DATA MINING ANALISIS TINGKAT KECELAKAAN LALU LINTAS DENGAN MENGGUNAKAN METODE *ASSOCIATION RULE*

SKRIPSI

Oleh :
Endisko Panjaitan
160210105

TEKNIK INFORMATIKA

FAKULTAS TEKNIK DAN KOMPUTER

UNIVERSITAS PUTERA BATAM

2020

DATA MINING ANALISIS TINGKAT KECELAKAAN LALU LINTAS DENGAN MENGGUNAKAN METODE *ASSOCIATION RULE*

SKRIPSI

Oleh :
Endisko Panjaitan
160210105

TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
2020

SURAT PERNYATAAN ORISINALITAS

Yang bertanda tanga dibawah ini :

Nama : Endisko Panjaitan

NPM : 160210105

Fakultas : Teknik dan Komputer

Program Studi : Teknik Informatika

Menyatakan bahwa penelitian yang saya buat dengan judul:

Data Mining Analisis Tingkat Kecelakaan Lalu Lintas Dengan Menggunakan Metode *Association Rule*. Adalah hasil karya sendiri dan bukan duplikasi dari karya orang lain. Sepengetahuan saya didalam naskah penelitian ini tidak terdapat karya ilmiah atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis dikutip didalam naskah ini dan disebutkan dalam sumber kutipan dan daftar pustaka.

Apabila ternyata di dalam naskah peneitian ini dapat dibuktikan terdapat unsur-unsur PLAGIASI, saya bersedia naskah penelitian ini digugurkan dan penelitian yang saya peroleh dibatalkan, serta diproses sesuai dengan peraturan undang-undang yang berlaku.

Demikan penyataan ini saya buat dengan sebenarnya tanpa ada paksaan dari siapapun.

Batam,02 Februari 2020

Endisko Panjaitan
160210105

DATA MINING ANALISIS TINGKAT KECELAKAAN LALU LINTAS DENGAN MENGGUNAKAN METODE *ASSOCIATION RULE*

SKRIPSI

Untuk memenuhi salah satu
syarat memperoleh gelar sarjana

Oleh
Endisko Panjaitan
160210105

Telah disetujui oleh pembimbing
pada tanggal seperti tertera dibawah ini
Batam, 21 Februari 2020

Sunarsan Sitohang, S.Kom.,M.TI.
Pembimbing

ABSTRAK

Kecelakaan lalu lintas adalah salah satu penyebab kematian yang cukup besar. Banyak faktor yang menyebabkan terjadinya kecelakaan lalu lintas ini. Namun sedikit analisis yang dapat dilakukan untuk mengatasi faktor penyebab kecelakaan tersebut. Penelitian ini menerapkan metode data mining untuk memproses data kecelakaan lalu lintas menggunakan metode association *rules* dengan algoritma *apriori* untuk meneliti faktor-faktor penyebab kecelakaan tersebut. *Associtioan rules* sangat cocok untuk mencari keterhubungan antara faktor-faktor penyebab kecelakaan lalu lintas dengan kategori ringan, sedang, berat dan rugi material saja. Pada penelitian ini, digunakan nilai *minimum support* (*min_supp*) 1.0 dan nilai *confidance minimum* (*min_conf*) sebesar 100% untuk menghasilkan suatu hubungan yang kuat antara faktor-faktor kecelakaan lalu lintas. Dengan pola kombinasi 4 iterasi/pengulangan (k 4-*itemset*) diketahui bahwa tingkat keterhubungan kecelakaan terbesar yaitu jenis kecelakaan tabrak "Depan Belakang", dengan kondisi pengendara "Lengah", di waktu "Pagi Hari" (06:01-10:00 wib),dengan usia pengendaran "17 tahun-30 tahun (Dewasa)", dari profesi pengendara yaitu "Karyawan", dan menyebabkan kecelakaan "Ringan" dengan tingkat *Conffidance* 100%. Tingkat kecelakaan paling sering terjadi di kota Batam adalah dengan tingkatan "Ringan" yaitusebanyak 764 kali kejadian atau 81% dari total kecelakaan lalu lintas yang ada di kota Batam.

Kata Kunci: Algoritma Apriori; Association Rules; Data Mining, Kecelakaan Lalu Lintas.

ABSTRACT

Batam City is one of the cities in Indonesia which is also famous for its industrial cities and tourist attractions. So that it can attract the attention of job seekers and tourists to come to this city. The development of transportation will lead to the rapid growth of traffic problems. This transportation problem will cause an accident. Traffic accidents are a significant cause of death. Many factors cause this traffic accident. However, little analysis can be done to overcome the factors causing the accident. This study applies data mining methods to process traffic accident data using the association rules method with a priori algorithm to examine the factors that cause these accidents. The association rules are very suitable to look for the relationship between the factors causing traffic accidents with only mild, moderate, heavy and material loss categories. In this study, a minimum support value (min_supp) of 1.0 and a minimum confidence value (min_conf) of 100% is used to produce a strong relationship between traffic accident factors. With a combination pattern of 4 iterations / repetitions (k 4-itemset) it is known that the highest level of connectedness of accidents is that the factors that often occur that have a strong relationship that causes accidents is the type of crash accident "Front Rear", with the condition of the driver "Medium", in the time "Morning" (06: 01-10: 00 WIB), with the age of "17 years-30 years (Adult)", from the driver's profession that is "Employees", and cause "Light" accidents with a 100% confidence rate. The most common accident rate in the city of Batam is the level of "light" which is 764 times or 81% of the total traffic accidents in the city of Batam.

Keywords: Apriori Algorithm; Association Rules; Data Mining, Traffic Accidents.

KATA PENGANTAR

Shalom.

Puji syukur penulis ucapkan kepada Tuhan Yesus Kristus yang telah melimpahkan segala rahmat dan karuniaNya. Sehingga penulis dapat menyelesaikan laporan tugas akhir yang merupakan salah satu persyaratan untuk menyeledsaikan program studi strata satu (S1) pada program studi Teknik Informatika Universitas Putera Batam.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Karena itu, kritik dan saran akan senantias saya terima dengan senang hati. Dengan segala keterbatasan, penulis menyadari pula bahwa skripsi ini takkan terwujud tanpa bantuan, bimbingan, dorongan berbagai pihak. Untuk itu, dengan segala kerendahan hati, penulis menyampaikan ucapan terima kasih kepada :

1. Rektor Univesitas Putera Batam;
2. Ketua Program Studi Teknik Informatika;
3. Bapak Sunarsan Sihotang, S.Kom.,M.TI. selaku pembimbing Skripsi Pada Program Studi Teknik Informatika Universitas Putera Batam;
4. Dosen dan staff Universitas Putera Batam;
5. Bapak Muclis Nadjar, S.H.,S.I.K. selaku komisaris polisi Polresta Barelang;
6. Bapak B.M Napitupulu selaku inspektur polisi dua kanit laka Polresta Barelang;
7. Bapak Hidayat S.E dan ibu Rosenta selaku Staff administrasi lalu lintas Polresta Barelang;
8. Ibu tercinta dan seluruh keluarga yang selalu mendukung sampai saat ini;
9. Teman-teman seperjuangan “Kelompok Selamanya”, Nanda Afrianda, Bintang, Rini, Dewi, dan Enilda.

Semoga Tuhan Yesus Kristus membalas kebaikan dan selalu mencurahkan hidayah serta taufik-Nya, Amin.

Batam,20-02-2020

Endisko Panjaitan

DAFTAR ISI

Halaman

HALAMAN SAMPUL
HALAMAN JUDUL
SURAT PERNYATAAN ORISINALITAS
LEMBAR PENGESAHAN
ABSTRAK i
ABSTRACT ii
KATA PENGANTAR iii
DAFTAR ISI iv
DAFTAR GAMBAR vii
DAFTAR TABEL viii
DAFTAR RUMUS ix
BAB I PENDAHULUAN 1
1.1 Latar Belakang 1
1.2 Identifikasi Masalah 5
1.3 Batasan Masalah 5
1.4 Rumusan Masalah 5
1.5 Tujuan Penelitian 6
1.6 Manfaat Penelitian 6
1.6.1 Manfaat Teoritis 6
1.6.2 Manfaat Praktis 6
BAB II KAJIAN PUSTAKA 22
2.1 *Knowladge Discovery in Database* (KDD) 22
2.2 *Data Mining* 24
2.3 *Association Rule* 25
2.4 *Algoritma Apriori* 26
BAB III METODE PENELITIAN 36
3.1 Desain Penelitian 36
3.2 Teknik Pengumpulan Data 37
3.2.1 *Observasi* 37

3.2.2 Wawancara 38
3.2.3 Studi Pustaka 38
3.3 Operasional Variabel 38
3.3.1 Varibel Tipe/Jenis Kecelakaan 38
3.3.2 Kondisi Pengendara 39
3.3.3 Waktu Kejadian Kecelakaan 39
3.3.4 Usia Pengemudi/Pengendara 40
3.3.5 Profesi Pengemudi /Pengendara 40
3.3.6 Tingkat kecelakaan 40
3.4 Metode Analisis 41
3.4.1 Analisis Masalah 41
3.4.2 Analisis Data 41
3.4.3 Pengolahan data 41
3.4.4 Analisis Apriori 42
3.5 Lokasi dan Jadwal Penelitian 42
BAB IV HASIL DAN PEMBAHASAN 43
4.1 Hasil 43
4.1.1 Analisa Data 43
4.1.2 Identifikasi Dengan Penerapan *Association Rule* 47
4.1.3 Hasil Pengujian Perhitungan Manual 56
4.1.4 Pengujian Software Weka 57
4.2 Pembahasan 65
4.2.1 Tingkat Kecelakaan Tahun 2018 – 2019 65
4.2.2 Penyebab Terjadinya Kecelakaan 65
BAB V SIMPULAN DAN SARAN 67
5.1 Simpulan 67
5.2 Saran 67
DAFTAR PUSTAKA 69
LAMPIRAN x
Lampiran 1 Pendukung Penelitian x

Lampiran 2 Daftar Riwayat Hidup xi
Lampiran 3 Surat Keterangan Penelitian xii
Lampiran 4 Data Hasil Olahan Peneliti xiv
Lampiran 5 Bukti Originalitas Penelitian xxxiii

DAFTAR GAMBAR

Halaman

Gambar 2.4 Kerangka Pemikiran 21
Gambar 3.1 Desain Penelitian 22
Gambar 4.1 Pengolohan Data MS.Excel 46 58
Gambar 4.2 Tampilan Awal WEKA 47 59
Gambar 4.3 Memasukkan Data Pada WEKA 59
Gambar 4.4 Memilih Data Yang Akan Di Proses 60
Gambar 4.5 Memulai Pengolahan Data 60
Gambar 4.6 Memilih Metode *Apriori* 61
Gambar 4.7 Penentuan *Min_Supp* dan *Min_Conf* 61
Gambar 4.8 Hasil Perhitungan WEKA..... 62
Gambar 4.9 Hasil Perhitungan WEKA..... 62
Gambar 4.10 Hasil Perhitungan WEKA 2-*itemset*..... 63
Gambar 4.11 Hasil Perhitungan WEKA 3-*itemset*..... 63
Gambar 4.12 Hasil Perhitungan WEKA 4-*itemset*..... 64
Gambar 4.13 Hasil Perhitungan WEKA 5-*itemset*..... 64
Gambar 4.14 Hasil Perhitungan WEKA 6-*itemset*..... 64

DAFTAR TABEL

Halaman

Tabel 4.1 Tingkat Kecelakaan Lalu lintas Tahun 2018-2019 44
Tabel 4.2 Jenis Dan Kode Variabel Kecelakaan 45
Tabel 4.3 Data Set Pertama Kejadian Kecelakaan 46
Tabel 4.4 Perhitungan Banyak Item 48
Tabel 4.5 Hasil Perhitungan Support 1-Itemset 49
Tabel 4.6 Hasil Support 1-itemset 49
Tabel 4.7 Perhitungan Support 2-itemset 50
Tabel 4.8 Hasil Perhitungan 3-*itemset* 51
Tabel 4.9 Hasil Perhitungan 4-i*temset* 52
Tabel 4.10 Hasil Support 5-*itemset* 53
Tabel 4.11 Confedence Dari Seluruh *itemset*. 54
Tabel 4.12 Final Asosiasi 55

DAFTAR RUMUS

Halaman

Rumus 2.4.1 Mencari Support 1-*itemset*.....**Error! Bookmark not defined.**
Rumus 2.4.1 Mencari Support 2-*itemset*.....**Error! Bookmark not defined.**
Rumus 2.4.3 Mencari *Confidence*. .....**Error! Bookmark not defined.**

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kota Batam adalah salah satu kota di Indonesia yang mempunyai letak yang strategis berada dijalur pelayaran internasional dan juga berbatasan dengan dua Negara yaitu Negara Singapura dan Negara Malaysia. Selain itu terkenal juga dengan kota industri dan tempat wisatanya. Sehingga dapat menarik perhatian para pencari kerja dan juga para wisatawan untuk datang ke kota ini. Dengan hal tersebut kota Batam menjadi salah satu kota dengan pertumbuhan jumlah penduduk yang sangat pesat di Indonesia. Dinas Kependudukan dan Catatan Sipil kota Batam menyatakan bahwa pertahun 2015 jumlah penduduk mencapai 1.037.187 jiwa (Wikipedia, 2020). Dan sampai saat ini bukan tidak ada kemungkinan jumlah penduduk semakin bertambah.

Seiring pesatnya pertumbuhan penduduk di kota Batam setiap tahunnya yang juga menyebabkan kebutuhan alat transportasi juga akan sangat meningkat. Dan secara tidak langsung akan memperbesar timbulnya permasalahan transportasi. (Alimuddin, Tungadi, & Saharuna, 2018) Permasalahan transportasi adalah bukan hanya terbatasnya pada batasan kelengkapan transportasi yang ada, namun juga dipengaruhi oleh aspek-aspek lainya, yakni pendapatan rendah, perpindahan dari desa ke kota yang cepat, terbatasnya sumber daya manusia, disiplin berkendaraan yang rendah, dan lemahnya perencanaan dan pengendalian dalam berkendara, sehingga aspek-aspek tersebut memperkuat masalah transportasi. Transportasi yang baik dapat

dikategorikan , apabila kecepatan standart, tidak ada kemacetan, aman, bebas kecelakaan atau tidak ada pelanggaran dalam berkendara dan saling ada rasa menghargai antara sesama pengendara.

Perkembangan transportasi akan menimbulkan pesatnya pertumbuhan permasalahan lalu lintas. Menurut pasal 1 ke 24 UU/22 tahun 2009 dalam (Fazrijal, 2014) mengatakan bahwa kecelakaan lalu lintas itu adalah peristiwa di jalan raya yang tidak diduga dan tidak disengaja melibatkan antara dua atau lebih kendaraan saling bertabrakan atau kendaraan dengan tanpa pengguna jalan lainnya, dimana mengakibatkan adanya salah satu korban atau bahkan dua-duanya mengalami luka ringan, luka berat, meninggal dunia dan hanya mengalami kerugian harta benda saja. Adapun faktor-faktor yang menyebabkan terjadinya kecelakaan lalu lintas di antaranya Kondisi jalan, kondisi kendaraan, cuaca (alam), kemacetan Lalu lintas, kondisi pengendara, pelanggaran dan jumlah kendaraan yang dari hari ke hari semakin meningkat sangatlah memicu meningkatnya jumlah kecelakaan lalu lintas.

Kecelakaan adalah salah satu faktor kematian yang begitu besar di dunia dan bahkan menduduki peringkat ketiga penyumbang terbesar angka kematian . Berdasarkan data dari World Health Organization (WHO), angka kecelakaan untuk setiap tahunnya sebanyak 1.300.000 jiwa melayang, hal ini sebabkan oleh kecelakaan lalu lintas yang ada di dunia (Alimuddin et al., 2018).

Jumlah kecelakaan di kota Batam sangatlah besar, secara keseluruhan dilihat dari data Kecelakaan Lalu Lintas yang ada di POLRESTA BARELANG (Barelang, 2019)dari tahun 2015-2019 bila rata-ratakan bisa mencapai satu sampai dua

kecelakaan setiap harinya. Data statistik jumlah kecelakaan lalu lintas di kota Batam menunjukkan peningkatan diperiode 2015-2016 dan ditahun selanjutanya mengalami penurunan. Pada periode 2015-2016 meningkat dari angka 486 kasus kecelakaan pada tahun 2015 menjadi 733 kasus kecelakaan pada tahun 2016. Tahun 2017-2019 mengalami penurunan yaitu 604 kasus kecelakaan pada tahun 2017, 517 kasus kecelakaan pada tahun 2018 dan 424 kasus kecelakaan perseptember 2019.

Kecelakaan ini dapat melibatkan berbagai macam jenis kendaraan mulai dari roda dua, roda tiga , roda empat dan lain-lain, pelaku/korban dari berbagai usia. Dari data yang sudah ada meski mengalami penurunan dari tahun ke tahun, tapi disini semua pihak perlu melakukan langkah-langkah pencegahan agar tingkat kecelakaan semakin menurun lagi dengan mengetahui dengan jelas faktor kecelakaan, dan membatasi pengendara dibawah umur.

Ada banyak faktor terjadinya kecelakaan lalu lintas ini. Namun analisis yang dilakukan untuk hal ini sangatlah sedikit. Maka sebuah teknologi *data mining* berkembang untuk menjawab tantangan ini. Data *mining* adalah mengidentifikasi informasi yang memiliki manfaat dan juga pengetahuan yang tergantung pada database yang besar dengan menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* (Koko Handoko, 2018). Untuk menganalisis segala masalah yang telah ditemukan, pertama yang harus dilakukan yaitu menganalisis data-data kecelakaan yang ada dengan metode ilmiah sehingga mendapatkan pengetahuan atau informasi tingkat kecelakaan lalu lintas. Mencermati data, tujuan dapat disesuaikan , dan target dapat dipilih berdasakan tujuan data tersebut. Karena

tidak semua data yang di dapat relevan, tepat, dan akurat. Maka dilakukan percematan data yang merupakan dimana proses data mining akan dimulai.

Dalam hal ini data mining yang digunakan untuk menentukan atau mencari tingkat kecelakaan lalu lintas yang berhubungan antar varibel adalah *association rule* dan algoritma *apriori*. *Association rule* biasanya mengenali tingkah laku dari sebuah kejadian khusus atau proses hubungan asosiasi muncul pada setiap kejadian. Adapun metode apriori sering dipakai untuk pemecahan masalah (Vanich Sajee, 2015). Penggunaan algoritma *apriori* dinilai efektif dalam menentukan tingkat kecelakaan lalu lintas (Aprianti, Hafizd, & Rizani, 2017). Sebuah sistem yang mampu memberikan kemampuan pemecahan masalah maupun kemampuan pengkomunikasian untuk masalah dengan kondisi semi terstruktur dan tak terstruktur, dimana tidak seorang pun tahu secara pasti bagaimana keputusan seharusnya dibuat. Berdasarkan hal diatas, maka perlu di lakukan penelitian analisis kecelakaan lalu lintas jalan raya yang mampu menyediakan sebuah sistem yang dapat memenuhi kebutuhan para pemegang kebijaksanaan. Adapun penelitian yang sudah dilakukan sebelumnya oleh Heru ditriansyah (Prasetio, 2018) tentang Analisis Kecelakaan Lalu Lintas Pada Ruas Jalan Batu Ampar Kota Batam Dengan Metode Accident Rate (*Analysis Of Traffic Accidents At Batu Ampar, Batam With Accident Rate Method*) yang memaparkan tingkat kecelakaan di sekitar ruas jalan Batu Ampar saja, tetapi pada penelitian ini penulis akan membahas tingkat kecelakaan seluruh yang ada di kota Batam.

1.2 Identifikasi Masalah

Berdasarkan latar belakang penelitian diatas maka masalah dapat di identifikasi menjadi :

1. Tingkat kecelakaan lalu lintas kota Batam pada tahun 2018-2019.
2. Masyarakat belum memahami tentang faktor-faktor kecelakaan lalu lintas.
3. Masyarakat masih menyepelekan tantang pengendara di bawah umur.

1.3 Batasan Masalah

Analisis kecelakaan lalu lintas jalan raya di kota batam merupakan penelitian yang cakupannya sangat luas, maka dari itu di tetapkan batasan masalah serta asumsi meliputi :

1 Lokasi studi adalah jaringan jalan yang tercatat di Polresta Barelang Batam terdapat peristiwa kecelakaan lalu lintas.

2 Data kecelakaan menggunakan data tahun 2018-2019 yang di peroleh dari Polresta Barelang Batam.

3 Menggunakan Data Mining Metode Association Rule Dengan Algoritma Apriori untuk menentukan faktor dan hubungan terjadinya kecelakaan lalu lintas jalan raya di Kota Batam.

1.4 Rumusan Masalah

Berdasarkan identifikasi masalah maka di buat suatu rumusan masalah yaitu:

1 Bagaimana mengetahui tingkat kecelakaan di kota Batam tahun 2018-2019 yang dilihat pada beberapa aspek yang sudah ditentukan?

2 Bagaimana mengetahui faktor yang sering terjadi atau tertinggi suatu kecelakaan di kota Batam?

1.5 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah :

1 Mengetahui, menganlisis tingkat kecelakan dari tahun 2018-2019 di kota Batam.

2 Mengetahui, menganalisis data dan menentukan faktor penyebab terjadinya kecelakaan di kota Batam.

1.6 Manfaat Penelitian

Manfaat penelitian ini terbagi menjadi dua yaitu :

1.6.1 Manfaat Teoritis

Adapun manfaat penelitian ini secara teoritis adalah

1. Kegunaan secara teoritis hasil penelitian ini diharapkan dapat memberikan tambahan pengetahuan dan referensi bagi kajian ilmu pemerimtahan, kususnya yang berkaitan dengan kebijakan pemerintah dibagian transportasi massal.

1.6.2 Manfaat Praktis

Adapun manfaat penelitian ini secara praktis

1. Dapat menambah dan memperkokoh teori yang telah ada tentang data mining.

2. Sebagai bahan referensi untuk membantu masyarakat dan pihak kepolisian yang membutuhkan informasi tentang tingkat kecelakaan setiap tahunnya di Kota Batam
3. Diharapkan hasil penelitian ini berguna bagi Universitas Putera Batam sebagai bahan bacaan dan referensi bagi setiap mahasiswa lainya.

BAB II

KAJIAN PUSTAKA

2.1 *Knowladge Discovery in Database* (KDD)

Knowledge Discovery in Database (KDD) merupakan proses analisa terstruktur untuk memperoleh informasi yang benar, baru, bermanfaat dan menemukan pola dari data yang besar dan komplek. *Data mining* menjadi inti dari proses *Knowledge Discovery in Database* (KDD), yaitu dengan menggunakan algoritma tertentu untuk mengeksplorasi setiap data, membuat model dan menemukan pola yang belum diketahui (Zanuardi & Suprayitno, 2018).

Pada dasarnya ada enam elemen yang paling esensial dalam teknik pencarian informasi / pengetahuan dalam *Knowledge Discovery in Database* (KDD) yaitu:

1. Menyelesaikan sejumlah besar data.
2. Dibutuhkan efesiensi berkaitan dengan jumlah data.
3. Mengutamakan keakuratan/kepastian.
4. Membutuhkan penggunaan bahasa tingkat tinggi.
5. Menggunakan beberapa bentuk dari pembelajaran otomatis.
6. Menghasilkan hasil yang menarik.

Knowledge Discovery in Database (KDD) terdapat beberapa langkah dalam proses pengerjaannnya yaitu :

1. *Selection* (Seleksi Data)

Selection data perlu dilskukan sebelum tahap penggalian informasi pada *Knowledge Discovery in Database* (KDD) dimulai, setelah selesai data akan dipakai untuk pemrosesan *data mining*.

2. *Preprocessing/Cleaning* (Pemilihan Data)

Proses ini dilakukan untuk membuang penggandaan data, memeriksa inkonsisten data dan memperbaiki kesalahan pada data-data, seperti kesalahan pada pencetakan. Juga dilakukan proses enrichment, yaitu proses memperkaya data yang sudah ada dengan data/informasi lain yang relevan dan diperlukan untuk *Knowledge Discovery in Database* (KDD), seperti data/informasi eksternal.

3. *Transformation* (Transformasi)

Pada tahap ini yang akan dilakukan adalah mentrasformasi bentuk data yang belum mempunyai entitas yang jelas pada bentuk data yang *valid* untuk diproses pada *data mining.*

4. Data Mining

Data Mining (DM) (El Houby, 2014) didefinisikan sebagai "ekstraksi yang tersembunyi informasi prediktif dari database besar. Ini adalah langkah inti dari proses yang lebih luas, yang disebut penemuan pengetahuan dalam database. Proses ini mencakup penerapan beberapa metode *preprocessing* yang bertujuan memfasilitasi aplikasi dari algoritma *data mining* dan metode *postprocessing* yang ditujukan di memperbaiki dan meningkatkan pengetahuan yang ditemukan.

5. *Interpratation/evaluation* (Interpretasi/Evaluasi)

Proses penerjemahan pola yang dihasilkan dari *data mining*. Pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap *evaluation* adalah bagian dari proses KDD yang mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya.

2.2 *Data Mining*

Data mining adalah mengidentifikasi informasi yang memiliki manfaat dan juga pengetahuan yang tergantung pada database yang besar dengan menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning (Koko Handoko, 2018). Tahap-tahap pada data mining dapat diilustrasikan seperti berikut:

1. Pembersihan data (untuk membuang data yang tidak konsisten dan *noise*)
2. Integrasi data (penggabungan data dari beberapa sumber)
3. Transformasi data (data diubah menjadi bentuk yang sesuai untuk di-*mining*)
4. Aplikasi teknik *Data Mining*
5. Evaluasi pola yang ditemukan (untuk menemukan yang menarik/bernilai).
6. Presentasi pengetahuan (dengan teknik visualisasi)

Pengelompokan Data *Mining* dibagi menjadi beberapa kelompok (Maulana & Fajrin, 2018) yaitu :

1. Deskripsi

Deskripsi merupakan cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data yang dimiliki.

2. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali *variable* target estimasi lebih ke arah numerik daripada ke arah kategori. Model yang dibangun menggunakan *record* lengkap yang menyediakan nilai *variable* target sebagai nilai prediksi.

3. Prediksi

Prediksi adalah menerka sebuah nilai yang belum diketahui dan juga memperkirakan nilai untuk masa mendatang.

4. Klasifikasi

Dalam klasifikasi terdapat target *variable* kategori, misal penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu tinggi, sedang, dan rendah.

5. Pengklasteran

Merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan.

6. Asosiasi

Asosiasi bertugas menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja.

2.3 *Association Rule*

Metode asosiasi termasuk salah satu metode dalam data mining yang banyak dipergunakan saat ini. Metode yang di gunakan adalah *Association Rule* menggunakan algoritma Apriori. Analisis Assosiasi (*association rule*) adalah teknik *data mining* untuk menemukan aturan assosiasif antara suatu kombinasi item. Analisis asosiasi dikenal juga sebagai salah satu teknik *data mining* yang menjadi

dasar dari berbagai teknik data *mining* lainnya. Khususnya salah satu tahap dari analisis asosiasi yang di sebut analisis pola frekuensi tinggi (*frguent pattern mining*) menarik perhatian banyak peneliti untuk menghasilkan algoritma yang efisien (Alimuddin et al., 2018).

Penting tidaknya suatu aturan asosiasif dapat di ketahui dengan dua parameter *support* (nilai penunjang) yaitu presentase kombinasi item tersebut dalam database dan *confidence* (nilai kepastian) yaitu kuatnya hubungan antar item dalam aturan asosiasif.

2.4 *Algoritma Apriori*

Apriori adalah algoritma klasik yang digunakan dalam data menambang untuk penemuan aturan asosiasi di Indonesia data transaksional. Algoritma dirancang untuk bekerja dengan sejumlah besar data yang mewakili urutan barang atau acara (ROȘCA & RĂDOIU, 2015). Algoritma *aprior*i digunakan untuk mencari *frequent itemset* yang memenuhi minimal *support* kemudian mendapatkan rule yang memenuhi minimal *confidence* dari *frequent itemset* tadi (Wijayanti, 2017). Secara umum pada iterasi ke-*k* dilakukan proses sebagai berikut :

1. Membentuk kandidat itemset. Kandidat *k-itemset* dibentuk dari kombinasi (*k*-1)-itemset yang didapat dari iterasi sebelumnya *(join*). Satu ciri dari algoritma *Apriori* adalah adanya pemangkasan (*prune*) kandidat *k-itemset* yang *subset*-nya berisi *k*-1 item tidak termasuk dalam pola frekuensi tinggi dengan panjang *k*-1.

2. Menghitung *support* dari tiap kandidat k-itemset. *Support* dari tiap kandidat k-itemset didapat dengan men-scan database untuk menghitung jumlah transaksi yang memuat semua item di dalam kandidat *k-itemset* tsb. Ini adalah juga ciri dari algoritma *apriori* dimana diperlukan

   penghitungan dengan scan seluruh database sebanyak *k-itemset* terpanjang.
3. Pola frekuensi tinggi. Pola frekuensi tinggi yang memuat *k* item atau k *itemset* ditetapkan dari kandidat *k-itemset* yang *support*nya lebih besar dari *minimum support.*
4. Bila tidak didapat pola frekuensi tinggi baru maka seluruh proses dihentikan. Bila tidak, maka *k* ditambah satu dan kembali ke bagian 1 (satu). Metodologi dasar analisis asosiasi terbagi menjadi dua tahap yaitu :

a. Analisa pola frekuensi tinggi

Tahap ini mencari kombinasi item yang memenuhi syarat minimum dari nilai support dalam database. Nilai support sebuah item diperoleh dengan rumus 2.4.1 dan 2.4.2:

Support(W) = $\frac{\Sigma\ transaksi\ yang\ mengandung\ W}{\Sigma\ \ transaksi\ yang\ telah\ terjadi}$ .....Rumus 2.4.1

Sedangkan pada dua itemset atau lebih diperoleh dengan rumus:

Support(W,X)= $\frac{\Sigma\ transaksi\ yang\ mengandung\ W\ dan\ X}{\Sigma\ total\ transaksi\ yang\ telah\ terjadi}$ .....Rumus 2.4.2

b. Pembentukan aturan asosiasi

Setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan assosiatif yang memenuhi syarat minimum untuk confidence dengan menghitung confidence aturan assosiatif W_X.

Nilai confidence dari aturan W_X diperoleh dari rumus 2.4.3

Confidence(W >X)= $\frac{\Sigma\ transaksi\ yang\ mengandung\ W\ dan\ X}{\Sigma\ total\ transaksi\ yang\ mengandung\ W}$ .....Rumus 2.4.3

2.4 Sofware Pendukung

Weka (*Waikato Environtment for Knowledge Analysis*) adalah *tools open source* (berlisensi GPL) yang berisi kumpulan algoritma *machine learning* dan data *preprocessing*. Weka dapat digunakan mulai dari tahap *preprocessing*, proses, evaluasi sampai dengan visualiasi. Berbagai algoritma dapat dibandingkan untuk memilih yang terbaik untuk masalah yang akan dipecahkan. Basis weka adalah Java sehingga bersifat *multiplatform* dan dapat digunakan sebagai *library* dalam aplikasi Java. Weka menyediakan tiga interface untuk mengolah dataset (Wibisono, Komputer, Pendidikan, & Beta, 2013).

1. *Explorer* digunakan untuk mencari algoritma yang paling cocok untuk data. Semua data di *load* ke memori sehingga dapat cepat diproses, tapi hanya dapat digunakan untuk data dengan jumlah terbatas.
2. *Eksperimenter* digunakan untuk mencari parameter yang cocok. Mirip dengan *explorer* tetapi prosesnya dapat diautomatisasi. Eksperimen ukuran besar (multi machine dengan RMI) dapat dilakukan dengan *interface* ini.

3. *Knowledge flow* digunakan untuk memproses data stream. Konfigurasi proses dapat diatur dengan mudah dan dapat menangani data berukuran besar. Mendukung *incremental learning*.
4. Simple CLI (*Command Line Interface*) menggunakan *command line* untuk *interface*.

2.3 Penelitian terdahulu

Adapun penelitian terdahulu yang sudah dilakukan oleh para peneliti-peneliti sebelumnya yang terkait dengan penelitian ini adalah

Wardiman Alimuddin (Alimuddin et al., 2018) menyatakan pada penelitannya yang berjudul tentang Analisis Tingkat Kecelakaan Lalu Lintas dengan Metode *Association Rule* enggunakan Algoritma *Apriori* bahwa *k*ecelakaan lalu lintas merupakan salah satu penyebab kematian yang cukup besar. Banyak faktor yang menyebabkan terjadinya kecelakaan lalu lintas. Namun tidak banyak analisis yang dilakukan terhadap faktor penyebab kecelakaan tersebut. Metode *Association rules* dengan algoritma apriori dapat digunakan untuk meneliti faktor-faktor penyebab kecelakaan tersebut. *Association rules* sangat sesuai untuk mencari keterhubungan antara faktor-faktor penyebab kecelakaan dengan kategori kecelakaan rugi material, ringan, sedang, dan berat. Pada penelitian ini, digunakan nilai *support minimum (min_supp)* 0.1 dan nilai *confidence minimum (min_conf)* 50% untuk menghasilkan pola keterhubungan yang kuat antar faktor-faktor kecelakaan lalu lintas. Dengan pola kombinasi 4 iterasi/pengulangan *(K 4-itemset)* diketahui bahwa tingkat

keterhubungan kecelakaan terbesar yaitu pada kecelakaan dengan skala ringan yaitu sebesar 638 kali kejadian atau 68% dari total kecelakaan lalu lintas yang terjadi di Kabupaten Pinrang.

Rusmin Saragih (Saragih, 2017) menyatakan penanggulangan kecelakaan lalu lintas dapat dilakukan dengan melakukan langkah – langkah pendekatan seperti sosialisasi dan pengawasan terhadap rambu – rambu serta trafik pada jalan raya. Untuk dapat mencari solusi penanggulangan yang tepat maka diperlukan informasi mengenai kecelakaan lalu lintas yang terjadi. Dengan mengamati variabel – variabel yang berkaitan dari para pelaku kecelakaan lalu lintas maka dapat diperoleh kriteria dari para pelaku kecelakaan lalu lintas yang dominan dari relasi variabel – variabel pelaku LAKA seperti pendidikan, umur dan variabel lainnya. Apriori diterapkan pada data pelaku LAKA dari beberapa tahun terakhir untuk dapat memperoleh informasi kaitan antara variabel pelaku LAKA sehingga dapat menjadi referensi terhadap penanggulangan dan pencegahan kecelakaan lalu lintas. Output apriori terhadap data pelaku laka menyajikan informasi relasi pola dari variabel para pelaku LAKA.

NUNUNG FADYLAH (Fauziah, 2017) menyebutkan faktor penyebab tingginya angka kecelakaan berasal dari kondisi eksisting ruas jalan, perilaku pengemudi dan kondisi lingkungan disekitar. Analisa ini menggunakan data ekunder dari Polrestabes Surabaya berupa data kecelakaan tahun 2010-2015 yang terjadi di kota Surabaya dan Data *Inventory.* Data tersebut di analisa mencari angka kecelakaan tertinggi yang kemudian di identifikasi sebagai ruas rawan kecelakaam *(Blacksite)* dan menentukan titik rawan kecelakaan *(Blackspot)* dengan metode pengumpulan

data primer dan sekunder. Sedangkan analisa penyebab dan penanggulan kecelakaan menggunakan metode analisa deskriptif dari data *Inventory*, Hasil analisa menunjukkan bahwa klasifikasi kecelakaan yang paling banyak terjadi berdasarkan korban adalah luka ringan, berdasarkan hari adalah hari jumat, berdasarkan jam adalah jam 06.00-12.00, berdasarkan bulan adalah bulan Desember, berdasarkan cuaca adalah cuaca cerah, kendaraan SPM vs SPM. Titik rawan kecelakaan Jl. A. Yani pada STA 000+200 dan STA 1000+1200, Jl. Kenjeran *Blackspot* terletak pada STA 1000+1200 dan STA 1600+1800, Jl. Ngagel *Blackspot* terletak pada STA 000+200 dan STA 400+600, Jl.Diponegoro *Blackspot* terletak pada STA 450+600, Jl. Demak *Blackspot* terletak pada STA 1200+1350, Jl.Ir.Sukarno *Blackspot* terletak pada STA 2400+2600, Jl. Kedung Cowek *Blackspot* terletak pada STA 050+100.Penyebab kecelakaan berasal dari kondisi eksisting jalan yang kurang mewadai, penanggulan berupa perbaikan jalan.

Arvian Januardi (Zanuardi & Suprayitno, 2018) dalam penelitiannya mengatakan studi bertujuan menganalisa karakteristik kecelakaan lalu lintas di Jl. Ahmad Yani Surabaya meliputi analisa situasi, identifikasi *blackspot*, faktor pemicu dan karakteristik kecelakaan. Metode yang digunakan adalah *Knowledge Discovery in Databases* dengan *data mining* didukung aplikasi SPSS dan RapidMiner. *Baseline* data diambil dari rekaman laka lantas Polrestabes Surabaya tahun 2015 s.d Agustus 2017. Hasil analisa menunjukkan peningkatan angka kecelakaan di Jl. Ahmad Yani paska dibangunnya *frontage* disebabkan oleh faktor kedisiplinan pengguna jalan. Tiga lokasi utama *blackspot* adalah sekitar Royal Plaza, pertigaan Margorejo dan

Rumah Sakit Angkatan Laut (RSAL). Analisa karakteristik kecelakaan mendapati beberapa hasil antara lain : 85,9% kecelakaan jamak; potensi kecelakaan terbesar pada hari Senin pukul 20:00-21:00; pihak terlibat kecelakaan didominasi sepeda motor (69,61%) ; dan 75,9% tingkat fatalitas masih rendah. Biaya kecelakaan meliputi biaya korban dan kerugiaan materil selalu meningkat setiap tahun. Peningkatan terbesar tejadi pada kerugian materil dari tahun 2015 ke 2016 (466,49%). Perhitungan total biaya kecelakaan selama periode pengamatan mencapai Rp. 10,78 Triliun.

Kecelakaan lalu lintas di jalan raya masih menjadi penyumbang tingginya angka kematian di Indonesia, sehingga menjadi perhatian khusus bagi kepolisian di negara ini. Tahapan awal pencegahan kecelakaan lalu lintas adalah dengan mengetahui faktor-faktor penyebab kecelakaan lalu lintas yang diperoleh melalui analisa data kecelakaan. Analisa tersebut dapat dilakukan dengan data mining, yaitu *K-Means Clustering. K-Means Clustering* mengelompokkan data menjadi beberapa *cluster* sesuai karakteristik data tersebut. Data kecelakaan lalu lintas dibagi menjadi 2 dataset, yakni dataset 1 dan dataset 2. Hasil *cluster* penerapan *K-means clustering* terhadap dataset 1 dan dataset 2 kemudian dilakukan pengujian *silhoutte oefficient* untuk mencari hasil *cluster* dengan kualitas terbaik. Pengujian *silhoutte coefficient* secara berurutan menghasilkan *distance measure* paling optimal yakni *clustering* dengan 4 *cluster* untuk dataset 1 dan *clustering* dengan 2 *cluster* untuk dataset 2. Selain memperoleh *cluster* dengan kualitas terbaik, penganalisaan data juga menghasilkan beberapa informasi kecelakaan lalu lintas yang sering terjadi, yakni

faktor penyebab dan korban kecelakaan adalah pengemudi, umur korban adalah 9 sampai 28 tahun, dan keadaan korban kecelakaan adalah luka ringan. Hasil dari penelitian dari Winda Aprianti (Aprianti & Permadi, 2018).

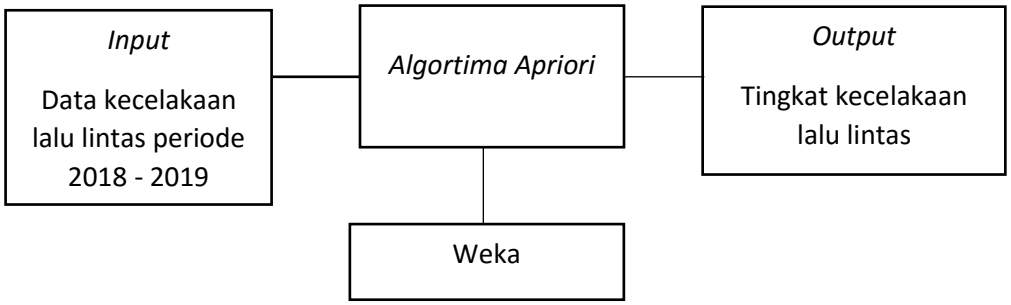
Gom Gom E P Manalu (Manalu et al., 2011) Pada penelitian ini dibahas masalah kecelakaan lalu lintas yang terjadi di Kota Tebing Tinggi. Tujuan dari penelitian ini adalah untuk mengetahui karakterisitik kecelakaan, mengetahui lokasi rawan kecelakaan *(black spot)* di Kota Tebing Tinggi, mengetahui besaran biaya kecelakaan.Metode yang digunakan untuk mengetahui karateristik kecelakaan adalah metode karateristik. Metode yang digunakan untuk menentukan lokasi rawan kecelakaan antara lain metode frekuensi dan metode tingkat kecelakaan. Metode yang digunakan untuk mengetahui besaran biaya kecelakaan dengan Metode *Groos Output*. Daerah / lokasi rawan kecelakaan *(black spot)* adalah suatu lokasi dimana tingkat kecelakaan tinggi dengan kejadian kecelakaan berulang dalam suatu ruang dan rentang waktu yang relatif sama yang diakibatkan oleh suatu penyebab tertentu. Data yang digunakan adalah data jumlah kecelakaan yang terjadi di Kota Tebing Tinggi yang terjadi pada tahun 2007-2011 yang dikelompokkan berdasarkan karateristik.Gambaran komposisi karakteristik kecelakaan adalah sebagai berikut : Hari Sabtu sebanyak 101 kecelakaan (21,12%), waktu terang sebanyak 247 kecelakaan (51,67%), kecelakaan fatal sebanyak 295 kecelakaan (61,71%), tabrakan depan-depan sebanyak 193 kecelakaan (40,37%), sepeda motor sebanyak 552 unit (63,59%), luka ringan sebanyak 410 orang (41,22%), laki-laki sebanyak 999 orang (78,84%), usia diatas 46 tahun sebanyak 255 orang (20,12%), pekerjaan pegawai

swasta/karyawan sebanyak 363 orang (28,65%). Hasil analisis juga menunjukkan bahwa terdapat 4 titik lokasi rawan kecelakaan (black spot) di segmen ruas jalan Kol. Yos Sudarso yaitu Sta 74+800, Sta 75+00, Sta 75+300 dan Sta 75+50 . Hasil analisis juga didapat biaya korban kecelakaan lalu lintas dengan total biaya Rp. 11.293.830.194 dan biaya kecelakaan lalu lintas dengan total biaya Rp. 14.175.152.920.

Agus Sasmito Ariwibowo (Aribowo, 2015) menjelaskan dalam penelitiannya yang berjudul Analisa Asosiatif Data Mining untuk Mengetahui Pola Kecelakaan Lalu Lintas bahwa data kecelakaan lalu lintas dapat diolah menjadi informasi yang penting bagi Kepolisian. Informasi diperoleh dengan cara menganalisis data kecelakaan lalu lintas untuk mengetahui apakah ada hubungan antara terjadinya kecelakaan untuk merek kendaraan dan jenis kendaraan. Penelitian ini menerapkan metode data mining untuk memproses kecelakaan lalu lintas data dengan menggunakan teknik data mining yang disebut Metode *Apriori*. *Apriori* metode digunakan untuk mengidentifikasi pola kecelakaan berdasarkan merek dan jenis kendaraan. asilnya digunakan untuk memperkirakan apakah ada korelasi antara kejadian kecelakaan lalu lintas dengan merek dan jenis kendaraan tertentu. Hasilnya dapat membantu Kepolisian untuk mengetahui apakah ada korelasi antara kejadian kecelakaan lalu lintas dengan jenis dan merek kendaraan.

2.4 Kerangka Pemikiran

Kerangka pemikiran dalam penelitian ini adalah :



Gambar 2.4 Kerangka Pemikiran
Sumber: Data Penelitian,2019

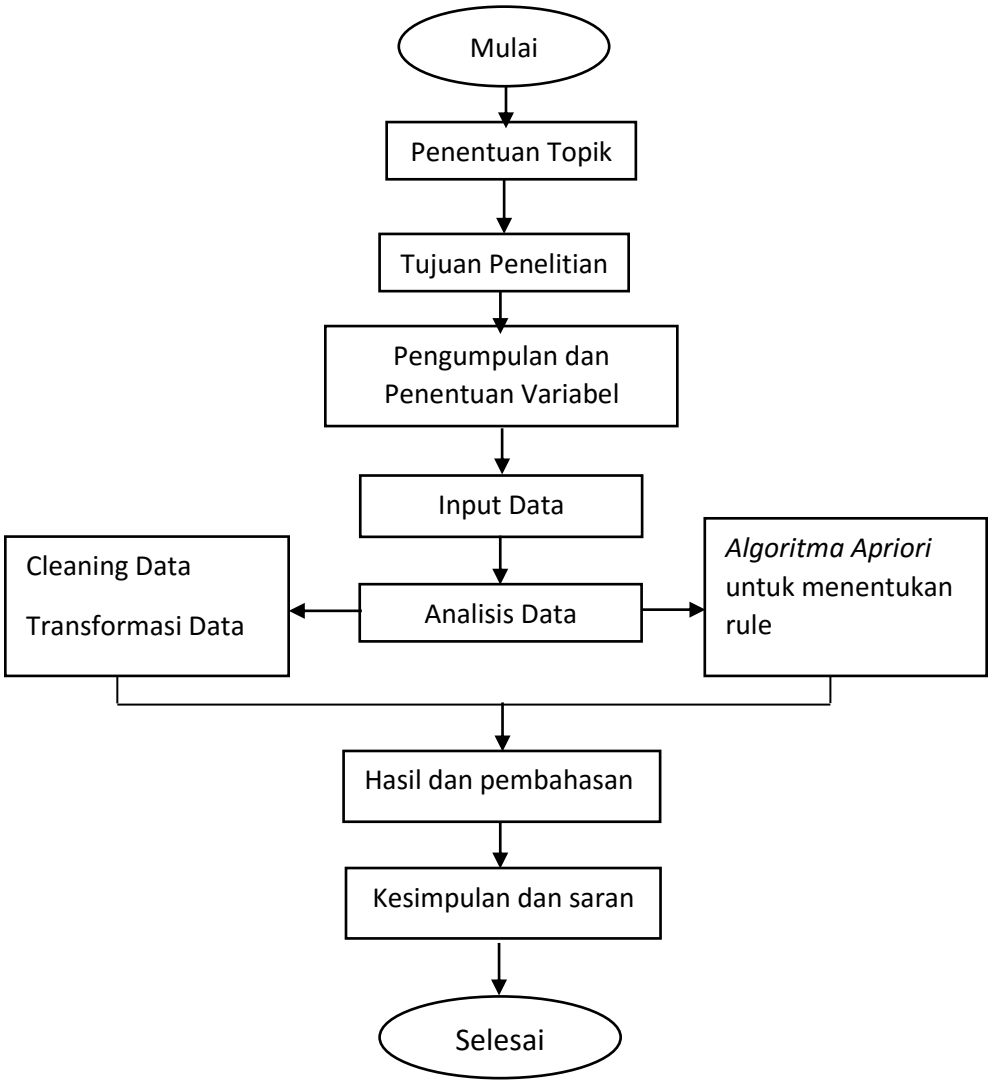
Kerangka Pemikiran pada gambar 2.4 ini menjelaskan bahwa tingkat kecelakaan lalu lintas di kota Batam periode tahun 2015 sampai tahun 2019 dan varibel-variabel yang sudah di tetapkan di BAB selanjutnya digunakan sebagai *input*, yang kemudian akan dikelola dengan menggunakan metode *Algoritma Apriori* dengan menggunakan *software* pendukung yaitu Weka. Sehingga menghasilkan *out put* yaitu seberapa besar tingkat kecelakaan lalu lintas di kota Batam, hubungan faktor kecelakaan terhadap tingkat kecelakaan dan usia rentan mengalami kecelakaan.

BAB III

METODE PENELITIAN

3.1 Desain Penelitian

Adapun langkah-langkah dalam penelitian ini yang digunakan untuk memperoleh sebuah tujuan sebagai berikut:



Gambar 3.1 Desain Penelitian
Sumber: Data Penelitian, 2019

Dalam gambar 3.1 Desain penelitian, menjelaskan analisis ini penelitian akan diawali dari penentuan topik, tujuan penelitian, pengumpulan dan penentuan, variabel, *input* data. Setelah itu dilakukan analisis data yang meliputi analisis *frequent itemset*. Selanjutnya adalah proses data *mining* tahap ini akan dilakukan proses *frequent itemset* dan *candidate generation itemset* secara bertahap untuk menghasilkan nilai *support* dan nilai *confidance* menggunakan algoritma *apriori* dengan bantuan software pendukung Weka. Kemudian masuk ke hasil dan pembahasan, saran dan kesimpulan.

3.2 Teknik Pengumpulan Data

Beberapa teknik pengumpulan data yang di pakai dalam penelitian ini adalah

3.2.1 *Observasi*

Observasi adalah teknik pengumpulan data yang tidak dilakukan pada orang saja, tetapi juga padaobyek-obyek alam lain juga. Teknik *observasi* ini digunakan bila, penelitian berkenaan dengan perilaku manusia, proses kerja, gejala-gejala alam dan bila responden yang diamati terlalu besar (Prof.Dr.Sugiyono, 2017).

Teknik *observasi* (pengamatan) dilakukan pengumpulan data tentang kecelakaan lalu lintas ke instansi kepolisian. Data kecelakaan ini secara umum tersedia pada direktorat lalu lintas Poltabes Barelang kota Batam unit laka lantas. Berdasarkan pengamatan yang sudah dilakukana pada instansi kepolisian, data kecelakaan lalu lintas hanya diolah dengan statistik penjumlahan kejadian kecelakaan biasa. Dari kecelakaan ini mengandung banyak informasi yang dapat mengungkap penyebab terjadinya kecelakaan lalu lintas.

3.2.2 Wawancara

Wawancara aadalah suatu cara pengumpulan data yang digunakan untuk mendapatkan informasi langsung dari sumbernya (Dr.Sudaryono, 2014). Wawancara ini dilakukan untuk mendengarkan asumsi dari instansi kepolisian dimana kepolisian ini merupakan pihak yang terjun langsung dalam kejadian lalu lintas diwilayah kota Batam. Penulis melakukan wawancara bersama Bapak MB. NAPITUPULU selaku inspektur polisi dua kanit laka Polresta Barelang.

3.2.3 Studi Pustaka

Dalam hal ini yang dilakukan adalah mengumpulkan data dengan mempelajari masalah yang berhubungan dengan objek yang diteliti serta bersumber dari buku-buku pedoman, literatur yang disusun oleh para ahli untuk melengkapi data yang diperlukan dalam penelitian.

3.3 Operasional Variabel

Dalam operasional variabel ini, ada 5 variabel *condition* dan 4 variabel *decision* yang dipakai dalam penelitian ini. Varibel penelitian adalah sesuatu yang jadi objek penelitian atau dapat juga disebut sebagai faktor-faktor yang menyebabkan adanya peristiwa kecelakaan lalu lintas ini. Dibawah ini adalah varibel penelitian yang akan di teliti:

3.3.1 Varibel Tipe/Jenis Kecelakaan

Varibel ini adalah model kecelakaan yang terjadi di lokasi dengan pembagian sebagai berikut :

1. Di ruas jalan, pejalan kaki di tabrak kendaraan

2. Tabrakan tunggal.
3. Di ruas persimpangan, tabrakan dua atau lebih kendaraan
4. Tabrakan depan – belakan
5. Tabrakan depan – depan
6. Tabrakan depan – samping
7. Tabrakan samping – samping
8. Tabrakan kendaraan dengan benda yang berada diruas jalan

3.3.2 Kondisi Pengendara

Di dalam variabel ini menyatakan kondisi pengendara sangat erat kaitannya dengan kecelakaan, variabelnya adalah:

1. Lengah
2. Lelah / Mengantu
3. Tidak tertip
4. Pengaruh obat dan Minuman
5. Dan lain-lain

3.3.3 Waktu Kejadian Kecelakaan

Dalam penelitian ini ada 5 waktu yang sudah di bagi yaitu:

1. Dini hari = 00:01 – 06:29 wib
2. Pagi hari = 06:30 – 10:00 wib
3. Siang hari = 10:01 – 14:59 wib
4. Sore hari = 15:00 – 18:30 wib
5. Malam hari = 18:31 – 24:00 wib

3.3.4 Usia Pengemudi/Pengendara

Variabel ini menyatakan usia rentan korban dalam kecelakaan lalu lintas. Variabel ini di bagi menjadi 4 sebagai berikut:

1. 0 – 15 tahun
2. 16 – 30 tahun
3. 31 – 45 tahun
4. 46 sampai seterusnya

3.3.5 Profesi Pengemudi /Pengendara

Variabel ini menyakatan profesi yang sering mengalami kecelakaan , adapun beberapa profesi yang sebagai sampel yaitu:

1. TNI
2. POLRI
3. Karyawan
4. Pelajar/Mahasiswa
5. Supir
6. Pedagang

3.3.6 Tingkat kecelakaan

Variabel tingkat kecelakaan adalah varibel yang di pakai sebagai varibel *decision*, varibel ini akan menghasilkan seberapa besar tingkat kecelakaan lalu lintas, variabel ini di bagi menjadi 4 tingkatan yaitu :

1. Berat
2. Sedang

3. Ringan
4. Rugi materi saja

3.4 Metode Analisis

3.4.1 Analisis Masalah

Analisa Masalah adalah menentukan analisa pada setiap permaslahan yang ada agar dapat menentukan solusi atau strategi yang cocok untuk menyelesaikan masalah.

3.4.2 Analisis Data

Teknik analisis data bertujuan menguraikan dan memecahkan masalah yang berdasarkan data yang di peroleh. Analisis yang digunakan adalah analisis data kualitatif. Analisis data kualitatif adalah upaya yang dilakukan dengan jalan mengumpulkan, memilah-milah, menklasifikasikan, dan mencatat yang dihasikan catatan lapangan serta memberikan kode agar sumber datanya tetap di telusuri.

3.4.3 Pengolahan data

Pengolahan data diartikan sebagai proses mengartikan data-data lapangan yang sesuai dengan tujuan, rancangan dan sifat penelitian. Metode pengolahan data dalam penelitian ini adalah yaitu :

1. Reduksi Data adalah mengurangi atau memilah-milah data yang sesuai dengan topik dimana data tersebut dihasilkan dari penelitian.
2. Koding data adalah penyesuaian data diperoleh dalam melakukan penelitian lapangan dengan pokok pada permasalahan dengan cara memberi kode-kode tertentu pada setiap data tersebut.

3.4.4 Analisis Apriori

Algoritma Apriori di gunakan untuk mencari *frequent itemset* yang memenuhi *min_supp* kemudian mendapatkan *rule* yang memenuhi konfigurasi *minimum* dari *frequent itemset* tadi. Cara kerja metode ini adalah

1. Tentukan *minimum support*
2. Iterasi 1 : hitung item-item dari *support* (transaksi yang memuat seluruh item) dengan men-scan database untuk 1-*itemset*, setelah 1-*itemset* didapatkan, dari 1-*itemset* apakah diatas *minimum support*, apabila telah memenuhi *minimum support*, 1-*itemset* tersebut akan menjadi pola *frequent* tinggi,
3. Iterasi 2 : untuk mendapatkan 2-*itemset*, harus dilakukan kombinasi dari *k-itemset* sebelumnya, kemudian scan database lagi untuk hitung item-item yang memuat *support*. *Itemset* yang memenuhi *minimum support* akan dipilih sebagai pola *frequent* tinggi dari kandidat
4. Tetapkan nilai *k-itemset* dari *support* yang telah memenuhi *minimum support* dari *k-itemset*
5. Lakukan proses untuk iterasi selanjutnya hingga tidak ada lagi k-*itemset* yang memenuhi *minimum support*.

3.5 Lokasi dan Jadwal Penelitian

Penelitian ini dilaksanakan di Poltabes Barelang yang berada di Jl. Sudirman No.4, Sukajadi, Kec.Batam Kota, Kota Batam, Kepulauan Riau dengan subjek laporan tahunan kecelakaan lalu lintas jalan raya Poltabes Barelang tahun 2019, dan dilakukan dari bulan oktober 2019 sampai november 2019.