



Implementasi Data Mining Pada Data Pelanggaran Lalu Lintas di Lampung Utara Menggunakan Algoritma *K-Means* (Studi Kasus Kejaksaan Negeri Lampung Utara)

Elin Sadar Wasih¹, Rustam², Sidik Rahmatullah³, Sigit Mintoro⁴, Ngajiyanto⁵

¹²³⁴⁵Sistem Informasi, STMIK Dian Cipta Cendikia Kotabumi, Lampung Utara, 34581,

E-mail: elinsw94@gmail.com¹, sidik@dcc.ac.id², sigitmintoro76@gmail.com³,

rustamdcc89@gmail.com⁴, ngajiy@gmail.com⁵

Abstrak

Banyaknya data pelanggaran lalu lintas menyebabkan terjadinya penumpukan data pada instansi. Jika data tersebut tidak diolah dengan baik maka data tersebut hanya akan menjadi tumpukan data yang kurang bermanfaat. Penumpukan data tersebut terjadi pada salah satu instansi di bidang hukum yang ada di Lampung Utara, yaitu Kejaksaan Negeri Lampung Utara. Maka dari itu di perlukan suatu pengolahan data dengan menggunakan data mining. suatu teknik pengelompokan data dengan menggunakan Algoritma *K-Means*. Algoritma *K-Means* digunakan dalam pembentukan sebuah *cluster*. Mengimplementasikan Algoritma *K-Means* untuk mengelompokkan data pelanggaran lalu lintas sehingga mempermudah bagi Kejaksaan Negeri Lampung Utara dalam menentukan tindak pelanggaran lalu lintas yang paling banyak terjadi dan Mempermudah bagian pelayanan tilang dalam mengidentifikasi pengelompokkan data pelanggaran lalu lintas di Lampung Utara ke dalam 2 *cluster* (*banyak dan sedikit*). berdasarkan dataset yang ada untuk memonitoring keselamatan pengendara di jalan raya agar dapat dilakukan evaluasi. Hasil Penelitian didapat Pelanggaran lalu lintas paling banyak terjadi didominasi oleh jenis pelanggaran pasal 291 (1) yaitu pengendara sepeda motor yang tidak menggunakan helm dan pasal 291 (2) yaitu pengendara sepeda motor yang membiarkan penumpangnya tidak menggunakan helm. Hasil perhitungan menggunakan excel dan pengujian menggunakan *Rapidminer* menghasilkan hasil yang sama. Pada *cluster* 1 terdapat 9 pasal, artinya pelanggaran yang paling banyak terjadi ada dalam 9 pasal tersebut. Pada *cluster* 2 terdapat 15 pasal, artinya pelanggaran yang paling sedikit terjadi ada dalam 15 pasal tersebut. Pada aplikasi *Rapidminer* kita dapat melihat hasil *clustering* berupa plot grafik, graph view dan chart scatter.

Kata kunci— Pelanggaran lalu lintas, Data Mining, *K-Means*, *cluster*, *RapidMiner*.

Abstract

The large number of traffic violation data causes the accumulation of data at the agency. If the data is not processed properly then the data will only become a pile of data that is less useful. The accumulation of data occurs in one of the agencies in the field of law in North Lampung, namely the North Lampung District Attorney. Therefore in need of a data processing using data mining. a data grouping technique using the *K-Means* Algorithm. *K-Means* algorithm is used in the formation of a cluster. Implementing the *K-Means* Algorithm to group traffic violation data so that it makes it easier for the North Lampung District Attorney to



determine the most frequent traffic violations and Facilitate the ticketing service department in identifying the grouping of traffic violation data in North Lampung into 2 clusters (many and large). a little). based on existing datasets to monitor the safety of motorists on the highway so that an evaluation can be carried out. The results showed that the most traffic violations were dominated by types of violations of Article 291 (1), namely motorcycle riders who did not use helmets and Article 291 (2), namely motorcycle riders who let their passengers not use helmets. The results of calculations using excel and testing using Rapidminer produces the same result. In cluster 1 there are 9 articles, meaning that the most violations occurred in these 9 articles. In cluster 2 there are 15 articles, meaning that the least violations occurred in those 15 articles. In the Rapidminer application we can see the results of clustering in the form of graph plots, graph views and scatter charts.

Keywords—Pure Milk, Cut Price Promotion, Data Mining, Expectation-Maximization Algorithm, RapidMiner

I. PENDAHULUAN

Pelanggaran lalu lintas semakin meningkat setiap tahunnya seiring dengan meningkatnya pertumbuhan jumlah pengguna kendaraan. Hal ini dikarenakan lalu lintas merupakan sarana yang erat kaitannya dengan masyarakat. Hampir disetiap waktu masyarakat terus melekat dengan berbagai aktivitas dan kepentingan yang berhubungan dengan lalu lintas. Dalam melakukan aktivitasnya masyarakat lebih memilih menggunakan kendaraan pribadi dibandingkan kendaraan umum. Masyarakat menganggap penggunaan kendaraan pribadi lebih efektif dibandingkan kendaraan umum. Tetapi, disisi lain banyak masyarakat yang belum sadar akan kedispilinan berlalu lintas. Masalah pelanggaran lalu lintas yang semakin meningkat akibat kecerobohan manusia ini memiliki dampak yang sangat serius, salah satunya adalah kecelakaan yang merenggut korban jiwa. Undang - undang Nomor 22 Tahun 2009 tentang Lalu Lintas dan Angkutan Jalan, mengungkapkan kecelakaan lalu lintas adalah suatu peristiwa di jalan yang tidak diduga dan tidak disengaja melibatkan

kendaraan dengan atau tanpa pengguna jalan lain yang mengakibatkan korban manusia dan/atau kerugian harta benda. (Pemerintah RI 2009)

Rendahnya edukasi serta minim pemahaman tentang peraturan lalu lintas menyebabkan banyak masyarakat yang melanggar. Meningkatnya jumlah pelanggar lalu lintas menyebabkan meningkatnya data pelanggaran lalu lintas. Korps Lalu Lintas Kepolisian Republik Indonesia (Korlantas Polri) telah menerbitkan 1,77 juta bukti pelanggaran (tilang) lalu lintas hingga Oktober 2021. Dari jumlah itu, sebanyak 793.821 tilang atau 44,89% merupakan pelanggaran ringan. Sebanyak 746.153 tilang atau 42,22% merupakan pelanggaran berat. Kemudian, 227.819 tilang atau 12,89% merupakan pelanggaran sedang. (Pahlevi 2021)

Banyaknya data pelanggaran lalu lintas menyebabkan terjadinya penumpukan data pada instansi. Jika data tersebut tidak diolah dengan baik maka data tersebut hanya akan menjadi tumpukan data yang kurang bermanfaat. Penumpukan data tersebut

terjadi pada salah satu instansi di bidang hukum yang ada di Lampung Utara, yaitu Kejaksaan Negeri Lampung Utara. Maka dari itu di perlukan suatu pengolahan data dengan menggunakan data mining. Data mining digunakan untuk pengolahan data dalam skala besar. Dalam pengolahan data pelanggaran lalu lintas ini perlu dilakukan suatu teknik pengelompokan data dengan menggunakan Algoritma *K-Means*. Algoritma *K-Means* digunakan dalam pembentukan sebuah *cluster*. Mengimplementasikan Algoritma *K-Means* untuk mengelompokkan data pelanggaran lalu lintas sehingga mempermudah bagi Kejaksaan Negeri Lampung Utara dalam menentukan tindak pelanggaran lalu lintas yang paling banyak terjadi dan Mempermudah bagian pelayanan tilang dalam mengidentifikasi pengelompokkan data pelanggaran lalu lintas di Lampung Utara ke dalam 2 *cluster* (*banyak dan sedikit*). berdasarkan dataset yang ada untuk memonitoring keselamatan pengendara di jalan raya agar dapat dilakukan evaluasi.

II. METODOLOGI PENELITIAN

a. Data Mining

Menurut Urban dalam bukunya "Decisions support systems and intelligent systems" data mining adalah istilah yang digunakan untuk menggambarkan penemuan pengetahuan dalam database, data mining adalah proses yang menggunakan teknik statistik, matematika, buatan dan pembelajaran mesin untuk mengekstrak dan untuk mengidentifikasi informasi terkait dan pengetahuan terkait dari basis data yang besar. Data mining adalah proses yang menggunakan teknik statistik, matematika, buatan, dan pembelajaran mesin untuk mengekstrak dan mengidentifikasi informasi yang berguna dan pengetahuan

terkait dari berbagai database (Sugiyarti et al. 2018).

Metodologi yang digunakan pada penelitian ini adalah KDD (Knowledge Discovery in Database). (Kurniawan and Jajuli n.d.) Tahapan pada metodologi KDD ada 6. Berikut ini merupakan tahapan metodologi KDD.

1. Selection Data

Pada tahap pertama ini dilakukannya terlebih dahulu pemilihan data atau pemilihan attribute apa saja yang dibutuhkan, karena tidak semua attribute yang ada dalam database tersebut dibutuhkan semuanya. Atribut yang digunakan adalah atribut yang sesuai dengan tujuan dari penelitian ini. Data yang digunakan pada penelitian ini adalah data pelanggaran lalu lintas tahun 2021.

2. Preprocessing Data

Teknik yang digunakan pada data preprocessing adalah data cleaning. Data cleaning merupakan suatu Langkah yang harus dilakukan untuk menghilangkan data noise ataupun missing values. Tujuan dari dilakukannya tahap data cleaning adalah menghindari dari suatu data yang masih mempunyai nilainya yang tidak sesuai atau nilai yang error dalam setiap attribute yang ada.

3. Transformation Data

Pada tahap data transformation dilakukannya proses normalisasi. Tujuan dari dilakukannya proses ini adalah untuk mengubah bentuk nilai setiap attribute sesuai dengan algoritma yang nantinya akan digunakan. Karena, tidak semua tipe data dapat dilakukan menggunakan metode yang diinginkan. Maka dari itu tipe data ini harus disesuaikan dengan kebutuhan yang ada dengan cara mentransformasikan data sesuai dengan tujuan kita.



4. Data Mining

Pada tahap data mining, metode atau algoritma yang digunakan adalah *K-Means Clustering*. Metode *K-Means* adalah salah satu metode dalam fungsi *Clustering* atau pengelompokan. Menurut (Larose, 2005) *clustering* mengacu pada pengelompokan data, observasi atau kasus berdasar kemiripan objek yang diteliti. Sebuah *cluster* adalah suatu kumpulan data yang mirip dengan lainnya atau ketidakmiripan data pada kelompok lain

5. Evaluation Data

Pada tahap evaluation ini, dilakukan analisis terhadap hasil dari proses pembelajaran data. Evaluasi ini dilakukan untuk mengetahui pemodelan yang diterapkan apakah sudah sesuai dan cocok diterapkan pada kasus penelitian ini serta sesuai dengan tujuan yang ingin dicapai dari penelitian ini. Pada tahap ini di peroleh data yang relevan dari proses data mining dengan metode *K-Means Clustering*.

6. Knowledge

Setelah suatu hasil Analisa sudah didapatkan, maka selanjutnya hasil dari penelitian kita dapat menjadi pengetahuan bagi banyak orang, terkhusus untuk pemerintahan setempat yang berada di Kabupaten Lampung Utara itu sendiri. Setelah diketeahui, pemerintah setempat dapat langsung bertindak dengan tujuan untuk mengurangi tindak pelanggaran lalu lintas yang menyebabkan kecelakaan yang sering terjadi di Kabupaten Lampung Utara.

b. *K-Means*

Algoritma *K-Means* adalah salah satu algoritma yang menggunakan metode partisi. *K-Means* adalah algoritma *clustering* yang membagi masing - masing item data ke dalam satu *cluster*. Data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu

cluster/kelompok dan data yang memiliki karakteristik yang berbeda dikelompokkan dengan *cluster*/kelompok yang lain sehingga data yang berada dalam satu *cluster*/kelompok memiliki tingkat variasi yang kecil. Adapun tujuan pengelompokan data ini adalah untuk meminimalkan fungsi objektif yang diatur dalam proses pengelompokan, yang pada umumnya berusaha meminimalkan variasi di dalam suatu kelompok dan memaksimalkan variasi antar kelompok.(Purba, Siawin, and . 2019)

Berikut merupakan langkah-langkah pengolahan data menggunakan algoritma *K-Means* :

1. Menentukan jumlah *cluster* k yang dibentuk
2. Tentukan k centroid (titik pusat *cluster*) secara acak. Penentuan centroid awal dilakukan secara acak berdasarkan data-data yang tersedia sebanyak jumlah *cluster* yang sudah ditentukan pada langkah pertama.
3. Menghitung jarak setiap objek ke masing-masing centroid dari setiap *cluster*. Hitung dengan rumus *Euclidian Distance* jarak yang paling terdekat dengan data terhadap titik pusat. Rumus *Euclidian* sebagai berikut :

$$\text{dist}(x,y)=$$

$$\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots (x_n - y_n)^2}$$

dimana :

$\text{dist}(x,y)$ = Jarak data ke x ke

pusat *cluster* y

x_1 = Data ke 1 pada atribut data x

y_1 = Titik pusat ke 1 pada atribut

ke y

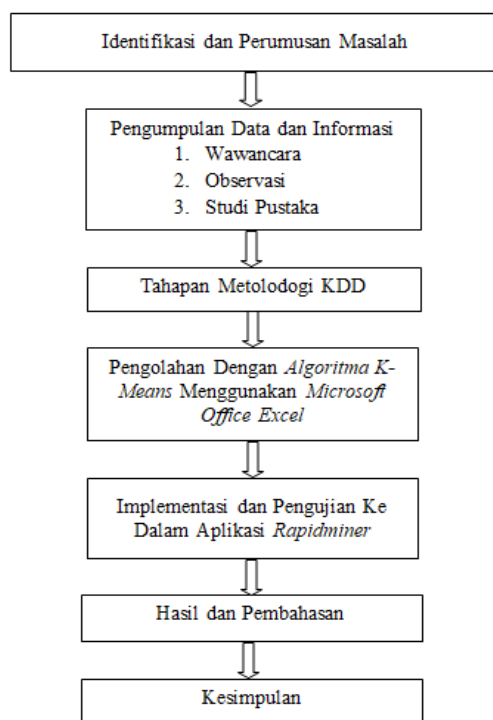
4. Mengelompokkan setiap data ke dalam *cluster* dengan jarak terkecil. Suatu data akan menjadi anggota sebuah *cluster* apabila jarak data ke pusat *cluster* bernilai paling kecil.
5. Memperbarui pusat *cluster* dengan cara

menghitung nilai rata-rata dari data yang menjadi anggota pada *cluster* tersebut.

6. Ulangi langkah 1, 2, 3 sampai tidak ada data yang pindah ke *cluster* yang lain atau sampai semua anggota *cluster* tidak lagi berubah

c. Kerangka Berpikir Penelitian

Tahapan metode penelitian tertuang dalam kerangka berpikir dibawah ini seperti pada Gambar 1.



Gambar 1. Kerangka Berpikir

d. Rapid Miner

Menurut (Fatmawati, 2016) Rapid Miner merupakan perangkat lunak yang dibuat oleh Dr. Markus Hofmann dari Institute of Technology Blanchardstown dan Raid Klinkenberg dari rapid-i.com dengan tampilan GUI (Graphical User Interface) sehingga memudahkan pengguna dalam menggunakan perangkat lunak ini (Srisulistiwati, Khaerudin, and ... 2021).



Gambar 2. Rapid Miner Menu Design

III. HASIL DAN PEMBAHASAN

a. Selection Data

Pada tahapan ini dilakukan pemilihan data atau atribut apa saja yang dibutuhkan karena tidak semua atribut akan digunakan, pemilihan atribut akan disesuaikan dengan tujuan penelitian ini

Tabel 1. Data sample Pelanggaran Lalu Lintas Bulan Januari 2021



Nomor Perkara	Tanda Nomor Kendaraan	Pelanggar	Sidang	Putusan	Jenis Kendaraan	Bukti	Dakwaan
1/Pid.LL/2021/PN Kbu	BE3861JC	Tulo Suwardi, Spd	08-Jan-21	Denda Rp 49,000;	Sepeda Motor	SIM C	288 (1) UULLAJ JO 106 (5) UULLAJ
2/Pid.LL/2021/PN Kbu	BE8554CU	Hendra Wiatno	08-Jan-21	Denda Rp 149,000 ;	Truk	STN K	286 UULLAJ 106 (3) UULLAJ
3/Pid.LL/2021/PN Kbu	BE5860QT	Sukidi	08-Jan-21	Denda Rp 39,000;	Sepeda Motor	SIM C	293 (2) UULLAJ JO 107 (2) UULLAJ
4/Pid.LL/2021/PN Kbu	BE4205JF	Candra Fauri	08-Jan-21	Denda Rp 49,000;	Sepeda Motor	STN K	291 (2) UULLAJ JO 106 UULLAJ
5/Pid.LL/2021/PN Kbu	BE4501JT	Febri Yolanda	08-Jan-21	Denda Rp 49,000;	Sepeda Motor	STN K	288 (2) UULLAJ JO 106 (5) UULLAJ
6/Pid.LL/2021/PN Kbu	BE3658WA	Wiko Hadi Aprian	08-Jan-21	Denda Rp 39,000;	Sepeda Motor	SIM C	285 (1) UULLAJ JO 106 (3) UULLAJ
7/Pid.LL/2021/PN Kbu	BE2782ACP	Anggun Oktaviani	08-Jan-21	Denda Rp 39,000;	Sepeda Motor	STN K	293 (2) UULLAJ JO 107 (2) UULLAJ
8/Pid.LL/2021/PN Kbu	BE4388JN	Marlina	08-Jan-21	Denda Rp 49,000;	Sepeda Motor	STN K	281 UULLAJ JO 77 (1) UULLAJ
9/Pid.LL/2021/PN Kbu	BE6471JO	Widya Cahya L	08-Jan-21	Denda Rp 49,000;	Sepeda Motor	STN K	281 UULLAJ JO 77 (1) UULLAJ
10/Pid.LL/2021/PN Kbu	BE9910C	Feri Thomas	08-Jan-21	Denda Rp 149,000 ;	Truk	SIM BII UM UM	286 UULLAJ 106 (3) UULLAJ
11/Pid.LL/2021/PN Kbu	BE9641	Arifin	08-Jan-21	Denda	Truk	SIM	305



21/PN Kbu	AV		Jan-21	Rp 149,000 ;		BII UM UM	UULLAJ JO 162 UULLAJ
12/Pid.LL/20 21/PN Kbu	BE3741 KV	Alex	08-Jan-21	Denda Rp 49,000;	Sepeda Motor	STN K	281 UULLAJ JO 77 (1) UULLAJ
13/Pid.LL/20 21/PN Kbu	BE9819 BU	Ardawi	08-Jan-21	Denda Rp 149,000 ;	Truk	STN K	286 UULLAJ 106 (3) UULLAJ
14/Pid.LL/20 21/PN Kbu	BE9150 WB	Tukidi	08-Jan-21	Denda Rp 149,000 ;	Truk	STN K	286 UULLAJ 106 (3) UULLAJ
15/Pid.LL/20 21/PN Kbu	B3550U RG	Dedi Setiawan	08-Jan-21	Denda Rp 49,000;	Sepeda Motor	STN K	287 (4) UULLAJ JO 106 UULLAJ
16/Pid.LL/20 21/PN Kbu	B9602S YT	Edi Winarno	08-Jan-21	Denda Rp 149,000 ;	Truk	STN K	288 (2) UULLAJ JO 106 (5) UULLAJ
17/Pid.LL/20 21/PN Kbu	BE9332 AJ	Iwan Wahyudi	08-Jan-21	Denda Rp 124,000 ;	Truk	SIM BII UM UM	278 UULLAJ JO 57 (3) UULLAJ
18/Pid.LL/20 21/PN Kbu	BE9613 AJ	Dedi Yusman	08-Jan-21	Denda Rp 149,000 ;	Truk	SIM C	285 (1) UULLAJ JO 106 (3) UULLAJ
19/Pid.LL/20 21/PN Kbu	BE3558 WF	Roby	08-Jan-21	Denda Rp 39,000;	Sepeda Motor	SIM C	285 (1) UULLAJ JO 106 (3) UULLAJ
20/Pid.LL/20 21/PN Kbu	D9363Y U	Anudi	08-Jan-21	Denda Rp 49,000;	Sepeda Motor	STN K	288 (2) UULLAJ JO 106 (5) UULLAJ
21/Pid.LL/20 21/PN Kbu	B6639K OV	Ambriyanto	08-Jan-21	Denda Rp 39,000;	Sepeda Motor	SIM C	285 (1) UULLAJ JO 106 (3) UULLAJ
22/Pid.LL/20 21/PN Kbu	BE2115I L	Sunarmo	08-Jan-	Denda Rp	Sepeda Motor	SIM C	285 (1) UULLAJ JO



			21	39,000;			106 (3) UULLAJ
--	--	--	----	---------	--	--	-------------------

Tabel 2. Data Pelanggaran Lalu Lintas Setelah Dilakukan Preprocessing Data

b. Preprocessing Data

Nomor Perkara	Dakwaan	Denda
1/Pid.LL/2021/PN Kbu	288 (1) UULLAJ JO 106 (5) UULLAJ	Denda Rp 49,000;
2/Pid.LL/2021/PN Kbu	286 UULLAJ 106 (3) UULLAJ	Denda Rp 149,000;
3/Pid.LL/2021/PN Kbu	293 (2) UULLAJ JO 107 (2) UULLAJ	Denda Rp 39,000;
4/Pid.LL/2021/PN Kbu	291 (2) UULLAJ JO 106 UULLAJ	Denda Rp 49,000;
5/Pid.LL/2021/PN Kbu	288 (2) UULLAJ JO 106 (5) UULLAJ	Denda Rp 49,000;
6/Pid.LL/2021/PN Kbu	285 (1) UULLAJ JO 106 (3) UULLAJ	Denda Rp 39,000;
7/Pid.LL/2021/PN Kbu	293 (2) UULLAJ JO 107 (2) UULLAJ	Denda Rp 39,000;
8/Pid.LL/2021/PN Kbu	281 UULLAJ JO 77 (1) UULLAJ	Denda Rp 49,000;
9/Pid.LL/2021/PN Kbu	281 UULLAJ JO 77 (1) UULLAJ	Denda Rp 49,000;
10/Pid.LL/2021/PN Kbu	286 UULLAJ 106 (3) UULLAJ	Denda Rp 149,000;
11/Pid.LL/2021/PN Kbu	305 UULLAJ JO 162 UULLAJ	Denda Rp 149,000;
12/Pid.LL/2021/PN Kbu	281 UULLAJ JO 77 (1) UULLAJ	Denda Rp 49,000;
13/Pid.LL/2021/PN Kbu	286 UULLAJ 106 (3) UULLAJ	Denda Rp 149,000;
14/Pid.LL/2021/PN Kbu	286 UULLAJ 106 (3) UULLAJ	Denda Rp 149,000;
15/Pid.LL/2021/PN Kbu	287 (4) UULLAJ JO 106 UULLAJ	Denda Rp 49,000;
16/Pid.LL/2021/PN Kbu	288 (2) UULLAJ JO 106 (5) UULLAJ	Denda Rp 149,000;
17/Pid.LL/2021/PN Kbu	278 UULLAJ JO 57 (3) UULLAJ	Denda Rp 124,000;
18/Pid.LL/2021/PN Kbu	285 (1) UULLAJ JO 106 (3) UULLAJ	Denda Rp 149,000;
19/Pid.LL/2021/PN Kbu	285 (1) UULLAJ JO 106 (3) UULLAJ	Denda Rp 39,000;
20/Pid.LL/2021/PN Kbu	288 (2) UULLAJ JO 106 (5) UULLAJ	Denda Rp 49,000;
21/Pid.LL/2021/PN Kbu	285 (1) UULLAJ JO 106 (3) UULLAJ	Denda Rp 39,000;
22/Pid.LL/2021/PN Kbu	285 (1) UULLAJ JO 106 (3) UULLAJ	Denda Rp 39,000;
23/Pid.LL/2021/PN Kbu	288 (2) UULLAJ JO 106 (5) UULLAJ	Denda Rp 49,000;
24/Pid.LL/2021/PN Kbu	288 (2) UULLAJ JO 106 (5) UULLAJ	Denda Rp 49,000;
25/Pid.LL/2021/PN Kbu	288 (2) UULLAJ JO 106 (5) UULLAJ	Denda Rp 49,000;
26/Pid.LL/2021/PN Kbu	288 (2) UULLAJ JO 106 (5) UULLAJ	Denda Rp 49,000;
27/Pid.LL/2021/PN Kbu	285 (1) UULLAJ JO 106 (3) UULLAJ	Denda Rp 39,000;
28/Pid.LL/2021/PN Kbu	288 (2) UULLAJ JO 106 (5) UULLAJ	Denda Rp 49,000;
29/Pid.LL/2021/PN Kbu	288 (2) UULLAJ JO 106 (5) UULLAJ	Denda Rp 49,000;
30/Pid.LL/2021/PN Kbu	293 (2) UULLAJ JO 107 (2) UULLAJ	Denda Rp 39,000;
31/Pid.LL/2021/PN Kbu	285 (1) UULLAJ JO 106 (3) UULLAJ	Denda Rp 39,000;
32/Pid.LL/2021/PN Kbu	288 (2) UULLAJ JO 106 (5) UULLAJ	Denda Rp 49,000;

c. Transformation Data

Tabel 3. Data Setelah Ditransformasi

Pasal	Jumlah	Denda
278 UULLAJ JO 57 (3) UULLAJ	8	124.000
279 UULLAJ JO 58 (2) UULLAJ	9	49.000



280 UULLAJ JO 68 (1) UULLAJ	91	49.000
281 UULLAJ JO 77 (1) UULLAJ	205	49.000
285 (1) UULLAJ JO 106 (3) UULLAJ	120	39.000
285 (2) UULLAJ JO 106 (3) UULLAJ	9	49.000
286 UULLAJ JO 106 (3) UULLAJ	8	149.000
287 (1) UULLAJ JO 106 (4) UULLAJ	147	49.000
287 (2) UULLAJ JO 106 (4) UULLAJ	7	49.000
287 (4) UULLAJ JO 106 UULLAJ	11	49.000
287 (5) UULLAJ JO 106 (4) UULLAJ	18	39.000
288 (1) UULLAJ JO 106 (5) UULLAJ	110	49.000
288 (2) UULLAJ JO 106 (5) UULLAJ	141	49.000
288 (3) UULLAJ JO 106 (5) UULLAJ	2	74.000
289 UULLAJ JO 106 (6) UULLAJ	8	149.000
291 (1) UULLAJ JO 106 UULLAJ	211	49.000
291 (2) UULLAJ JO 106 UULLAJ	204	49.000
293 (2) UULLAJ JO 107 (2) UULLAJ	111	39.000
297 UULLAJ JO 115 UULLAJ	12	39.000
303 UULLAJ JO 137 (4) UULLAJ	18	149.000
305 UULLAJ JO 162 UULLAJ	6	149.000
306 UULLAJ JO 162 UULLAJ	10	149.000
307 UULLAJ JO 169 (1) UULLAJ	47	149.000
308 UULLAJ JO 173 UULLAJ	4	99.000

e. Data Mining K-Means

Setelah melakukan transformasi data maka tahapan selanjutnya adalah data mining dengan Algoritma *K-Means*. Tahap pertama yang kita lakukan adalah dengan menentukan titik centroid awal. Titik centroid awal dapat dipilih secara acak.

Tabel 4. Titik Centroid Awal

Titik Centroid Awal		
C1	211	49.000
C2	2	74.000

Setelah ditentukan titik centroid awal, tahap selanjutnya adalah menghitung jarak setiap data ke centroid terdekat dengan menggunakan Euclidian Distance.

$\text{dist}(x,y)=$

$$\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2}$$

1. Iterasi 1

Setelah diperoleh nilai dari masing masing *cluster* selanjutnya adalah menentukan jarak terpendek antara C1 dan C2 dengan pusat data *clustering*. Nilai yang telah ditentukan melalui jarak terpendek akan dipisahkan ke dalam kolom *cluster 1* dan *cluster 2* dan diberi tanda * untuk menandakan bahwa nilai tersebut masuk ke dalam *cluster* sesuai jarak terpendek. Hasil tersebut dapat dilihat pada tabel dibawah ini.

Tabel 5. Hasil *Cluster* Pada Iterasi 1

C1	C2	Jarak Terpendek	Cluster 1	Cluster 2
216.41	50.358	50.35871		
16448	71325	325		*
202	25.961	25.96150		
	50997	997		*
120	92.444	92.44457		
	57799	799		*
6	204.53			
	36158	6	*	



91.547 80172	123.08 1274	91.54780 172	*	
202	25.961 50997	25.96150 997		*
226.29 40565	75.239 61722	75.23961 722		*
64	147.13 93897	64	*	
204	25.495 09757	25.49509 757		*
200	26.570 66051	26.57066 051		*
193.25 88937	38.483 76281	38.48376 281		*
101	110.85 57621	101	*	
70	141.23 03084	70	*	
210.48 99047	0	0		*
226.29 40565	75.239 61722	75.23961 722		*
0	210.48 99047	0	*	
7	203.54 11506	7	*	
100.49 87562	114.48 14395	100.4987 562		*
199.25 10979	36.400 54945	36.40054 945		*
217.36 8351	76.687 67828	76.68767 828		*
228.08 98946	75.106 59092	75.10659 092		*
224.50 16704	75.425 45989	75.42545 989		*
192.08 33153	87.464 27842	87.46427 842		*
212.95 30465	25.079 87241	25.07987 241		*

Dari hasil perhitungan Iterasi 1 maka akan dilakukan perhitungan untuk menemukan titik centroid baru. Titik centroid baru diperoleh dengan mengitung rata-rata angka

yang termasuk ke dalam *cluster* tersebut. Berikut ini akan disajikan tabel untuk melihat data yang masuk ke dalam *cluster* 1 dan 2 untuk kemudian dilakukan perhitungan dalam menentukan pusat *cluster* baru

Tabel 6. Titik Centroid baru untuk Iterasi 2

Cluster Baru	
C1	C2
156.125	16.75
46.5	94.625

2. Iterasi 2

Perhitungan akan terus dilakukan hingga tidak ada lagi *cluster* yang berubah. Ikuti langkah-langkah perhitungan seperti pada iterasi 1. Berikut ini merupakan hasil perhitungan pada iterasi 2.

Tabel 7. Hasil Perhitungan Iterasi 2

C1	C2	Jarak Terpendek	Cluster 1	Cluster 2
167.1 744	30.65 05	30.650499 6		*
147.1 462	46.27 854	46.278538 5		*
65.17 297	87.14 759	65.172967	*	
48.93 89	193.7	48.938896 9	*	
36.89 533	117.2 804	36.895333 4	*	
147.1 462	46.27 854	46.278538 5		*
180.1 312	55.07 452	55.074523 4		*
9.461 27	138.0 098	9.4612697 4	*	
149.1 46	46.65 515	46.655151 1		*
145.1 465	45.98 59	45.985901 4		*
138.3 285	55.63 904	55.639043 2		*



46.19 27	103.8 133	46.192701	*	
15.33 022	132.3 62	15.330219 3	*	
156.5 591	25.35 652	25.356520 4		*
180.1 312	55.07 452	55.074523 4		*
54.93 192	199.5 362	54.931918 1	*	
47.94 023	192.7 283	47.940229 7	*	
45.74 402	109.4 404	45.744022 8	*	
144.3 2	55.82 744	55.827440 6		*
172.0 022	54.38 937	54.389365 9		*
181.7 794	55.42 746	55.427458 2		*
178.4 902	54.79 236	54.792363 7		*
149.7 148	62.22 301	62.223011 2		*
160.9 294	13.47 973	13.479730 2		*

Tabel 8. Titik Centroid Baru Untuk Iterasi 3

Cluster Baru	
C1	C2
148.8888889	11.8
46.77777778	97.66666667

3. Iterasi 3

Pada iterasi ke 3 dilakukan perhitungan yang sama seperti iterasi 1 dan iterasi 2 sehingga diperoleh hasil sebagai berikut :

Tabel 9. Hasil Perhitungan Cluster Iterasi 3

C1	C2	Jarak Terpendek	Cluster 1	Cluster 2
160.6 641	26.60 61	26.606097 9		*
139.9 065	48.74 715	48.747148 1		*
57.93	92.95	57.931526	*	

153	743	2		
56.15 51	199.2 352	56.155098 3	*	
29.91 758	123.0 813	29.917582 3	*	
139.9 065	48.74 715	48.747148 1		*
174.0 663	51.47 379	51.473790 5		*
2.916 534	143.6 923	2.9165343 9	*	
141.9 063	48.90 281	48.902806 1		*
137.9 068	48.67 324	48.673241 6		*
131.1 198	58.99 337	58.993370 6		*
38.95 233	109.5 978	38.952329 2	*	
8.195 904	138.0 619	8.1959038 2	*	
149.3 901	25.61 545	25.615446 7		*
174.0 663	51.47 379	51.473790 5		*
62.15 085	205.0 587	62.150851 9	*	
55.15 59	198.2 657	55.155895 8	*	
38.67 896	115.2 494	38.678957 2	*	
137.1 097	58.66 701	58.667007 6		*
166.0 761	51.70 639	51.706393 3		*
175.6 89	51.65 996	51.659956 6		*
172.4 515	51.36 488	51.364882 1		*
144.3 285	62.24 268	62.242679 2		*
154.0 128	7.913 14	7.9131395 7		*



Tabel 10. Titik Centroid Baru Untuk Iterasi

4	
Cluster Baru	
C1	C2
148.8888889	11.8
46.77777778	97.66666667

4. Iterasi 4

Pada iterasi 4 diperoleh hasil sebagai berikut:

Tabel 11. Tabel Hasil Perhitungan Iterasi 4

C1	C2	Jarak Terpendek	Cluster 1	Cluster 2
160.66 40925	26.60 60979	26.60609 788		*
139.90 65385	48.74 71481	48.74714 806		*
57.931 5262	92.95 74335	57.93152 62	*	
56.155 09827	199.2 35249	56.15509 827	*	
29.917 58226	123.0 81346	29.91758 226		
139.90 65385	48.74 71481	48.74714 806		*
174.06 62567	51.47 37905	51.47379 053		*
2.9165 34389	143.6 92326	2.916534 389	*	
141.90 62897	48.90 28061	48.90280 61		*
137.90 67944	48.67 32416	48.67324 156		*
131.11 97737	58.99 33706	58.99337 063		*
38.952 32921	109.5 9783	38.95232 921	*	
8.1959 03825	138.0 61886	8.195903 825	*	
149.39 00769	25.61 54467	25.61544 673		*
174.06 62567	51.47 37905	51.47379 053		*
62.150	205.0	62.15085	*	

85193	58734	193		
55.155	198.2	55.15589	*	
89578	65692	578		
38.678	115.2	38.67895	*	
95718	49372	718		
137.10	58.66	58.66700		*
96704	70076	757		
166.07	51.70	51.70639		*
6139	63933	333		
175.68	51.65	51.65995		*
89788	99566	655		
172.45	51.36	51.36488		*
14603	48821	208		
144.32	62.24	62.24267		*
85432	26792	918		
154.01	7.913	7.913139		*
28261	13957	565		

Dari hasil perhitungan iterasi 4 diketahui bahwa posisi *cluster* tidak lagi berubah. Maka perhitungan sudah bisa dihentikan. Untuk lebih jelasnya akan disajikan posisi *cluster* pada iterasi 1 sampai iterasi 4 pada tabel dibawah ini.

Tabel 12. Posisi *cluster* pada iterasi 1 – 4

Iterasi 1	Iterasi 2	Iterasi 3	Iterasi 4
C2	C2	C2	C2
C2	C2	C2	C2
C2	C1	C1	C1
C1	C1	C1	C1
C1	C1	C1	C1
C2	C2	C2	C2
C2	C2	C2	C2
C1	C1	C1	C1
C2	C2	C2	C2
C2	C2	C2	C2
C2	C2	C2	C2
C1	C1	C1	C1
C1	C1	C1	C1
C2	C2	C2	C2
C2	C2	C2	C2
C1	C1	C1	C1
C1	C1	C1	C1
C1	C1	C1	C1
C2	C2	C2	C2



C2	C2	C2	C2
C2	C2	C2	C2
C2	C2	C2	C2
C2	C2	C2	C2
C2	C2	C2	C2

Dapat diambil kesimpulan bahwa hasil pengelompokkan *cluster* seperti pada tabel dibawah ini.

Tabel 13. Hasil Pengelompokkan Data *Cluster*

No	Cluster	Keterangan	Hasil
1.	C1	Banyak	9 items
2.	C2	Sedikit	15 items

Berdasarkan hasil *clustering* menggunakan Algoritma *K-Means* didapat hasil bahwa jenis pelanggaran yang paling banyak terjadi (C1) ada 9 items yang terdiri dari :

Tabel 14. Pasal dan Jenis Pelanggaran Pada C1

No	Pasal	Jenis Pelanggaran
1.	280 UULLAJ JO 68 (1) UULLAJ	Setiap orang yang mengemudikan Kendaraan Bermotor di Jalan yang tidak dipasang Tanda Nomor Kendaraan Bermotor yang ditetapkan oleh Kepolisian Negara Republik Indonesia
2.	281 UULLAJ JO 77 (1) UULLAJ	Setiap orang yang mengemudikan Kendaraan Bermotor di Jalan yang tidak memiliki Surat Izin Mengemudi
3.	285 (1) UULLAJ JO 106 (3) UULLAJ	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
4.	287 (1) UULLAJ JO 106 (4) UULLAJ	Setiap orang yang mengemudikan Kendaraan Bermotor di Jalan yang melanggar aturan perintah atau larangan yang dinyatakan dengan Rambu Lalu Lintas
5.	288 (1) UULLAJ JO 106 (5) UULLAJ	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
6.	288 (2) UULLAJ JO 106 (5) UULLAJ	Setiap orang yang mengemudikan Kendaraan Bermotor di Jalan yang tidak dapat menunjukkan Surat Izin Mengemudi yang sah
7.	291 (1) UULLAJ JO	Setiap orang yang mengemudikan



	106 UULLAJ	Sepeda Motor tidak mengenakan helm standar nasional Indonesia
8.	291 (2) UULLAJ JO 106 UULLAJ	Setiap orang yang mengemudikan Sepeda Motor yang membiarkan penumpangnya tidak mengenakan helm
9.	293 (2) UULLAJ JO 107 (2) UULLAJ	Setiap orang yang mengemudikan Sepeda Motor di Jalan tanpa menyalakan lampu utama pada siang hari

Pelanggaran lalu lintas paling banyak terjadi didominasi oleh jenis pelanggaran pasal 291 (1) yaitu pengendara sepeda motor yang tidak menggunakan helm dan pasal 291 (2) yaitu pengendara sepeda motor yang membiarkan penumpangnya tidak menggunakan helm.

Tabel 15. Pasal dan Jenis Pelanggaran Pada Cluster 2

No	Pasal	Jenis Pelanggaran
1.	278 UULLAJ JO 57 (3) UULLAJ	Setiap orang yang mengemudikan Kendaraan Bermotor beroda empat atau lebih di Jalan yang tidak dilengkapi dengan perlengkapan berupa ban cadangan, segitiga pengaman, dongkrak, pembuka roda, dan peralatan pertolongan pertama pada kecelakaan
2.	279 UULLAJ JO	Setiap orang yang mengemudikan

	58 (2) UULLAJ	Kendaraan Bermotor di Jalan yang dipasang perlengkapan yang dapat mengganggu keselamatan berlalu lintas
3.	285 (2) UULLAJ JO 106 (3) UULLAJ	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 kaca
4.	286 UULLAJ 106 (3) UULLAJ	Setiap orang yang mengemudikan Kendaraan Bermotor beroda empat atau lebih di Jalan yang tidak memenuhi persyaratan laik jalan
5.	287 (2) UULLAJ JO 106 (4) UULLAJ	Setiap orang yang mengemudikan Kendaraan Bermotor di Jalan yang melanggar aturan perintah atau larangan yang dinyatakan dengan Alat Pemberi Isyarat Lalu Lintas



6.	287 (4) UULLAJ JO 106 UULLAJ	Setiap orang yang mengemudikan Kendaraan Bermotor di Jalan yang melanggar ketentuan mengenai penggunaan atau hak utama bagi Kendaraan Bermotor yang menggunakan alat peringatan dengan bunyi dan sinar
7.	287 (5) UULLAJ JO 106 (4) UULLAJ	Setiap orang yang mengemudikan Kendaraan Bermotor di Jalan yang melanggar aturan batas kecepatan paling tinggi atau paling rendah
8.	288 (3) UULLAJ JO 106 (5) UULLAJ	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
9.	289 UULLAJ JO 106 (6) UULLAJ	Setiap orang yang mengemudikan Kendaraan Bermotor atau Penumpang yang duduk di samping Pengemudi yang tidak mengenakan sabuk keselamatan
10.	297 UULLAJ JO 115 UULLAJ	Setiap orang yang mengemudikan Kendaraan Bermotor berbalapan di Jalan

11.	303 UULLAJ JO 137 (4) UULLAJ	Setiap orang yang mengemudikan mobil barang untuk mengangkut orang kecuali dengan alasan
12.	305 UULLAJ JO 162 UULLAJ	Setiap orang yang mengemudikan Kendaraan Bermotor yang mengangkut barang khusus yang tidak memenuhi ketentuan tentang persyaratan keselamatan, pemberian tanda barang, Parkir, bongkar dan muat, waktu operasi dan rekomendasi dari instansi terkait
13.	306 UULLAJ JO 162 UULLAJ	Setiap orang yang mengemudikan kendaraan angkutan barang yang tidak dilengkapi surat muatan dokumen perjalanan
14.	307 UULLAJ JO 169 (1) UULLAJ	Setiap orang yang mengemudikan Kendaraan Bermotor Angkutan Umum Barang yang tidak mematuhi ketentuan mengenai tata cara pemuatan, daya angkut, dimensi kendaraan
15.	308 UULLAJ JO 173 UULLAJ	Tidak memiliki izin menyelenggarakan angkutan orang dalam trayek

Pelanggaran lalu lintas yang paling sedikit terjadi yaitu pada pelanggaran lalu lintas pasal 288 (3) yaitu kendaraan yang tidak dilengkapi surat keterangan uji berkala dan

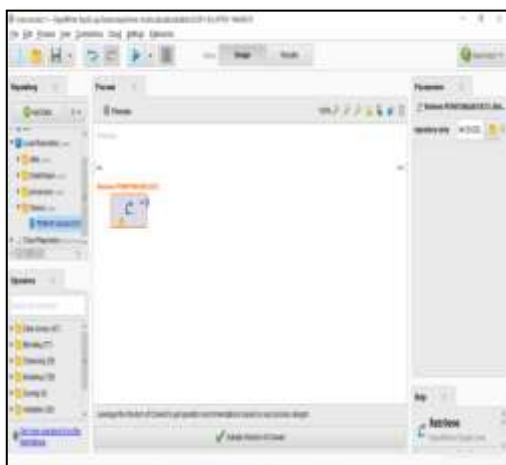
pasl 308 yaitu kendaraan yang tidak memiliki izin menyelenggarakan angkutan orang dalam trayek.

1.2 Hasil Implementasi Rapid Miner

Row No.	Pasal	Jumlah	Denda
1	279 UALLAJ	8	124
2	279 UALLAJ	9	40
3	280 UALLAJ	91	40
4	281 UALLAJ	205	40
5	285 (1) UALL	120	38
6	285 (2) UALL	9	40
7	286 UALLAJ	8	148
8	287 (1) UALL	147	40
9	287 (2) UALL	7	40
10	287 (4) UALL	11	40
11	287 (5) UALL	18	39
12	288 (1) UALL	110	40
13	288 (2) UALL	141	40
14	288 (3) UALL	2	74
15	289 UALLAJ	8	149

Gambar 3. Tampilan Data Excel Yang Sudah Diimpor

Simpan data excel yang sudah berhasil diimpor akan tampil seperti gambar diatas. Data yang tampil hanya data yang diblok pada saat select data. Setelah data sudah tampil maka kita kembali ke menu design untuk memasukkan operators.



Gambar 4. Memasukkan Data Excel Ke Design



Gambar 5. Tampilan Read Data Excel dan Metode K-Means

Row No.	H	Pasal	Cluster	Jumlah	Denda
1	1	279 UALLAJ	cluster_1	8	124
2	2	279 UALLAJ	cluster_1	9	40
3	3	280 UALLAJ	cluster_5	91	40
4	4	281 UALLAJ	cluster_5	205	40
5	5	285 (1) UALL	cluster_5	120	38
6	6	285 (2) UALL	cluster_1	9	40
7	7	286 UALLAJ	cluster_1	8	148
8	8	287 (1) UALL	cluster_5	147	40
9	9	287 (2) UALL	cluster_1	7	40
10	10	287 (4) UALL	cluster_1	11	40
11	11	287 (5) UALL	cluster_1	18	39
12	12	288 (1) UALL	cluster_5	110	40
13	13	288 (2) UALL	cluster_5	141	40
14	14	288 (3) UALL	cluster_1	2	74
15	15	289 UALLAJ	cluster_1	8	149

Gambar 6. Tampilan Hasil Clustering Pada Example Set

IV. KESIMPULAN

Dari hasil penelitian didapat kesimpulan sebagai berikut :

1. Pelanggaran lalu lintas paling banyak terjadi didominasi oleh jenis pelanggaran pasal 291 (1) yaitu pengendara sepeda motor yang tidak menggunakan helm dan pasal 291 (2) yaitu pengendara sepeda motor yang



membiarkan penumpangnya tidak menggunakan helm.

2. Hasil perhitungan menggunakan excel dan pengujian menggunakan *Rapidminer* menghasilkan hasil yang sama. Pada *cluster* 1 terdapat 9 pasal, artinya pelanggaran yang paling banyak terjadi ada dalam 9 pasal tersebut. Pada *cluster* 2 terdapat 15 pasal, artinya pelanggaran yang paling sedikit terjadi ada dalam 15 pasal tersebut. Pada aplikasi *Rapidminer* kita dapat melihat hasil *clustering* berupa plot grafik, graph view dan chart scatter. Dari hasil tersebut dapat dilihat secara jelas data-data yang masuk ke dalam *cluster* yang terdekat.

V. UCAPAN TERIMA KASIH

Kami Sebagai Peneliti Mengucapkan Terimakasih Kepada Kejaksaan Negeri Lampung Utara Yang Memberikan izin untuk Pengolahan data Pelanggaran Lalu Lintas, dan STMIK Dian Cipta Cendikia Kotabumi yang mendukung dalam penelitian ini.

VI. DAFTAR PUSTAKA

- Kurniawan, Titus, and Mohamad Jajuli. "Clustering Data Kecelakaan Lalu Lintas Di Kecamatan Cileungsi Menggunakan Metode K-Means." 6(1): 1–12.
- Pahlevi, Reza. 2021. "Jumlah Tilang Lalu Lintas Capai 1,77 Juta Hingga Oktober 2021 | Databoks." www.databoks.katadata.co.id.
- Pemerintah RI. 2009. "Undang - Undang Republik Indonesia Nomor 22 Tahun 2009 Tentang Lalu Lintas Dan Angkutan Jalan." 1: 12–42.
- Purba, Windania, Willy Siawin, and . Hardi. 2019. "Implementasi Data

Mining Untuk Pengelompokan Dan Prediksi Karyawan Yang Berpotensi Phk Dengan Algoritma K-Means Clustering." *Jurnal Sistem Informasi dan Ilmu Komputer Prima (JUSIKOM PRIMA)* 2(2): 85–90.

- Srisulistiowati, D B, M Khaerudin, and ... 2021. "Sistem Informasi Prediksi Penjualan Alat Tulis Kantor Dengan Metode Fp-Growth (Studi Kasus Toko Koperasi Sekolah Bina Mulia)." *JSI (Jurnal ...)*: 243–56.
- Sugiyarti, E. et al. 2018. "Decision Support System of Scholarship Grantee Selection Using Data Mining." *International Journal of Pure and Applied Mathematics* 119(15).