# KLASIFIKASI TINGKAT LUKA KORBAN KECELAKAAN LALU LINTAS DENGAN METODE *CHI-SQUARE AUTOMATIC INTERACTION DETECTION (CHAID) WITH SYNTHETIC MINORITY OVERSAMPLING TECHNIQUE (SMOTE)*

(Studi Kasus : Data Kecelakaan Lalu Lintas di Kabupaten POLRES Sleman Tahun 2016 sampai 2017)

## TUGAS AKHIR




Diusulkan oleh:

**Bana Ali Fikri**

14611082

**JURUSAN STATISTIKA**

**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM**

**UNIVERSITAS ISLAM INDONESIA**

**YOGYAKARTA**

**2018**

# KLASIFIKASI TINGKAT LUKA KORBAN KECELAKAAN LALU LINTAS DENGAN METODE *CHI-SQUARE AUTOMATIC INTERACTION DETECTION (CHAID) WITH SYNTHETIC MINORITY OVERSAMPLING TECHNIQUE (SMOTE)*

(Studi Kasus : Data Kecelakaan Lalu Lintas di Kabupaten POLRES Sleman Tahun 2016 sampai 2017)

## TUGAS AKHIR

Diajukan Sebagai Salah Satu Syarat Untuk memperoleh Gelar Sarjana Jurusan Statistika




Diusulkan oleh:

**Bana Ali Fikri**

14611082

**JURUSAN STATISTIKA**

**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM**

**UNIVERSITAS ISLAM INDONESIA**

**YOGYAKARTA**

**2018**

## HALAMAN PERSETUJUAN PEMBIMBING
## TUGAS AKHIR

| | | |
|---|---|---|
| Judul | : | Klasifikasi Tingkat Luka Korban Kecelakaan Lalu Lintas dengan Metode *Chi-Square Automatic Interaction Detection (CHAID) with Synthetic Minority Oversampling Technique (Smote)* (Studi Kasus : Data Kecelakaan Lalu Lintas di Kabupaten POLRES Sleman Tahun 2016 sampai 2017) |
| Nama Mahasiswa | : | Bana Ali Fikri |
| Nomor Mahasiswa | : | 14 611 082 |

**TUGAS AKHIR INI TELAH DIPERIKSA DAN DISETUJUI UNTUK DIUJIKAN**

Yogyakarta, 7 April 2018

**Pembimbing**

**(Muhammad Muhajir, S.Si., M.Sc.)**

## HALAMAN PENGESAHAN

## TUGAS AKHIR

**KLASIFIKASI TINGKAT LUKA KORBAN KECELAKAAN LALU LINTAS DENGAN METODE *CHI-SQUARE AUTOMATIC INTERACTION DETECTION (CHAID) WITH SYNTHETIC MINORITY OVERSAMPLING TECHNIQUE (SMOTE)***

(Studi Kasus : Data Kecelakaan Lalu Lintas di Kabupaten POLRES Sleman Tahun 2016 sampai 2017

**Nama Mahasiswa : Bana Ali Fikri**

**Nomor Mahasiswa : 14 611 082**

**TUGAS AKHIR INI TELAH DIUJIKAN**

**PADA TANGGAL 7 APRIL 2018**

| **Nama Penguji** | **Tanda Tangan** |
|---|---|
| 1. Epha Diana Supandi, S.Si, M.Sc | ..... |
| 2. Prof. Akhmad Fauzy, M.Si., Ph.D | ..... |
| 3. Muhammad Muhajir, S.Si., M.Sc. | ..... |

Mengetahui,

Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

**Drs. Allwar, M.Sc., Ph.D**



iii

# KATA PENGANTAR



*Assalamu'alaikum Warahmatullahi Wabarakatuh*

*Alhamdulillah* Penulis ucapkan kepada Allah SWT, karena berkat limpahan Rahmat, Taufik, Hidayah, serta Inayah-Nya Penulis dapat menyelesaikan tugas akhir yang berjudul "**KLASIFIKASI TINGKAT LUKA KORBAN KECELAKAAN LALU LINTAS DENGAN METODE *CHI-SQUARE AUTOMATIC INTERACTION DETECTION (CHAID) WITH SYNTHETIC MINORITY OVERSAMPLING TECHNIQUE (SMOTE)*** (Studi Kasus : Data Kecelakaan Lalu Lintas di Kabupaten POLRES Sleman Tahun 2016 sampai 2017" dengan baik. Shalawat serta salam Penulis haturkan kepada Nabi Muhammad SAW beserta keluarga, sahabat, dan umatnya.

Penulis menyadari bahwa penulisan tugas akhir ini banyak memperoleh bantuan dari berbagai pihak, baik berupa saran, kritik, bimbingan maupun bantuan lainnya. Oleh karena itu, pada kesempatan ini Penulis menyampaikan ucapan terima kasih kepada:

1. Bapak Drs. Allwar, M.Sc., Ph.D., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam yang telah memberikan izin penelitian.
2. Bapak Dr. RB Fajriya Hakim, S.Si., M.Si., selaku Ketua Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Indonesia yang juga telah banyak membantu dan memberi ilmu dan wawasan baru kepada Penulis.
3. Bapak Prof. Akhmad Fauzy, S.Si., M.Si, Ph.D., selaku dosen pembimbing Akademik.
4. Bapak Muhammad Muhajir., S.Si., M.Sc., selaku dosen pembimbing Tugas Akhir, atas bimbingan beliau pada penelitian ini.

5. POLRES Kabupaten Sleman, yang telah mengijinkan penulis untuk melakukan penelitian dan pengambilan data penelitian.
6. Kedua orang tua Penulis, beserta seluruh keluarga besar Penulis yang luar biasa dan selalu memberikan dukungan, semangat, serta selalu mendo'akan Penulis agar diberi kemudahan dan kelancaran dalam penyusunan tugas akhir ini
7. Semua pihak yang telah membantu yang tidak dapat Penulis sebutkan satu persatu. Semoga Allah SWT selalu memberi rahmat dan anugerah-Nya kepada mereka semua tanpa henti. Aamiin.

Semoga semua pihak yang memberi dukungan dan bantuan mendapat balasan dari Allah SWT. Penulis menyadari bahwa dalam tugas akhir ini masih banyak kekurangan karena keterbatasan pengetahuan dan kemampuan yang dimiliki oleh Penulis. Oleh karena itu, kritik dan saran yang bersifat membangun diharapkan demi kesempurnaan penyusunan tugas akhir ini. Semoga tugas akhir ini bermanfaat bagi semua pihak.

*Wassalamu'alaikum Warahmatullahi Wabarakatuh*

Sleman, 3 April 2018

Penulis

# DAFTAR ISI

HALAMAN JUDUL ..... i
HALAMAN PERSETUJUAN PEMBIMBING ..... ii
HALAMAN PENGESAHAN ..... iii
KATA PENGANTAR ..... iv
DAFTAR ISI ..... vi
DAFTAR TABEL ..... ix
DAFTAR GAMBAR ..... x
DAFTAR ISTILAH ..... xi
DAFTAR SIMBOL ..... xiii
DAFTAR LAMPIRAN ..... xiv
PERNYATAAN ..... xv
INTISARI ..... xvi
*ABSTRACT* ..... xvii

BAB I PENDAHULUAN ..... 1
1.1 Latar Belakang Masalah ..... 1
1.2 Rumusan Masalah ..... 3
1.3 Batasan Masalah ..... 3
1.4 Tujuan Penelitian ..... 4
1.5 Manfaat Penelitian ..... 4
BAB II TINJAUAN PUSTAKA ..... 5

BAB III LANDASAN TEORI ..... 8
3.1 Definisi Kecelakaan Lalu Lintas ..... 8
3.2 Pengertian Kendaraan Bermotor ..... 8
3.3 Jenis Kecelakaan Lalu Lintas ..... 8

3.4 Faktor-Faktor Penyebab Kecelakaan Lalu Lintas ..... 11
3.5 *SMOTE* ..... 14
3.5 *Data Mining* ..... 14
3.6 Klasifikasi ..... 15
3.7 Data Training dan Data Testing ..... 16
3.8 Analisis *CHAID* ..... 16
3.9 Algoritma *CHAID* ..... 17
3.10 Koreksi Bonferroni ..... 21
3.11 Diagram Pohon Klasifikasi *CHAID* ..... 22
3.12 Evaluasi Klasifikasi ..... 23

BAB IV METODE PENELITIAN ..... 25
4.1 Populasi dan Sampel Penelitian ..... 25
4.2 Variabel dan Definisi Operasional Variabel ..... 25
4.3 Metode Pengumpulan Data ..... 28
4.4 Metode Analisis Data ..... 28
4.5 Tahapan Penelitian ..... 28

BAB V PEMBAHASAN ..... 30
5.1 Gambaran Umum Data Kecelakaan Lalu Lintas ..... 30
5.2 Pembagian Data *Training* dan Data *Testing* ..... 37
5.3 Analisis *CHAID* ..... 37
5.4 Tingkat Akurasi dan Presisi Metode *CHAID* ..... 40
5.5 Analisis *CHAID* dengan *SMOTE* ..... 41
5.6 Tingkat Akurasi dan Presisi Metode *CHAID* dengan *SMOTE* . 53
5.7 Validasi Sistem ..... 54

BAB VI PENUTUP ..... 55
6.1 Kesimpulan ..... 55

6.2 Saran ..... 56

DAFTAR PUSTAKA ..... 57

LAMPIRAN ..... 61

# DAFTAR TABEL

**Tabel 3.1** Ilustrasi Pasangan Penggabungan Variabel ..... 19
**Tabel 3.2** *Confusion Matrix*..... 24
**Tabel 4.1** Variabel Data..... 25
**Tabel 5.1** Tingkat Akurasi dan Presisi Metode *CHAID*..... 39
**Tabel 5.2** Tingkat Akurasi dan Presisi Metode *CHAID* dengan *SMOTE* ..... 53
**Tabel 5.3** Validasi Sistem..... 54

# DAFTAR GAMBAR

**Gambar 3.1** Ilustrasi *SMOTE* ..... 15
**Gambar 3.2** Diagram Pohon dalam Analisis *CHAID* ..... 22
**Gambar 4.1** Diagram Alir Algoritma *CHAID* ..... 29
**Gambar 5.1** Kecelakaan Lalu lintas Berdasarkan Waktu Kejadian ..... 30
**Gambar 5.2** Kecelakaan Berdasarkan Tempat Kejadian ..... 31
**Gambar 5.3** Kecelakaan Lalu Lintas Berdasarkan Jenis Kecelakaan ..... 31
**Gambar 5.4** Kecelakaan Lalu Lintas Berdasarkan Kendaraan Terlibat ..... 32
**Gambar 5.5** Kecelakaan Lalu Lintas Berdasarkan Umur ..... 33
**Gambar 5.6** Kecelakaan Lalu Lintas Berdasarkan Profesi ..... 33
**Gambar 5.7** Kecelakaan Lalu Lintas Berdasarkan Kepemilikan SIM ..... 34
**Gambar 5.8** Kecelakaan Lalu Lintas Berdasarkan Jenis Kelamin ..... 35
**Gambar 5.9** Kecelakaan Lalu Lintas Berdasarkan Pendidikan ..... 35
**Gambar 5.10** Kecelakaan Lalu Lintas Berdasarkan Tingkat Luka Korban ..... 36
**Gambar 5.11** Grafik Jumlah Kecelakaan di Musim Hujan dan Kemarau ..... 37
**Gambar 5.12** Pohon Keputusan Metode *CHAID* ..... 39

# DAFTAR ISTILAH

| | |
|---|---|
| POLRES | : Polisi Resort |
| DIY | : Daerah Istimewa Yogyakarta |
| *CHAID* | : *Chi-Squared Automatic Interaction Detection* |
| CART | : *Classification And Regression Trees* |
| BPS | : Badan Pusat Statistik |
| FMIPA | : Fakultas Matematika dan Ilmu Pengetahuan Alam |
| WHO | : *World Health Organization* |
| UNIROW | : Universitas PGRI Ronggolawe |
| UII | : Universitas Islam Indonesia |
| UNY | : Universitas Negeri Yogyakarta |
| IPK | : Indeks Prestasi Kumulatif |
| QUEST | : *Quick, Unbiased and Efficient. Statistical Tree* |
| KDD | : *Knowledge Discovery in Database* |
| SPSS | : *Statistical Package for the Social Sciences* |
| MATLAB | : *Matrix Laboratory* |
| SAS | : *Statistical Analysis System* |
| LB | : Luka Berat |

| | |
|---|---|
| LR | : Luka Ringan |
| MD | : Meninggal Dunia |
| Mbl | : Mobil |
| Spm | : Sepeda Motor |
| *SMOTE* | : *Synthetic Minority Oversampling Technique* |

# DAFTAR SIMBOL

| | | |
|---|---|---|
| M | : | Pengali Bonferroni |
| c | : | Banyaknya kategori variabel independent awal |
| r | : | Banyaknya kategori variabel independent setelah penggabungan |
| α | : | *Alpha* |
| $O_{ij}$ | : | Banyaknya pengamatan pada baris ke-i dan kolom ke-j (frekuensi yang diperoleh) |
| $E_{ij}$ | : | Nilai harapan pengamatan pada baris ke-i dan kolom ke-j (frekuensi yang diharapkan) |

# DAFTAR LAMPIRAN

**Lampiran 1** Data Kecelakaan Lalu Lintas ..... 61
**Lampiran 2** *Output* hasil klasifikasi *CHAID* ..... 62

## PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang sebelumnya pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 3 April 2018




Penulis



XV

# KLASIFIKASI TINGKAT LUKA KORBAN KECELAKAAN LALU LINTAS DENGAN METODE *CHI-SQUARE AUTOMATIC INTERACTION DETECTION (CHAID) WITH SYNTHETIC MINORITY OVERSAMPLING TECHNIQUE (SMOTE)*

(Studi Kasus : Data Kecelakaan Lalu Lintas di Kabupaten POLRES Sleman Tahun 2016 sampai 2017

Oleh : Bana Ali Fikri

Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Islam Indonesia

## INTISARI

Kecelakaan lalu lintas adalah peristiwa dijalan yang tidak disangka-sangka dan tidak disengaja melibatkan kendaraan dengan atau tanpa pemakai jalan lainnya mengakibatkan korban manusia atau kerugian harta benda. Kecelakaan lalu lintas menurut BPS tahun 2015 mencapai 98.9 ribu kasus. Angka tersebut meningkat 3.19 persen dibandingkan tahun sebelumnya yang mencapai 95.5 ribu kasus, sementara di Kabupaten Sleman angka kecelakaan pada semester pertama tahun 2017 sebagai kabupaten paling banyak kecelakaan di provinsi DIY sebanyak 627 kasus, maka perlu dilakukan penelitian yang bertujuan untuk mengurangi korban kecelakaan lalu lintas yang mengalami luka berat, meninggal dunia dan dapat meminimalisir atau mengurangi jumlah kecelakaan lalu lintas di POLRES Kabupaten Sleman. Salah satu alat untuk menganalisis adalah dengan menggunakan klasifikasi metode *CHAID*. *CHAID* adalah suatu teknik iteratif yang menguji satu-persatu variabel independen yang digunakan dalam klasifikasi, dan menyusunnya berdasarkan pada tingkat signifikansi statistik *chi-square* terhadap variabel dependennya. Berdasarkan hasil yang diperoleh *output* metode *CHAID* terdapat 46 segmen atau aturan. Segmen yang direkomendasikan terhadaap pihak kepolisian adalah segmen segmen 1, segmen 2, segmen 4, segmen 5, segmen 8, segmen 11, segmen 14, segmen 15, segmen 17, segmen 20, segmen 23, segmen 25, segmen 30, segmen 31, segmen 34, segmen 37, segmen 42, segmen 45. Model klasifikasi metode *CHAID* memiliki akurasi sebesar 81.5% dan memiliki proporsi yang cukup seimbang, maka model tersebut sudah cukup baik.

Kata Kunci : Kecelakaan Lalu Lintas, *CHAID*

# *CLASSIFICATION SEVERITY OF THE VICTIMS OF TRAFFIC ACCIDENTS WITH CHI-SQUARE AUTOMATIC INTERACTION DETECTION (CHAID) WITH SYNTHETIC MINORITY OVERSAMPLING TECHNIQUE (SMOTE)*

*(Case Study: Traffic Accident Data at POLRES Sleman District Year 2016 to 2017*

By : Bana Ali Fikri

Department of Statistics Faculty of Mathematics and Science

Islamic University of Indonesia

## *ABSTRACT*

*Traffic accidents are unexpected and unintentional street events involving vehicles with or without other road users resulting in human casualties or property losses. Traffic accidents by BPS in 2015 reached 98.9 thousand cases. The number increased by 3.19 percent compared to the previous year which reached 95.5 thousand cases, while in Sleman district the number of accidents in the first half of 2017 as the district most accidents in the province of Yogyakarta as many as 627 cases, it is necessary to conduct research that aims to reduce traffic casualties victims severely injured, died and can minimize or reduce the number of traffic accidents in POLRES Sleman District. One tool to analyze is to use CHAID method classification. CHAID is an iterative technique that tests one by one independent variables used in classification, and arrange it based on the level of statistical significance of chi-square to its dependent variable. Based on the results obtained CHAID method output there are 46 segments or rules. The segments recommended by the police are segment 1, segment 2, segment 4, segment 5, segment 8, segment 11, segment 14, segment 15, segment 17, segment 20, segment 23, segment 25, segment 30, segment 31, segment 34, segment 37, segment 42, segment 45. CHAID method classification model has an accuracy of 81.5% and has a fairly balanced proportion, then the model is good enough.*

*Keywords: Traffic Accident, CHAID*

# BAB I
# PENDAHULUAN

## 1.1 Latar Belakang

Transportasi merupakan alat yang memiliki tujuan untuk mengangkut manusia, hewan dan barang ketempat tujuan. Definisi lain dari transportasi adalah memindahkan manusia, hewan ataupun barang dari tempat asal ketempat tujuannya dengan bantuan berupa suatu alat yang dapat digerakan oleh mesin atau makhluk hidup. (Sora, 2015).

Di dunia transportasi yang paling sering digunakan adalah kendaraan bermotor. Kendaraan bermotor menjadi favorit oleh manusia pada saat ini, karena kemudahanya dalam menggunakan alat transportasi tersebut dan dapat dimiliki dengan uang yang tidak begitu banyak. Banyaknya kendaraan bermotor yang semakin bertambah sementara jumlah jalan tidak bertambah akan berdampak buruk pada peningkatan kecelakaan lalu lintas.

Berdasarkan WHO pada tahun 2013 ada 1,25 juta orang meninggal karena kecelakaan lalu lintas, hal ini menunjukan bahwa setiap satu orang terbunuh setiap 25 detik. Angka tersebut merupakan angka yang tinggi. Berdasarkan Data Badan Pusat Statistik (BPS) disebutkan pada tahun 2015, jumlah kecelakaan lalu lintas mencapai 98,9 ribu kasus. Angka tersebut meningkat 3,19 persen dibandingkan tahun sebelumnya yang mencapai 95,5 ribu kasus. Indonesia berada di peringkat ketiga di Asia di bawah Tiongkok dan India dengan total 38.279 total kematian akibat kecelakaan lalu lintas di tahun 2015. Berdasarkan data Indonesia menduduki peringkat ketiga tetapi dilihat dari presentase statistik dari jumlah populasi, Indonesia berada di peringkat pertama dengan angka kematian 0,015 persen dari jumlah populasi di bawah Tiongkok dengan presentase 0,018 persen dan India 0,017 persen. (Situmorang, 2016).

Daerah Istimewa Yogyakarta salah satu provinsi yang memiliki kepadatan penduduk yang tinggi sebesar 1174 jiwa/km$^2$. DIY masuk 4 besar dengan angka

kepadatan penduduk yang tinggi di Indonesia setelah DKI Jakarta, Jawa Barat dan Banten. Kepadatan penduduk yang tinggi tentunya memiliki dampak terhadap berbagai sektor. Sektor transportasi menjadi dampak terjadinya kepadatan penduduk yang tinggi. Pada tahun 2016 dari 1 Januari sampai 28 september jumlah kendaraan bermotor ada 84.312 unit. Kabupaten Sleman menjadi kabupaten terbanyak penambahan sepeda motor baru sebanyak 25.537 unit dan roda empat sebanyak 6.018 unit. Kabupaten Sleman sebagai kabupaten terbanyak penambahan kendaraan bermotor baru dan kabupaten paling padat penduduk kedua setelah kota Yogyakarta sebanyak 2031 jiwa/km$^2$ tentunya menimbulkan dampak buruk terhadap peningkatan kecelakaan lalu lintas. Angka kecelakaan pada semester pertama di tahun 2017 kabupaten sleman sebagai kabupaten paling banyak terjadi kecelakaan sebanyak 627 kasus (Nto, 2017).

Kecelakaan tentunya memiliki sebab dan tidak terjadi karena kebetulan. Oleh karena itu, sebab kecelakaan harus dianalisis dan dapat ditemukan, supaya tindakan korektif kepada penyebab itu bisa dilakukan serta dengan upaya preventif lebih lanjut kecelakaan bisa dicegah. Kecelakaan merupakan tindakan tidak direncanakan dan tidak terkendali, ketika aksi dan reaksi objek, bahan, atau radiasi menyebabkan cedera atau kemungkinan cedera (Heinrich, 1980).

Klasifikasi merupakan suatu proses yang memiliki tujuan untuk mempelajari fitur dari sebuah atau sekelompok data yang sudah diketahui kelasnya dan menghasilkan sebuah model klasifikasi yang dapat digunakan untuk memprediksi kelas atau label data yang baru (Liu, 2012). Di Statistika ada beberapa metode klasifikasi, diantaranya *CHAID* (*Chi- Square Automatic Interaction Detection*), *KNN* (K- *Nearest Neighbour*), *CART* (*Clasification and Regression Tree*), dan *Naïve Bayes*. Dalam penelitian ini akan digunakan metode *CHAID*.

Metode *CHAID* adalah metode yang relative baru. Dr. G. V. Kass adalah orang yang memperkenalkan metode ini untuk pertama kali pada sebuah artikel berjudul *"An Exploratory Technique for Investigating Large Quantities of Categorial Data"* dalam buku *Applied Statistics* di tahun 1980. Metode ini hanya efektif jika diterapkan pada

data dengan pengamatan yang sangat banyak (Du Toit et al, 1986). *CHAID* adalah sebuah metode untuk mengklasifikasikan data kategori dimana tujuan dari prosedurnya adalah untuk membagi rangkaian data menjadi subgrup-subgrup berdasarkan pada variabel dependennya (Lehmann dan Eherler, 2001). Metode ini nantinya akan menghasilkan *output* berupa diagram pohon, maka metode ini lebih mudah diinterpretasikan dikarenakan peneliti dapat melihat secara langsung proses pemisahan dan penggabungan varibel independen. Berkaitan dengan penelitian ini, akan diklasifikasikan data kecelakaan lalu lintas di POLRES Kabupaten Sleman tahun 2016 sampai 2017 dengan menggunakan metode *CHAID*.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang tersebut maka permasalahan yang dapat diidentifikasi penulis dalam penelitian kali ini adalah:

1. Bagaimana gambaran umum pada data korban kecelakaan lalu lintas di POLRES Kabupaten Sleman tahun 2016 sampai 2017?
2. Bagaimana klasifikasi yang terbentuk pada tingkat luka korban kecelakaan lalu lintas di POLRES Kabupaten Sleman 2016 sampai 2017?

## 1.3 Batasan Masalah

Supaya pembahasan dalam penelitian ini tidak meluas, maka peneliti memberikan batasan sebagai berikut:

1. Data pada penelitian ini yang digunakan adalah data kecelakaan lalu lintas kendaraan bermotor dan diperoleh dari POLRES Kabupaten Sleman tahun 2016 sampai 2017.
2. Metode yang digunakan pada penelitian ini adalah *CHAID*
3. Variabel yang digunakan waktu kejadian perkara, tempat kejadian perkara, jenis kecelakaan, kendaraan terlibat, usia korban, profesi korban, kepemilikan SIM korban, jenis kelamin korban, pendidikan korban dan tingkat luka korban.

## 1.4 Tujuan Penelitian

Tujuan dari penelitian adalah sebagai berikut:

1. Mengetahui gambaran umum pada data korban kecelakaan lalu lintas di POLRES Kabupaten Sleman tahun 2016 sampai 2017.
2. Mengetahui klasifikasi yang terbentuk pada tingkat luka korban kecelakaan lalu lintas di POLRES Kabupaten Sleman 2016 sampai 2017.

## 1.5 Manfaat Penelitian

Adapun manfaat penelitian adalah sebagai berikut:

1. Sebagai bahan evaluasi pihak POLRES Kabupaten Sleman dalam pengendalian dan pengaturan lalu lintas yang terjadi di wilayah hukumnya.
2. Sebagai bahan pertimbangan pihak terkait dalam pengambilan kebijakan yang tepat mengenai kecelakaan lalu lintas.

# BAB II
# TINJAUAN PUSTAKA

Penelitian yang dilakukan oleh peneliti melihat beberapa penelitian terdahulu yang masih berhubungan sebagai referensi. Adapun beberapa literature yang mendukung penelitian ini sebagai berikut:

Pramana pada tahun 2011 melakukan penelitian mengenai klasifikasi dengan metode *CHAID* dan penerapannya pada klasifikasi alumni FMIPA UNY. Tujuan penelitian ini membandingakan hasil klasifikasi metode *CHAID* dengan analisis regresi logistik biner. Hasil penelitian adalah metode *CHAID* dianggap lebih unggul karena dapat membagi alumni membagi beberapa segmen, sementara regresi logistik biner hanya mengklasifikasikan alumni menjadi kategori lulus tepat waktu dan tidak tepat waktu.

Menurut Oktafianto tahun 2017 dalam jurnal berjudul Implementasi Metode *CHAID* Pada Klasifikasi Indeks Prestasi Kumulatif Mahasiswa FMIPA UNIROW menerangkan bahwa sarjana dikatakan lulus jika telah menempuh seluruh beban belajar yang ditetapkan dan memiliki Indeks Prestasi Kumulatif (IPK) lebih besar atau sama dengan 2.00. Mahasiswa yang sudah dikatakan lulus tentunya menginginkan segera bekerja sesuai dengan keahlianya. Akan tetapi, pada saat memasuki dunia kerja mahasiswa disyaratkan mempunyai IPK minimal 3.00. Dalam kenyataan dan fakta yang diperoleh menginformasikan bahwa tidak sedikit mahasiswa yang memperoleh IPK di bawah standar 3.00. Hal tersebut dikarenakan berbagai macam karakteristik profil mahasiwa yang mempengaruhi perbedaan capaian pembelajaran lulusan yang ditargetkan dengan IPK. Tujuan dari penelitian ini yaitu untuk mengimplementasikan metode *CHAID* (*Chi-Squared Automatic Interaction Detection*) pada klasifikasi IPK mahasiswa FMIPA tahun 2007 - 2015guna mengetahui karakteristik klasifikasi mahasiswa yang mempunyai nilai IPK di atas standar 3.00. Karakteristik profil

mahasiswa yang digunakan yaitu jenis kelamin, asal daerah, usia, jalur masuk, prodi, dan masa studi. Hasil yang diperoleh dengan menggunakan analisis *CHAID* adalah tiga *node* atau kelompok. Kelompok mahasiswa yang memiliki IPK di atas 3.51 berjenis kelamin perempuan dan merupakan prodi biologi mempunyai tingkat keberhasilan sebesar 70.6%.

Penelitian yang dilakukan oleh Nina dan Abadyo tahun 2013 mengenai *CHAID* untuk mengklasifikasi status mahasiswa setelah lulus perkuliahan studi kasus pada alumnus prodi matematika jurusan matematika. FMIPA Universitas Negeri Malang tahun 2007-2012. Penelitian ini memiliki tujuan untuk mengklasifikasi status mahasiswa setelah lulus perkuliahan dengan menggunakan metode *CHAID*. Hasil dari penelitian ini adalah untuk alumnus yang bekerja kurang dari atau sama dengan 6 bulan memiliki karakteristik yaitu alumnus dengan status belum menikah ketika yudisium yang mempunyai IPK lebih dari 3.00 dan usia ketika bekerja pertama kali adalah kurang dari atau sama dengan 25 tahun. Kategori mendapatkan pekerjaan lebih dari 6 bulan memiliki karakteristik alumnus berstatus belum menikah ketika yudisium yang mempunyai IPK lebih dari 2.75 hingga kurang dari atau sama dengan 3.00, sedangkan untuk kategori yang tidak bekerja memiliki karakteristik alumnus yang sudah menikah ketika yudisium.

Nuriyah pada tahun 2013 melakukan penelitian tentang perbandingan metode *CHAID* dan *CART* dalam menentukan klasifikasi alumni UIN Sunan Kalijaga berdasarkan masa studi. Hasil dari penelitian adalah metode *CHAID* menghasilkan 5 segmen dimana ada 3 variabel independen yang signifikan terhadap model. Metode *CHAID* menyebutkan segmen yang memiliki ketepatan waktu masa studi terbesar adalah segmen dimana yang alumninya memiliki IPK antara $3 \leq IPK \leq 3.50$ dan berasal dari luar DIY. Sedangkan *CART* menghasilkan 6 segmen yang signifikan terhadap model dimana ada 7 variabel independen yang signifikan terhadap model. Metode *CART* menyebutkan segmen yang memiliki ketepatan waktu masa studi terbesar adalah alumni yang memiliki IPK $3 \leq IPK \leq 3.50$.

Menurut Zia tahun 2016 dalam skripsi yang berjudul Klasifikasi Rencana Tindak Lanjut Menggunakan Metode *CHAID* dan Metode *QUEST* menerangkan bahwa penggunaan metode *CHAID* pada penelitian ini menghasilkan persentase tertinggi pengunjung penderita diabetes mellitus di puskesmas kecamatan Siwalan yang dirawat jalan sebesar 92% yaitu pada segmen ke-3 dengan karakteristik dirawat jalan karena alasan ada riwayat keluarga kadar insulin normal dengan kondisi awal ringan. Sedangkan untuk metode *QUEST* sebesar 83.6% yaitu pada segmen ke-1 dengan karakteristik dirujuk karena alasan kondisi awal yang berat. Persentase tertinggi pengunjung penderita diabetes mellitus di Puskesmas Kecamatan Siwalan yang dirujuk pada metode *CHAID* sebesar 79.5% yaitu pada segmen ke-1 dengan karakteristik dirujuk karena alasan kondisi awal yang berat. Sedangkan untuk metode *QUEST* sebesar 79.7% yaitu pada segmen ke-1 dengan karakteristik dirujuk karena alasan kondisi awal yang berat.

Penelitian yang dilakukan oleh Onder dan Uyar pada tahun 2017 terkait analisis menggunakan metode *CHAID* dengan tujuan untuk mengetahui karakteristik tingkat keberhasilan siswa. Kelompok studi terdiri dari 4229 siswa yang belajar di sekolah menengah di Burdur. Data dikumpulkan dengan menggunakan kuesioner pada tahun akademik 2014-2015. Variabel indepeden dalam penelitian ini adalah jenis kelamin, pendidikan orang tua, pekerjaan orang tua, tempat tinggal, jumlah saudara kandung dan pendapatan per bulan, sementara variabel dependennya adalah tingkat keberhasilan siswa. Hasil menunjukan karakteristik orang tua yang paling menentukan untuk tingkat keberhasilan siswa yaitu status pendidikan ayah.

# BAB III
# LANDASAN TEORI

## 3.1 Definisi Kecelakaan Lalu Lintas

Pengertian kecelakaan lalu lintas berdasarkan ketentuan yang ditetapkan dalam pasal 93 Peraturan Pemerintah Nomor 43 tahun 1993 ayat 1 adalah:

*"Suatu peristiwa dijalan yang tidak disangka-sangka dan tidak disengaja melibatkan kendaraan dengan atau tanpa pemakai jalan lainnya mengakibatkan korban manusia atau kerugian harta benda".*

## 3.2 Pengertian Kendaraan Bermotor

Kendaraan Bermotor adalah semua kendaraan beroda beserta gandengannya yang digunakan di semua jenis jalan darat, dan digerakkan oleh peralatan teknik berupa motor atau peralatan lainnya yang berfungsi untuk mengubah suatu sumber daya energi tertentu menjadi tenaga gerak kendaraan bermotor yang bersangkutan, termasuk alat-alat berat dan alat-alat besar yang dalam operasinya menggunakan roda dan motor dan tidak melekat secara permanen serta kendaraan bermotor yang dioperasikan di air. (*Pasal 1 Angka 13 UU Nomor 28 Tahun 2009 Tentang Pajak Daerah Dan Retribusi Daerah*).

## 3.3 Jenis Kecelakaan Lalu Lintas

Berdasarkan pasal 93 Peraturan Pemerintah No.43 Tahun 1993 disebutkan bahwa pengkategorian jenis dan bentuk kecelakaan ada lima, yaitu:

### 3.3.1 Kecelakaan Berdasarkan Korban Kecelakaan

Menurut Peraturan Pemerintah RI No.43 Tahun 1993 pengkategorian berdasarkan korban kecelakaan menjadi tiga, yaitu:

1. Kecelakaan korban meninggal dunia merupakan korban yang dipastikan meninggal dikarenakan kecelakaan lalu lintas dalam waktu paling lama 30 hari setelah kecelakaan tersebut.
2. Kecelakaan korban luka berat merupakan korban yang mengalami luka-luka sampai menderita cacat tetap atau harus dirawat dalam jangka waktu 30 hari setelah kecelakaan tersebut.
3. Korban luka ringan adalah korban yang tidak disebutkan pada korban meninggal dan korban luka berat atau tidak masuk dalam pengertian diatas.

### 3.3.2 Kecelakaan Berdasarkan Lokasi Kejadian

Kecelakaan tidak ada yang menduga kapan terjadinya, bisa terjadi dimanapun. Kecelakaan dapat terjadi baik di jalan lurus, jalan menikung, tanjakan, turunan, di pegunungan, perkotaan ataupun pedesaan. Dalam penelitian ini lokasi kejadian dikategorikan mengacu pada UU No.38 Tahun 2004 pengelompokan jalan berdasarkan statusnya sebagai berikut:

1. Jalan nasional adalah jalan arteri dan jalan kolektor dalam sistem jaringan jalan primer yang menghubungkan antaribukota provinsi, dan jalan strategis nasional, serta jalan tol.
2. Jalan provinsi adalah jalan kolektor dalam sistem jaringan jalan primer yang menghubungkan ibukota provinsi dengan ibukota kabupaten/kota, atau antaribukota kabupaten/kota, dan jalan strategis provinsi.
3. Jalan kabupaten adalah jalan lokal dalam sistem jaringan jalan primer yang menghubungkan ibukota kabupaten dengan ibukota kecamatan, antaribukota kecamatan, ibukota kabupaten dengan pusat kegiatan lokal, antarpusat kegiatan lokal, serta jalan umum dalam sistem jaringan jalan sekunder dalam wilayah kabupaten, dan jalan strategis kabupaten.
4. Jalan kota merupakan jalan umum dalam sistem jaringan jalan sekunder yang menghubungkan antarpusat pelayanan dalam kota,

menghubungkan pusat pelayanan dengan persil, menghubungkan antarpersil, serta menghubungkan antarpusat permukiman yang berada di dalam kota.

5. Jalan desa adalah jalan umum yang menghubungkan kawasan dan/atau antarpermukiman di dalam desa, serta jalan lingkungan.

### 3.3.3 Kecelakaan Berdasarkan Waktu Terjadinya Kecelakaan

Waktu terjadinya kecelakaan dikategorikan menjadi dua kategori yaitu padat kendaraan dan sepi kendaraan. Waktu padat kendaraan pada pukul 06.00-08.00, 12.00-13.30, dan16.00-18.00, sementara dikatakan waktu sepi kendaraan selain waktu padat kendaraan. (Srinadi, 2016).

### 3.3.4 Kecelakaan Berdasarkan Posisi Kecelakaan

PT Jasa Marga dalam Sulistyono, (1998) mengelompokkan jenis tabrakan atau kecelakaan yang melatar belakangi terjadinya kecelakaan lalu lintas menjadi :

1. Tabrakan depan-depan merupakan jenis tabrakan yang terjadi antara dua kendaraan yang tengah melaju dimana keduanya saling beradu muka dari arah yang berlawanan, yaitu bagian depan kendaraan yang satu dengan bagian depan kendaraan lainnya.
2. Tabrakan depan-samping merupakan jenis tabrakan yang terjadi antara dua kendaraan yang tengah melaju dimana bagian depan kendaran yang satu menabrak bagian samping kendaraan lainnya.
3. Tabrakan depan-belakang merupakan jenis tabrakan yang terjadi antara dua kendaraan yang tengah melaju dimana bagian depan kendaraan yang satu menabrak bagian belakang kendaraan di depannya dan kendaraan tersebut berada pada arah yang sama.

4. Tabrakan samping-samping merupakan jenis tabrakan yang terjadi antara dua kendaraan yang tengah melaju dimana bagian samping kendaraan yang satu menabrak bagian yang lain.
5. Menabrak penyeberang jalan merupakan jenis tabrakan yang terjadi antara kendaraan yang tengah melaju dan pejalan kaki yang sedang menyeberang jalan.
6. Tabrakan sendiri merupakan jenis tabrakan dimana kendaraan yang tengah melaju mengalami kecelakaan tunggal atau sendiri.
7. Tabrakan beruntun merupakan jenis tabrakan dimana kendaraan yang tengah melaju menabrak mengakibatkan terjadinya kecelakaan yang melibatkan lebih dari dua kendaraan secara beruntun.
8. Menabrak obyek tetap adalah jenis tabrakan dimana kendaraan yang tengah melaju menabrak obyek tetap dijalan.

### 3.3.5 Kecelakaan Berdasarkan Jumlah Kendaraan Yang Terlibat

Kecelakaan berdasarkan kendaraan terlibat dibagi menjadi dua, yaitu kecelakaan tunggal dilakukan oleh satu kendaraan, kecelakaan ganda dilakukan oleh dua atau lebih kendaraan termasuk didalamnya kecelakaan beruntun (Simanjuntak, 2013).

## 3.4 Faktor-Faktor Penyebab Kecelakaan Lalu Lintas

Terjadinya kecelakaan lalu lintas bisa disebabkan berbagai macam faktor, bisa disebabkan faktor internal dan eksternal. Maksud faktor internal adalah kecelakaan yang disebabkan oleh manusia, sementara faktor eksternal meliputi faktor jalan, faktor kendaraan dan faktor cuaca (Afidah dan Susilaningrum, 2012). Penjelasan secara detail menganai faktor penyebab kecelakaan sebagai berikut:

### 3.4.1 Faktor Manusia

Faktor manusia menjadi faktor utama penyebab terjadinya kecelakaan lalu lintas. Faktor manusia yang telah di catat oleh pihak kepolisian yaitu

kepemilikan SIM dan jenis pekerjaan korban. Pada penelitian ini jenis kelamin korban dikategorikan menjadi laki-laki dan perempuan.

Jenjang pendidikan adalah tahap dalam pendidikan yang berkelanjutan yang ditetapkan berdasarkan tingkat perkembangan para peserta didik, keluasan bahan pengajaran, dan tujuan pendidikan yang dicantumkan dalam kurikulum (KBBI, 2016). Berdasarkan Undang Undang No.20 tahun 2000 disebutkan bahwa jenjang pendidikan formal terdiri dari Sekolah Dasar, Sekolah Menengah Pertama, Sekolah Menengah Atas/Sekolah Menengah Kejuruan, Perguruan Tinggi.

Pengkategorian jenis pekerjaan berdasarkan BPS pada tahun 2011 meliputi mahasiswa/pelajar, wiraswasta/wirausaha, PNS/TNI/POLRI, pengemudi dan lainya, maka pada penelitian ini kategori jenis pekerjaan korban menjadi mahasiswa/pelajar, wiraswasta/wirausaha, PNS/TNI/POLRI, pengemudi dan lainya.

Usia Korban dikategorikan menjadi usia anak-anak dan remaja (1-15 tahun), usia muda (16-35 tahun), usia dewasa (36-55 tahun) dan usia lanjut lebih besar dari 55 tahun (Afidah dan Susilaningrum, 2011).

Surat Izin Mengemudi atau sering disebut SIM adalah bukti registrasi dan identifikasi yang diberikan oleh Polri kepada seseorang yang telah memenuhi persyaratan administrasi, sehat jasmani dan rohani, memahami peraturan lalu lintas dan terampil mengemudikan kendaraan bermotor. Setiap orang yang mengemudikan Kendaraan Bermotor di Jalan wajib memiliki Surat Izin Mengemudi sesuai dengan jenis Kendaraan Bermotor yang dikemudikan (UU No.22 Tahun 2009). Berdasarkan pengertian di atas menunjukan pengendara yang memiliki SIM dapat menjadi tolak ukur dalam berkendara, namun hal tersebut tidak menjadi jaminan tidak akan terjadi kecelakaan bagi pengendara atau pengemudi yang memiliki SIM. Pada penelitian ini kepemilikian SIM dikategorikan menjadi memiliki SIM (pengendara korban kecelakaan lalu lintas atau korban yang terlibat dalam

kecelakaan lalu lintas memiliki SIM) dan tidak memiliki SIM (pengendara korban kecelakaan lalu lintas atau korban yang terlibat dalam kecelakaan lalu lintas tidak memiliki SIM.

### 3.4.2 Faktor Kendaraan

Pada zaman modern seperti sekarang ini kendaraan menjadi sarana angkutan yang sangat penting, dikarenakan bisa meringankan aktivitas manusia dengan cepat, tidak repot dan hemat. Akan tetapi ketika kendaraan tidak dirawat dengan benar nantinya terdapat bagian-bagian kendaraan yang rusak seperti ban pecah, rem tidak berfungsi dengan baik, lampu utama mati, kelelahan logam yang menyebabkan bagian kendaraan patah, dan berbagai permasalahan lainya. Secara keseluruhan faktor kendaraan sangat terkait dengan teknologi yang digunakan dan perawatan yang dilakukan terhadap kendaraan.

### 3.4.3 Faktor Jalan

Menurut Undang-Undang Republik Indonesia Nomor 38 Tahun 2004 jalan adalah prasarana transportasi darat yang meliputi segala bagian jalan, termasuk bangunan pelengkap dan perlengkapannya yang diperuntukkan bagi lalu lintas, yang berada pada permukaan tanah, di atas permukaan tanah, di bawah permukaan tanah dan/atau air, serta di atas permukaan air, kecuali jalan kereta api, jalan lori, dan jalan kabel.

Untuk keperluan pengaturan penggunaan dan pemenuhan kebutuhan angkutan, jalan dibagi dalam beberapa kelas yang didasarkan pada kebutuhan transportasi, pemilihan moda secara tepat dengan mempertimbangkan keunggulan karakteristik masing-masing moda, perkembangan teknologi kendaraan bermotor, muatan sumbu terberat kendaraan bermotor serta konstruksi jalan. Pengelompokkan jalan menurut muatan sumbu yang disebut juga kelas jalan adalah berdasarkan Undang-

Undang Republik Indonesia No 14 Tahun 1992 tentang Lalu Lintas dan Angkutan jalan.

### 3.4.4 Faktor Cuaca

Faktor cuaca dapat menyebabkan kecelakaan. Hujan akan sangat mempengaruhi kinerja kendaraan seperti jalan menjadi lebih licin maka jarak pengereman menjadi lebih jauh, tidak seperti pada saat kering dan jarak pandang menjadi tidak begitu jelas karena alat pembersih kaca tidak bisa berfungsi dengan baik.

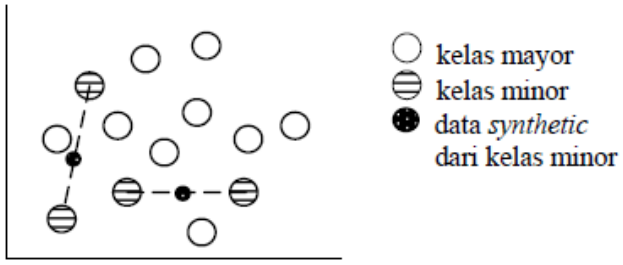
## 3.5 *SMOTE*

*SMOTE* (*Synthetic Minority Oversampling Technique*) diperkenalkan oleh Nithes V. Chawla (Chawla, 2002) untuk pertama kalinya. Pendekatan tersebut bekerja dengan membuat "*synthetic*" data atau data replikasi dari data minor.

Ketidakseimbangan data terjadi apabila jumlah objek suatu kelas data lebih banyak jika dibandingkan dengan kelas lain. Kelas mayor adalah data yang objeknya lebih banyak sementara lainnya disebut kelas minor. Penggunaan data tidak seimbang memiliki pengaruh yang sangat besar untuk membuat model pada hasil model yang diperoleh. Pengolahan algoritma yang tidak memperhatikan ketidakseimbangan data akan condong diliputi oleh kelas mayor dan mengabaikan kelas minor. Metode *SMOTE* sebagai salah satu solusi dalam menangani ketidakseimbangan data dengan prinsip menambah jumlah data kelas minor supaya setara dengan kelas mayor dengan cara membangkitkan data buatan. Data sintesis atau buatan tersebut cara membuatnya berdasarkan k-tetangga terdekat (*k-nearest neighbor*). Jumlah k-tetangga terdekat ditentukan dengan mempertimbangkan kemudahan dalam melaksanakannya (Chawla, 2002).

Metode *SMOTE* bekerja dengan mencari *k nearest neighbors* (yaitu ketetanggaan data) untuk setiap data di kelas minor, setelah itu buat *synthetic data* sebanyak prosentase duplikasi yang diinginkan antara data minor dan *k nearest*

*neighbors* yang dipilih secara random. Ilustrasi distribusi data setelah diterapkan metode *SMOTE* dapat dilihat pada Gambar 3.1 (Chawla, 2002).




**Gambar 3.1** Ilustrasi *SMOTE*

Pada pembentukan data synthetic yang baru, ada 2 jenis perhitungan terhadap *nearest neighbor*, untuk data nominal dan data *continues*.

Data *continues* :

- Dihitung perbedaan untuk setiap atribut antara *minority sample* dengan salah satu dari *k nearest neighboornya.*
- Perbedaan ini dikalikan dengan nilai *random* antara 1 dan 0
- Hasilnya ditambahkan dengan nilai *minority sample*, inilah hasil pembuatan *feature vector* yang baru (*synthetic minority class)* yang baru

Data *nominal* :

- Diambil voting antara *minority sample* dan *nearest neighbor*nya. Jika tidak ada *majority class*, maka pilihlah nilai atribut pada *minority sample* tersebut.
- Nilai tersebut ditandai menjadi *synthetic minority class* yang baru.

## 3.6 *Data Mining*

*Data mining* merupakan suatu proses menemukan hubungan yang berarti, pola dan kecenderungan dengan memeriksa sekumpulan data besar/banyak data yang tersimpan dalam penyimpanan dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika (Larose, 2005). Secara sederhana *data mining* merupakan ekstraksi informasi atau pola yang penting atau menarik dari data yang ada di *database* yang besar. *Data mining* dikenal juga dengan istilah *Knowledge Discovery in Database* (KDD) (Susanto, 2014). *Data mining* mengacu pada proses untuk

menambang (*mining*) pengetahuan dari sekumpulan data yang sangat besar (Rani, 2015).

Teknik dalam data mining memungkinkan untuk membangun sebuah model untuk mengenali pola data yang lain selain dalam basis data yang tersimpan. Selain itu, data mining dapat juga digunakan untuk prediksi maupun klasifikasi yang bertujuan agar terbentuk pola universal dari data-data yang ada (Haryati dkk, 2015).

*Data mining* merupakan proses penambangan atau penemuan terhadap informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar. Beberapa karakteristik *data mining*, antara lain (Davies dan Beynon, 2004) :

1. *Data mining* berhubungan dengan penemuan sesuatu yang tersembunyi dan pola tertentu yang tidak diketahui sebelumnya.
2. *Data mining* biasanya menggunakan data yang sangat besar. Biasanya data yang besar digunakan untuk membuat hasil lebih dipercaya.
3. *Data mining* berguna untuk membuat keputusan yang kritis, terutama dalam strategi.

Ada 3 teknik yang populer dalam data mining (Haryati dkk, 2015) :

1. *Association Rule Mining*

   *Association Rule Mining* merupakan teknik dalam data mining untuk menentukan hubungan antara kombinasi atribut. Contohnya adalah analisa pembelian disuatu supermarket menghasilkan strategi dalam menggunakan diskon untuk kombinasi barang tertentu.
2. *Clustering*

   *Clustering* digunakan untuk memberikan label pada kelas data yang belum diketahui sehingga sering disebut *unsupervised learning.* Prinsip *clustering* adalah memaksimalkan kesamaan antar cluster.
3. Klasifikasi

   Hasil dalam klasifikasi adalah adanya variabel kategori. Contohnya adalah asal mahasiswa yang dapat diklasifikasikan menjadi berasal dari Provinsi Jawa Tengah dan Provinsi di luar Jawa Tengah.

## 3.7 Klasifikasi

Klasifikasi merupakan suatu proses yang memiliki tujuan untuk mempelajari fitur dari sebuah atau sekelompok data yang sudah diketahui kelasnya dan menghasilkan sebuah model klasifikasi yang dapat digunakan untuk memprediksi kelas atau label data yang baru (Liu, 2012). Klasifikasi adalah proses dari mencari suatu himpunan model (fungsi) yang dapat mendeskripsikan dan membedakan kelas-kelas data atau konsep-konsep, dengan tujuan dapat menggunakan model tersebut untuk memprediksi kelas dari suatu objek yang kelasnya belum diketahui (Rani, 2015). Di Statistika ada beberapa metode klasifikasi, diantaranya *CHAID* (*Chi- Square Automatic Interaction Detection*), *KNN* (K- *Nearest Neighbour*), *CART* (*Clasification and Regression Tree*), dan *Naïve Bayes*. Dalam penelitian ini akan digunakan metode *CHAID.*

## 3.8 Data *Training* dan Data *Testing*

Pada umumnya data dibagi menjadi dua, yaitu data training dan data testing. Algoritma klasifikasi (seperti *decision tree, bayesian, neural network,* SVM) menggunakan data *training* untuk membentuk sebuah model klasifikasi. Model ini merupakan representasi pengetahuan yang nantinya akan digunakan untuk mengetahui sejauh mana hasil klasifikasi dapat melakukan klasifikasi dengan benar. Oleh karena itu, data yang ada terdapat pada data *testing* semestinya tidak terdapat pada data *training* sehingga dapat diketahui model klasifikasi dapat melakukan klasifikasi dengan baik dan benar. Pembagian antara data *training* dan data *testing* tidak mengikat akan tetapi supaya variasi dalam model tidak terlalu besar maka dianjurkan data *training* memiliki perbandingan lebih besar dengan data *testing.* Pada umumnya 2/3 dari total data dijadikan sebagai data *training* ssementara sisanya dijadikan sebagai data *testing.* Penelitian yang menghasilkan keakuratan model klasifikasi optimum dengan proporsi 80:20 dan 90:10 untuk data *training* dan data *testing* (Paratu, 2003).

## 3.9 Analisis *CHAID*

Seseorang bernama Dr. G. V. Kass pada tahun 1980 untuk pertama kalinya memperkenalkan metode *CHAID* pada sebuah artikel dengan judul "*An Exploratory Technique for investigating Large Quantities of Categorical Data"* dalam buku *Applied Statistics. Automatic Interaction Detection (AID)* adalah sebutan metode tersebut yang dikenal (Permana, 2011). Menurut Galagher (2000), *CHAID* adalah suatu teknik iteratif yang menguji satu-persatu variabel independen yang digunakan dalam klasifikasi, dan menyusunnya berdasarkan pada tingkat signifikansi statistik *chi-square* terhadap variabel dependennya.

Metode *CHAID* akan lebih berarti pada data yang besar, tidak disarankan pada data yang berukuran kecil. Banyak data minimal yang digunakan adalah 500 (Du Toit et al, 1986).

Metode *CHAID* merupakan salah satu metode untuk mengklasifikasikan data kategori dengan tata cara membagi rangkaian data subgroup-subgrup berdasarkan pada variabel dependenya. (Lehmann dan Eherler, 2001).

Dalam metode *CHAID* data yang digunakan dibedakan menjadi variabel terikat atau dependen dan variabel bebas atau independen. Variabel-variabel independen dalam metode *CHAID* dibedakan menjadi tiga bentuk, yaitu:

1. Monotonik adalah variabel independen yang kategori di dalamnya bisa digabungkan atau dikombinasikan oleh *CHAID* hanya jika keduanya berdekatan satu sama lain atau mengikuti urutan aslinya (data ordinal). Misalnya : usia, pendapatan dan jenjang pendidikan.
2. Bebas adalah variabel independen yang kategori di dalamnya bisa digabungkan atau dikombinasikan apabila keduanya berdekatan ataupun berjauhan (data nominal). Misalnya : pekerjaan, status pernikahan, dan jenis kelamin.
3. Mengambang (*floating*) adalah kategori - kategori di dalamnya bisa dianggap seperti monotonik kecuali untuk kategori yang terakhir (*missing value)*, yang bisa dikombinasikan dengan kategori manapun.

*CHAID* memiliki kegunaan untuk membentuk segmentasi yang membagi sebuah sampel menjadi dua atau lebih kelompok yang berbeda berdasarkan sebuah kriteria tertentu. Selanjutnya dilanjutkan dengan membagi kelompok tersebut menjadi kelompok yang lebih kecil sesuai variabel variabel independen yang lain. Proses tersebut akan berhenti ketika tidak terdapat variabel-variabel independen yang signifikan secara statistik. (Kunto dan Hasana, 2006).

## 3.10 Algoritma *CHAID*

Algoritma *CHAID* secara garis besar dibagi menjadi tiga tahap, tahap pertama Penggabungan (*merging*), kedua Pemisahan (*Splitting*) dan yang ketiga Penghentian (*Stopping*). Diagram pohon dimulai dari *root node* (node akar) melalui tiga tahap tersebut pada setiap *node* yang terbentuk dan secara berulang.

### 3.10.1 Penggabungan (*Merging)*

Pada tahap pertama adalah proses pemeriksaan signifikansi dari tiap-tiap kategori variabel independen terhadap variabel dependen. Tahap penggabungan secara detail untuk setiap variabel independen dalam menggabungkan kategori-kategori non-signifikan seperti berikut :

1. Buat tabel kontingensi dua arah untuk masing-masing variabel independen dengan variabel dependennya.
2. Langkah selanjutnya menghitung statistik *chi-square* untuk setiap pasang kategori yang dapat dipilih untuk digabung menjadi satu, untuk menguji kebebasannya dalam sebuah sub tabel kontingensi 2 x J yang dibentuk oleh sepasang kategori tersebut dengan variabel dependennya yang mempunyai sebanyak J kategori. Adapun langkah uji *chi-square* seperti berikut :

   a. Hipotesis

      $H_0$ : $P_{ij} = P_{i\bullet}\, P_{\bullet j}$ (Variabel i dan variabel j independen atau tidak terdapat hubungan antara variabel i dan variabel j)

$H_1$ : $P_{ij} \neq P_{i\bullet}\, P_{\bullet j}$ (Variabel i dan variabel j dependen atau terdapat hubungan antara variabel i dan variabel j)

b. Tentukan $\alpha = 0.05$

c. Tentukan daerah penolakan,

yaitu $\chi^2_{\text{hitung}} > \chi^2_{0.05;(b-1)(k-1)}$ atau *p-value* < α

d. Mencari $\chi^2 = \sum_{IJ} \frac{(O_{IJ} - E_{IJ})^2}{E_{IJ}}$ (3.1)

e. Keputusan

$H_0$ ditolak jika $\chi^2_{\text{hitung}} > \chi^2_{\text{tabel}}$ atau *p-value* < α

f. Pengambilan kesimpulan

Bila $\chi^2$ masuk daerah penolakan, Ho ditolak

Bila $\chi^2$ tidak masuk dalam daerah penolakan, Ho diterima

Sebagai contoh sebuah variabel independen $X_i$ adalah variabel monotonic dengan *a* kategori, dimana *i*=1,2,...a. Variabel dependen *Y* memiliki *b* kategori. Agar dapat mengetahui kategori variabel independen yang tidak signifikan dipasangkan masing-masing kategori pada variabel independen dengan variabel dependen. Banyaknya pasangan yang mungkin adalah kombilasi *b* dari *a*.

**Tabel 3.1 Ilustrasi Pasangan Penggabungan Variabel**

| Kategori 1 | Kategori 2 | Kategori 3 |
|---|---|---|
| $X_1$ | $X_2$ | $P_{1,2}$ |
| $X_1$ | $X_3$ | $P_{1,3}$ |
| . | . | . |
| . | . | . |
| $X_a$ | $X_1$ | $Pa,_1$ |
| . | . | . |
| . | . | . |
| $X_a$ | $X_{a-1}$ | $P_{a,a-1}$ |

3. Hitung *p-value* berpasangan bersamaan, untuk masing-masing nilai *chi-square* berpasangan. Gabungkan sebuah pasangan kategori yang paling mirip (yaitu pasangan yang mempunyai nilai *chi-square* berpasangan terkecil dan *p-value* terbesar) menjadi sebuah kategori tunggal, diantara pasangan-pasangan yang tidak signifikan dan selanjutnya diteruskan kelangkah berikutnya (langkah nomor 4).

   Misalnya pada tabel 3.1 apabila terdapat pasangan dengan nilai *p-value* lebih besar dari taraf signifikansi, maka pasangan tersebut digabungkan. Misal pasangan kategori $X_1$ dan $X_2$ tidak signifikan, maka pasangan tersebut digabungkan menjadi satu variabel baru $X_{1,2.}$
4. Kategori baru yang terbentuk setelah digabung dengan kategori lainnya dalam variabel independen diperiksa kembali kesignifikansiannya. Apabila terdapat pasangan yang belum signifikan, maka langkah 3 diulangi lagi. Apabila ternyata sudah semua signifikan lanjut ke langkah selanjutnya.
5. Hitung *p-value* terkoreksi Bonferroni didasarkan pada tabel yang telah digabung.

### 3.10.2 Pemisahan (*Splitting*)

Pada tahap *splitting* akan memilih variabel independen yang akan digunakan sebagai pemisah *node* (*split node*) yang paling baik. Proses pemilihan dikerjakan dengan cara mebandingkan *p-value* pada tiap-tiap variabel independen. Adapun langkah *Splitting* sebagai berikut:

1. Langkah yang pertama adalah memilih varibel independen yang mempunyai *p-value* paling kecil atau paling signifikan yang nantinya digunakan sebagai *split node*.
2. Apabila nilai *p-value* kurang dari sama dengan spesifikasi *alpha*, maka *split node* menggunakan variabel independen ini. Apabila tidak terdapat variabel independen dengan nilai *p-value* yang signifikan maka tidak

perlu dilakukan *split* dan *node* ditentukan sebagai terminal *node* (akhir *node*).

### 3.10.3 Penghentian (*Stopping*)

Pembentukan subkelompok berikunya, dengan mengulangi langkah penggabungan. Pada tahap *stopping* akan dilakukan penghentian proses pertumbuhan dengan peraturan sebagai berikut:

1. Variabel independen yang signifikan menunjukkan perbedaan terhadap variabel dependen sudah tidak ada lagi.
2. Proses pertumbuhan akan terhenti ketika pohon sekarang mencapai batas nilai mmaksimum pohon dari spesifikasi. Contoh peneliti menetapkan batas kedalaman pertumbuhan pohon klasifikasi adalah 3, maka pertumbuhan pohon klasifikasi terhenti ketika pertumbuhan pohon sudah mencapai kedalaman 3.
3. *Node* tidak akan di-*split* apabila ukuran dari *child node* kurang dari nilai ukuran *child node* minimum spesifikasi, atau berisi pengamatan-pengamatan dengan banyak yang terlalu sedikit. Contoh ditetapkan ukuran minimal *child node* adalah 50, maka node tersebut tidak akan dipecah ketika *splitting* mengasilkan ukuran *child node* kurang dari 50.

## 3.11 Koreksi Bonferroni (*Bonferroni Correction*)

Pengertian Koreksi Bonferroni adalah "suatu proses koreksi yang digunakan ketika beberapa uji statistik untuk kebebasan atau ketidakbebasan dilakukan secara bersamaan" (Kunto dan Hasna, 2006). Perbandingan berganda biasa digunakan dalam koreksi bonferroni.

Menurut Gallagher (2000) bahwa pengali Bonferroni untuk tiap-tiap jenis variabel independen adalah sebagai berikut:

1. Variabel independen Monotonik

$$M = \binom{c-1}{r-1} \tag{3.2}$$

dimana :

M = Pengali Bonferroni

c = banyaknya kategori variabel independen awal

r = banyaknya kategori variabel independen setelah penggabungan

2. Variabel independen Bebas

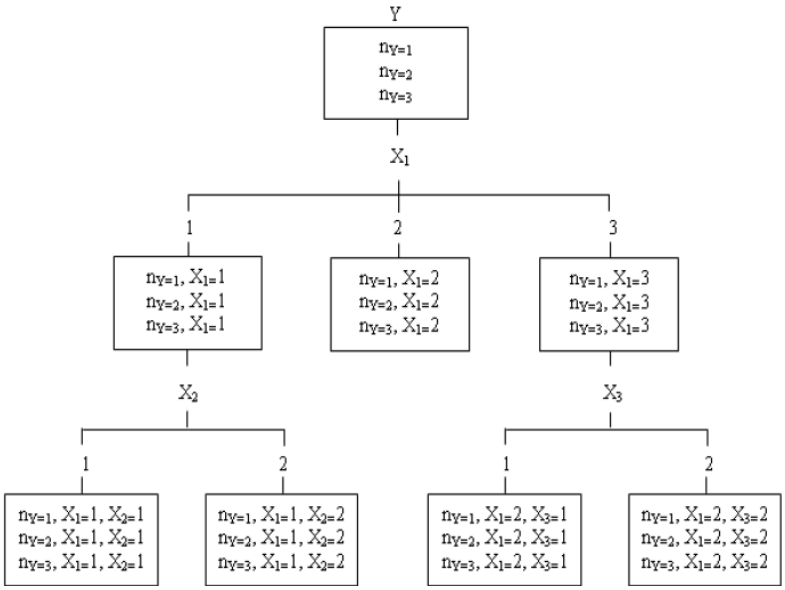
$$M = \sum_{i=0}^{r-1} (-1)^i \frac{(r-1)^c}{i!\,(r-i)!} \tag{3.3}$$

3. Variabel independen Mengambang *(Floating*)

$$M = \binom{c-2}{r-2} + r = \binom{c-2}{r-1} \tag{3.4}$$

### 3.12 Diagram Pohon Klasifikasi *CHAID* (*CHAID Classification Tree*)

Diagram pohon klasifikasi yang menggambarkan pembentukan segmen adalah output dari analisis *CHAID*. Diagram *CHAID* terdiri dari batang pohon (*tree trunk)* dengan membagi (*split)* menjadi lebih kecil berupa cabang-cabang (*brances).* Pada gambar 3.1 adalah gambar diagram pohon *CHAID.*




**Gambar 3.2** Diagram Pohon dalam Analisis *CHAID*
Sumber : Lehmann dkk, 2001.

Menurut Myers (Kunto dan Hasana, 2006), “diagram pohon *CHAID* mengikuti aturan “dari atas ke bawah” (*Top-down stopping rule*), dimana diagram pohon disusun mulai dari kelompok induk (*parent node),* berlanjut di bawahnya sub kelompok (*child node)* yang berturutturut dari hasil pembagian kelompok induk berdasarkan kriteria tertentu”. *Terminal node* adalah *node* pada ujung pohon yang tidak ada percabangan lagi. Sub kelompok dari sampel yang diteliti dan berisi keseluruhan sampel dan frekuensi absolut $n_i$ untuk setiap kategori yang disusun adalah gambaran dari masing-masing *node* pada diagram pohon ini.

Dalam penelitian Pramana (2011) berdasarkan (Lehman dan Eherler, 2001) istilah kedalaman (*depth*) dalam analisis *CHAID* sering disebutkan yang memiliki makna banyaknya tingkatan *node-node* sub kelompok sampai ke bawah pada *node* sub kelompok yang terakhir. Kedalaman pertama, sampel dibagi oleh $X_1$ sebagai variabel independen paling baik untuk variabel dependen sesuai uji *chi-square*. Frekuensi variabel *Y* merupakan informasi dari masing-masing *node,* sebagai variabel dependen. Variabel dependen tersebut merupakan bagian dari sub kelompok yang dihasilkan berdasarkan kategori yang disebutkan *($X_1$)*. Kedalaman kedua, (*node* $X_2$ dan $X_3$) merupakan pembagian dari $X_1$ (untuk *node* ke-1 dan ke-3). Dengan menggunaka cara yang sama, sampel berikutnya dibagi oleh variabel penjelas yang lain, yaitu $X_2$ dan $X_3$, dan selanjutnya menjadi sub kelompok pada *node* ke-4, 5, 6, dan 7 (Lehmann dan Eherler, 2001). Persentase responden untuk setiap kategori dari variabel dependen, dan jumlah total responden untuk tiap-tiap *node* ditampilkan pada masing-masing *node.*

### 3.13 Evaluasi Klasifikasi

Model klasifikasi dapat diketahui tingkat ketelitianya dengan menggunakan tiga buah parameter, yaitu akurasi, *recall*, dan presisi. Akurasi merupakan ukuran dari seberapa bagus model bisa mengklasifikasikan dokumen dengan benar. Presisi merupakan tingkat ketelitian atau ketepatan dalam pengklasifikasian sementara r*ecall* memiliki fungsi untuk mengukur proporsi positif aktual yang benar diidentifikasi (Witten dan Frank, 2005). Berikut rumus akurasi, presisi dan *recall* :

$$\text{Akurasi} = \frac{TP+TN}{TP+FP+TN+FN} \quad (3.5)$$

$$\text{Presisi} = \frac{TP}{TP+FP} \quad (3.6)$$

$$\textit{Recall} = \frac{TP}{TP+FP} \quad (3.7)$$

dimana:

TP : *true positive*

TN : *true negative*

FP : *false positive*

FN : *false negative*

Parameter akurasi, *recall*, dan presisi ketika memiliki nilai yang seimbang maka model klasifikasi dikatakan baik. Model klasifikasi yang akurat belum tentu baik apabila nilai *recall* dan presisinya rendah. C*onfusion matrix* dapat juga digunakan untuk melakukan evaluasi kinerja dari *classifier*. *Confusion matrix* merupakan alat ukur berbentuk matrix yang digunakan untuk menperoleh jumlah ketepatan klasifikasi terhadap kelas dengan algoritma yang dipakai (Witten dan Frank, 2005). Berikut adalah bentuk C*onfusion matrix* :

**Tabel 3.2** *Confusion matrix*

| | *Predicted Positive* | *Predicted Negative* |
|---|---|---|
| *Actual Positive* | *True Positive (TP)* | *False Negative (FN)* |
| *Actual Negative* | *False Positive (FP)* | *True Negative (TN)* |

# BAB IV
# METODOLOGI PENELITIAN

## 4.1 Populasi dan Sampel Penelitian

Dalam penelitian ini menggunakan data sekunder berupa sampel data kecelakaan lalu lintas tahun 2016 sampai 2017 di POLRES Kabupaten Sleman**,** sementara populasinya adalah data kecelakaan lalu lintas di POLRES Kabupaten Sleman.

## 4.2 Variabel dan Definisi Operasional Variabel

Pada penelitian ini ada beberapa variabel, adapun variabel yang digunakan adalah sebagai berikut:

**Tabel 4.1** Variabel Data

| Variabel | Keterangan |
|---|---|
| Y | Tingkat Luka Korban |
| $X_1$ | Waktu Kejadian |
| $X_2$ | Tempat Kejadian Perkara |
| $X_3$ | Jenis Kecelakaan |
| $X_4$ | Kendaraan Terlibat |
| $X_5$ | Umur |
| $X_6$ | Profesi |
| $X_7$ | Kepemilikan SIM |
| $X_8$ | Jenis Kelamin |
| $X_9$ | Pendidikan |

Adapun Definisi Operasional Variabel seperti berikut:

| | | |
|---|---|---|
| Tingkat Luka Korban | : | Akibat dari kecelakaan lalu lintas yang dibagi menjadi:<br>1. Meninggal Dunia<br>2. Luka Berat<br>3. Luka Ringan |
| Waktu Kejadian | : | Waktu terjadinya kecelakaan lalu lintas dibagi menjadi:<br>1. Sepi (pukul 06.00-08.00, 12.00-13.30 dan 16.00-18.00)<br>2. Padat (pukul 08.01-11.59, 13.31-15.59 dan 18.01-05.59). |
| Tempat Kejadian Perkara | : | Tempat terjadinya kecelakaan lalu lintas yang di kategorikan menjadi:<br>1. Jalan Nasional<br>2. Jalan Provinsi<br>3. Jalan Kabupaten<br>4. Jalan Desa |
| Jenis Kecelakaan | : | Jenis kecelakaan yang terjadi dikategorikan menjadi :<br>1. Depan Depan<br>2. Depan Samping<br>3. Depan Belakang<br>4. Samping Samping<br>5. Tabrak Lari<br>6. Laka Tunggal |
| Kendaraan Terlibat | : | Kendaraan yang terlibat pada saat terjadinya kecelakaan yang terdiri dari |

1. Sepeda Motor.
2. Mobil
3. Truk/Bus

Umur Korban : Usia korban kecelakaan yang di kategorikan menjadi:

1. Anak-Anak dan Remaja (usia 1-15 tahun)
2. Muda (usia 16-35 tahun)
3. Dewasa ( 36-55 tahun)
4. Lansia (>55 tahun)

Profesi Korban : Profesi korban kecelakaan:

1. Pelajar/Mahasiswa
2. Wiraswasta/Wirausaha
3. PNS, TNI/POLRI
4. Pengemudi
5. Lain-lain

Kepemilikan SIM Korban : Kepemilikan SIM korban kecelakaan:

1. Memiliki
2. Tidak Memiliki

Jenis Kelamin Korban : Jenis kelamin korban kecelakaan:

1. Laki-laki
2. Perempuan

Pendidikan Korban : Pendidikan korban kecelakaan:

1. SD (Sekolah Dasar)
2. SMP (Sekolah Menengah Pertama)
3. SMA (Sekolah Menengah Atas)
4. PT (perguruan Tinggi)

## 4.3 Metode Pengumpulan Data

Penelitian ini data yang digunakan adalah data skunder yang diperoleh dari POLRES Kabupaten Sleman DIY. Data tersebut merupakan data kecelakaan lalu lintas tahun 2016 sampai 2017 di Kabupaten Sleman.

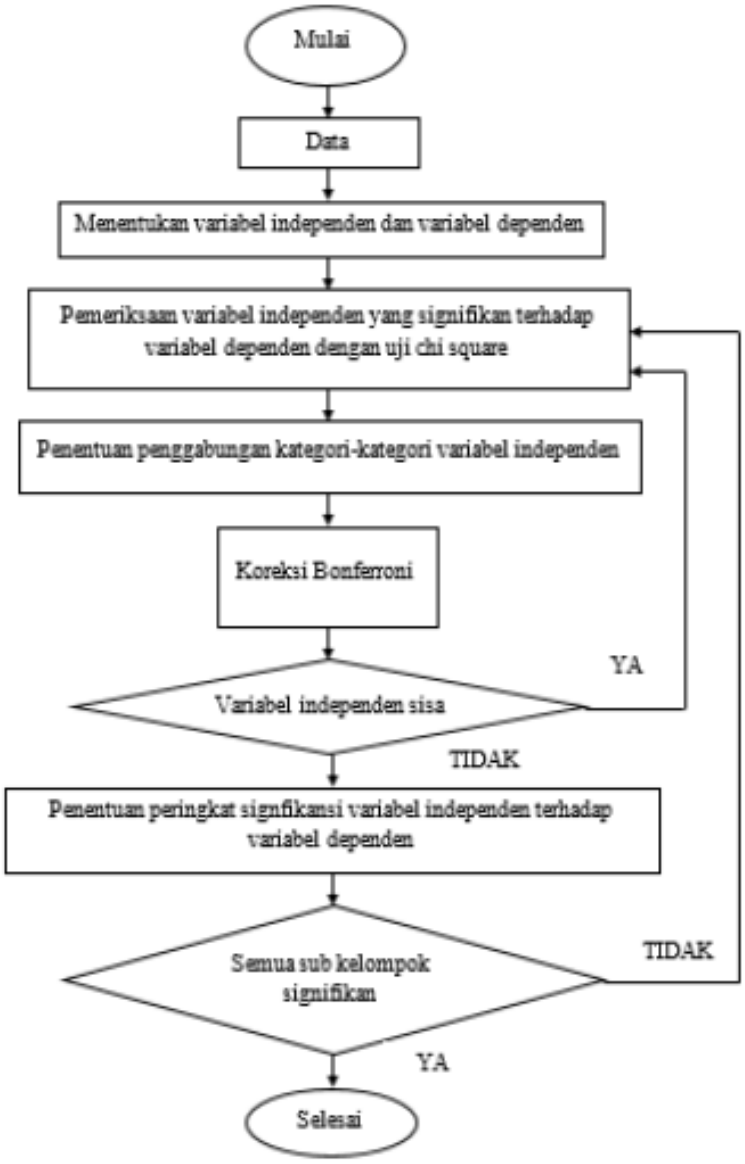
## 4.4 Metode Analisis Data

Pada penelitian ini ada beberapa software yang digunakan diantaranya *Microsoft excel* 2013 dan *Spina*. Beberapa metode yang digunakan dalam penelitian ini, yaitu:

1. Analisis Deskriptif, digunakan untuk menggambarkan secara umum data kecelakaan lalu lintas di POLRES Kabupaten Sleman tahun 2016 sampai 2017.
2. *CHAID*, digunakan untuk mengklasifikasikan data kecelakaan lalu lintas di POLRES Kabupaten Sleman tahun 2016 sampai 2017.

## 4.4 Tahapan Penelitian

Pada penelitian ini tahapan penelitian yang digunakan seperti gambar 4.1.

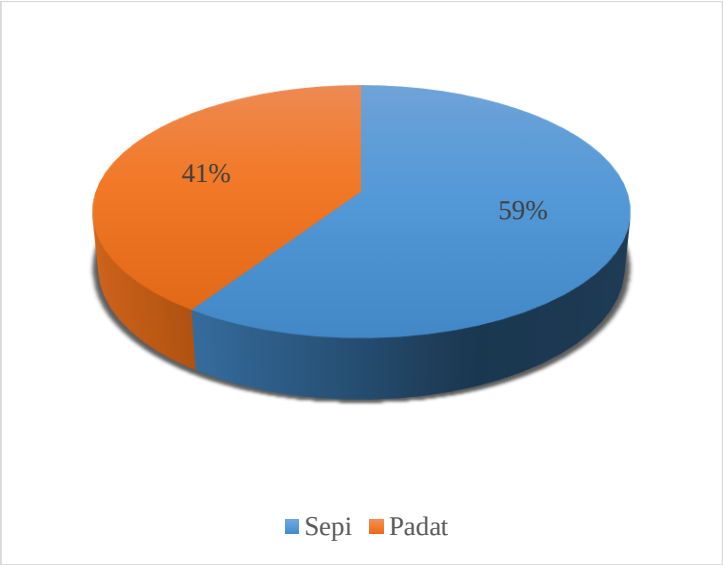



**Gambar 4.1** Diagram Alir Algoritma *CHAID*

# BAB V
# PEMBAHASAN

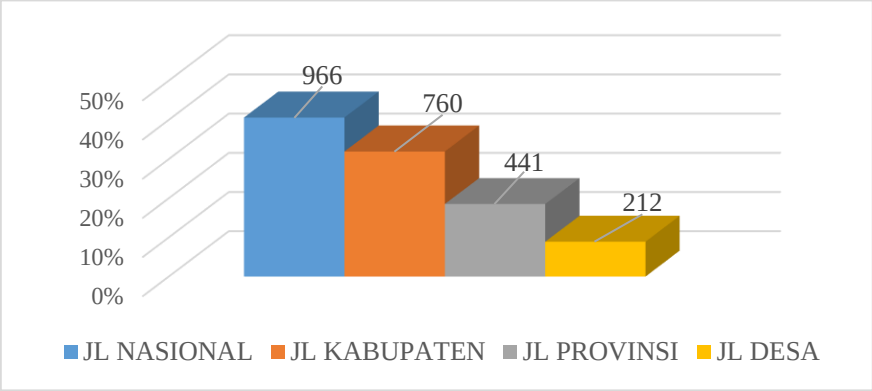
## 5.1 Gambaran Umum Data Kecelakaan Lalu Lintas

Analisis deskriptif pada penelitian ini adalah untuk melhat gambaran secara umum dari data kecelakaan lalu lintas di POLRES Kabupaten Sleman Tahun 2016 sampai 2017.



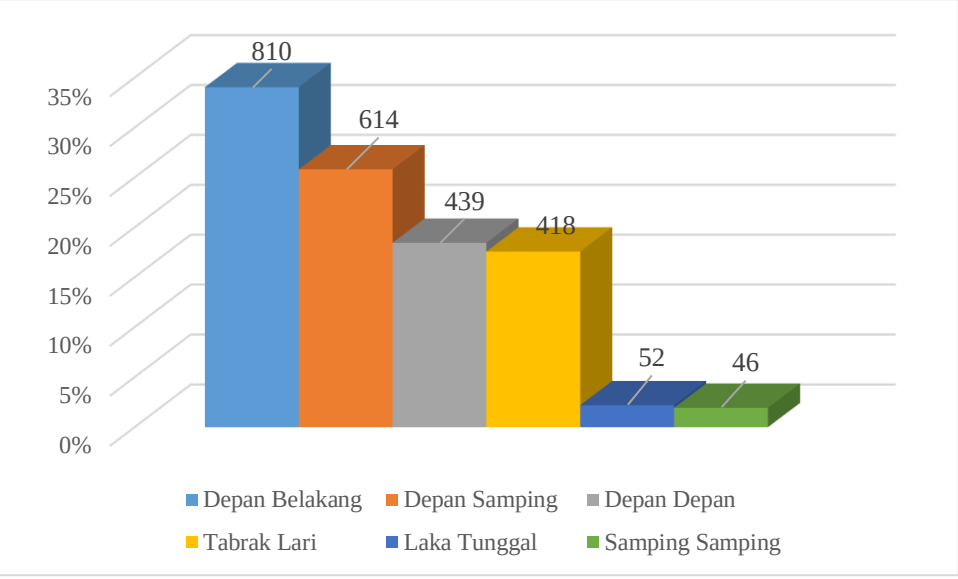

**Gambar 5.1** Kecelakaan Lalu lintas Berdasarkan Waktu Kejadian

Berdasarkan gambar 5.1 dapat dilihat bahwa kecelakaan lalu lintas berdasarkan waktu kejadian dibedakan menjadi waktu padat dan sepi. Kecelakaan lalu lintas berdasarkan waktu kejadian terjadi paling banyak atau didominasi pada waktu sepi yaitu sebanyak 1412 kejadian atau sebesar 59 persen, sementara pada waktu padat terjadi kecelakaan sebanyak 967 kejadian atau sebesar 41 persen.



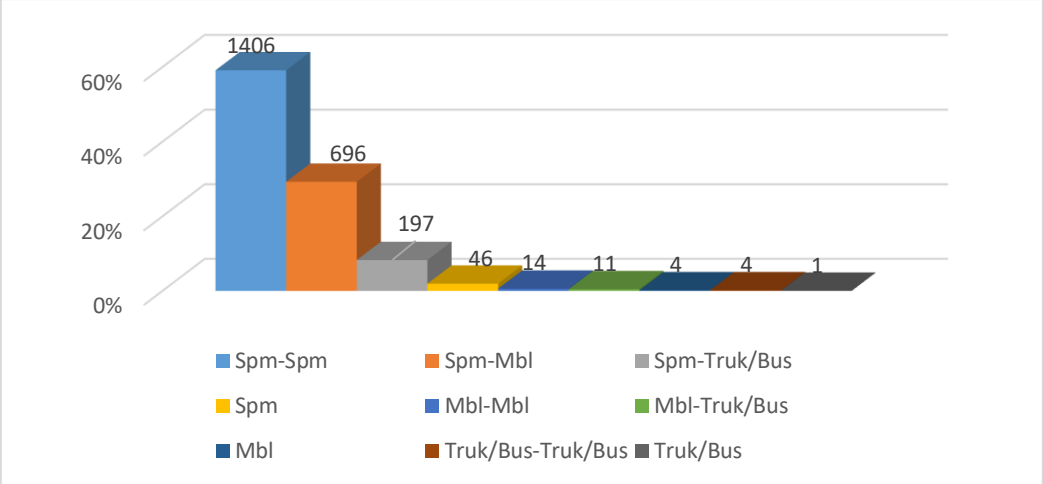

**Gambar 5.2** Kecelakaan Berdasarkan Tempat Kejadian

Pada gambar 5.2 terlihat bahwa tempat terjadinya kecelakaan lalu lintas terjadi pada jalan nasional, jalan kabupaten, jalan provinsi dan jalan desa. Jalan nasional adalah jalan yang paling banyak terjadi kecelakaan sebanyak 966 kejadian atau sebesar 41 persen diurutan kedua jalan kabupaten sebanyak 760 kejadian atau sebesar 32%, diurutan ketiga jl provinsi sebanyak 441 kejadian atau sebesar 19%, di jalan desa terjadi kecelakaan paling sedikit sebanyak 212 kejadian atau sebesar 9 persen.



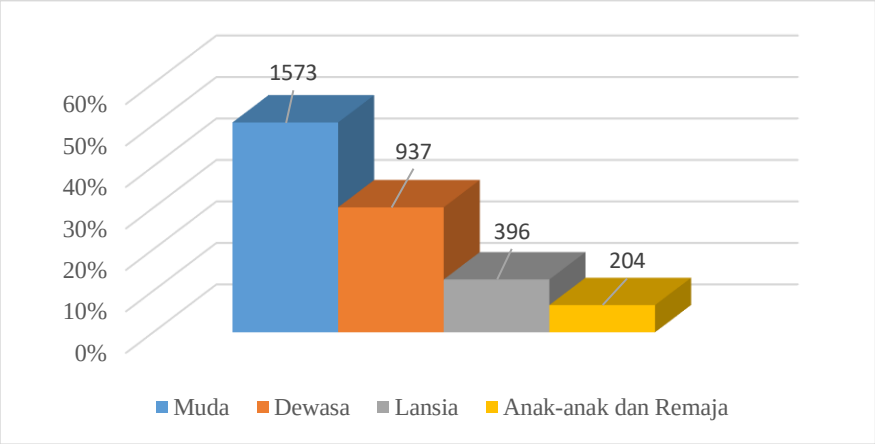

**Gambar 5.3** Kecelakaan Lalu Lintas Berdasarkan Jenis Kecelakaan

Gambar di atas (gambar 5.3) menunjukan grafik kecelakaan lalu lintas berdasarkan jenis kecelakaan. Jenis kecelakaan paling banyak terjadi adalah depan belakang sebanyak 810 kejadian atau sebesar 34 persen, diurutan kedua adalah depan samping sebanyak 614 kejadian atau sebesar 26 persen, diurutan ketiga dan keempat adalah depan depan dan tabrak lari memiliki persentase yang sama yaitu 18 persen. Posisi paling akhir yaitu samping samping, jenis kecelakaan tersebut sebanyak 46 kejadian.



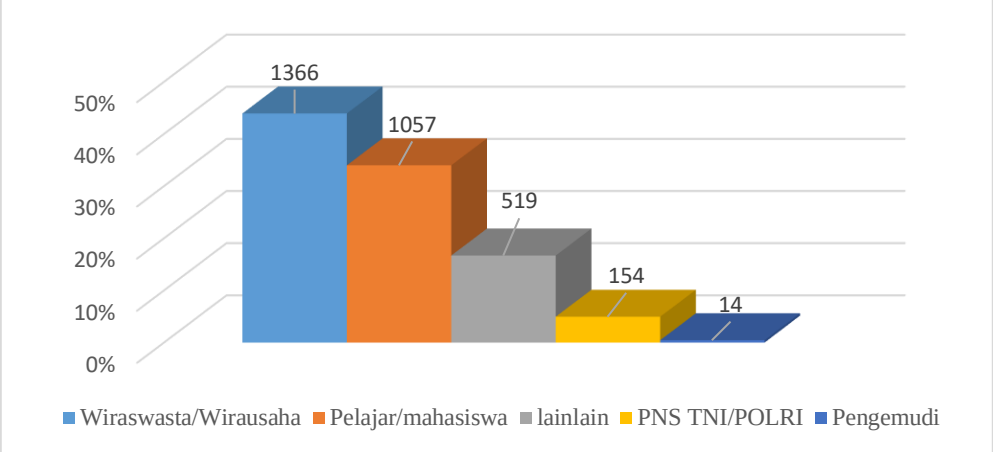

**Gambar 5.4** Kecelakaan Lalu Lintas Berdasarkan Kendaraan Terlibat

Berdasarkan gambar 5.4 dapat dilihat bahwa kecelakaan lalu lintas terjadi paling dominan melibatkan spm-spm dengan kejadian sebanyak 1406 kejadian atau sebesar 59 persen, diurutan kedua melibatkan spm-mbl sebanyak 696 kejadian atau 29 persen, diurutan ketiga melibatkan spm-truk/bus dengan jumlah 197 kejadian atau sebesar 8 persen, sementara spm dan mbl-mbl berada di urutan empat dan lima dengan selisih hanya 1 persen. Kendaraan yang terlibat selanjutnya memiliki jumlah yang sangat kecil maka ketika berbentuk persentase menjadi 0 persen yaitu mbl-truk/bus, mbl dan truk/bus-truk/bus, sementara truk/bus sebagai kendaraan terlibat paling sedikit hanya 1 kejadian.



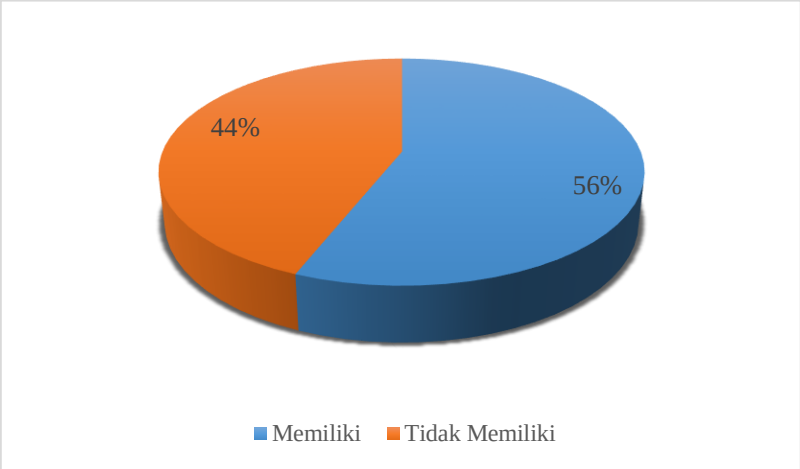

**Gambar 5.5** Kecelakaan Lalu Lintas Berdasarkan Umur

Berdasarkan gambar 5.5 kecelakaan lalu lintas berdasarkan umur dibedakan menjadi muda, dewasa, lansia, anak-anak dan remaja. Muda menjadi umur paling banyak terjadi kecelakaan lalu lintas sebanyak 1573 kejadian, lebih dari setengah nya kecelakaan lalu lintas melibatkan umur muda, selanjutnya diikuti dewasa sebanyak 937 kejadian atau sebanyak 30 persen, diurutan ketiga lansia sebanyak (13%) 394 kejadian dan paling sedikit adalah anak-anak dan remaja hanya sebanyak (7%) 204 kejadian.



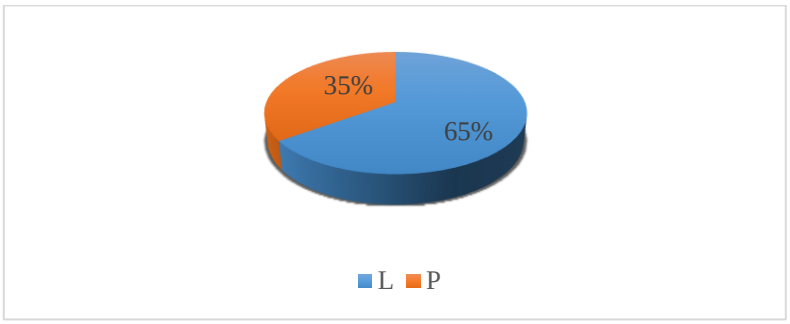

**Gambar 5.6** Kecelakaan Lalu Lintas Berdasarkan Profesi

Gambar 5.6 menunjukan bahwa kecelakaan lalu lintas berdasarkan profesi paling banyak adalah wiraswasta/wirausaha sebanyak 1366 kejadian atau sebesar 44 persen, diurutan kedua pelajar/mahasiswa sebanyak 1057 kejadian atau sebesar 34 persen, diikuti profesi lain-lain sebesar 17 persen, diurutan keempat PNS, TNI/POLRI sebesar 5 persen dan paling sedikit berprofesi sebagai sopir atau pengemudi 14 kejadian atau dalam bentuk persen sebesar 0 persen.



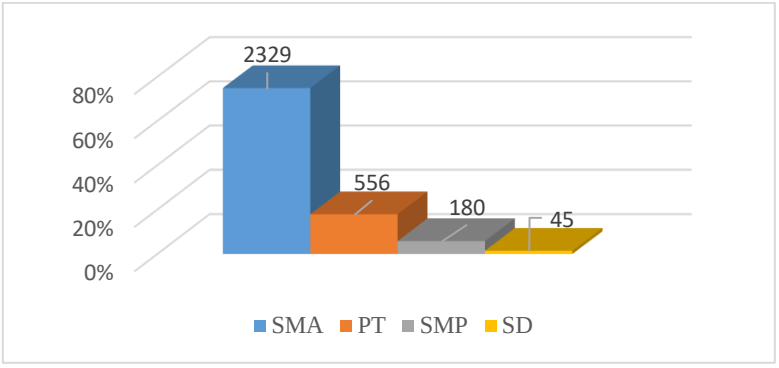

**Gambar 5.7** Kecelakaan Lalu Lintas Berdasarkan Kepemilikan SIM

Gambar 5.7 menampilkan grafik kecelakaan lalu lintas berdasarkan kepemilikan SIM. Kepemlikian SIM dibedakan menjadi memiliki dan tidak memiliki. Grafik menunjukan bahwa kecelakaan lalu lintas terjadi paling banyak justru pada pengguna kendaraan bermotor yang memiliki SIM sebanyak 1743 kejadian atau sebesar 56 persen. SIM cara memperoleh tidak mudah dan gampang harus memenuhi syarat-syarat tertentu dan melakukan ujian SIM yang terdiri dari ujian teori dan ujian praktek. Ujian teori terdiri dari 30 soal, dinyatakan lulus apabila benar 21 soal, setelah dinyatakan lulus maka dapat melakukan ujian praktek SIM berupa mengendarai kendaraan bermotor pada lintasan khusus, setelah semuanya lulus maka berhak untuk mendapatkan SIM. Diurutan kedua pengguna kendaraan bermotor yang tidak memiliki SIM sebanyak 1367 kejadian atau sebesar 44 persen.




**Gambar 5.8** Kecelakaan Lalu Lintas Berdasarkan Jenis Kelamin

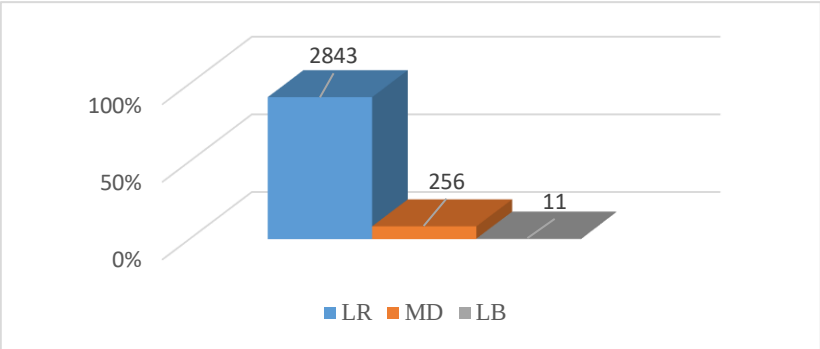
Berdasarkan gambar 5.8 kecelakaan lalu lintas berdaasarkan jenis kelamin paling banyak terjadi pada jenis kelamin lak-laki sebesar 983 kejadian 65 persen lebih dari setengahnya kecelakaan lalu lintas melibatkan jenis kelamin laki-laki, selanjutnya diurutan kedua jenis kelamin perempuan hanya sebesar 35 persen.




**Gambar 5.9** Kecelakaan Lalu Lintas Berdasarkan Pendidikan

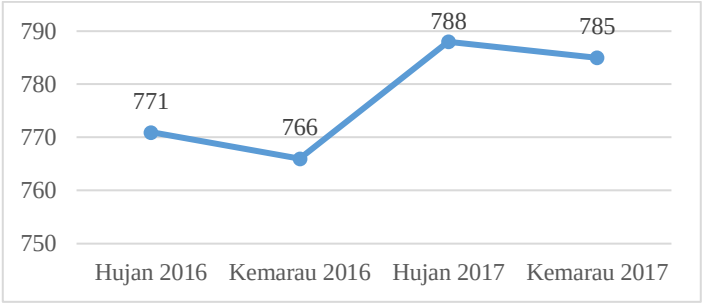
Kecelakaan lalu lintas berdasarkan pendidikan sebagai penyumbang paling menonjol atau terbanyak adalah berpendidikan SMA sebanyak 2329 kejadian atau sebesar 75 persen, diurutan kedua pengguna kendaraan bermotor dengan pendidikan PT dengan selisih yang cukup jauh dengan pengguna kendaraan bermotor berpendidikan SMA, yakni dengan selisih 1773 kejadian, sementara jumlah kejadian pada urutan kedua sebanyak 556 kejadian atau sebesar 18 persen. Diurutan ketiga

berpendidikan SMP sebesar 6 persen dan paling sedikit berpendidikan SD hanya sebesar 1 persen, pernyataan diatas sesuai dengan gambar 5.9.




**Gambar 5.10** Kecelakaan Lalu Lintas Berdasarkan Tingkat Luka Korban

Kecelakaan lalu lintas memiliki dampak yang buruk, akibat dari kecelakaan lalu lintas mengakibatkan kerugian materi dan korban, tingkat luka korban kecelakaan lalu lintas di kategorikan menjadi 3 kategori, yakni luka ringan, luka berat dan meninggal dunia. Pada gambar 5.10 dapat dilihat bahwa kecelakaan lalul lintas paling dominan mengakibatkan korban mengalami luka ringan sebesar 91 persen, diurutan kedua korban mengalami meninggal dunia sebesar 8 persen, sementara luka berat sebagai luka yang dialami korban kecelakaan paling sedikit sebesar 11 kejadian atau dalam bentuk persentase sebesar 0 persen.




**Gambar 5.11** Grafik Jumlah Kecelakaan di Musim Hujan dan Kemarau

Berdasarkan grafik di atas dapat dilihat bahwa kecelakaan pada saat musim hujan memiliki jumlah yang lebih banyak di bandingkan dengan musim kemarau, hal

ini dapat dilihat di musim hujan tahun 2016 sebanyak771 kejadian dan di tahun 2016 sebanyak 788 kejadian, sementara pada musim kemarau di tahun 2016 sebanyak 766 dan pada tahun 2017 sebanyak 785.

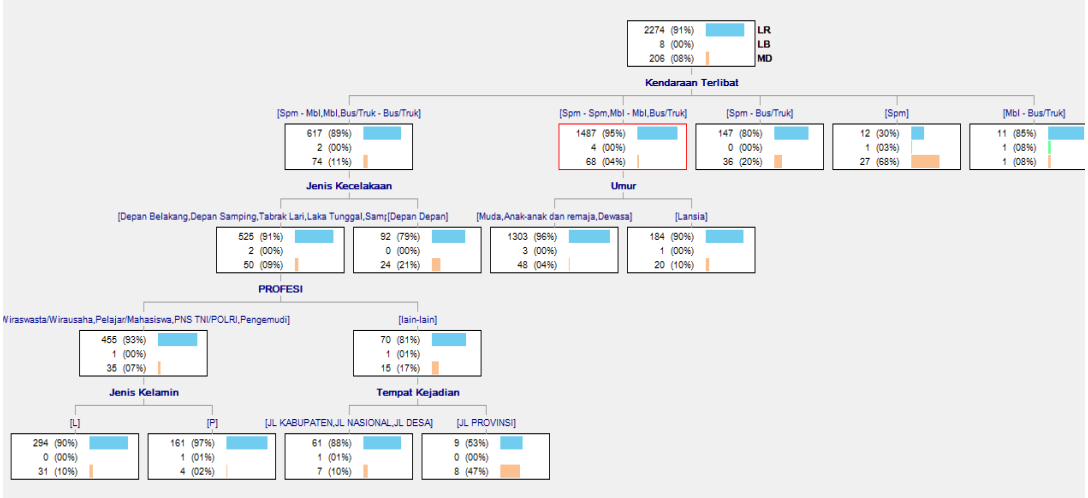
### 5.2 Pembagian Data *Training* dan Data *Testing*

Pada penelitian ini, data yang digunakan adalah data kecelakaan lalu lintas di POLRES Kabupaten Sleman yang berjumlah 3110 data. Data di atas nantinya akan dibagi menjadi dua yaitu sebagai data *training* dan data *testing*. Dalam penelitian ini penelitian dibagi menjadi 80% untuk data *training* dan 20% untuk data *testing*. Pembagian data tersebut bebas ditentukan sendiri oleh peneliti sesuai dengan tujuan penelitian, sementara penentuan pembagian data *training* dan data *testing* pada penelitian ini mengacu pada penelitian Estoatnowo tahun 2012.

### 5.3 Analisis *CHAID*

Analisis yang digunakan dalam penelitian ini adalah metode *CHAID*. *Output* dari metode *CHAID* akan membentuk pohon keputusan atau aturan klasifikasi yang bisa digunakan untuk mengetahui klasifikasi dari data kecelakaan lalu lintas di POLRES kabupaten Sleman tahun 2016 sampai tahun 2017**.** Aplikasi atau *Software* yang digunakan yaitu *Sipina*, *software* tersebut *open source* jadi peneliti dapat menggunakanya tanpa harus membayar *license*.

Hasil *output* dari metode *CHAID* (gambar 5.1) berupa pohon klasifikasi dengan *node* 16 buah, *node* akhir terdapat 4 buah dan depth atau kedalaman sebanyak 4. *Node* paling atas diketahui jumlah total data tingkat luka 3110 korban yang terdiri dari 2274 korban (91%) dengan tingkat luka ringan, 8 korban (00%) dengan tingkat luka parah dan 206 korban (08%) dengan tingkat luka meninggal dunia. Variabel independen yang signifikan dan dapat membentuk pohon adalah variabel Kendaraan Terlibat, Jenis Kelamin, Umur, Profesi dan Tempat Kejadian.




**Gambar 5.12** Pohon Keputusan Metode *CHAID*

Berdasarkan klasifikasi metode *CHAID* (gambar 5.11) terbentuk 10 segmen yang berbeda, sebagai berikut:

1. Jika kendaraan terlibat Spm-Mbl, Mbl, Bus/Truk-Bus/Truk dan jenis kecelakaan depan belakang, depan samping, tabrak lari, laka tunggal, samping-samping dan profesi wiraswasta/wirausaha, pelajar.mahasiswa, PNS TNI/POLRI, pengemudi dan jenis kelamin L maka dari data diperoleh luka ringan sebanyak 294 korban dan meninggal dunia sebanyak 31 korban. Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia.
2. Jika kendaraan terlibat Spm-Mbl, Mbl, Bus/Truk-Bus/Truk dan jenis kecelakaan depan belakang, depan samping, tabrak lari, laka tunggal, samping-samping dan profesi wiraswasta/wirausaha, pelajar.mahasiswa, PNS TNI/POLRI, pengemudi dan jenis kelamin P maka dari data diperoleh luka ringan sebanyak 161 korban, luka berat 1 korban dan meninggal dunia sebanyak 4 korban. Sehingga korban pada segmen ini diduga akan mengalami luka berat dan meninggal dunia.
3. Jika kendaraan terlibat Spm-Mbl, Mbl, Bus/Truk-Bus/Truk dan jenis kecelakaan depan belakang, depan samping, tabrak lari, laka tunggal, samping-samping dan profesi lain-lain dan tempat kejadian jl kabupaten, jl nasional, jl desa maka dari data diperoleh luka ringan sebanyak 61 korban, luka berat 1 korban dan meninggal

dunia sebanyak 7 korban. Sehingga korban pada segmen ini diduga akan mengalami luka berat dan meninggal dunia.

4. Jika kendaraan terlibat Spm-Mbl, Mbl, Bus/Truk-Bus/Truk dan jenis kecelakaan depan belakang, depan samping, tabrak lari, laka tunggal, samping-samping dan profesi lain-lain dan tempat kejadian jl provinsi maka dari data diperoleh luka ringan sebanyak 9 korban dan meninggal dunia sebanyak 8 korban. Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia.
5. Jika kendaraan terlibat Spm-Mbl, Mbl, Bus/Truk-Bus/Truk dan jenis kecelakaan depan depan maka dari data diperoleh luka ringan sebanyak 92 korban dan meninggal dunia sebanyak 8 korban. Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia.
6. Jika kendaraan terlibat Spm-Spm, Mbl-Mbl, Bus/Truk dan umur muda,anak-anak dan remaja, dewasa maka dari data diperoleh luka ringan sebanyak 1303 korban, luka berat sebanyak 4 korban dan meninggal dunia sebanyak 48 korban. Sehingga korban pada segmen ini diduga akan mengalami luka berat dan meninggal dunia.
7. Jika kendaraan terlibat Spm-Spm, Mbl-Mbl, Bus/Truk dan umur lansia maka dari data diperoleh luka ringan sebanyak 184 korban, luka berat sebanyak 1 korban dan meninggal dunia sebanyak 20 korban. Sehingga korban pada segmen ini diduga akan mengalami luka berat dan meninggal dunia.
8. Jika kendaraan terlibat Spm-Bus/Truk maka dari data diperoleh luka ringan sebanyak 147 korban dan meninggal dunia sebanyak 36 korban. Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia.
9. Jika kendaraan terlibat Spm maka dari data diperoleh luka ringan sebanyak 12 korban, luka berat 1 korban dan meninggal dunia 27 korban. Sehingga korban pada segmen ini diduga akan mengalami luka berat dan meninggal dunia.
10. Jika kendaraan terlibat Mbl-Bus/Truk maka dari data diperoleh luka ringan sebanyak 11 korban, luka berat 1 korban dan meninggal dunia 1 korban. Sehingga korban pada segmen ini diduga akan mengalami luka berat dan meninggal dunia.

### 5.4 Tingkat Akurasi dan Presisi Metode *CHAID*

Berdasarkan 10 aturan atau segmen di atas maka dilakukan perhitungan nilai akurasi dan presisi supaya dapat mengetahui sejauh mana model dalam mengklasifikasikan dengan benar. Berikut adalah tabel tingkat akurasi dan presisi dengan mengunakan metode *CHAID*

**Tabel 5.1 Tingkat Akurasi dan Presisi Metode *CHAID***

| **Data** | **Tingkat Luka** | **LR** | **LB** | **MD** | |
|---|---|---|---|---|---|
| ***Training*** | LR | 2262 | 0 | 12 | |
| | LB | 7 | 0 | 1 | |
| | MD | 179 | 0 | 27 | |
| | Persentase | 92,4% | 0% | 68% | 92% |
| **Data** | **Tingkat Luka** | **LR** | **LB** | **MD** | |
| ***Testing*** | LR | 563 | 0 | 6 | |
| | LB | 3 | 0 | 0 | |
| | MD | 38 | 0 | 12 | |
| | Persentase | 93% | 0% | 67% | 92,4% |

Akurasi $= \frac{563+0+12}{563+0+6+3+0+0+38+0+12} = 92.4\%$

Presisi LR $= \frac{563}{563+3+38} = 93\%$

Presisi LB $= \frac{0}{0+0+0} = 0\%$

Presisi MD $= \frac{12}{6+0+12} = 67\%$

Berdasarkan Tabel 5.1 dapat dilihat bahwa klasifikasi yang terbentuk dengan menggunakan metode *CHAID*, dengan persentase data *training* 80% dan data *testing* 20%, diperoleh tingkat akurasi sebesar 92.4%

Padat tabel 5.1 dapat dilihat tingkat akurasi sebesar 92.4% sudah cukup baik, tetapi model dalam memprediksi masih terdapat kelas data yaitu luka berat yang tidak ada satupun yang benar. Hal ini terjadi karena ketidakseimbangan kelas. Permasalahan ketidakseimbangan kelas adalah kondisi dimana jumlah kelas memiliki perbedaan yang tinggi atau bisa dikatakan pengamatan pada suatu kelas kurang dengan kelas lainya. Pada pohon keputusan (gambar 5.11) paling atas kelas data luka ringan sebanyak 2274, luka berat sebanyak 8 dan meninggal dunia sebanyak 206. Kelas data luka ringan jumlahnya sangat banyak jika dibandingkan dengan jumlah kelas data luka berat dan meninggal dunia. Permasalahan ini bisa mempengaruhi kinerja prediksi model dikarenakan model akan lebih condong memprediksi kelas yang memiliki jumlah lebih besar dari data sampel. Proses *SMOTE* (*Synthetic Minority Over-sampling Technique*) akan dilakukan pada penelitian ini yang nantinya akan digunakan untuk membuat kelas data menjadi seimbang. Penggunaan teknik *SMOTE* menghasilkan hasil yang baik dan cara yang efektif untuk menangani ketidakseimbangan kelas yang mengalami *overfitting* pada teknik *oversampling* untuk memproses kelas minoritas (positif) (Chawla, 2002).

Kelas data yang tidak seimbang dapat dilihat dari jumlahnya. Dalam penelitian ini kelas data luka ringan sebanyak 2274, luka berat sebanyak 8 dan meninggal dunia sebanyak 206. Kelas luka ringan memiliki jumlah yang jauh lebih banyak jika dibandingkan dengan luka berat dan meninggal dunia. Kelas data yang memiliki jumlah sedikit akan diduplikasi supaya menjadi seimbang melalui tahap *SMOTE*. Proses duplikasi data pada data yang memiliki jumlah sedikit dari *oversampling* sebanyak 100% dan penerapan sebanyak 10 kali maka diperoleh jumlah kelas data pada luka ringan sebanyak 2843, luka berat sebanyak 1408 dan meninggal dunia sebanyak 2048.

### 5.5 Analisis *CHAID* dengan *SMOTE*

Langkah berikutnya melakukan analisis dengan metode *CHAID* dengan menggunakan data yang sudah dilakukan *SMOTE* dengan persentas leraning dan

*testing* masing-masing 80% dan 20%. Hasil pohon keputusan dengan data setelah *SMOTE* terlihat pada lampiran 2.

Berdasarkan pohon keputusan (lampiran 2) maka dapat disimpulkan bahwa seluruh variabel independen signifikan dan terbentuk 78 segmen yang berbeda, sebagai berikut:

1. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan depan belakang dan kendaraan terlibat Spm-Mbl, Mbl-Bus/Truk dan profesi wirausaha/wiraswasta dan tempat kejadian jl kabupaten, jl desa maka dari data diperoleh luka ringan sebanyak 9 (35%) dan meninggal dunia 17 (65%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 1).
2. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan depan belakang dan kendaraan terlibat Spm-Mbl, Mbl-Bus/Truk dan profesi wirausaha/wiraswasta dan tempat kejadian jl nasional, jl provinsi maka dari data diperoleh luka ringan sebanyak 23 (05%) dan luka berat 408 (94%) dan meninggal dunia 2 (00%). Sehingga korban pada segmen ini diduga akan mengalami luka berat. (Segmen 2).
3. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan depan belakang dan kendaraan terlibat Spm-Mbl, Mbl-Bus/Truk dan profesi pelajar/mahasiswa, PNS TNI/POLRI maka dari data diperoleh luka ringan sebanyak 35 (90%) dan meninggal dunia 4 (10%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 3).
4. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan depan belakang dan kendaraan terlibat Spm-Mbl, Mbl-Bus/Truk dan profesi lain-lain maka dari data diperoleh luka ringan sebanyak 12 (20%) dan meninggal dunia 39 (80%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 4).
5. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan depan belakang dan kendaraan terlibat Spm-Spm, Mbl-Mbl dan tempat kejadian jl kabupaten, jl provinsi, jl desa dan waktu kejadian padat maka dari data diperoleh luka ringan

sebanyak 60 (41%) dan meninggal dunia 87 (59%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 5).

6. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan depan belakang dan kendaraan terlibat Spm-Spm, Mbl-Mbl dan tempat kejadian jl kabupaten, jl provinsi, jl desa dan waktu kejadian sepi maka dari data diperoleh luka ringan sebanyak 69 (75%) dan luka berat 3(03) meninggal dunia 20 (22%). Sehingga korban pada segmen ini diduga akan mengalami luka berat dan meninggal dunia. (Segmen 6).
7. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan depan belakang dan kendaraan terlibat Spm-Spm, Mbl-Mbl dan tempat kejadian jl nasional maka dari data diperoleh luka ringan sebanyak 68 (97%), luka berat sebanyak 1 (01%) dan meninggal dunia 1 (01%). Sehingga korban pada segmen ini diduga akan mengalami luka berat dan meninggal dunia. (Segmen 7).
8. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan depan belakang dan kendaraan terlibat Spm-Bus/Truk maka dari data diperoleh luka ringan sebanyak 26 (39%) dan meninggal dunia 41 (61%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 8).
9. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan depan samping dan waktu kejadian padat dan profesi wiraswasta/wirausaha, pelajar/mahasiswa, PNS TNI/POLRI, pengemudi maka dari data diperoleh luka ringan sebanyak 108 (99%) dan meninggal dunia 1 (01%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 9).
10. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan depan samping dan waktu kejadian padat dan profesi lain-lain maka dari data diperoleh luka ringan sebanyak 23 (74%) dan meninggal dunia 8 (26%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 10).
11. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan depan samping dan waktu kejadian sepi dan tempat kejadian jl kabupaten dan kendaraan terlibat Spm – Spm maka dari data diperoleh luka ringan sebanyak 25 (08%), luka berat

sebanyak 272 (87%) dan meninggal dunia 14 (05%). Sehingga korban pada segmen ini diduga akan mengalami luka berat dan meninggal dunia. (Segmen 11).

12. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan depan samping dan waktu kejadian sepi dan tempat kejadian jl nasional, jl desa dan kendaraan terlibat Spm - Mbl maka dari data diperoleh luka ringan sebanyak 20 (56%), luka berat sebanyak 1 (03%) dan meninggal dunia 15 (42%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 12).
13. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan depan samping dan waktu kejadian sepi dan tempat kejadian jl nasional, jl desa dan kendaraan terlibat Spm - Spm maka dari data diperoleh luka ringan sebanyak 50 (93%), luka berat sebanyak 2 (04%) dan meninggal dunia 2 (04%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 13).
14. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan depan samping dan waktu kejadian sepi dan tempat kejadian jl nasional, jl desa dan kendaraan terlibat Spm – Bus/Truk maka dari data diperoleh luka ringan sebanyak 4 (21%), dan meninggal dunia 15 (79%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 14).
15. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan depan samping dan waktu kejadian sepi dan tempat kejadian jl provinsi dan umur muda maka dari data diperoleh luka ringan sebanyak 14 (3%), luka berat sebanyak 440 (95%) dan meninggal dunia 8 (2%). Sehingga korban pada segmen ini diduga akan mengalami luka berat dan meninggal dunia. (Segmen 15).
16. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan depan samping dan waktu kejadian sepi dan tempat kejadian jl provinsi dan umur anak- anak & remaja, dewasa maka dari data diperoleh luka ringan sebanyak 18 (95%), dan meninggal dunia 1 (05%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 16).
17. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan tabrak lari, depan-depan, samping-samping dan tempat kejadian jl kabupaten, jl provinsi dan

kendaraan terlibat Spm-Mbl, Spm-Bus/Truk dan pendidikan SMA maka dari data diperoleh luka ringan sebanyak 46 (18%), luka berat sebanyak 1 (00%) dan meninggal dunia 203 (81%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 17).

18. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan tabrak lari, depan-depan, samping-samping dan tempat kejadian jl kabupaten, jl provinsi dan kendaraan terlibat Spm-Mbl, Spm-Bus/Truk dan pendidikan SD, PT, SMP maka dari data diperoleh luka ringan sebanyak 29 (63%), dan meninggal dunia 17 (37%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 18).
19. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan tabrak lari, depan-depan, samping-samping dan tempat kejadian jl kabupaten, jl provinsi dan kendaraan terlibat Spm-Spm, Bus/Truk-Bus/Truk dan umur muda, dewasa maka dari data diperoleh luka ringan sebanyak 136 (80%), dan meninggal dunia 35 (20%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 19).
20. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan tabrak lari, depan-depan, samping-samping dan tempat kejadian jl kabupaten, jl provinsi dan kendaraan terlibat Spm-Spm, Bus/Truk-Bus/Truk dan umur anak & remaja, lansia maka dari data diperoleh luka ringan sebanyak 34 (39%), dan meninggal dunia 54 (61%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 20).
21. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan tabrak lari, depan-depan, samping-samping dan tempat kejadian jl nasional, jl desa dan kendaraan terlibat Spm-Mbl, Spm-Bus/Truk maka dari data diperoleh luka ringan sebanyak 37 (69%), dan meninggal dunia 17 (31%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 21).
22. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan tabrak lari, depan-depan, samping-samping dan tempat kejadian jl nasional, jl desa dan kendaraan

terlibat Spm-Spm, Mbl-Mbl, Mbl-Bus/Truk maka dari data diperoleh luka ringan sebanyak 98 (99%), dan meninggal dunia 1 (01%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 22).

23. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan laka tunggal dan jenis kelamin laki-laki maka dari data diperoleh luka ringan sebanyak 6 (06%), dan meninggal dunia 100 (94%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 23).
24. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan laka tunggal dan jenis kelamin perempuan maka dari data diperoleh luka ringan sebanyak 4 (67%), dan meninggal dunia 2 (33%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 24).
25. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm – Mbl dan jenis kecelakaan depan belakang, depan-depan dan umur muda dan jenis kelamin laki-laki maka dari data diperoleh luka ringan sebanyak 57 (24%), dan meninggal dunia 185 (76%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 25).
26. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm-Mbl dan jenis kecelakaan depan belakang, depan-depan dan umur muda dan jenis kelamin perempuan maka dari data diperoleh luka ringan sebanyak 33 (59%), dan meninggal dunia 23 (41%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 26).
27. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm Mbl dan jenis kecelakaan depan belakang, depan-depan dan umur anak-anak & remaja, dewasa, lansia maka dari data diperoleh luka ringan sebanyak 93 (63%), dan meninggal dunia 54 (37%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 27).
28. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm-Mbl dan jenis kecelakaan depan samping, tabrak lari, laka tunggal, samping-samping dan umur muda, anak & remaja, dewasa dan tempat kejadian jl kabupaten maka dari data

diperoleh luka ringan sebanyak 35 (83%), dan meninggal dunia 7 (17%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 28).

29. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm-Mbl dan jenis kecelakaan depan samping, tabrak lari, laka tunggal, samping-samping dan umur muda, anak & remaja, dewasa dan tempat kejadian jl nasional, jl desa maka dari data diperoleh luka ringan sebanyak 112 (98%), dan meninggal dunia 2 (02%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 29).
30. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm-Mbl dan jenis kecelakaan depan samping, tabrak lari, laka tunggal, samping-samping dan umur muda, anak & remaja, dewasa dan tempat kejadian jl provinsi maka dari data diperoleh luka ringan sebanyak 22 (48%), dan meninggal dunia 24 (52%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 30).
31. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm-Mbl dan jenis kecelakaan depan samping, tabrak lari, laka tunggal, samping-samping dan umur lansia dan tempat kejadian jl kabupaten maka dari data diperoleh luka ringan sebanyak 4 (12%), dan meninggal dunia 29 (88%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 31).
32. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm-Mbl dan jenis kecelakaan depan – samping, tabrak lari, laka tunggal, samping-samping dan umur lansia dan tempat kejadian jl provinsi, jl nasional, jl desa maka dari data diperoleh luka ringan sebanyak 13 (62%), dan meninggal dunia 8 (38%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 32).
33. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm-Spm, Mbl-mbl, Mbl-Bus/Truk, Bus/Truk-Bus/Truk dan tempat kejadian jl kabupaten umur muda maka dari data diperoleh luka ringan sebanyak 138 (95%), luka berat sebanyak 1 (1%)dan meninggal dunia 6 (4%). Sehingga korban pada segmen ini diduga akan mengalami luka berat dan meninggal dunia. (Segmen 33).

34. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm-Spm, Mbl-mbl, Mbl-Bus/Truk, Bus/Truk-Bus/Truk dan tempat kejadian jl kabupaten umur anak-anak & remaja, lansia dan jenis kelamin laki-laki maka dari data diperoleh luka ringan sebanyak 30 (29%) dan meninggal dunia 75 (71%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 34).
35. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm-Spm, Mbl-mbl, Mbl-Bus/Truk, Bus/Truk-Bus/Truk dan tempat kejadian jl kabupaten umur anak-anak & remaja, lansia dan jenis kelamin perempuan maka dari data diperoleh luka ringan sebanyak 19 (95%) dan meninggal dunia 1 (5%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 35).
36. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm-Spm, Mbl-mbl, Mbl-Bus/Truk, Bus/Truk-Bus/Truk dan tempat kejadian jl kabupaten umur dewas dan waktu kejadian padat maka dari data diperoleh luka ringan sebanyak 38 (83%) dan meninggal dunia 8 (17%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 36).
37. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm–Spm, Mbl-mbl, Mbl-Bus/Truk, Bus/Truk-Bus/Truk dan tempat kejadian jl kabupaten umur dewas dan waktu kejadian sepi maka dari data diperoleh luka ringan sebanyak 53 (47%) dan meninggal dunia 59 (53%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 37).
38. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm-Spm, Mbl-Mbl, Mbl-Bus/Truk, Bus/Truk-Bus/Truk dan tempat kejadian jl provinsi dan jenis kecelakaan depan belakang, depan-depan maka dari data diperoleh luka ringan sebanyak 77 (83%) dan meninggal dunia 16 (17%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 38).
39. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm-Spm, Mbl-mbl, Mbl-Bus/Truk, Bus/Truk-Bus/Truk dan tempat kejadian jl provinsi dan jenis kecelakaan depan samping maka dari data diperoleh luka ringan sebanyak 47

(58%) dan meninggal dunia 34 (42%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 39).

40. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm-Spm, Mbl-Mbl, Mbl-Bus/Truk, Bus/Truk-Bus/Truk dan tempat kejadian jl provinsi dan jenis kecelakaan tabrak lari maka dari data diperoleh luka ringan sebanyak 31 (100%). (Segmen 40).
41. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm-Spm, Mbl-Mbl, Mbl-Bus/Truk, Bus/Truk-Bus/Truk dan tempat kejadian jl desa maka dari data diperoleh luka ringan sebanyak 91 (99%) dan luka berat 1 (01%). Sehingga korban pada segmen ini diduga akan mengalami luka berat. (Segmen 41).
42. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm-Bus/Truk dan tempat kejadian jl nasional, jl kabupaten dan jenis kecelakaan depan belakang, depan samping maka dari data diperoleh luka ringan sebanyak 43 (20%) dan meninggal dunia 168 (80%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 42).
43. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm-Bus/Truk dan tempat kejadian jl nasional, jl kabupaten dan jenis kecelakaan tabrak lari, depan-depan, samping-samping maka dari data diperoleh luka ringan sebanyak 17 (65%) dan meninggal dunia 9 (35%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 43).
44. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm-Bus/Truk dan tempat kejadian jl provinsi, jl desa maka dari data diperoleh luka ringan sebanyak 22 (85%) dan meninggal dunia 4 (15%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 44).
45. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm dan profesi wirausaha/wiraswasta, pelajar/mahasiswa maka dari data diperoleh luka ringan sebanyak 4 (2%) dan meninggal dunia 159 (98%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 45).

46. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm dan profesi PNS TNI/POLRI, lain - lain maka dari data diperoleh luka ringan sebanyak 3 (60%), luka berat 1 (20%) dan meninggal dunia 1 (20%). Sehingga korban pada segmen ini diduga akan mengalami luka berat dan meninggal dunia. (Segmen 46).

Metode *CHAID* dapat meminimalisir atau meminimumkan jumlah korban yang mengalami luka berat dan meninggal dunia, maka peneliti memberikan rekomendasi karakteristik korban kepada POLRES Kabupaten Sleman seperti berikut:

1. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan depan belakang dan kendaraan terlibat Spm-Mbl, Mbl-Bus/Truk dan profesi wirausaha/wiraswasta dan tempat kejadian jl kabupaten, jl desa maka dari data diperoleh luka ringan sebanyak 9 (35%) dan meninggal dunia 17 (65%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 1).
2. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan depan belakang dan kendaraan terlibat Spm-Mbl, Mbl-Bus/Truk dan profesi wirausaha/wiraswasta dan tempat kejadian jl nasional, jl provinsi maka dari data diperoleh luka ringan sebanyak 23 (05%) dan luka berat 408 (94%) dan meninggal dunia 2 (00%). Sehingga korban pada segmen ini diduga akan mengalami luka berat. (Segmen 2).
3. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan depan belakang dan kendaraan terlibat Spm-Mbl, Mbl-Bus/Truk dan profesi lain-lain maka dari data diperoleh luka ringan sebanyak 12 (20%) dan meninggal dunia 39 (80%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 4).
4. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan depan belakang dan kendaraan terlibat Spm-Spm, Mbl-Mbl dan tempat kejadian jl kabupaten, jl provinsi, jl desa dan waktu kejadian padat maka dari data

diperoleh luka ringan sebanyak 60 (41%) dan meninggal dunia 87 (59%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 5).

5. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan depan belakang dan kendaraan terlibat Spm-Bus/Truk maka dari data diperoleh luka ringan sebanyak 26 (39%) dan meninggal dunia 41 (61%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 8).
6. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan depan samping dan waktu kejadian sepi dan tempat kejadian jl kabupaten dan kendaraan terlibat Spm-Spm maka dari data diperoleh luka ringan sebanyak 25 (08%), luka berat sebanyak 272 (87%) dan meninggal dunia 14 (05%). Sehingga korban pada segmen ini diduga akan mengalami luka berat dan meninggal dunia. (Segmen 11).
7. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan depan samping dan waktu kejadian sepi dan tempat kejadian jl nasional, jl desa dan kendaraan terlibat Spm-Bus/Truk maka dari data diperoleh luka ringan sebanyak 4 (21%), dan meninggal dunia 15 (79%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 14).
8. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan depan samping dan waktu kejadian sepi dan tempat kejadian jl provinsi dan umur muda maka dari data diperoleh luka ringan sebanyak 14 (3%), luka berat sebanyak 440 (95%) dan meninggal dunia 8 (2%). Sehingga korban pada segmen ini diduga akan mengalami luka berat dan meninggal dunia. (Segmen 15).
9. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan tabrak lari, depan-depan, samping-samping dan tempat kejadian jl kabupaten, jl provinsi dan kendaraan terlibat Spm-Mbl, Spm-Bus/Truk dan pendidikan SMA maka dari data diperoleh luka ringan sebanyak 46 (18%), luka berat sebanyak 1 (00%) dan meninggal dunia 203 (81%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 17).

10. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan tabrak lari, depan-depan, samping-samping dan tempat kejadian jl kabupaten, jl provinsi dan kendaraan terlibat Spm-Spm, Bus/Truk-Bus/Truk dan umur anak & remaja, lansia maka dari data diperoleh luka ringan sebanyak 34 (39%), dan meninggal dunia 54 (61%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 20).
11. Jika kepemilikan SIM tidak memiliki SIM dan jenis kecelakaan laka tunggal dan jenis kelamin laki-laki maka dari data diperoleh luka ringan sebanyak 6 (06%), dan meninggal dunia 100 (94%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 23).
12. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm-Mbl dan jenis kecelakaan depan belakang, depan-depan dan umur muda dan jenis kelamin laki-laki maka dari data diperoleh luka ringan sebanyak 57 (24%), dan meninggal dunia 185 (76%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 25).
13. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm-Mbl dan jenis kecelakaan depan samping, tabrak lari, laka tunggal, samping-samping dan umur muda, anak & remaja, dewasa dan tempat kejadian jl provinsi maka dari data diperoleh luka ringan sebanyak 22 (48%), dan meninggal dunia 24 (52%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 30).
14. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm-Mbl dan jenis kecelakaan depan samping, tabrak lari, laka tunggal, samping-samping dan umur lansia dan tempat kejadian jl kabupaten maka dari data diperoleh luka ringan sebanyak 4 (12%), dan meninggal dunia 29 (88%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 31).
15. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm-Spm, Mbl-Mbl, Mbl-Bus/Truk, Bus/Truk-Bus/Truk dan tempat kejadian jl kabupaten umur anak-anak & remaja, lansia dan jenis kelamin laki-laki maka dari data

diperoleh luka ringan sebanyak 30 (29%) dan meninggal dunia 75 (71%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 34).

16. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm-Spm, Mbl-Mbl, Mbl-Bus/Truk, Bus/Truk-Bus/Truk dan tempat kejadian jl kabupaten umur dewas dan waktu kejadian sepi maka dari data diperoleh luka ringan sebanyak 53 (47%) dan meninggal dunia 59 (53%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 37).
17. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm-Bus/Truk dan tempat kejadian jl nasional, jl kabupaten dan jenis kecelakaan depan belakang, depan samping maka dari data diperoleh luka ringan sebanyak 43 (20%) dan meninggal dunia 168 (80%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 42).
18. Jika kepemilikan SIM memiliki SIM dan kendaraan terlibat Spm dan profesi wirausaha/wiraswasta, pelajar/mahasiswa maka dari data diperoleh luka ringan sebanyak 4 (2%) dan meninggal dunia 159 (98%). Sehingga korban pada segmen ini diduga akan mengalami meninggal dunia. (Segmen 45).

### 5.6 Tingkat Akurasi dan Presisi Metode *CHAID* dengan *SMOTE*

Model yang baik adalah model yang memiliki akurasi dan presisi yang tinggi di bawah ini adalah tabel akurasi dan presisi metode *CHAID* setelah *SMOTE*:

**Tabel 5.2 Tingkat Akurasi dan Presisi Metode *CHAID* dengan SMOTE**

| **Data** | **Tingkat Luka** | **LR** | **LB** | **MD** | |
|---|---|---|---|---|---|
| ***Training*** | LR | 1798 | 62 | 428 | |
| | LB | 10 | 1120 | 1 | |
| | MD | 311 | 24 | 1285 | |
| | Persentase | 84.9% | 92.9% | 75% | 83.4% |

| **Data** | **Tingkat Luka** | **LR** | **LB** | **MD** | |
|---|---|---|---|---|---|
| ***Testing*** | LR | 420 | 20 | 115 | |
| | LB | 6 | 270 | 1 | |
| | MD | 84 | 7 | 337 | |
| | Persentase | 82.4% | 90.9% | 74.4% | 81.5% |

Akurasi $= \frac{420+270+337}{420+20+115+6+270+1+84+7+337} = 81.5\%$

Presisi LR $= \frac{420}{420+6+84} = 82.4\%$

Presisi LB $= \frac{270}{20+270+7} = 90.9\%$ Presisi MD $= \frac{337}{115+1+337} = 74.4\%$

Berdasarkan tabel di atas (tabel 5.2) dapat dilihat bahwa klasifikasi yang terbentuk dengan menggunakan metode *CHAID* setelah *SMOTE*, dengan persentase data *training* 80% dan data *testing* 20%, diperoleh tingkat akurasi sebesar 81.5% dan memiliki presisi yang sudak cukup baik.

### 5.7 Validasi Sistem

Pada tahap validasi system, maka akan dibandingkan *output* dari metode *CHAID, Improve CHAID* dan C4.5. Dibawah ini tabel perbandinganya:

**Tabel 5.3 Validasi Sistem**

| ***Output*** | *CHAID* | *Improve CHAID* | C4.5 |
|---|---|---|---|
| **Akurasi** | 81.5% | 75.6% | 77.7% |

Berdasarkan tabel di atas dapat diambil kesimpulan bahwa metode *CHAID* memiliki nilai akurasi yang berbeda (lebih tinggi) dengan metode lain atau dapat dikatakan metode *CHAID* memiliki akurasi yang lebih baik atau optimal untuk mengklasfikiasikan data dalam penelitian ini.

# BAB VI
# PENUTUP

## 6.1 Kesimpulan

Berdasarkan hasil analisis dan pembahasan yang telah dijelaskan di bab sebelumnya oleh peneliti, maka dapat ditarik kesimpulan yaitu :

1. Gambaran umum pada data kecelakaan lalu lintas di POLRES Sleman pada tahun 2016 sampai 2017 terjadi kecelakaan sebanyak 2379 kejadian. Kecelakaan lalu lintas yang memiliki jumlah paling banyak melibatkan usia muda, jenis kelamin laki-laki, pendidikan SMA, memiliki SIM, berprofesi wiraswasta/wirausaha, waktu kejadian di waktu sepi, tempat kejadian di jl nasional, jenis kecelakaan depan belakang, kendaraan terlibat sepeda motor dengan sepeda motor dan tingkat luka ringan,
2. Rekomendasi karakteristik korban kepada POLRES Kabupaten Sleman untuk mengurangi korban yang mengalami luka berat dan meninggal dunia adalah segmen 1, segmen 2, segmen 4, segmen 5, segmen 8, segmen 11, segmen 14, segmen 15, segmen 17, segmen 20, segmen 23, segmen 25, segmen 30, segmen 31, segmen 34, segmen 37, segmen 42, segmen 45.

## 6.2 Saran

1. Pada penelitian selanjutnya, diharapkan dapat menggunakan metode *decision tree* lainya misalnya *CART, QUEST*, *Improve C4.5* dan metode lainnya untuk membandingkan tingkat akurasi dan presisi klasifikasi untuk memperoleh model terbaik.
2. Pada penelitian selanjutnya dapat ditambahkan variabel lain seperti cuaca, kondisi jalan, tempat kejadian di persimpangan atau di jalan lurus, kecepatan kendaraan bermotor dan lainya.

3. Bagi pihak kepolisian atau pihak terkait perketat proses pembuatan SIM, perlu adanya rambu-rambu lalu lintas atau penerangan pada jalan yang kurang penerangan, berikan peringatan atau seruan pada saat musim hujan supaya lebih hati-hati dan cek selalu perlengkapan kendaraan bermotor, berikan penyuluhan atau pemahaman cara berkendara yang baik dan benar kepada pengguna jalan yang memiliki umur muda (16 sampai 35 tahun) dan juga khususnya pengendara sepeda motor.

# DAFTAR PUSTAKA


Afidah dan Susilaningrum. 2011. Pola Tingkat Keparahan Korban Kecelakaan Lalu Lintas dengan Menggunakan Regresi Logistik Multinomial.

Badan Pusat Statistik. 2015. Kepadatan Penduduk menurut Provinsi, 2000-2015. https://www.bps.go.id/linkTableDinamis/view/id/842. Diakses pada 26 Januari 2018.

Badan Pusat Provinsi Statistik Daerah Istimewa Yogyakarta, 2018. Kepadatan Penduduk menurut Kabupaten/Kota di D.I.Yogyakarta. https://www.bps.go.id/linkTableDinamis/view/id/842. Diakses pada 30 Januari 2018.

B, Liu. *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publisher. 2012.

Data Kecelakaan Lalu Lintas Tahun 2016 sampai 2017. Polres Sleman. Kepolisian Republik Indonesia Daerah Istimewa Yogyakarta.

Davies, and Paul Beynon. 2004. "*Database System Third Edition*". Palgrave macmillan, New York.

Du Toit, S. H. C., A. G. W. Steyn & R. H. Stumph. (1986). *Graphical Exploratory Data Analysis.*New York : Springer-Verlag.

Gallagher, C.A. (2000). *An Iterative Approach to Classification Analysis*. www.casact.org/library/ratemaking/90dp237.pdf. Diakses tanggal 25 Januari 2018.

Heinrich, H. W. 1980. *Industrial Accident Prevention. McGraw-Hill Book Company*. New York

Haryati, Siska., Sudarsono, Aji., dan Suryana, Eko. 2015. Implementasi Data Mining untuk Memprediksi Masa Studi Mahasiswa menggunakan Algoritma C4.5 (Studi Kasus : Universitas Dehasen Bengkulu). *Jurnal Media Infotama*. Vol. 11 (2). 130-138.

Kass GV. 1980. *An Explorasi Technique for Investigating Large Quantities of Categorical Data*. App. Statist 29 (2) :119-127.

Kunto, Y.S.dan Hasana, S.N. 2006. Analisis *CHAID* Sebagai Alat Bantu Statistika Untuk Segmentasi Pasar, jurnal Manajemen, Vol. 1 No. 2. Surabaya : Universitas Kristen Petra.

Larose D, T., 2005, *Discovering knowledge in data : an introduction to data mining*, Jhon Wiley & Sons Inc.

Lehmann, T. dan Eherler, D. 2001. *Responder Profiling with CHAID and Dependency Analysis*. www.informatik.unifreiburg.de/~ml/ecmlpkdd/ WSProceedings/w10/lehmann.pdf. Diakses tanggal 26 Januari 2018.

Nina dan Abadyo. 2012. *CHAID* Untuk Mengklasifikasi Status Mahasiswa Setelah Lulus Perkuliahan (Studi Kasus Pada Alumnus Prodi Matematika Jurusan Matematika. Fmipa Universitas Negeri Malang Tahun 2007-2012). Malang: Jurnal Matematika. Vol.1, No.2 tahun 2012.

Nto. 2017. Angka Kecelakaan Lalu Lintas di DIY Tahun 2017 Menurun. http://jogja.tribunnews.com/2017/09/21/angka-kecelakaan-lalu-lintas-di-diy-tahun-2017-menurun. Diakses pada 26 Januari 2018.

Nuriyah. 2013. Perbandingan Metode *Chi-Square Automatic Interaction detection (CHAID)* dan *Classification and Regression Tree (CART)* Dalam Menentukan Klasifikasi Alumni UIN Sunan Kalijaga Berdasarkan Masa Studi.

Oktafianto, K. 2017. Implementasi Metode *Chi-Squared Automatic Interaction Detection* pada Klasifikasi Indeks Prestasi Kumulatif Mahasiswa FMIPA UNIROW. *Technology Science and Engineering Journal, Vol 1 No 1 February 2017.*

Paratu, Lintang Gustika. 2013. Analisis pada Pelanggan Telekomunikasi menggunakan Algoritma C4.5. *Tugas Akhir.* Program Studi Statistika Jurusan Matematika FMIPA UGM. Yogyakarta.

Peraturan Pemerinta Republik Indonesia No 43 tahun 1993 Tentang Prasarana dan Lalu Lintas Jalan.

Pramana, H. 2011. Klasifikasi dengan Metode *CHAID* (*Chi-Squared Automatic Interaction Detectin*) dan Penerapannya pada Klasifikasi Alumni FMIPA UNY. Skripsi. FMIPA, Universitas Negeri Yogyakarta.

Pramudiono, I 2007. *Pengantar Data mining : Menambang Permata Pengetahuan di Gunung Data.* http://www.ilmu computer.org/wp-content/uploads/ 2006/08/iko-datamining.

Rani, Larissa Navia. 2015. Klasifikasi Nasabah menggunakan Algoritma C4.5 sebagai Dasar Pemberian Kredit. *Jurnal KomTekInfo Fakultas Ilmu Komputer.* Vol. 2 (2).

Santosa, B. (2007). *Data Mining Teknik Pemanfaatan Data Untuk Keperluan Bisnis.* Yogyakarta: Graha Ilmu

Situmorang. 2016. WHO: Angka Kecelakaan Lalu Lintas di Indonesia Tertinggi se-Asia. http://entertainment.analisadaily.com/read/who-angka-kecelakaan-lalu-lintas-di-indonesia-tertinggi-se-asia/240063/2016/05/29. Diakses pada 31 Januari 2018.

Sora, N. 2015. Pengertian Transportasi dan Fungsinya maupun Contohnya. http://www.pengertianku.net/2015/12/pengertian-transportasi-dan-fungsinya-maupun-contohnya.html. Diakses tanggal 31 Januari 2018.

Sulistyono, S. 1998. Karakteristik Kecelakaan Lalu Lintas (Studi Kasus: Jalan Tol Surabaya-Gempol, Jawa Timur), Prosiding Simposium FSTPT I di ITB. Bandung.

Susanto, Heri. 2014. Data Mining untuk Memprediksi Prestasi Siswa berdasarkan Sosial Ekonomi, Motivasi, Kedisiplinan dan Prestasi Masa Lalu. *Jurnal Pendidikan Vokasi*. Vo. 4 (2). 222-231.

Undang- Undang Republik Indonesia Nomor 28 Tahun 2007. Tentang Ketentuan Umum Dan Tata Cara Perpajakan.

Undang-Undang Republik Indonesia No 22 tahun 2009. Tentang Lalu Lintas dan Angkutan Jalan.

V. N. Chawla, K. W. Bowyer, L. O. Hall dan W. P. Kegelmeyer, Journal of Artificial Intelligence Research SMOTE: Synthetic Minority Over-Sampling Technique. http://arxiv.org/pdf/1106.1813.pdf, 2002. Diunduh pada 28 januari 2018.

Witten, I. H. dan Frank, E. 2005. *Data Mining : Pratical Machine Learning Tools and Techniques Second Edition*. San Francisco: Morgan Kauffman.

World Health Organization .2017. Road Traffic Deaths. http://www.who.int /gho/road_safety/mortality/en/. Diakses pada 27 Januari 2018.

Zia, Q, E. 2016. Klasifikasi Rencana Tindak Lanjut Menggunakan Metode *CHAID* (*Chi-Square Automatic Interaction detection*) dan Metode *QUEST* (*Quick Unbiased Efficient Statistical Trees*) (studi kasus: Pengunjung Penderita Diabetes Militus Tahun 2010-2015 di Puskesmas Kecamatan Siwalan Kabupaten Pekalongan). Skripsi. FMIPA, Universitas Islam Indonesia.

# LAMPIRAN

**Lampiran 1.** Data Kecelakaan Lalu Lintas

| | Waktu Kejadian | Tempat Kejadian | Jenis Kecelakaan | Kendaraan Terlibat | Umur | Profesi | Kepemilikan SIM | Jenis Kelamin | Pendidikan | Tingkat Luka |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Padat | JL KABUPATEN | Depan Belakang | Spm - Mbl | Muda | Wiraswasta/Wirausaha | Tidak Memiliki | L | SMA | LR |
| 2 | Sepi | JL NASIONAL | Depan Samping | Spm - Spm | Muda | Pelajar/Mahasiswa | Tidak Memiliki | P | SMA | LR |
| 3 | Sepi | JL NASIONAL | Depan Samping | Spm - Spm | Anak-anal | Pelajar/Mahasiswa | Tidak Memiliki | P | SD | LR |
| 4 | Sepi | JL NASIONAL | Depan Samping | Spm - Spm | Muda | Wiraswasta/Wirausaha | Tidak Memiliki | P | SMA | LR |
| 5 | Sepi | JL NASIONAL | Depan Samping | Spm - Spm | Muda | Wiraswasta/Wirausaha | Tidak Memiliki | P | SMA | LR |
| 6 | Sepi | JL NASIONAL | Depan Samping | Spm - Spm | Dewasa | Wiraswasta/Wirausaha | Tidak Memiliki | P | SMA | LR |
| 7 | Sepi | JL NASIONAL | Depan Samping | Spm - Spm | Muda | Wiraswasta/Wirausaha | Memiliki | L | SMA | LR |
| 8 | Padat | JL NASIONAL | Depan Samping | Spm - Mbl | Muda | Wiraswasta/Wirausaha | Memiliki | L | SMA | LR |
| 9 | Padat | JL NASIONAL | Depan Samping | Spm - Mbl | Muda | Wiraswasta/Wirausaha | Memiliki | P | SMA | LR |
| 10 | Sepi | JL PROVINSI | Tabrak Lari | Spm - Mbl | Lansia | lain-lain | Memiliki | L | SMA | LR |
| 11 | Sepi | JL NASIONAL | Depan Samping | Spm - Spm | Dewasa | lain-lain | Memiliki | L | SMA | LR |
| 12 | Sepi | JL NASIONAL | Depan Samping | Spm - Mbl | Lansia | lain-lain | tidak Memiliki | L | SMA | LB |
| 13 | Padat | JL NASIONAL | Depan Samping | Spm - Spm | Muda | Pelajar/Mahasiswa | tidak Memiliki | L | PT | LR |
| 14 | Padat | JL NASIONAL | Depan Samping | Spm - Spm | Muda | Wiraswasta/Wirausaha | Tidak Memiliki | L | SMA | LR |
| 15 | Sepi | JL DESA | Depan Belakang | Spm - Spm | Muda | Pelajar/Mahasiswa | Tidak Memiliki | L | SMA | LR |
| 16 | Sepi | JL NASIONAL | Depan Samping | Spm - Spm | Dewasa | Wiraswasta/Wirausaha | Tidak Memiliki | L | SMA | LR |
| 17 | Sepi | JL NASIONAL | Depan Samping | Spm - Spm | Muda | Wiraswasta/Wirausaha | Memiliki | L | SMA | LR |
| 18 | Sepi | JL NASIONAL | Tabrak Lari | Spm - Spm | Muda | Wiraswasta/Wirausaha | Memiliki | L | SMA | LR |
| 19 | Sepi | JL NASIONAL | Tabrak Lari | Spm - Spm | Dewasa | Wiraswasta/Wirausaha | Memiliki | L | SMA | LR |
| 20 | Padat | JL KABUPATEN | Depan Samping | Spm - Spm | Muda | Wiraswasta/Wirausaha | Tidak Memiliki | P | SMA | LR |
| 21 | Sepi | JL DESA | Depan Samping | Spm - Spm | Muda | Pelajar/Mahasiswa | Tidak Memiliki | L | SMA | LR |
| 22 | Sepi | JL DESA | Depan Samping | Spm - Spm | Dewasa | PNS TNI/POLRI | Memiliki | L | SMA | LR |
| 23 | Sepi | JL PROVINSI | Depan Belakang | Spm - Mbl | Muda | Pelajar/Mahasiswa | Memiliki | L | PT | LR |
| 24 | Sepi | JL PROVINSI | Depan Depan | Spm - Spm | Anak-anal | Pelajar/Mahasiswa | Tidak Memiliki | P | SD | LR |
| 25 | Padat | JL DESA | Depan Depan | Spm - Spm | Muda | Wiraswasta/Wirausaha | Memiliki | L | SMA | LR |
| 26 | Padat | JL DESA | Depan Depan | Spm - Spm | Dewasa | lain-lain | Memiliki | L | PT | LR |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 3083 | Padat | JL DESA | Depan Depan | Spm - Spm | Muda | lain-lain | Memiliki | P | SMA | LR |
| 3084 | Sepi | JL NASIONAL | Depan Samping | Spm - Spm | Muda | lain-lain | Tidak memiliki | P | SMA | LR |
| 3085 | Padat | JL KABUPATEN | Depan Belakang | Spm - Spm | Lansia | lain-lain | Memiliki | P | SMA | LR |
| 3086 | Sepi | JL KABUPATEN | Depan Depan | Spm - Spm | Dewasa | lain-lain | Tidak memiliki | P | SMA | LR |
| 3087 | Sepi | JL NASIONAL | Depan Samping | Spm - Mbl | Muda | Wiraswasta/Wirausaha | Memiliki | L | SMA | LR |
| 3088 | Padat | JL NASIONAL | Depan Belakang | Spm - Spm | Muda | Pelajar/Mahasiswa | Memiliki | L | SMA | LR |
| 3089 | Padat | JL NASIONAL | Depan Belakang | Spm - Spm | Muda | Pelajar/Mahasiswa | Memiliki | P | SMA | LR |
| 3090 | Padat | JL NASIONAL | Depan Belakang | Spm - Spm | Muda | Pelajar/Mahasiswa | Memiliki | P | SMA | LR |
| 3091 | Sepi | JL NASIONAL | Depan Belakang | Spm - Mbl | Muda | Pelajar/Mahasiswa | Tidak memiliki | L | PT | LR |
| 3092 | Sepi | JL DESA | Depan Belakang | Spm - Spm | Dewasa | Wiraswasta/Wirausaha | Memiliki | L | SMA | LR |
| 3093 | Padat | JL DESA | Tabrak Lari | Spm - Mbl | Dewasa | Wiraswasta/Wirausaha | Memiliki | L | SMA | LR |
| 3094 | Sepi | JL PROVINSI | Depan Depan | Spm - Mbl | Muda | Pelajar/Mahasiswa | Memiliki | L | PT | LR |
| 3095 | Padat | JL NASIONAL | Depan Samping | Spm - Mbl | Dewasa | Wiraswasta/Wirausaha | Tidak memiliki | L | SMA | LR |
| 3096 | Padat | JL NASIONAL | Tabrak lari | Spm - Mbl | Muda | Pelajar/Mahasiswa | Tidak memiliki | L | SMA | LR |
| 3097 | Sepi | JL PROVINSI | Depan Samping | Spm - Mbl | Muda | Pelajar/Mahasiswa | Memiliki | L | PT | LR |
| 3098 | Sepi | JL PROVINSI | Depan Belakang | Spm - Spm | Muda | Pelajar/Mahasiswa | Memiliki | L | PT | LR |
| 3099 | Sepi | JL DESA | Depan Belakang | Spm - Spm | Muda | Wiraswasta/Wirausaha | Memiliki | L | SMA | LR |
| 3100 | Sepi | JL PROVINSI | Depan Depan | Spm - Mbl | Muda | Pelajar/Mahasiswa | Tidak memiliki | L | PT | LR |
| 3101 | Padat | JL DESA | Depan Samping | Spm - Spm | Dewasa | Wiraswasta/Wirausaha | Memiliki | P | SMA | LR |
| 3102 | Padat | JL DESA | Depan Samping | Spm - Spm | Muda | Pelajar/Mahasiswa | Memiliki | P | PT | LR |
| 3103 | Sepi | JL NASIONAL | Tabrak lari | Spm - Spm | Dewasa | lain-lain | Memiliki | P | SMA | LR |
| 3104 | Sepi | JL KABUPATEN | Depan Samping | Spm - Spm | Lansia | Wiraswasta/Wirausaha | Memiliki | L | SMA | LR |
| 3105 | Sepi | JL KABUPATEN | Depan Samping | Spm - Spm | Muda | Wiraswasta/Wirausaha | Memiliki | L | SMA | LR |
| 3106 | Padat | JL NASIONAL | Tabrak lari | Spm - Spm | Muda | Pelajar/Mahasiswa | Memiliki | L | PT | LR |
| 3107 | Padat | JL NASIONAL | Depan Samping | Spm - Spm | Dewasa | Wiraswasta/Wirausaha | Memiliki | L | SMA | LR |
| 3108 | Padat | JL NASIONAL | Depan Samping | Spm - Spm | Dewasa | Wiraswasta/Wirausaha | Memiliki | L | SMA | LR |
| 3109 | Sepi | JL NASIONAL | Depan Belakang | Spm - Mbl | Dewasa | lain-lain | Tidak memiliki | L | SMA | LR |
| 3110 | Sepi | JL KABUPATEN | Depan Belakang | Spm - Mbl | Muda | Pelajar/Mahasiswa | Tidak memiliki | L | SMA | LR |
| 3111 | Sepi | JL NASIONAL | Depan Samping | Spm - Bus/Truk | Muda | Wiraswasta/Wirausaha | Tidak memiliki | L | SMA | MD |

**Lampiran 2.** *Output* Metode *CHAID* **(**Pohon Keputusan)