

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB II

LANDASAN TEORI

Landasan teori merupakan panduan untuk membahas tentang teori pemecahan masalah yang dihadapi. Dalam hal ini akan dikemukakan beberapa teori yang berhubungan dan relevan dengan masalah yang dibahas dalam penelitian ini. Teori-teori yang menjadi landasan dalam penulisan tugas akhir ini antara lain teori kejahatan pelanggaran lalu lintas, *Data Mining*, metode *Clustering*, dan algoritma *K-Means*.

2.1 Pelanggaran Lalu Lintas

Peraturan pada dasarnya dibuat dengan tujuan untuk mempermudah kehidupan manusia. Bila tidak ada peraturan dan rambu-rambu lalu lintas, maka setiap hari pengguna jalan tidak mau mengindahkan kepentingan orang lain. Namun demikian, meskipun peraturan sudah dibuat ternyata tidak ada jaminan akan dipatuhi. Masalah kedisiplinan berlalu lintas yang buruk merupakan fenomena yang terjadi di kota-kota besar di negara-negara sedang berkembang (Yuwono 2012).

Pelanggaran merupakan suatu tindakan yang tidak sesuai dengan aturan yang ada, baik dalam norma masyarakat atau hukum yang berlaku. Dalam konteks ini pelanggaran lalu lintas adalah suatu tindakan baik sengaja ataupun tidak sengaja melakukan perbuatan untuk tidak mematuhi aturan-aturan lalu lintas yang berlaku. Pelanggaran lalu lintas adalah pelanggaran terhadap persyaratan administrasi atau pelanggaran terhadap persyaratan teknis oleh pemakai kendaraan bermotor sesuai ketentuan peraturan perundangan lalu lintas yang berlaku. Penindakan pelanggaran lalu lintas adalah tindakan hukum yang ditujukan kepada pelanggar peraturan lalu lintas yang dilakukan oleh petugas kepolisian Republik Indonesia secara edukatif maupun secara yuridis.

Tindakan edukatif adalah tindakan yang diberikan oleh petugas kepolisian Republik Indonesia berupa pemberian teguran dan peringatan dengan cara simpatik terhadap para pelanggar lalu lintas, sedangkan secara yuridis adalah penindakan dengan menggunakan tilang dan atau menggunakan berita acara singkat/ sumir/

tipiring atau dengan berita acara biasa terhadap pelanggaran yang berpotensi atau memiliki bobot sangat fatal (berat) dan dapat merusak fasilitas umum (putusnya jembatan dan lain-lain) serta melakukan penyidikan terhadap kecelakaan lalu lintas.

Pelanggaran lalu lintas tertentu atau yang sering disebut dengan tilang merupakan kasus dalam ruang lingkup hukum pidana yang diatur dalam UU Berdasarkan Undang-Undang Nomor 22 Tahun 2009 tentang Lalu Lintas dan Angkutan Jalan. Pelanggaran lalu lintas atau tilang yang biasanya sering terjadi adalah pelanggaran terhadap pasal 281 dan 288 ayat 1 mengenai kelengkapan surat kendaraan SIM dan STNK serta pasal 287 ayat 1 mengenai pelanggaran rambu lalu lintas (Suhariyanto 2015).

2.1.1 Bentuk-Bentuk Pelanggaran Lalu Lintas



Gambar 2.1 Pasal 287 Jo 106 (4)a UU LLAJ Ayat 1

Bentuk – bentuk pelanggaran lalu lintas diantaranya sebagai berikut: Undang-Undang Republik Indonesia Nomor 22 Tahun 2009 Tentang Lalu Lintas Dan Angkutan Jalan Bab XX Tentang Ketentuan Pidana (Undang-Undang RI No 22 Tahun 2009 Tentang Lalu Lintas Dan Angkutan Jalan 2009).

Ketentuan Pidana Pelanggaran Lalu Lintas

a. Pasal 267 Ayat 3 Jo 106 (4) UU LLAJ

Pelanggar yang tidak dapat hadir sebagaimana dimaksud pada ayat (2) dapat menitipkan denda kepada bank yang ditunjuk oleh pemerintah.

b. Pasal 280 Jo 68 (1) UU LLAJ

Setiap orang yang mengemudikan kendaraan bermotor di jalan yang tidak dipasang tanda nomor kendaraan bermotor yang ditetapkan oleh Kepolisian Negara Republik Indonesia sebagaimana dimaksud dalam Pasal 68 ayat (1) dipidana dengan pidana kurungan paling lama 2 (dua) bulan atau denda paling banyak Rp500.000,00 (lima ratus ribu rupiah).

c. Pasal 281 Jo 77 (1) UU LLAJ

Setiap orang yang mengemudikan kendaraan bermotor di jalan yang tidak memiliki Surat Izin Mengemudi sebagaimana dimaksud dalam Pasal 77 ayat (1) dipidana dengan pidana kurungan paling lama 4 (empat) bulan atau denda paling banyak Rp1.000.000,00 (satu juta rupiah).

d. Pasal 285 Jo 106 (3) a UU LLAJ

1. Pasal 285 Ayat 1 Jo 106 (3) a UU LLAJ

Setiap orang yang mengemudikan sepeda motor di jalan yang tidak memenuhi persyaratan teknis dan laik jalan yang meliputi kaca spion, klakson, lampu utama, lampu rem, lampu penunjuk arah, alat pemantul cahaya, alat pengukur kecepatan, knalpot, dan kedalaman alur ban sebagaimana dimaksud dalam Pasal 106 ayat (3) *juncto* Pasal 48 ayat (2) dan ayat (3) dipidana dengan pidana kurungan paling lama 1 (satu) bulan atau denda paling banyak Rp250.000,00 (dua ratus lima puluh ribu rupiah).

2. Pasal 285 Ayat 2 Jo 106 (3) a UU LLAJ

Setiap orang yang mengemudikan kendaraan bermotor beroda empat atau lebih di Jalan yang tidak memenuhi persyaratan teknis yang meliputi kaca spion, klakson, lampu utama, lampu mundur, lampu tanda batas dimensi badan kendaraan, lampu gandengan, lampu rem, lampu penunjuk arah, alat pemantul cahaya, alat pengukur kecepatan, kedalaman alur ban, kaca depan, spakbor, bumper, penggandengan, penempelan, atau penghapus

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

kaca sebagaimana dimaksud dalam Pasal 106 ayat (3) juncto Pasal 48 ayat (2) dipidana dengan pidana kurungan paling lama 2 (dua) bulan atau denda paling banyak Rp500.000,00 (lima ratus ribu rupiah).

e. Pasal 287 Jo 106 (4) a UU LLAJ

1. Pasal 287 Ayat 1 Jo 106 (4) a UU LLAJ

Setiap orang yang mengemudikan kendaraan bermotor di jalan yang melanggar aturan perintah atau larangan yang dinyatakan dengan Rambu Lalu Lintas sebagaimana dimaksud dalam Pasal 106 ayat (4) huruf a atau Marka Jalan sebagaimana dimaksud dalam Pasal 106 ayat (4) huruf b dipidana dengan pidana kurungan paling lama 2 (dua) bulan atau denda paling banyak Rp500.000,00 (lima ratus ribu rupiah).

2. Pasal 287 Ayat 2 Jo 106 (4) a UU LLAJ

Setiap orang yang mengemudikan kendaraan bermotor di jalan yang melanggar aturan perintah atau larangan yang dinyatakan dengan Alat Pemberi Isyarat Lalu Lintas sebagaimana dimaksud dalam Pasal 106 ayat (4) huruf c dipidana dengan pidana kurungan paling lama 2 (dua) bulan atau denda paling banyak Rp500.000,00 (lima ratus ribu rupiah).

3. Pasal 287 Ayat 3 Jo 106 (4) a UU LLAJ

Setiap orang yang mengemudikan kendaraan bermotor di jalan yang melanggar aturan gerakan lalu lintas sebagaimana dimaksud dalam Pasal 106 ayat (4) huruf d atau tata cara berhenti dan Parkir sebagaimana dimaksud dalam Pasal 106 ayat (4) huruf e dipidana dengan pidana kurungan paling lama 1 (satu) bulan atau denda paling banyak Rp250.000,00 (dua ratus lima puluh ribu rupiah).

f. Pasal 288 Jo 106 (4) a UU LLAJ

1. Pasal 288 Ayat 1 Jo 106 (4) a UU LLAJ

Setiap orang yang mengemudikan kendaraan bermotor di jalan yang tidak dilengkapi dengan Surat Tanda Nomor Kendaraan Bermotor atau Surat Tanda Coba Kendaraan Bermotor yang ditetapkan oleh Kepolisian Negara Republik Indonesia sebagaimana dimaksud dalam Pasal 106 ayat (5)a

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

dipidana dengan pidana kurungan paling lama 2 (dua) bulan atau denda paling banyak Rp500.000,00 (lima ratus ribu rupiah).

2. Pasal 288 Ayat 2 Jo 106 (4) a UU LLAJ

Setiap orang yang mengemudikan kendaraan bermotor di jalan yang tidak dapat menunjukkan Surat Izin Mengemudi yang sah sebagaimana dimaksud dalam Pasal 106 ayat (5) huruf b dipidana dengan pidana kurungan paling lama 1 (satu) bulan dan/atau denda paling banyak Rp250.000,00 (dua ratus lima puluh ribu rupiah).

3. Pasal 288 Ayat 3 Jo 106 (4) a UU LLAJ

Setiap orang yang mengemudikan mobil penumpang umum, mobil bus, mobil barang, kereta gandengan, dan kereta tempelan yang tidak dilengkapi dengan surat keterangan uji berkala dan tanda lulus uji berkala sebagaimana dimaksud dalam Pasal 106 ayat (5) huruf c dipidana dengan pidana kurungan paling lama 2 (dua) bulan atau denda paling banyak Rp500.000,00 (lima ratus ribu rupiah).

g. Pasal 289 Jo 106 (6) UU LLAJ

Setiap orang yang mengemudikan kendaraan bermotor atau penumpang yang duduk di samping Pengemudi yang tidak mengenakan sabuk keselamatan sebagaimana dimaksud dalam Pasal 106 ayat (6) dipidana dengan pidana kurungan paling lama 1 (satu) bulan atau denda paling banyak Rp250.000,00 (dua ratus lima puluh ribu rupiah).

h. Pasal 291 Jo 106 (8) UU LLAJ

1. Pasal 291 Ayat 1 Jo 106 (8) UU LLAJ

Setiap orang yang mengemudikan sepeda motor tidak mengenakan helm standar nasional Indonesia sebagaimana dimaksud dalam Pasal 106 ayat (8) dipidana dengan pidana kurungan paling lama 1 (satu) bulan atau denda paling banyak Rp250.000,00 (dua ratus lima puluh ribu rupiah).

2. Pasal 291 Ayat 2 Jo 106 (8) UU LLAJ

Setiap orang yang mengemudikan sepeda motor yang membiarkan penumpangnya tidak mengenakan helm sebagaimana dimaksud dalam Pasal 106 ayat (8) dipidana dengan pidana kurungan paling lama 1 (satu)

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

bulan atau denda paling banyak Rp250.000,00 (dua ratus lima puluh ribu rupiah).

i. Pasal 293 Jo 107 (1) UU LLAJ

1. Pasal 293 Ayat 1 Jo 107 (1) UU LLAJ

Setiap orang yang mengemudikan kendaraan bermotor di jalan tanpa menyalakan lampu utama pada malam hari dan kondisi tertentu sebagaimana dimaksud dalam Pasal 107 ayat (1) dipidana dengan pidana kurungan paling lama 1 (satu) bulan atau denda paling banyak Rp250.000,00 (dua ratus lima puluh ribu rupiah).

2. Pasal 293 Ayat 2 Jo 107 (1) UU LLAJ

Setiap orang yang mengemudikan sepeda motor di jalan tanpa menyalakan lampu utama pada siang hari sebagaimana dimaksud dalam Pasal 107 ayat (2) dipidana dengan pidana kurungan paling lama 15 (lima belas) hari atau denda paling banyak Rp100.000,00 (seratus ribu rupiah).

j. Pasal 306 Jo 168 (1) UU LLAJ

Setiap orang yang mengemudikan kendaraan angkutan barang yang tidak dilengkapi surat muatan dokumen perjalanan sebagaimana dimaksud dalam Pasal 168 ayat (1) dipidana dengan pidana kurungan paling lama 1 (satu) bulan atau denda paling banyak Rp250.000,00 (dua ratus lima puluh ribu rupiah).

k. Pasal 307 Jo 169 (1) UU LLAJ

Setiap orang yang mengemudikan kendaraan bermotor angkutan umum barang yang tidak mematuhi ketentuan mengenai tata cara pemuatan, daya angkut, dimensi kendaraan sebagaimana dimaksud dalam Pasal 169 ayat (1) dipidana dengan pidana kurungan paling lama 2 (dua) bulan atau denda paling banyak Rp500.000,00 (lima ratus ribu rupiah).

2.1.2 Dampak Pelanggaran Lalu Lintas

Tentunya dari permasalahan yang terjadi pada kondisi lalu lintas di Indonesia telah menimbulkan berbagai masalah khususnya menyangkut permasalahan lalu lintas. Permasalahan tersebut, seperti:

1. Tingginya angka kecelakaan lalu lintas baik pada persimpangan lampu lalu lintas maupun pada jalan raya.

2. Keselamatan para pengendara dan para pejalan kaki menjadi terancam.
3. Kemacetan lalu lintas akibat dari masyarakat yang enggan untuk berjalan kaki atau memanfaatkan sepeda ontel.
4. Kebiasaan melanggar peraturan lalu lintas yang biasa kemudian menjadi budaya melanggar peraturan.

2.1.3 Plat No Kendaraan Untuk Provinsi Riau

Setiap kendaraan dilengkapi dengan *primary* yang bersifat unik untuk masing-masing kendaraan. Dari kombinasi angka dan huruf pada kendaraan yang di Indonesia, 2 atau 3 karakter huruf terakhir merupakan kode untuk wilayah masing-masing kabupaten/ kota dari setiap provinsi. Dan dibawah ini merupakan kode wilayah untuk provinsi Riau.

1. A, T, Q, J, N : Kota Pekanbaru
2. B dan V : Kabupaten Indragiri Hilir
3. C dan I : Kabupaten Pelalawan
4. D dan E : Kabupaten Bengkalis
5. F, Z, dan O : Kabupaten Kampar
6. G dan L : Kabupaten Indragiri Hilir
7. K : Kabupaten Kuantan Singingi
8. M dan U : Kabupaten Rokan Hulu
9. P dan W : Kabupaten Rokan Hilir
10. R dan H : Kota Dumai
11. S dan Y : Kabupaten Siak
12. X : Meranti

2.1.4 Jenis Kendaraan

Dalam data pelanggaran lalu-lintas terdapat atribut jenis kendaraan. Atribut ini merupakan kode untuk kendaraan yang dikenakan tilang.

1. R2 : Kendaraan Roda Dua
2. R4 : Kendaraan Roda Empat
3. R6 : Kendaraan Roda Enam
4. R10 : Kendaraan Roda Sepuluh
5. R12 : Kendaraan Roda Dua Belas

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

6. MB : Mobil Bak

2.1.5 Kelompok Umur Pelanggar

Kelompok umur yang digunakan dalam penelitian ini adalah kategori umur dari Depkes RI yaitu:

1. Balita : 0-5 Tahun
2. Kanak-Kanak : 5-11 Tahun
3. Remaja Awal : 12-16 Tahun
4. Remaja Akhir : 17-25 Tahun
5. Dewasa Awal : 26-35 Tahun
6. Dewasa Akhir : 36-45 Tahun
7. Lansia Awal : 46-55 Tahun
8. Lansia Akhir : 56-65 Tahun
9. Manula : > 65 tahun

2.1.6 Pos Gurindam Wilayah Hukum Pekanbaru

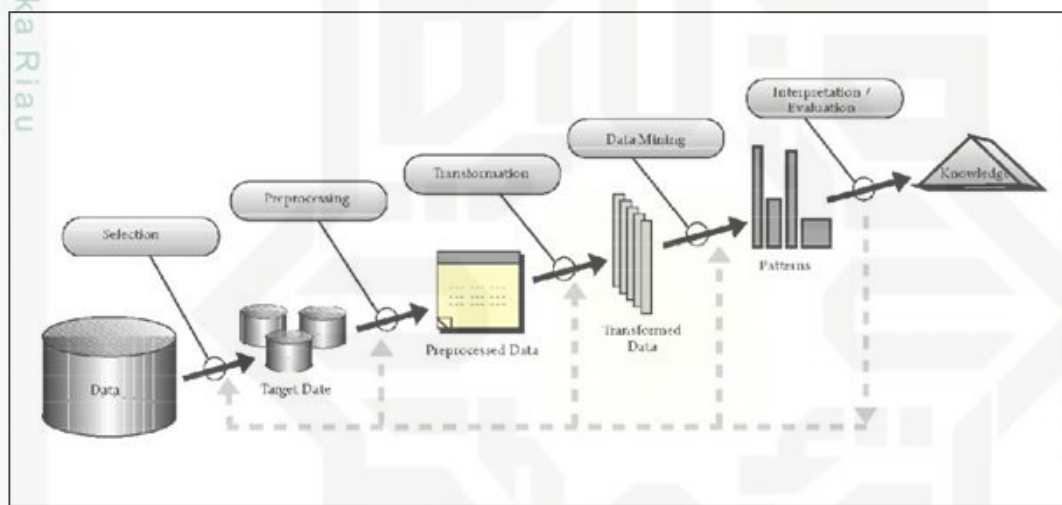
1. Pos Gurindam 1 : Jl. Sudirman, Depan Polda Riau
2. Pos Gurindam 2 : Jl. Sudirman- Tuangku Tambusai, Dibawah Flyover
3. Pos Gurindam 3 : Jl. Sudirman Arah Bandara
4. Pos Gurindam 4 : Jl. Soekarno- Hatta, Pasar Pagi Arengka
5. Pos Gurindam 5 : Labuh Baru, Payung Sekaki, Samping Mall Ska
6. Pos Gurindam 6 : Jl. Riau Ujung, Depan Polsek Payung Sekaki
7. Pos Gurindam 7 : Jl. Riau-Jl. Yos Sudarso
8. Pos Gurindam 8 : Jl. Ahmad Yani
9. Pos Gurindam 9 : Jl. Sudirman, Depan Mall Pekanbaru
10. Pos Gurindam 10 : Jl. Ahmad Yani- Jl. Cokroaminoto
11. Pos Gurindam 11 : Jl. Arifin Ahmad
12. Pos Gurindam 12 : Jl. HR. Soebrantas

2.2 Knowledge Discovery In Database (KDD)

KDD merupakan proses *nontrivial* dalam mengekstraksi data yang implisit, belum diketahui sebelumnya, dan berpotensi menjadi informasi yang berguna (Fayyad 1996). *Nontrivial* karena beberapa pencarian atau inferensi yang melibatkan bukan merupakan hasil komputasi secara langsung terhadap kuantitas

yang telah didefinisikan sebelumnya, seperti komputasi nilai rata-rata sekumpulan bilangan. Pola yang ditemukan harus valid terhadap data baru pada suatu tingkat kepastian tertentu. Pola-pola tersebut harus dapat menjadi suatu deskripsi atau gambaran tentang suatu pengetahuan yang secara potensial berguna dan menguntungkan bagi pengguna atau tugas tertentu. Akhirnya, pola-pola tersebut juga harus dapat dipahami dan dimengerti, walaupun terdapat kemungkinan tidak dapat secara langsung dan harus melewati beberapa proses dahulu.

Tahapan- Tahapan dari KDD



Gambar 2.2 Tahapan Tahapan dari KDD (Fayyad, 1996)

Berikut merupakan tahapan dari KDD (Fayyad 1996), Penjelasan tiap tahapan KDD dari gambar 2.2 diatas adalah sebagai berikut (Irmayantii 2016):

2.2.1 Seleksi Data

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang akan digunakan untuk proses *Data Mining*, disimpan dalam suatu berkas, terpisah dari basis data operasional.

2.2.2 Preprocessing / Cleaning

Sebelum proses *Data Mining* dapat dilaksanakan, perlu dilakukan tahapan *preprocessing* pada data yang menjadi fokus KDD. Tahapan *preprocessing* mencakup antara lain membersihkan data dari data yang *noise* seperti membuang duplikasi data, memeriksa data yang inkonsisten, memeriksa *missing value*,

2.3 Data Mining

Data Mining merupakan proses keempat dalam KDD (*Knowledge Discovery In Database*). *Data Mining* adalah proses utama dalam mencari pengetahuan yang tersembunyi dalam *database*, adapun pengertian *Data Mining* menurut beberapa ahli adalah sebagai berikut:

Menurut (Tan, Steinbach dan Kumar 2006), mendefinisikan *Data Mining* sebagai proses untuk mendapatkan informasi yang berguna dari gudang basis data yang besar. *Data Mining* juga dapat diartikan sebagai pengekstrakan informasi baru yang diambil dari bongkahan data besar yang membantu dalam pengambilan keputusan. Menurut *Data Mining* adalah suatu proses menemukan hubungan yang berarti, pola, dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika (Larose 2007).

Selain definisi di atas beberapa definisi juga diberikan seperti, “*Data Mining* adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual.” (Pramudiono, 2006). Menurut (Han dan Kamber 2006) secara sederhana *Data Mining* mengacu kepada mengekstrak atau “menambang” pengetahuan dari sekumpulan besar data. Menambang dalam hal ini bukan diibaratkan sebagai menambang emas atau tambang pasir, tetapi lebih diibaratkan sebagai “*Knowledge Mining From Data*” atau lebih ringkasnya menambang pengetahuan.

Menurut (Hermawati dan Astuti 2013) *Data Mining* adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (*machine learning*) untuk menganalisis dan mengekstraksi pengetahuan (*knowledge*) secara otomatis. *Data Mining* bukanlah suatu bidang yang sama sekali baru. Salah satu kesulitan untuk mendefinisikan *Data Mining* adalah kenyataan bahwa *Data Mining* mewarisi banyak aspek dan teknik dari bidang-bidang ilmu yang sudah mapan terlebih dahulu. Dan menunjukkan bahwa *Data Mining* memiliki akar yang panjang dari bidang ilmu seperti kecerdasan buatan (*artificial intelligent*), *machine learning*, statistik, *database*, dan juga *information retrieval* (Pramudiono, 2006).

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2.3.1 Karakteristik Data Mining

Data Mining juga memiliki beberapa karakteristik tertentu. Berikut karakteristik dari *Data Mining* (Davies dan Beynon 2004):

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

2.3.2 Fungsi Data Mining

Menurut (MacLennan, Tang dan Crivat 2009), banyak fungsi dari *Data Mining* yang dapat digunakan, pada masalah tertentu fungsi *Data Mining* dapat digunakan untuk menjawab permasalahan yang dihadapi. Berikut fungsi *Data Mining* secara umum:

1. Klasifikasi (*Classification*)

Fungsi klasifikasi adalah untuk mengelompokkan suatu target *class* kedalam kategori yang telah dipilih. Algoritma untuk melakukan klasifikasi antara lain *nearest neighborhood*, pohon keputusan, *naive bayes*, *neural network* dan *support vector machines*.

2. *Clustering*

Fungsi dari *Clustering* adalah untuk mencari pengelompokan atribut ke dalam segmentasi-segmentasi berdasarkan similaritas. Algoritma yang digunakan dalam proses *Clustering* adalah *K-Means Clustering*.

3. *Association*

Fungsi dari *association* adalah untuk mencari keterkaitan antara atribut atau *item set*, berdasarkan jumlah *item* yang muncul dan *rule association* yang ada. Beberapa algoritma dalam *association rule* ini adalah *Apriori*, *FP-Growth*, *FOLD Growth*, dan *ECLAT*.

4. *Regretion*

Fungsi dari *regretion* hampir mirip dengan klasifikasi. Fungsi *regretion* adalah bertujuan untuk mencari prediksi dari suatu pola yang ada.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

5. *Forecasting*

Fungsi dari *forecasting* adalah untuk peramalan waktu yang akan *dating* berdasarkan *trend* yang telah terjadi diwaktu sebelumnya.

6. *Sequence Analysis*

Fungsi dari *sequence analysis* adalah untuk mencari pola urutan dari rangkaian kejadian.

7. *Deviation Analysis*

Fungsi dari *deviation analysis* adalah untuk mencari kejadian langka yang sangat berbeda dari keadaan normal (kejadian abnormal).

2.3.3 Tugas Data Mining

Tugas *Data Mining* secara garis besar dibagi menjadi dua kategori utama, yaitu (Tan, Steinbach dan Kumar 2006) :

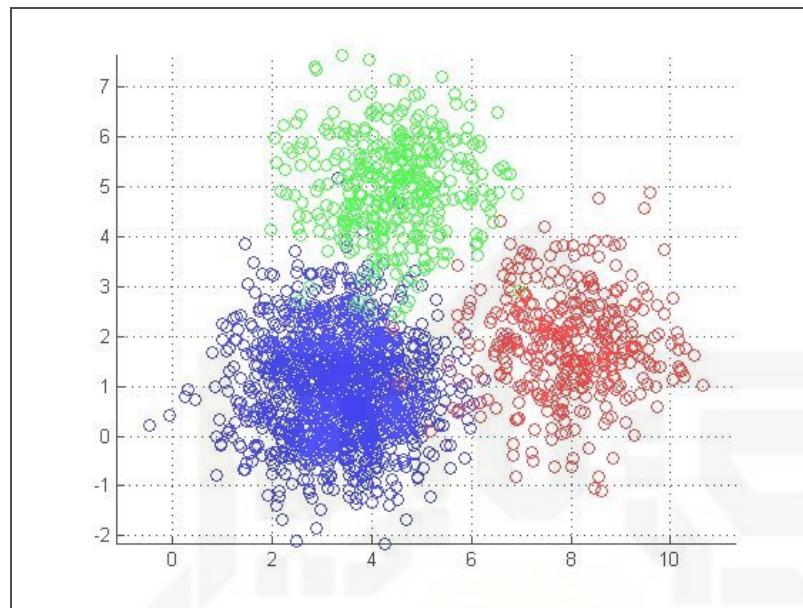
1. Tugas Prediktif

Tujuan utama dari tugas ini adalah untuk memprediksikan nilai dari atribut tertentu berdasarkan nilai dari atribut lainnya. Atribut yang diprediksi dikenal sebagai target atau *dependent variable*, sedangkan atribut yang digunakan untuk membuat prediksi disebut penjelas atau *independent variable*.

2. Tugas Deskriptif

Tujuan utama dari tugas ini adalah untuk memperoleh pola (*correlation, trend, cluster, trajectory, anomaly*) untuk menyimpulkan hubungan didalam data. Tugas deskriptif merupakan tugas *Data Mining* yang sering dibutuhkan pada teknik *post processing* untuk melakukan validasi dan menjelaskan hasil proses *Data Mining*.

2.4 Clustering



Gambar 2.3 Clustering

Menurut (Han dan Kamber, Data Mining: Concepts and Techniques 2006) (Han dan Kamber, Data Mining: Concepts and Techniques 2006), *Clustering* adalah proses pengelompokan kumpulan data menjadi beberapa kelompok sehingga objek di dalam satu kelompok memiliki banyak kesamaan dan memiliki banyak perbedaan dengan objek dikelompok lain. Perbedaan dan persamaannya biasanya berdasarkan nilai atribut dari objek tersebut dan dapat juga berupa perhitungan jarak. *Clustering* sendiri juga disebut *unsupervised classification*, karena *Clustering* lebih bersifat untuk dipelajari dan diperhatikan. *Cluster analysis* merupakan proses partisi satu set objek data ke dalam himpunan bagian. Setiap himpunan bagian adalah *cluster*, sehingga objek yang di dalam *cluster* mirip satu sama dengan yang lainnya, dan mempunyai perbedaan dengan objek dari *cluster* yang lain. Partisi tidak dilakukan dengan manual tetapi dengan algoritma *Clustering*. Oleh karena itu, *Clustering* sangat berguna dan bisa menemukan *group* yang tidak dikenal dalam data.

- Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2.5 Algoritma *K-Means*

Metode *K-Means* pertama kali diperkenalkan oleh MacQueen JB pada tahun 1976. Metode ini adalah salah satu metode *non hierarchi* yang umum digunakan. Metode ini termasuk dalam teknik penyekatan (*partition*) yang membagi atau memisahkan objek ke k daerah bagian yang terpisah. *K-Means* adalah salah satu metode data *Clustering* non hirarki yang mempartisi data ke dalam bentuk satu atau lebih *cluster* / kelompok, sehingga data yang memiliki karakteristik yang sama dikelompokkan dalam satu *cluster* yang sama dan data yang memiliki karakteristik berbeda dikelompokkan ke dalam kelompok lain. Ukuran kemiripan yang digunakan dalam *cluster* adalah fungsi jarak. Sehingga pemaksimalan kemiripan data didapatkan berdasarkan jarak terpendek antara data terhadap titik *centroid* (Asroni dan Adrian 2015).

K-Means merupakan salah satu algoritma *Clustering*. Tujuan algoritma ini yaitu untuk membagi data menjadi beberapa kelompok. Algoritma ini menerima masukan berupa data tanpa label kelas. Hal ini berbeda dengan *supervised learning* yang menerima masukan berupa vektor $(-x-1, y1), (-x-2, y2), \dots, (-x-i, yi)$, di mana x_i merupakan data dari suatu data pelatihan dan y_i merupakan label kelas untuk x_i .

Pada algoritma pembelajaran ini, komputer mengelompokkan sendiri data-data yang menjadi masukannya tanpa mengetahui terlebih dulu target kelasnya. Pembelajaran ini termasuk dalam *unsupervised learning*. Masukan yang diterima adalah data atau objek dan k buah kelompok (*cluster*) yang diinginkan. Algoritma ini akan mengelompokkan data atau objek ke dalam k buah kelompok tersebut. Pada setiap *cluster* terdapat titik pusat (*centroid*) yang merepresentasikan *cluster* tersebut.

Algoritma untuk melakukan *K-Means Clustering* adalah sebagai berikut (Han dan Kamber, Data Mining: Concepts and Techniques 2006):

1. Menentukan banyaknya *cluster*

Untuk melakukan *Clustering* dengan algoritma *K-Means* langkah yang pertama kali yaitu menentukan banyak *cluster* yang akan dibentuk. Pada penelitian ini, terdapat lima *cluster*.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2. Menentukan titik pusat (*centroid*)

Setelah menentukan banyak *cluster* yang akan dibentuk, langkah selanjutnya yaitu dengan menentukan titik pusat (*centroid*) dari tiap-tiap *cluster* dan pengambilan titik pusat dilakukan secara acak (*random*).

3. Menghitung jarak setiap objek ke titik pusat (*centroid*)

Langkah selanjutnya untuk melakukan *cluster* setelah menentukan titik pusat adalah dengan menghitung jarak setiap data dengan titik pusat yang sudah ditentukan sebelumnya. Rumus untuk menghitung jarak setiap objek ke titik pusat adalah :

$$D(X_j - C_j) = \sqrt{\sum_{j=0}^n (X_j - C_j)^2} \dots \dots \dots \text{Persamaan 2.1}$$

Keterangan

X = dataset pelanggaran lalu lintas, dataset yang digunakan untuk perhitungan *Data Mining* ini adalah kelompok umur, no polisi, jenis kendaraan, tempat terjadi pelanggaran dan jenis pelanggaran.

C = titik pusat (*centroid*) yang sudah ditentukan secara acak (*random*).

4. Kelompokkan data berdasarkan jarak terpendeknya antara data dengan *centroid* menjadi sebuah kelompok *cluster*.

5. Hitung rata-rata tiap kelompok *cluster* yang terbentuk untuk dijadikan sebagai *centroid* yang baru dan diulangi mencari jarak terpendek antara data dan *centroid* apabila *centroid* berubah dan perhitungan akan berhenti apabila *centroid* tidak mengalami perubahan.

$$C_{kj} = \frac{\sum_{i=1}^p x_{ij}}{p} \dots \dots \dots \text{Persamaan 2.2}$$

Keterangan

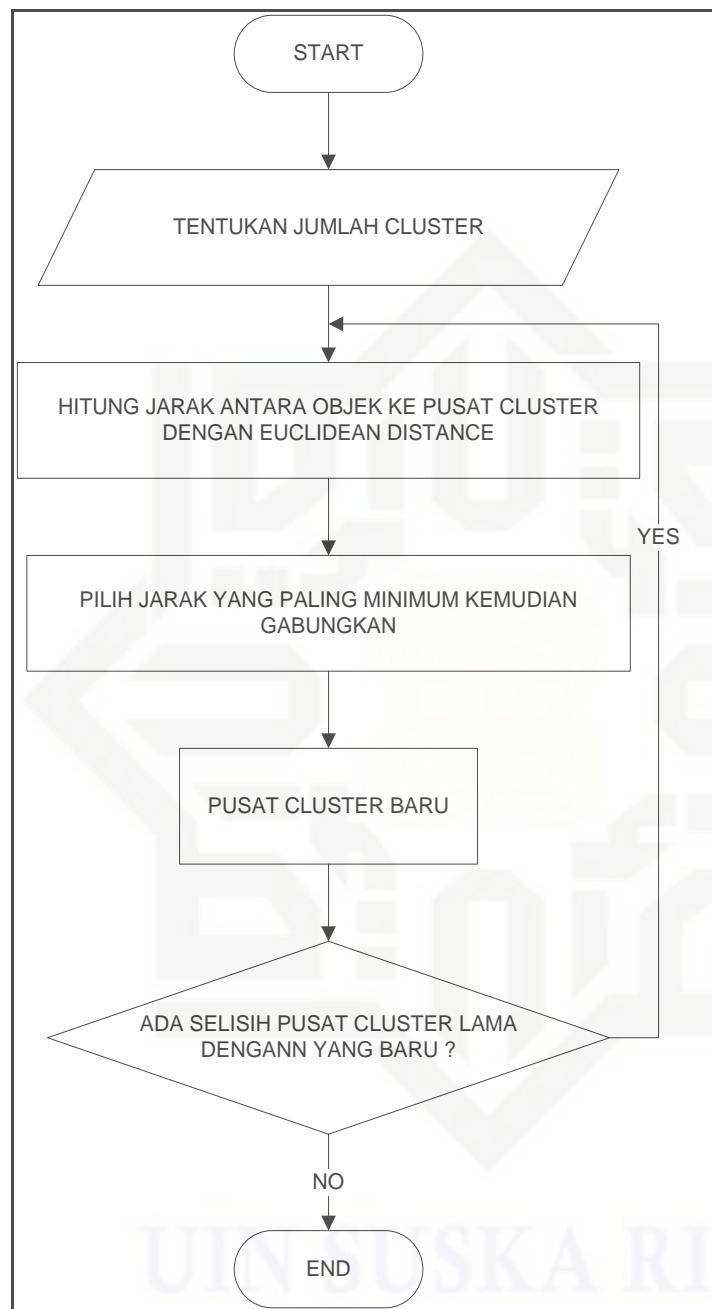
x_{ij} = \in kluster ke - k

p = banyaknya anggota kluster ke k

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2.5.1 Flowchart *Clustering* dengan algoritma *K-Means*



Gambar 2.4 Flowchart Algoritma K-Means

2.5.2 Contoh Penggunaan Algoritma *K-Means*

Dibawah ini merupakan contoh sederhana penggunaan algoritma *K-Means*.

Diketahui : jumlah data =12, jumlah atribut= 2.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Tabel 2.1 Contoh Data

NO	Kota /Kab	Luas Lahan	Produksi
1	Kab. Ponorogo	66,693.00	402,047.00
2	Kab. Trenggalek	31,136.00	182,848.00
3	Kab. Tulungagung	49,230.00	259,581.00
4	Kab. Blitar	50,577.00	289,494.00
5	Kab. Kediri	51,083.00	281,392.00
6	Kab. Malang	65,597.00	464,498.00
7	Kab. Lumajang	72,552.00	387,168.00
8	Kab. Jember	162,619.00	964,001.00
9	Kab. Banyuwangi	113,609.00	706,419.00
10	Kab. Bondowoso	61,330.00	329,557.00
11	Kab. Situbondo	48,902.00	290,954.00
12	Kab. Probolinggo	59,130.00	311,258.00

- 1 Penentuan Jumlah *Cluster*
 Jumlah *cluster* yang akan dibentuk berjumlah 3.
- 2 Penentuan Pusat Awal *Cluster*

Tabel 2.2 Penentuan Pusat Cluster

Data	Cluster	Luas Lahan	Produksi
Data ke-8	1	162.619	964.001
Data ke-7	2	72.552	387.168
Data ke-2	3	31.136	182.848

- 3 Perhitungan Jarak Pusat *Cluster*
 Perhitungan jarak ini menggunakan *Euclidean Distance* dengan rumus sebagai berikut:

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} \dots \dots \text{Persamaan 2.3}$$

Keterangan

X= dataset

Y= titik pusat *cluster* yang sudah ditentukan

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Tabel 2.3 Penentuan Jarak Terpendek Tiap Cluster

NO	Kota /Kab	Luas Lahan	Produksi	C1	C2	C3	Jarak Terpendek
1	Kab. Ponorogo	66,693	402,047	570,08	15,99101	222,0642	15,9910138
2	Kab. Trenggalek	31,136	182,848	792,14	208,4753	0	0
3	Kab. Tulungagung	49,23	259,581	713,49	129,701	78,83747	78,8374666
4	Kab. Blitar	50,577	289,494	683,75	100,1155	108,4035	100,115488
5	Kab. Kediri	51,083	281,392	691,67	107,9328	100,5425	100,542542
6	Kab. Malang	65,597	464,498	508,83	77,64213	283,7504	77,6421337
7	Kab. Lumajang	72,552	387,168	583,82	0	208,4753	0
8	Kab. Jember	162,619	964,001	0	583,8222	792,1413	0
9	Kab. Banyuwangi	113,609	706,419	262,20	321,8802	530,0268	262,203102
10	Kab. Bondowoso	61,33	329,557	642,48	58,69379	149,7839	58,6937868
11	Kab. Situbondo	48,902	290,954	682,59	99,07803	109,5561	99,0780314
12	Kab. Probolinggo	59,13	311,258	660,90	77,08747	131,426	77,087471

4 Pengelompokkan data berdasarkan jarak terpendek

Tabel 2.4 Pengelompokkan Berdasarkan Jarak Terpendek

No.	C1	C2	C3
1		1	
2			1
3			1
4		1	
5			1
6		1	
7		1	
8	1		
9	1		
10		1	
11		1	

No.	C1	C2	C3
12		1	

Keterangan

Angka 1 = jarak terpendek yang terletak pada Cn

5. Hitung rata-rata tiap *cluster* untuk dijadikan titik *cluster* baru

$$\begin{aligned} \text{C1 luas lahan} &= (X8 + X9) / 2 \\ &= (162,619 + 113,609) / 2 \\ &= 138,114 \end{aligned}$$

$$\begin{aligned} \text{C1 produksi} &= (X8 + X9) / 2 \\ &= (964,001 + 706,419) / 2 \\ &= 835,21 \end{aligned}$$

$$\begin{aligned} \text{C2 luas lahan} &= (X1 + X4 + X6 + X7 + X10 + X11 + X12) / 6 \\ &= (66,693 + 50,577 + 65,597 + 72,552 + 61,33 + 48,902 + 59,13) / 6 \\ &= 60,683 \end{aligned}$$

$$\begin{aligned} \text{C2 produksi} &= (X1 + X4 + X6 + X7 + X10 + X11 + X12) / 6 \\ &= (402,05 + 289,5 + 264,5 + 387,2 + 329,557 + 291 + 311,3) / 6 \\ &= 353,568 \end{aligned}$$

$$\begin{aligned} \text{C3 luas lahan} &= (X2 + X3 + X4 + X5) / 4 \\ &= 43,81633 \end{aligned}$$

$$\begin{aligned} \text{C3 produksi} &= (X2 + X3 + X4 + X5) / 4 \\ &= 241,2737 \end{aligned}$$

Sehingga didapatkan pusat *cluster* baru

Tabel 2.5 Pusat Cluster Baru

Cluster	Luas Lahan	Produksi
1	162.619	964.001
2	72.552	387.168
3	31.136	182.848

Lalu lanjutkan ke iterasi selanjutnya dengan menggunakan titik *cluster* yang baru. Iterasi akan berhenti jika pada iterasi ke-n, tidak ada pertukaran anggota sama sekali antar *cluster* (anggota masing-masing *cluster* sama dengan iterasi sebelumnya).

2.6 Penelitian Terkait

Sebelum melanjutkan penelitian ini, penulis melihat beberapa referensi untuk dijadikan acuan. Adapun beberapa penelitian terdahulu yang menjadi bahan dalam penelitian ini adalah:

Tabel 2.6 Penelitian Terkait

NO	JURNAL	JUDUL	PENULIS	KET
1	Kumpulan Jurnal Ilmu Komputer (KLIK) Volume 01 No 01 Sept 2014 ISSN: 2406-7857	<i>Clustering</i> Penentuan Potensi Kejahatan Di Kota Banjar Baru Dengan Metode <i>K-Means</i>	Sri Rahayu, Dodon T Nugrahadi, Fatma Indriani	Variabel-variabel yang digunakan adalah hukuman, bulan dan laporan. Sehingga penelitian ini bertujuan untuk mengkaji tahapan teknik <i>Clustering</i> dan membangun <i>Clustering</i> penentuan potensi kejahatan daerah di kota Banjarbaru.
2	Majalah Ilmiah Informasi Dan Teknologi (INTD) ISSN: 2339-210X Volume: III, No: I, Mei 2014	<i>Clustering</i> Pelanggaran Berkendaraan Menggunakan Algoritma K-Means Pada Polres Binjai	Relita Buaton, Fitri Nurhayati	Pengelompokan data dengan menggunakan metode <i>clustering</i> dengan algoritma K-Means yakni usia, jenis kendaraan dan jenis pelanggaran. Dengan pengelompokan objek diperoleh hasil yakni usia diantara 17 sampai denan 37 tahun, yang melakukan pelanggaran dengan tidak menggunakan sefty belt dan melanggar rambu lalu lintas lebih banyak menggunakan sepeda motor dan mobil
3	Citec Journal Vol 1 No 4 Agus 2014 ISSN: 2354-5771	Perbandingan Algoritma <i>K-Means</i> Dan EM Untuk <i>Clusterisasi</i> Nilai Mahasiswa Berdasarkan Asal Sekolah	Mardiani	membandingkan performa dari dua algoritma <i>Clustering</i> yaitu <i>K-Means</i> dan EM (<i>Expectation Maximation</i>), menunjukkan algoritma <i>K-Means</i> merupakan algoritma terbaik dibandingkan Algoritma EM, karena dilihat dari nilai-nilai koefisien silhouette yang telah didapatkan, nilai-nilai koefisien silhouette pada algoritma <i>K-Means</i> lebih banyak yang mendekati angka 1 dibandingkan algoritma EM.
4	PLOS ONE Feb 2014 Vol 9 Issue 2 e85884	Comparison of <i>K-Means</i> and Fuzzy C-Means Algorithm Performance for	Jiandong Yin, Hongzan Sun, Jiawen	membandingkan akurasi, durasi dan reproduktifitas 2 algoritma <i>Clustering</i> yaitu algoritma <i>K-Means</i> dan Fuzzy C-Means

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

		Automated Determination of the Arterial Input Fuction	Yang, Qiyong Guo	menggunakan data baik data simulasi maupun data klinis untuk deteksi AIF dengan tujuan mendapatkan perbaikan dalam kuantifikasi perfusi MRI. Hasil menunjukkan bahwa analisis <i>K-Means</i> menghasilkan pengelompokan yang lebih akurat dan memiliki hasil AIF yang kuat, meskipun membutuhkan waktu lebih lama untuk mengeksekusi dibandingkan dengan metode FCM.
5	International Journal Of Advanced Computer Science And Applications Vol 4 No 4 2013	Comparative Analysis of <i>K-Means</i> and Fuzzy C-Means Algorithms	Soumi Ghost, Sanjay Kumar Dubey	menyimpulkan secara keseluruhan bahwa algoritma <i>K-Means</i> lebih unggul dari algoritma fuzzy c-means. Dari hasil yang diperoleh dapat disimpulkan bahwa algoritma <i>K-Means</i> lebih baik daripada algoritma FCM. FCM menghasilkan hasil yang dekat dengan <i>Clustering K-Means</i> namun masih memerlukan waktu komputasi lebih banyak daripada <i>K-Means</i> karena perhitungan fuzzy dalam algoritma.
6	Journal of Agricultural Informatics Vol 6 No 3, 13-23, 2015 ISSN 2061-862X	Comparison of <i>K-Means</i> and Fuzzy C-Means Algorithms on Different <i>Cluster Structures</i>	Zeynel Cebeci, Figen Yildiz	merekomendasikan untuk menggunakan algoritma <i>K-Means</i> karena waktu komputasi yg lebih rendah dibandingkan dengan algoritma FCM dan <i>K-Means</i> akan menjadi pilihan yang baik untuk dataset besar karena kecepatan pelaksanaannya.

Masih minimumnya penelitian menggunakan metode *K-Means* terkhusus untuk studi kasus pelanggaran lalu lintas dengan jenis metode *Clustering*. Hal ini yang menjadikan penulis untuk melakukan penelitian dalam menentukan *Clustering* Pelanggaran Lalu Lintas di Kota Pekanbaru Menggunakan *K-Means Clustering*.