> NO --> umur3 --> GRD B	TRAFFIC ON COMMUTE
REGULER --> NO --> umur3 --> GRD C	TRAFFIC ON COMMUTE
REGULER --> NO --> umur4 --> GRD B	ON WORK
REGULER --> NO --> umur4 --> GRD C	ON WORK
REGULER --> NO --> umur4 --> GRD A	TRAFFIC ON COMMUTE
FUJI--> GRD B --> umur1	TRAFFIC ON COMMUTE
FUJI--> GRD B --> umur2 --> YES	TRAFFIC ON COMMUTE
FUJI--> GRD B --> umur3	TRAFFIC ON COMMUTE
FUJI--> GRD B --> umur4	TRAFFIC ON COMMUTE
FUJI--> GRD C --> umur2	TRAFFIC INSIDE FACT
FUJI--> GRD C --> umur3	TRAFFIC ON COMMUTE
FUJI--> GRD A --> umur1 --> YES	TRAFFIC ON COMMUTE
FUJI--> GRD A --> umur1 --> NO --> D	TRAFFIC ON COMMUTE

PENGUJIAN DAN PEMBAHASAN



TESTING C4.5 DENGAN 5% DATA TESTING

UJI ALGORITMA C4.5

TAMPIL PROSES

Id	Jenis_Kelamin	Umur	Tipe_Pekerjaan	Lama_Bekerja	Status	Pelatihan	Kecelakaan_Kerja	Klasifikasi
02031613145	L	umur3	GRD C	Senior	ISS	NO	TRAFFIC INSIDE	cocok
02051412261	L	umur1	GRD B	Junior	SWABINA	YES	TRAFFIC ON CALL	cocok
02101412431	P	umur1	GRD B	Junior	SWABINA	NO	TRAFFIC INSIDE	cocok
03011311211	L	umur4	GRD A	Senior	KONTRAKTOR	YES	TRAFFIC INSIDE	cocok
03021612401	L	umur1	GRD B	Junior	FUJI	YES	TRAFFIC ON CALL	cocok
04051312049	L	umur3	GRD B	Senior	REGULER	YES	TRAFFIC INSIDE	tidak cocok
04081113170	L	umur2	GRD C	Junior	SWABINA	YES	TRAFFIC INSIDE	tidak cocok
05011413803	L	umur2	GRD C	Junior	SWABINA	NO	ON WORK	cocok
050913111940	L	umur3	GRD A	Senior	FUJI	NO	TRAFFIC ON CALL	cocok
07031311139	L	umur3	GRD A	Senior	SWABINA	NO	ON WORK	tidak cocok

KECOCOKAN

AKURASI

ERROR RATE

PENGUJIAN DAN PEMBAHASAN

ALGORITMA
C4.5



Pembagian Data	Hasil
95% - 5%	70%
90% - 10%	60%
85% - 15%	46%
80% - 20 %	52%
75% - 25 %	46%

PENGUJIAN DAN PEMBAHASAN

ALGORITMA
C5.0



Pembagian Data	Hasil
95% - 5%	60%
90% - 10%	45%
85% - 15%	46%
80% - 20 %	32%
75% - 25 %	32%

1. Perangkat lunak dapat melakukan *pre processing*, pembangunan *tree* dan pengujian algoritma sehingga didapatkan pola prediksi kecelakaan kerja beserta angka keakuratannya.
2. *Rules* sebagai representasi dari pola kecelakaan kerja yang dihasilkan oleh Algoritma C5.0 lebih ringkas daripada *rules* yang dihasilkan oleh C4.5, yaitu 32 pola kecelakaan kerja.
3. Hasil keakuratan pola yang dihasilkan oleh algoritma C4.5 lebih tinggi dibanding dengan algoritma C 5.0, yaitu 70%.

1. Data yang dipakai dalam perangkat lunak ini masih sedikit dan bervariasi. Dalam perhitungan data mining, sebaiknya data harus dalam jumlah yang banyak karena mempengaruhi pola dan hasil keakuratan yang didapatkan. Untuk penelitian selanjutnya dapat menggunakan data yang lebih banyak.
2. Perangkat lunak ini belum bisa menampilkan hasil *decision tree* yang menarik. Untuk penelitian selanjutnya dapat menggunakan representasi data dalam bentuk *tree* atau *list rules* yang lebih baik sehingga memudahkan dalam testing data.
3. Pengembangan perangkat lunak ini dapat dilakukan dengan menambah menu input data baru demi manfaat yang lebih baik.

1. Bujlow, Tomasz, Tahir Riaz dkk. 2012. *A Method for Classification of Network Traffic on C5.0 Machine Learning Algorithm*. ICNC'12 : 2012 International Conference on Computing, Networking and Communication (ICNC). (pp 237-241). IEEE
2. Dunham, MH. 2003. *Data Mining Introductory and Andvanced Topics*. Upper Saddle River. NJ : Pearson Education, Inc
3. Gorunescu, Florin. 2011. *Data Mining: Concepts, Models, and Techniques*. Verlag Berlin Heidelberg. Springer
4. Han, J, Micheline K, Jian Pei. 2012. *Data Mining Concept and Techniques 3rd Edition*. Morgan Kaufman Publisher. USA
5. I. H. Witten and E. Frank. 2005. *Data Mining: Practical Machine Learning Tools and Techniques, 2nd Edition*. Morgan Kaufmann, San Francisco
6. Kantardzic, M. 2003. *Data Mining Concept Models, Methods, and Algorithm*. New Jersey, USA : A John Wiley & Sons
7. Moore, Samuel. A, Daniel M. D Addario, dkk. 2011. *Are Decision Trees Always Greener on the Open (Source) Side of the Fence?*. Sciweaver

8. Permatasari, Anggun. 2009. *Investigasi Kecelakaan Penyebranan Perlintasan KRL UI-Margonda Depok*. Tugas Akhir Fakultas Kesehatan Masyarakat. Universitas Indonesia
9. Polat, Kemal, Gunes, Salih. 2009. *A Novel Hybrid Intelligent Method Based on C4.5 Decision Tree Classifier and One-Against-All Approach for Multi-Class Classification Problem*. (pp 1587-1592). Elsevier
10. Quinlan, J. R. 1993. *C4.5: Programs for Machine Learning*. San Mateo CA : Morgan Kaufmann
11. Rivas, T, M. Paz, dkk. 2011. *Explaining and Predicting Workplace Accident Using Data-Mining Techniques*. (pp 739-747). Elsivier
12. Tan, Pang-Ning, Michael Steinbach, Vipin Kumar. 2006. Introduction of Data Mining. Boston : Pearson – Addison Wesley
13. Wu, Xindong, Kumar, Vipin, dkk. 2008. *Top 10 Algorithms in Data Mining*. Knowl Inf Systm (2008) 14 : 1-37. Springer-Verlag London Limited



Thank you...