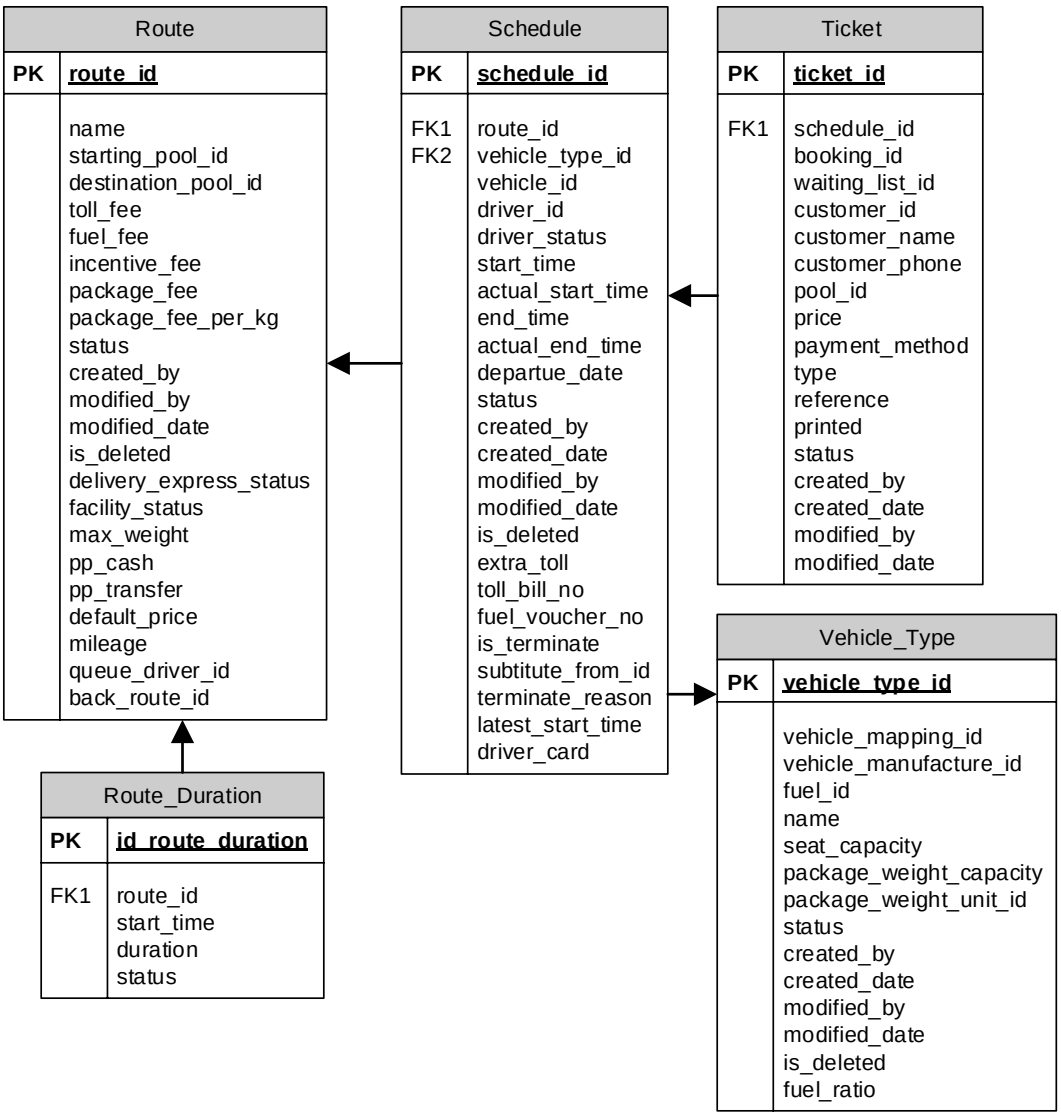
BAB 4
HASIL PENELITIAN DAN PEMBAHASAN

4.1 Analisis Masalah

Analisis masalah merupakan tahap menganalisis lalu menjabarkan permasalahan yang ada. Berdasarkan hasil pengamatan, diketahui bahwa penjadwalan kendaraan di PT. Citra Tiara Global dilakukan dengan cara melihat jumlah kendaraan yang pengemudi yang tersedia, pihak PT. Citra Tiara Global tidak melakukan penjadwalan dengan melihat pola konsumen dalam melakukan perjalanan dan hanya memperkirakan waktu tiba kendaraan di tujuan. Hal ini mengakibatkan terdapat beberapa waktu pemberangkatan yang tidak maksimal, terdapat pemberangkatan yang hanya membawa beberapa penumpang dan kursi yang tersedia tidak terisi penuh. Hal ini juga mengakibatkan pengeluaran pada setiap keberangkatan dapat lebih tinggi dibandingkan dengan pendapatannya.

4.2 Analisis Sumber Data

Sumber data yang didapat dalam penelitian ini merupakan data penjadwalan yang telah dilakukan di PT. Citra Tiara Global sejak bulan Januari 2016 hingga desember 2016. Terdapat beberapa tabel dari basis data yang akan digunakan dalam penelitian ini, skema relasi tabel yang terkait dapat dilihat pada Gambar 4.1



Gambar 4.1 Skema Relasi Basis Data

Penjelasan struktur tabel dari masing-masing tabel yang digunakan pada penellitian ini adalah sebagai berikut :

1. Tabel *Vehicle_Type*

Tabel *Vehicle_Type* merupakan tabel untuk menyimpan data kendaraan yang beroperasional di PT. Citra Tiara Global, *primary key* di tabel ini adalah

vehicle_type_id. Untuk lebih jelasnya struktur tabel *Vehicle_Type* dapat dilihat pada Tabel 4.1.

Tabel 4.1 *Vehicle_Type*

Nama Field	Tipe Data	Ukuran	Kunci	Keterangan
vehicle_type_id	integer	15	PK	Not Null
vehicle_mapping_id	integer	15		Not Null
vehicle_manufacture_id	integer	15		Not Null
fuel_id	integer	15		Not Null
name	varchar	50		Not Null
seat_capacity	Int	15		Not Null
package_weight_capacity	double			Not Null
package_weight_unit_id	integer	15		Not Null
status	varchar	20		Not Null
created_by	integer	15		Not Null
created_date	datetime			Not Null
modified_by	integer	15		Not Null
modified_date	datetime			Not Null
is_deleted	integer	15		Not Null
Fuel_ratio	double			Not Null

2. Tabel *Route*

Tabel *Route* merupakan tabel untuk menyimpan data rute tujuan pada PT. Citra Tiara Global, *primary key* di tabel ini adalah *route_id*. Untuk lebih jelasnya struktur tabel *route* dapat dilihat pada Tabel 4.2.

Tabel 4.2 *Route*

Nama Field	Tipe Data	Ukuran	Kunci	Keterangan
route_id	integer	5	PK	Not Null
Name	varchar	100		Not Null
starting_pool_id	varchar	5		Not Null
destination_pool_id	varchar	5		Not Null
toll_fee	double			Not Null
fuel_fee	double			Not Null
incentive_fee	double			Not Null
package_fee_per_kg	double			Not Null
Status	varchar	20		Not Null
created_by	integer	15		Not Null
created_date	datetime			Not Null
modified_by	integer	15		Not Null
modified_date	datetime			Not Null
is_deleted	integer	15		Not Null
delivery_express_status	varchar	50		Not Null
facility_status	varchar	50		Not Null

Nama Field	Tipe Data	Ukuran	Kunci	Keterangan
max_weight	double			Not Null
pp_cash	double			Not Null
pp_transfer	double			Not Null
default_price	double			Not Null
mileage	double			Not Null
queue_driver_id	varchar	20		Not Null
back_route_id	integer	15		Not Null

3. Tabel *Route_Duration*

Tabel *Route_Duration* merupakan tabel untuk menyimpan data perkiraan lama perjalanan ke tujuan pada PT. Citra Tiara Global, *primary key* di tabel ini adalah *id_route_duration.* Untuk lebih jelasnya struktur tabel *route_duration* dapat dilihat pada Tabel 4.3

Tabel 4.3 *Route_Duration*

Nama Field	Tipe Data	Ukuran	Kunci	Keterangan
id_route_duration	integer	5	PK	Not Null
route_id	integer	5	FK Reference Tabel *Route* (route_id)	Not Null
start_time	time			Not Null
duration	double			Not Null
Status	varchar	20		Not Null

4. Tabel *Schedule*

Tabel *schedule* merupakan tabel untuk menyimpan data penjadwalan pada PT. Citra Tiara Global, *primary key* di tabel ini adalah *schedule_id.* Untuk lebih jelasnya struktur tabel *schedule* dapat dilihat pada Tabel 4.4

Tabel 4.4 *Schedule*

Nama Field	Tipe Data	Ukuran	Kunci	Keterangan
schedule_id	varchar	20	PK	Not Null
route_id	integer	5	FK Reference Tabel *Route* (route_id)	Not Null
vehicle_type_id	integer	15	FK Reference Tabel *Vehicle_Type* (vehicle_type_id)	Not Null
driver_id	varchar	10		Not Null
driver_status	varchar	20		Not Null
start_time	time			Not Null
actual_start_time	time			Null
end_time	time			Null
actual_end_time	time			Null

Nama Field	Tipe Data	Ukuran	Kunci	Keterangan
departure_date	date			Null
Status	varchar	20		Null
created_by	integer	15		Not Null
created_date	datetime			Not Null
modified_by	integer	15		Not Null
modified_date	datetime			Null
is_deleted	integer	15		Null
extra_toll	integer	15		Null
toll_bill_no	varchar	20		Null
fuel_voucher_no	varchar	20		Null
is_terminate	varchar	10		Null
subtitute_from_id	varchar	11		Null
terminate_reason	varchar	100		Null
latest_start_time	varchar	20		Null
driver_card	varchar	50		Null

5. Tabel *Ticket*

Tabel *Ticket* merupakan tabel untuk menyimpan data pemesanan penumpang dan paket pada PT. Citra Tiara Global, *primary key* di tabel ini adalah *ticket_id.* Untuk lebih jelasnya struktur tabel *ticket* dapat dilihat pada Tabel 4.5

Tabel 4.5 *Ticket*

Nama Field	Tipe Data	Ukuran	Kunci	Keterangan
ticket_id	varchar	15	PK	Not Null
schedule_id	varchar	20	FK Reference Tabel *schedule* (schedule_id)	Not Null
booking_id	varchar	16		Not Null
waiting_list_id	varchar	10		Not Null
customer_id	varchar	16		Not Null
customer_name	varchar	50		Not Null
customer_phone	varchar	20		Not Null
pool_id	varchar	5		Not Null
Price	double			Not Null
payment_method	varchar	50		Not Null
type	varchar	50		Not Null
reference	varchar	50		Not Null
printed	integer	15		Not Null
status	varchar	20		Not Null
created_by	integer	15		Not Null
created_date	datetime			Not Null
modyfied_by	integer	15		Not Null
modified_date	datetime			Not Null

4.3 Analisis *Preprocessing Data*

Preprocessing data adalah hal yang harus dilakukan dalam proses *data mining*, karena tidak semua data atau atribut data dalam data digunakan dalam proses *data mining*. Proses ini dilakukan agar data yang akan digunakan sesuai dengan kebutuhan. Adapun tahapan-tahapan *preprocessing* dalam penelitian ini adalah sebagai berikut :

1. Ekstraksi Data

Ekstraksi data dapat diartikan sebagai proses pengambilan data dari sumber data dalam rangka untuk melanjutkan proses pengolahan data ke tingkat selanjutnya untuk menyimpan data hasil ekstrak tersebut. Dalam penelitian ini, data pada basis data di ekstrak, kemudian disimpan kedalam sebuah tabel terpisah agar memudahkan dalam proses pengolahan data.

2. Pemilihan Atribut

Pemilihan atribut adalah proses pemilihan atribut data yang akan digunakan, sehingga data tersebut dapat di proses sesuai dengan kebutuhan *data mining.* Dalam penelitian ini terdapat beberapa atribut yang akan digunakan berdasarkan tabel-tabel yang telah dijelaskan sebelumnya. Atribut yang akan digunakan pada penelitian ini dapat dilihat pada Tabel 4.6

Tabel 4.6 Pemilihan Atribut

Nama Atribut	Asal Tabel	Penjelasan
seat_capacity	vehicle_type	Merupakan atribut pembantu dalam proses pembersihan data. Tujuannya adalah untuk mengetahui jumlah kursi yang tersedia pada setiap kendaraan yang terjadwal.
starting_pool_id	route	Merupakan atribut untuk mengetahui asal keberangkatan. Atribut ini akan digunakan sebagai data latih karena asal keberangkatan berpengaruh terhadap jadwal keberangkatan
destination_pool_id	route	Merupakan atribut untuk mengetahui tujuan dari

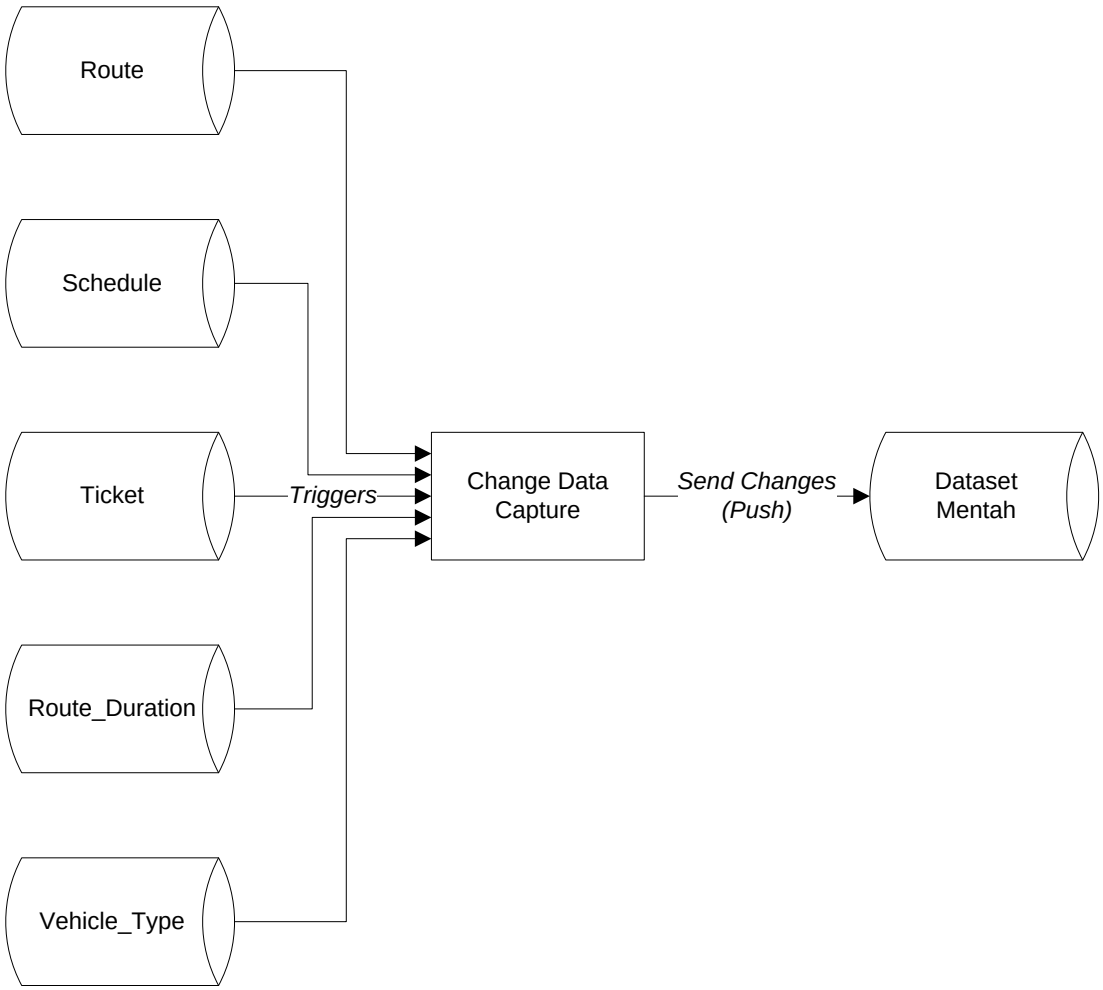
Nama Atribut	Asal Tabel	Penjelasan
		keberangkatan. Atribut ini akan digunakan sebagai data latih karena tujuan keberangkatan akan berpengaruh terhadap waktu tempuh untuk meprediksi lama perjalanan.
duration	route_duration	Merupakan atribut pembantu dalam proses pembersihan data. Tujuannya untuk memprediksi waktu tempu perjalanan.
start_time	schedule	Merupakan atribut untuk mengetahui jadwal keberangkatan yang telah ditentukan. Atribut ini akan dibandingkan dengan actual_start_time untuk mengetahui waktu selisih antara waktu jadwal ditentukan dengan waktu keberangkatan sesungguhnya.
actual_start_time	schedule	Merupakan atribut pembantu untuk mengetahui selisih antara waktu jadwal yang telah ditentukan dengan waktu keberangkatan sesungguhnya.
actual_end_time	schedule	Merupakan atribut pembantu untuk mengetahui selisih antara estimasi waktu tiba (*duration*) dengan waktu tiba ke tujuan yang sesungguhnya.
type	ticket	Merupakan atribut pembantu untuk mengetahui jumlah penumpang yang melakukan pemberangkatan pada setiap jadwal.
starting_pool_id	route	Merupakan atribut untuk mengetahui asal keberangkatan. Atribut ini akan digunakan sebagai data latih karena asal keberangkatan berpengaruh terhadap jadwal keberangkatan

Berdasarkan dari pemilihan atribut pada Tabel 4.6, maka *dataset* yang akan diambil dapat dilihat pada Tabel 4.7

Tabel 4.7 *Dataset* Mentah

Starting _pool_id	Destination_ pool_id	Start _time	Duration (Menit)	Actual_ start_time	Actual_ end_time	Seat_ capacity	Count (Type)
BN	DU	10:45:00	240	10:33:11	13:48:35	8	8
BN	DU	10:45:00	240	10:38:15	14:24:17	8	6
BN	DU	10:45:00	240	10:39:00	14:11:21	8	4
BN	DU	10:45:00	240	10:39:22	Null	8	2
BN	DU	10:45:00	240	10:39:26	Null	8	1
BN	DU	10:45:00	240	10:39:51	14:35:07	8	7
BN	DU	10:45:00	240	10:40:08	14:09:56	8	5
BN	DU	10:45:00	240	10:40:24	13:59:47	8	3
BN	DU	10:45:00	240	10:40:51	13:49:53	8	8
BN	DU	10:45:00	240	10:40:56	Null	8	8
BN	DU	10:45:00	240	10:40:58	13:48:41	8	8
BN	DU	10:45:00	240	10:41:15	14:15:30	8	8
BN	DU	10:45:00	240	10:41:32	Null	8	8
BN	DU	10:45:00	240	10:41:38	14:07:59	8	8
BN	DU	10:45:00	240	10:41:45	17:13:18	8	8
BN	DU	10:45:00	240	10:42:18	Null	8	7
BN	DU	10:45:00	240	10:42:29	Null	8	6
BN	DU	10:45:00	240	10:42:36	14:40:49	8	8
BN	DU	10:45:00	240	10:42:36	Null	8	7
BN	DU	10:45:00	240	10:42:38	14:20:47	8	8

Untuk mendapatkan data seperti pada Tabel 4.7 digunakan teknik *Change Data Capture*. Teknik yang digunakan adalah CDC *Push.* Teknik ini akan memasukkan data kedalam basis data apabila terjadi perubahan terhadap data yang diinginkan berdasarkan pemicu (*triggers)* yang diberikan. Pada penelitian ini pemicu yang diberikan adalah pada saat mencetak *manifest* atau surat jalan kendaraan dan pada saat kendaraan sampai di tujuan. Diagram alur proses CDC *push* pada penelitian ini dapat dilihat pada Gambar 4.2.



Gambar 4.2 Diagram Alir *Change Data Capture*

Penggunaan teknik CDC pada penelitian ini adalah untuk mengurangi waktu proses *query* untuk pengambilan dataset pada basis data operasional, karena data yang diperlukan berada pada beberapa tabel. Langkah yang dilakukan pada proses CDC ini adalah sebagai berikut :

a. Melakukan proses query terhadap tabel *route*, *schedule*, *ticket*, *route_duration*, *vehicle_type* dengan pemicu petugas menekan tombol cetak *manifest.*

b. Daftar atribut yang akan terisi saat petugas menekan tombol cetak *manifest* adalah *starting_pool_id*, *destination_pool_id*, *start_time*, *duration*,

actual_start_time, *seat_capacity*, dan jumlah penumpang, untuk atribut actual_end_time secara *default* akan diberikan nilai null.

c. Atribut actual_end_time akan terisi saat petugas menekan tombol sampai untuk jadwal keberangkatan yang kendaraannya telah sampai di tujuan.

3. Pembersihan Data

Pembersihan data merupakan proses menghilangkan *noise* dan menghilangkan data yang tidak relevan atau inkonsisten. Dalam penelitian ini, yaitu menghilangkan atribut yang bernilai null, menyamakan beberapa isi atribut untuk menyederhanakan keberagaman data. Tidak semua atribut dari *dataset* yang ada pada Tabel 4.7 akan digunakan. Atribut yang akan digunakan sebagai *dataset* dapat dilihat pada Tabel 4.8 dan contoh *dataset* yang telah dilakukan proses pembersihan data dapat dilihat pada Tabel 4.9.

Tabel 4.8 Atribut *Dataset*

Atribut	Penjelasan
starting_pool_id	Cabang / pool keberangkatan
destination_pool_id	Cabang / pool tujuan
penumpang	Bernilai angka sesuai dengan ketentuan yang telah dijelaskan pada BAB 2.angka tersebut berdasarkan persentase perhitungan antara atribut *seat_capacity* dengan *count(type).*
Jadwal	Atribut yang berisi jadwal keberangkatan yang telah ditentukan, dambil dari atribut *start_time*
delay_berangkat	Bernilai angka sesuai dengan ketentuan yang telah dijelaskan pada bab 2. Selisih didapat dari atribut *start_time* dengan *actual_start_time*.
delay_tiba	Bernilai angka sesuai dengan ketentuan yang telah dijelaskan pada bab 2. Selisih didapat dari atribut *start_time* ditambah dengan *duration* dengan *actual_end_time.* Atribut yang bernilai null dianggap terjadi delay di atas waktu yang telah ditentukan.

Tabel 4.9 *Dataset* hasil *preprocessing*

Starting _pool_id	Destination _pool_id	Penumpang	Jadwal	delay_ berangkat	delay_ tiba
DU	SM	3	6:00:00	0	4
DU	SM	2	8:00:00	1	2
DU	SM	3	9:00:00	1	4
DU	SM	3	10:00:00	1	4
DU	SM	1	11:00:00	1	2
DU	SM	2	12:00:00	1	2
DU	SM	3	13:00:00	1	1
DU	SM	2	14:00:00	1	0
DU	SM	2	15:00:00	2	0
DU	SM	1	16:00:00	2	0
DU	SM	2	5:00:00	0	4
DU	SM	3	6:00:00	1	0
DU	SM	2	11:00:00	1	2
DU	SM	1	12:00:00	1	0
DU	SM	3	13:00:00	1	2
DU	SM	2	16:00:00	1	0
DU	SM	2	17:00:00	2	0
DU	SM	3	18:00:00	1	0
DU	SM	3	20:00:00	1	0
DU	SM	1	21:00:00	0	0

4.4 Analisis Kasus dengan Metode Association Rule

Penelitian ini bertuuan untuk mendapatkan informasi mengenai prediksi penjadwalan keberangkatan travel di PT. Citra Tiara Global sesuai dengan data penjadwalan yang telah terjadi. Dalam menggali pola data yang digunakan sebagai *dataset* adalah data penjadwalan selama empat bulan kebelakang dengan menggunakan metode *data mining association rule,* dengan cara menemukan aturan asosiatif atau pola kombinasi antara asal keberangkata, tujuan keberangkatan, jadwal keberangkatan dan delay yang terjadi.

Algoritma yang digunakan pada penelitian ini adalah algoritma *apriori*. Algoritma ini menggnakan pengetahuan mengenai *frequent itemset* yang telah diketahui sebelumnya untuk memproses informasi selanjutnya. Adapun proses pengerjaan algoritma *apriori* adalah sebagai berikut :

1. Dari data hasil proses *preprocessing,* selanjutnya akan diproses untuk mencari data kandidat *itemset* (C_k).

2. Apabila kandidat *itemset* telah ditemukan, maka data kandidat *itemset* tersebut akan di simpan untuk dilakukan perhitungan nilai *support* dengan menggunakan persamaan 2.1 yang ada pada bab 2.
3. Setiap kandidat *itemset* yang memiliki nilai *support* lebih kecil dari nilai *minimum support* yang telah ditentukan, maka akan diabaikan dan tidak akan digunakan untuk proses selanjutnya. Nilai *minimum support* ditentukan sesuai dengan keinginan pengguna yang akan memproses data tersebut dengan rentang nilai antara 0-100%.
4. Kandidat *itemset* yang memilliki nilai *support* sama dengan atau lebih besar dari nilai *minimum support,* selanjutnya akan disimpan untuk dilakukan proses penggabungan untuk setiap *itemset* terpilih.
5. Selanjutnya dari hasil penggabungan *itemst,* hitung kembali nilai *support* yang dihasilkan dari setiap k-*itemset* tersebut dengan menggunakan persamaan 2.2 yang ada pada bab2.
6. Lakukan proses pemangkasan teradap *itemset* yang memiliki nilai *support* lebih kecil dari nilai *minimum support.*
7. Lakukan proses 4 sampai 6 secara berulang hingga tidak ada lagi *itemset* yang dapat dilakukan proses penggabungan ataupun pemangkasan, sehingga menghasilkan data *frequent itemset.*

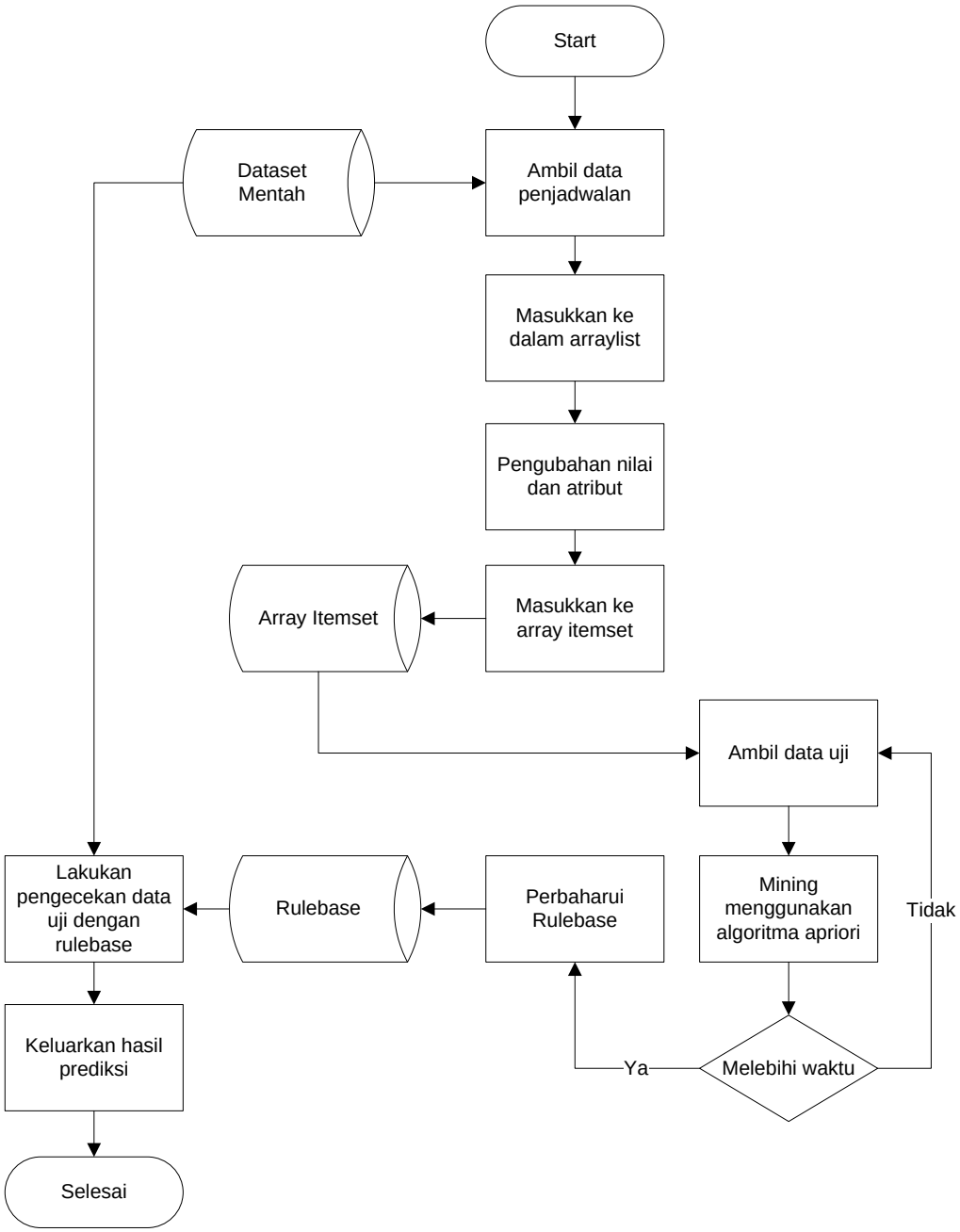
Setelah data *frequent itemset* ditemukan, proses selanjutnya adalah penentuan *association rule* yang terbentuk. Adapun proses pengerjaannya adalah sebagai berikut :

1. Dari setiap data *frequent itemset* yang telah didapat dari proses algoritma *apriori*, selanjutnya akan dilakukan proses perhitungan nilai *confidence* dengan menggunakan persamaan 2.3 yang ada pada bab 2 untuk menghasilkan data *association rule,* penentuan nilai *confidence* dapat ditentukkan oleh pengguna, dengan rentang nilai 0 sampai dengan 100%.
2. Apabila data *itemset* memiliki nilai *confidence* lebih kecil dari nilai *minimu confidence* yang ditentukkan pengguna, maka data *itemset* tersebut akan dipangkas. Sementara data *itemset* yang memenuhi standar nilai *confidence,*

akan dijadikan aturan untuk penentuan penjadwalan di PT. Citra Tiara Global

4.5 *Data Stream Mining*

Setelah dilakukan *preprocessing data*, tahap selanjutnya adalah *data stream mining*. Proses yang akan dilakukan sama dengan tahapan *preprocessing data* untuk mendapatkan rulebasenya, yang membedakannya adalah adanya *threshold* waktu yang tidak boleh dilewati. *Data stream mining* memiliki karakteristik pemprosesan data yang mengalir dengan cepat dan apabila tidak segera di proses akan menghilang, karena proses yang dilakukan berada di dalam memory tanpa menggunakan basis data, hal ini bermanfaat untuk memproses data dengan cepat. Diagram alir proses *data stream mining* dapat dilihat pada Gambar 4.3.



Gambar 4.3 Diagram Alir *Data Stream Mining*

Pada diagram alir yang ada pada Gambar 4.3 proses yang terjadi adalah sebagai berikut :

1. Proses pengambilan data dari basis data operasional yang telah melalui proses pemilihan atribut, dalam penelitian ini data yang diambil adalah data penjadwalan terbaru dengan rentang waktu empat bulan kebelakang.
2. Data yang telah diambil akan dimasukkan ke dalam *array* untuk dilakukan proses pengubahan sesuai dengan ketentuan yang telah ditetapkan oleh perusahaan perihal informasi yang dibutuhkan. Proses ini bertujuan untuk mengurangi keberagaman nilai yang ada pada basis data.
3. Data yang telah melalui proses perubahan terhadap nilai dan atribut akan dimasukkan kedalam basis data berbetuk array.
4. Proses selanjutnya adalah melakukan *data mining* menggunakan algoritma *apriori*, dalam proses ini diberikan batasan waktu proses dalam menemukan kombinasi item, hal ini bertujuan agar informasi yang dibutuhkan dapat diperoleh dengan cepat.
5. Apabila batas waktu yang telah ditentukkan telah terlewati, proses selanjutnya adalah memasukkan hasil dari mining tersebut kedalam basis data *rulebase*.
6. Langkah selanjutnya adalah melakukan pengujian terhadap data penjadwalan terbaru, dan akan dilakukan pengecekan dengan data yang ada pada *rulebase* untuk mendapatkan informasi prediksi.

4.6 Pengujian Model

Pengujian model dimaksudkan untuk mengetahui pemodelan yang telah dilakukan telah benar dalam langkah ataupun hasil yang diharapkan. Dalam pengujian ini digunakan contoh kasus menggunakan 100 data yang ada, data yang digunakan adalah data yang sudah melalui proses *preprocessing.* Pengujian dilakukan dengan membandingkan hasil perhitungan aplikasi yang dibuat dengan perhitungan manual menggunakan *excel* dan aplikasi *weka.* Langkah pengujiannya adalah sebagai berikut :

1. Perhitungan Manual

a. *Dataset* yang digunakan

Analisis data yang dilakukan pada contoh kasus ini adalah melihat aturan asosiasi berdasarkan waktu keberangkatan, keterlambatan keberangkatan, dan keberangkatan waktu tiba. Contoh kasus ini menggunakan 100 buah data yang sudah melalui tahap *preprocessing*. Adapun daftar transaksi jadwal yang dijadikan sebagai contoh kasus dapat dilihat pada Tabel 4.10

Tabel 4.10 Contoh *Dataset*

Asal	Tujuan	Jadwal Keberangkatan	Delay Berangkat	Delay Tiba
SM	TS	6:15:00	0	0
KN	TS	6:00:00	0	0
DU	BN	4:15:00	1	2
DU	SM	4:30:00	1	4
SM	DU	18:30:00	0	4
DU	SM	4:30:00	1	4
DU	FT	18:45:00	1	0
DU	FT	20:45:00	2	0
FT	DU	21:45:00	0	4
FT	DU	22:45:00	0	4
DU	FT	22:45:00	0	0
SM	DU	18:30:00	0	4
DU	SM	4:30:00	1	4
DU	SM	23:00:00	1	4
DU	SM	23:00:00	0	4
SM	DU	18:30:00	1	4
SM	DU	5:30:00	0	4
DU	SM	4:30:00	0	4
FT	DU	22:45:00	0	4
DU	FT	22:45:00	1	0
FT	DU	7:45:00	0	0
SM	DU	18:30:00	0	0
DU	SM	23:00:00	0	4
FT	DU	7:45:00	0	0

Berdasarkan *dataset* pada Tabel 4.10, diperoleh daftar item yang terdapat pada *dataset* tersebut yang dapat dilihat pada Tabel 4.11

Tabel 4.11 Daftar *Item*

Item	Type
SM	Starting_Pool (S)
KN	Starting_Pool (S)
DU	Starting_Pool (S)

Item	Type
FT	Starting_Pool (S)
TS	Starting_Pool (S)
SH	Starting_Pool (S)

Item	Type
TS	Destination_Pool (D)
BN	Destination_Pool (D)
SM	Destination_Pool (D)
DU	Destination_Pool (D)
FT	Destination_Pool (D)
KN	Destination_Pool (D)
6:15:00	Jadwal
6:00:00	Jadwal
4:15:00	Jadwal
4:30:00	Jadwal
18:30:00	Jadwal
18:45:00	Jadwal
20:45:00	Jadwal
21:45:00	Jadwal
22:45:00	Jadwal
23:00:00	Jadwal
5:30:00	Jadwal
7:45:00	Jadwal
19:45:00	Jadwal
11:45:00	Jadwal
15:45:00	Jadwal

Item	Type
8:15:00	Jadwal
13:15:00	Jadwal
18:15:00	Jadwal
8:00:00	Jadwal
13:00:00	Jadwal
18:00:00	Jadwal
11:15:00	Jadwal
16:15:00	Jadwal
21:15:00	Jadwal
11:00:00	Jadwal
16:00:00	Jadwal
21:00:00	Jadwal
4:45:00	Jadwal
0	Delay Start (DS)
1	Delay Start (DS)
2	Delay Start (DS)
0	Delay Destination (DD)
2	Delay Destination (DD)
4	Delay Destination (DD)
3	Delay Destination (DD)

b. Berdasarkan Tabel 4.11 terdapat 47 item unik, langkah selanjutnya adalah mencari *minimum support* untuk digunakan pada perhitungan selanjutnya. Pada kasus ini *minimum support* diambil berdasarkan rata-rata 1 *itemset* yaitu $\frac{\sum_1^n support\ count}{\sum_1^n nama\ item} = \frac{500}{47} = 10{,}63 \approx 11$ transaksi. Langkah selanjutnya adalah melakukan *scan* data transaksi yang ada pada Tabel 4.10 untuk mengetahui *support count* dan nilai *support* dari masing-masing item. Nilai *support* didapat menggunakan persamaan 2.1. berikut adalah hasil dari proses *scan* yang telah dilakukan.

Tabel 4.12 Kandidat *Itemset* 1 (C1)

Item	Type	*Support Count*	Nilai *Support*
SM	S	21	(21/100)*100 = 21%
KN	S	18	(18/100)*100 = 18%
DU	S	24	(24/100)*100 = 24%
FT	S	10	(1/100)*100 = 10%
TS	S	26	(26/100)*100 = 26%
SH	S	1	(1/100)*100 = 1%
TS	D	32	(32/100)*100 = 32%
BN	D	2	(2/100)*100 = 2%
SM	D	20	(20/100)*100 = 20%

Item	Type	*Support Count*	Nilai *Support*
DU	D	18	(18/100)*100 = 18%
FT	D	11	(11/100)*100 = 11%
KN	D	17	(17/100)*100 = 17%
6:15:00	Jadwal	2	(2/100)*100 = 2%
6:00:00	Jadwal	3	(3/100)*100 = 3%
4:15:00	Jadwal	2	(2/100)*100 = 2%
4:30:00	Jadwal	6	(6/100)*100 = 6%
18:30:00	Jadwal	5	(5/100)*100 = 5%
18:45:00	Jadwal	2	(2/100)*100 = 2%
20:45:00	Jadwal	3	(3/100)*100 = 3%
21:45:00	Jadwal	3	(3/100)*100 = 3%
22:45:00	Jadwal	6	(6/100)*100 = 6%
23:00:00	Jadwal	3	(3/100)*100 = 3%
5:30:00	Jadwal	2	(2/100)*100 = 2%
7:45:00	Jadwal	4	(4/100)*100 = 4%
19:45:00	Jadwal	1	(1/100)*100 = 1%
11:45:00	Jadwal	1	(1/100)*100 = 1%
15:45:00	Jadwal	1	(1/100)*100 = 1%
8:15:00	Jadwal	6	(6/100)*100 = 6%
13:15:00	Jadwal	7	(7/100)*100 = 7%
18:15:00	Jadwal	6	(6/100)*100 = 6%
8:00:00	Jadwal	8	(8/100)*100 = 8%
13:00:00	Jadwal	8	(8/100)*100 = 8%
18:00:00	Jadwal	7	(7/100)*100 = 7%
11:15:00	Jadwal	2	(2/100)*100 = 2%
16:15:00	Jadwal	2	(2/100)*100 = 2%
21:15:00	Jadwal	3	(3/100)*100 = 3%
11:00:00	Jadwal	2	(2/100)*100 = 2%
16:00:00	Jadwal	2	(2/100)*100 = 2%
21:00:00	Jadwal	2	(2/100)*100 = 2%
4:45:00	Jadwal	1	(2/100)*100 = 1%
0	DS	76	(76/100)*100 = 76%
1	DS	21	(21/100)*100 = 21%
2	DS	3	(3/100)*100 = 3%
0	DD	49	(49/100)*100 = 49%
2	DD	1	(1/100)*100 = 1%
4	DD	49	(49/100)*100 = 49%
3	DD	1	(1/100)*100 = 1%

c. Pada Tabel 4.12 dapat dilihat, terdapat *itemset* dengan *support count* lebih kecil dari *minimum support* yang telah ditentukan. Maka langkah selanjutnya adalah menghilangkan *itemset* yang memiliki *support count* lebih kecil dari *minimum support*, dapat dilihat pada Tabel 4.13.

Tabel 4.13 *Itemset* 1 (L1)

Item	Type	*Support Count*	Nilai *Support*
SM	S	21	(21/100)*100 = 21%
KN	S	18	(18/100)*100 = 18%
DU	S	24	(24/100)*100 = 24%
TS	S	26	(26/100)*100 = 26%
TS	D	32	(32/100)*100 = 32%
SM	D	20	(20/100)*100 = 20%
DU	D	18	(18/100)*100 = 18%
FT	D	11	(11/100)*100 = 11%
KN	D	17	(17/100)*100 = 17%
0	DS	76	(76/100)*100 = 76%
1	DS	21	(21/100)*100 = 21%
0	DD	49	(49/100)*100 = 49%
4	DD	49	(49/100)*100 = 49%

d. Setelah mendapatkan daftar *itemset* level 1, lakukan kombinasi 2 *itemset* untuk setiap item pada Tabel 4.13. Hitung kembali nilai *support* untuk masing-masing kombinasi yang terbentuk, sehingga dihasilkan seperti pada Tabel 4.14

Tabel 4.14 Kandidat *Itemset* 2 (C2)

Item	Type	Item	Type	*Support Count*	Nilai *Support*
SM	S	TS	D	14	(14/100)*100 = 14%
SM	S	DU	D	7	(7/100)*100 = 7%
SM	S	0	DS	18	(18/100)*100 = 18%
SM	S	1	DS	3	(3/100)*100 = 3%
SM	S	0	DD	12	(12/100)*100 = 12%
SM	S	4	DD	9	(9/100)*100 = 9%
KN	S	TS	D	17	(17/100)*100 = 17%
KN	S	DU	D	1	(1/100)*100 = 1%
KN	S	0	DS	17	(17/100)*100 = 17%
KN	S	0	DD	13	(13/100)*100 = 13%
KN	S	4	DD	5	(5/100)*100 = 5%
DU	S	SM	D	9	(9/100)*100 = 9%
DU	S	FT	D	11	(11/100)*100 = 11%
DU	S	KN	D	2	(2/100)*100 = 2%
DU	S	0	DS	8	(8/100)*100 = 8%
DU	S	1	DS	15	(15/100)*100 = 15%
DU	S	0	DD	13	(13/100)*100 = 13%
DU	S	4	DD	10	(10/100)*100 = 10%
TS	S	SM	D	11	(11/100)*100 = 11%
TS	S	KN	D	15	(15/100)*100 = 15%
TS	S	0	DS	22	(22/100)*100 = 22%
TS	S	1	DS	3	(3/100)*100 = 3%
TS	S	0	DD	7	(7/100)*100 = 7%

Item	Type	Item	Type	*Support Count*	Nilai *Support*
TS	S	4	DD	18	(18/100)*100 = 18%
TS	D	0	DS	29	(29/100)*100 = 29%
TS	D	1	DS	2	(2/100)*100 = 2%
TS	D	0	DD	23	(23/100)*100 = 23%
TS	D	4	DD	9	(9/100)*100 = 9%
SM	D	0	DS	14	(14/100)*100 = 14%
SM	D	1	DS	5	(5/100)*100 = 5%
SM	D	0	DD	5	(5/100)*100 = 5%
SM	D	4	DD	15	(15/100)*100 = 15%
DU	D	0	DS	17	(17/100)*100 = 17%
DU	D	1	DS	1	(1/100)*100 = 1%
DU	D	0	DD	6	(6/100)*100 = 6%
DU	D	4	DD	12	(12/100)*100 = 12%
FT	D	0	DS	3	(3/100)*100 = 3%
FT	D	1	DS	7	(7/100)*100 = 7%
FT	D	0	DD	11	(11/100)*100 = 11%
KN	D	0	DS	13	(13/100)*100 = 13%
KN	D	1	DS	4	(4/100)*100 = 4%
KN	D	0	DD	4	(4/100)*100 = 4%
KN	D	4	DD	12	(12/100)*100 = 12%
0	DS	0	DD	38	(38/100)*100 = 38%
1	DS	4	DD	37	(37/100)*100 = 37%
1	DS	0	DD	9	(9/100)*100 = 9%
2	DS	4	DD	11	(11/100)*100 = 11%

e. Setelah mendapatkan kombinasi dari 2 *itemset*, kemudian hilangkan kombinasi yang memiliki *nilai support* dibawah dari nilai *minimum support* yang telah ditentukan, sehingga diperoleh kombinasi 2 *itemset* yang sesuai pada Tabel 4.15.

Tabel 4.15 *Itemset* 2 (L2)

Item	Type	Item	Type	*Support Count*	Nilai *Support*
SM	S	TS	D	14	(14/100)*100 = 14%
SM	S	0	DS	18	(18/100)*100 = 18%
SM	S	0	DD	12	(12/100)*100 = 12%
KN	S	TS	D	17	(17/100)*100 = 17%
KN	S	0	DS	17	(17/100)*100 = 17%
KN	S	0	DD	13	(13/100)*100 = 13%
DU	S	FT	D	11	(11/100)*100 = 11%
DU	S	1	DS	15	(15/100)*100 = 15%
DU	S	0	DD	13	(13/100)*100 = 13%
DU	S	4	DD	10	(10/100)*100 = 10%
TS	S	SM	D	11	(11/100)*100 = 11%
TS	S	KN	D	15	(15/100)*100 = 15%
TS	S	0	DS	22	(22/100)*100 = 22%

Item	Type	Item	Type	*Support Count*	Nilai *Support*
TS	S	4	DD	18	(18/100)*100 = 18%
TS	D	0	DS	29	(29/100)*100 = 29%
TS	D	0	DD	23	(23/100)*100 = 23%
SM	D	0	DS	14	(14/100)*100 = 14%
SM	D	4	DD	15	(15/100)*100 = 15%
DU	D	0	DS	17	(17/100)*100 = 17%
DU	D	4	DD	12	(12/100)*100 = 12%
FT	D	0	DD	11	(11/100)*100 = 11%
KN	D	0	DS	13	(13/100)*100 = 13%
KN	D	4	DD	12	(12/100)*100 = 12%
0	DS	0	DD	38	(38/100)*100 = 38%
1	DS	4	DD	37	(37/100)*100 = 37%
2	DS	4	DD	11	(11/100)*100 = 11%
SM	S	TS	D	14	(14/100)*100 = 14%

f. Setelah mendapatkan kombinasi 2 *itemset*, lakukan langkah selanjutnya untuk melakukan kombinasi 3 *itemset* untuk setiap item yang ada padaTabel 4.15. Hitung kembali nilai *support* untuk masing-masing kombinasi yang terbentuk, sehingga dihasilkan seperti pada Tabel 4.16

Tabel 4.16 Kandidat *Itemset* 3 (C3)

Item	Type	Item	Type	Item	Type	*Support Count*	Nilai *Support*
SM	S	TS	D	0	DD	12	(12/100)*100 = 12%
SM	S	TS	D	0	DD	10	(10/100)*100 = 10%
SM	S	TS	D	1	DS	2	(2/100)*100 = 2%
SM	S	TS	D	4	DD	4	(4/100)*100 = 4%
SM	S	0	DS	0	DD	12	(12/100)*100 = 12%
SM	S	0	DS	4	DD	6	(6/100)*100 = 6%
SM	S	0	DS	DU	D	6	(6/100)*100 = 6%
SM	S	0	DD	DU	D	2	(2/100)*100 = 2%
SM	S	1	DS	4	DD	3	(3/100)*100 = 3%
SM	S	1	DS	DU	D	1	(1/100)*100 = 1%
SM	S	4	DD	DU	D	5	(5/100)*100 = 5%
TS	D	0	DS	0	DD	22	(22/100)*100 = 22%
TS	D	0	DS	KN	S	16	(16/100)*100 = 16%
TS	D	0	DS	4	DD	7	(7/100)*100 = 7%
TS	D	0	DD	KN	S	13	(13/100)*100 = 13%
TS	D	KN	S	4	DD	4	(4/100)*100 = 4%
TS	D	1	DS	4	DD	2	(2/100)*100 = 2%
0	DS	0	DD	KN	S	12	(12/100)*100 = 12%
0	DS	0	DD	DU	S	4	(4/100)*100 = 4%
0	DS	0	DD	FT	D	3	(3/100)*100 = 3%
0	DS	0	DD	TS	D	6	(6/100)*100 = 6%
0	DS	0	DD	SM	D	4	(4/100)*100 = 4%
0	DS	0	DD	KN	D	3	(3/100)*100 = 3%

Item	Type	Item	Type	Item	Type	*Support Count*	Nilai *Support*
0	DS	0	DD	DU	D	6	(6/100)*100 = 6%
0	DS	KN	S	4	DD	5	(5/100)*100 = 5%
0	DS	KN	S	DU	D	1	(1/100)*100 = 1%
0	DS	DU	S	FT	D	3	(3/100)*100 = 3%
0	DS	DU	S	SM	D	5	(5/100)*100 = 5%
0	DS	DU	S	4	DD	4	(4/100)*100 = 4%
0	DS	TS	S	SM	D	9	(9/100)*100 = 9%
0	DS	TS	S	KN	D	13	(13/100)*100 = 13%
0	DS	TS	S	4	DD	15	(15/100)*100 = 15%
0	DS	SM	D	4	DD	10	(10/100)*100 = 10%
0	DS	KN	D	4	DD	9	(9/100)*100 = 9%
0	DS	4	DD	DU	D	11	(11/100)*100 = 11%
0	DD	DU	S	FT	D	11	(11/100)*100 = 11%
0	DD	DU	S	1	DS	8	(8/100)*100 = 8%
0	DD	DU	S	SM	D	1	(1/100)*100 = 1%
0	DD	DU	S	KN	D	1	(1/100)*100 = 1%
0	DD	FT	D	1	DS	7	(7/100)*100 = 7%
0	DD	1	DS	TS	S	1	(1/100)*100 = 1%
0	DD	1	DS	SM	D	1	(1/100)*100 = 1%
0	DD	1	DS	KN	D	1	(1/100)*100 = 1%
0	DD	TS	S	SM	D	4	(4/100)*100 = 4%
0	DD	TS	S	KN	D	3	(3/100)*100 = 3%
KN	S	4	DD	DU	D	1	(1/100)*100 = 1%
DU	S	FT	D	1	DS	7	(7/100)*100 = 7%
DU	S	1	DS	SM	D	4	(4/100)*100 = 4%
DU	S	1	DS	KN	D	2	(2/100)*100 = 2%
SM	S	TS	D	0	DD	12	(12/100)*100 = 12%
SM	S	TS	D	0	DD	10	(10/100)*100 = 10%
SM	S	TS	D	1	DS	2	(2/100)*100 = 2%
SM	S	TS	D	4	DD	4	(4/100)*100 = 4%
SM	S	0	DS	0	DD	12	(12/100)*100 = 12%
SM	S	0	DS	4	DD	6	(6/100)*100 = 6%
SM	S	0	DS	DU	D	6	(6/100)*100 = 6%
SM	S	0	DD	DU	D	2	(2/100)*100 = 2%
SM	S	1	DS	4	DD	3	(3/100)*100 = 3%
SM	S	1	DS	DU	D	1	(1/100)*100 = 1%
SM	S	4	DD	DU	D	5	(5/100)*100 = 5%

g. Setelah mendapatkan kombinasi dari 3 *itemset*, kemudian hilangkan kombinasi yang memiliki *nilai support* dibawah dari nilai *minimum support* yang telah ditentukan, sehingga diperoleh kombinasi 3 *itemset* yang sesuai pada Tabel 4.17

Tabel 4.17 *Itemset* 3 (L3)

Item	Type	Item	Type	Item	Type	*Support Count*	Nilai *Support*
SM	S	TS	D	0	DD	12	(12/100)*100 = 12%
SM	S	TS	D	0	DD	10	(10/100)*100 = 10%
SM	S	TS	D	1	DS	2	(2/100)*100 = 2%
SM	S	TS	D	4	DD	4	(4/100)*100 = 4%
SM	S	0	DS	0	DD	12	(12/100)*100 = 12%
SM	S	0	DS	4	DD	6	(6/100)*100 = 6%
SM	S	0	DS	DU	D	6	(6/100)*100 = 6%
SM	S	0	DD	DU	D	2	(2/100)*100 = 2%
SM	S	1	DS	4	DD	3	(3/100)*100 = 3%
SM	S	1	DS	DU	D	1	(1/100)*100 = 1%
SM	S	4	DD	DU	D	5	(5/100)*100 = 5%

h. Setelah mendapatkan kombinasi 3 *itemset,* lakukan langkah selanjutnya untuk melakukan kombinasi 4 *itemset* untuk setiap item yang ada pada Tabel 4.17. Hitung kembali nilai *support* untuk masing-masing kombinasi yang terbentuk, sehingga dihasilkan seperti pada Tabel 4.18

Tabel 4.18 Kandidat *Itemset* 4 (C4)

Item	Type	Item	Type	Item	Type	Item	Type	*Support Count*	Nilai *Support*
SM	S	TS	D	0	DS	0	DD	10	(10/100)*100 = 10%
SM	S	TS	D	0	DS	4	DD	2	(2/100)*100 = 2%
SM	S	0	DD	0	DD	DU	D	2	(2/100)*100 = 2%
SM	S	0	DD	4	DD	DU	D	4	(4/100)*100 = 4%
TS	D	0	DD	0	DD	KN	S	12	(12/100)*100 = 12%
TS	D	0	DS	KN	S	4	DD	4	(4/100)*100 = 4%
0	DS	TS	S	TS	S	KN	D	3	(3/100)*100 = 3%
0	DS	DU	S	DU	S	FT	D	3	(3/100)*100 = 3%
0	DS	4	DS	4	DD	DU	D	1	(1/100)*100 = 1%
0	DS	KN	S	KN	D	4	DD	9	(9/100)*100 = 9%

i. Setelah mendapatkan kombinasi dari 4 *itemset*, kemudian hilangkan kombinasi yang memiliki *nilai support* dibawah dari nilai *minimum support*

yang telah ditentukan, sehingga diperoleh kombinasi 4 *itemset* yang sesuai pada Tabel 4.19

Tabel 4.19 *Itemset* 4 (L4)

Item	Type	Item	Type	Item	Type	Item	Type	*Support Count*	Nilai *Support*
TS	D	0	DD	0	DD	KN	S	12	(12/100)*100 = 12%

j. Berdasarkan hasil pada Tabel 4.19, tidak ada lagi kombinasi item yang dapat dibentuk sehingga proses pencarian *itemset* berhenti.

k. Setelah ditemukan *frequent itemset* yang sama atau lebih besar dari *minimum support* yang telah ditentukan, langkah selanjutnya adalah menghitung nilai *confidence* untuk kombinasi *itemset* pada Tabel 4.19 menggunakan persamaan 2.3. hasil perhitungan menggunakan pembulatan, dengan asumsi apabila dibelakang koma bernilai lebih besar atas sama dengan lima maka akan dibulatkan ke atas, sedangkan abapila angka dibelakang koma lebih kecil dari lima akan dibuatkan ke bawah. Kombinasi *itemset* tersebut dipilih, karena merupakan kombinasi yang lebih lengkap dibandingkan kombinasi item pada L_1 – L_3. Hasil perhitungan nilai *confidence* dapat dilihat pada Tabel 4.20

Tabel 4.20 Nilai *Confidence*

Item	Type	Item	Type	Item	Type	Support Count	Item	Type	Support Count	Nilai Confidence (%)
A		B		C		$A \cap B \cap C$	D		D	
TS	D	0	DS	0	DD	22	KN	S	18	55%
TS	D	0	DS	KN	S	16	0	DD	49	75%
TS	D	0	DD	KN	S	13	0	DS	76	92%
KN	S	0	DD	0	DS	12	TS	D	32	100%

l. Berdasarkan perhitungan pada Tabel 4.20 dapat dihasilkan pengetahuan sebagai berikut :

d. Apabila penumpang dengan tujuan TS dan tidak ada *delay* pada keberangkatan, dan *delay* pada kedatangan, maka penumpang tersebut berasal dari KN sebanyak 55%.

e. Apabila penumpang dengan tujuan TS dan tidak ada *delay* pada keberangkatan dan berasal dari KN, maka penumpang tersebut tidak akan mengalami *delay* kedatangan sebanyak 75%.

f. Apabila penumpang dengan tujuan TS dan tidak ada *delay* pada kedatangan dan berasal dari KN, maka penumpang tersebut tidak akan mengalami *delay* keberangkatan sebanyak 92%.

g. Apabila penumpang tersebut berasal dari KN dan tidak ada *delay* kedatangan, dan *delay* keberangkatan, maka penumpang tersebut akan menuju ke TS sebanyak 100%.

2. Perhitungan Aplikasi Weka

Aplikasi weka digunakan untuk membuktikan hasil dari perhitungan manual yang telah dilakukan. Langkah pembuktian yang dilakukan adalah sebagai berikut :

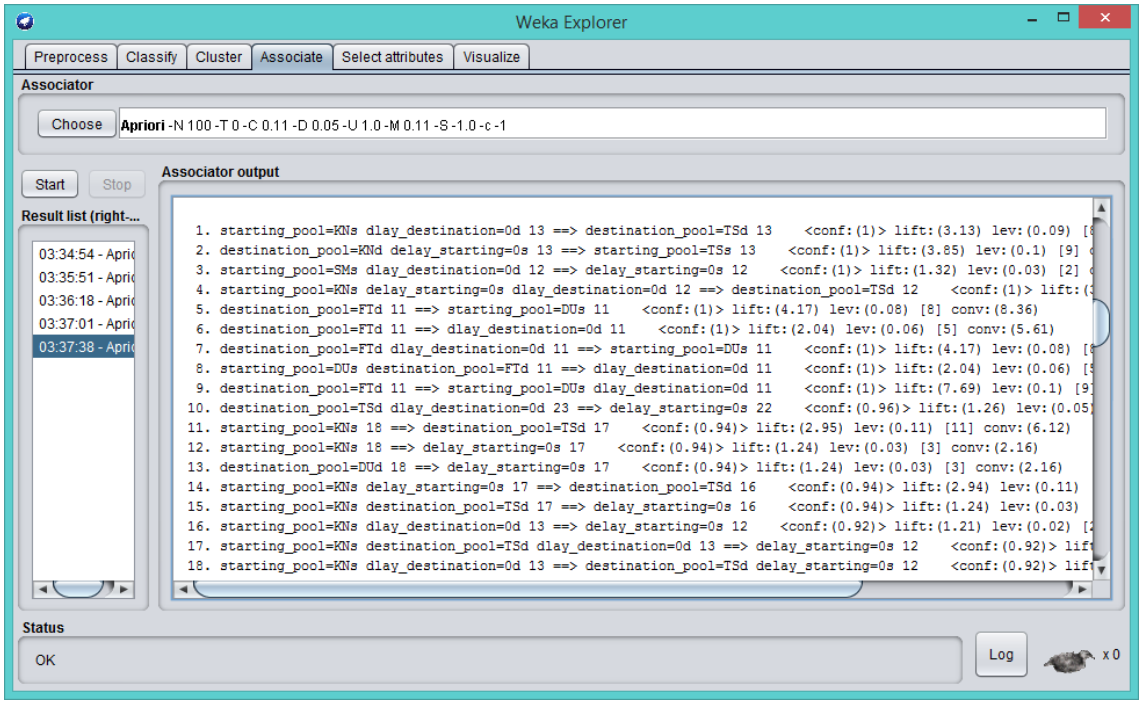
a. Mempersiapkan *dataset* sesuai dengan data contoh kasus. Adapun *dataset* yang digunakan dapat dilihat pada Tabel 4.21.

Tabel 4.21 Contoh *Dataset* Weka

No	***Dataset***
1	@relation jadwal
2	
3	@attribute starting_pool {SMs, KNs, DUs, FTs, TSs, SHs}
4	@attribute destination_pool {TSd, BNd, SMd, DUd, FTd, KNd}
5	@attribute jadwal {06:15:00, 06:00:00, 04:15:00, 04:30:00, 18:30:00, 18:45:00,
6	20:45:00, 21:45:00, 22:45:00, 23:00:00,05:30:00, 07:45:00, 19:45:00,
7	11:45:00,15:45:00, 08:15:00, 13:15:00, 18:15:00, 08:00:00, 13:00:00, 18:00:00,
8	11:15:00, 16:15:00, 21:15:00, 11:00:00, 16:00:00, 21:00:00, 04:45:00}
9	@attribute delay_starting {0s, 1s, 2s}
10	@attribute dlay_destination {0d, 2d, 3d, 4d}
11	
12	@data
13	SMs, TSd, 06:15:00, 0s, 0d
14	KNs, TSd, 06:00:00, 0s, 0d
15	DUs, BNd, 04:15:00, 1s, 2d
16	DUs, SMd, 04:30:00, 1s, 4d
17	SMs, DUd, 18:30:00, 0s, 4d
18	DUs, SMd, 04:30:00, 1s, 4d
19	DUs, FTd, 18:45:00, 1s, 0d

No	*Dataset*
20	DUs, FTd, 20:45:00, 2s, 0d
21	FTs, DUd, 21:45:00, 0s, 4d
22	FTs, DUd, 22:45:00, 0s, 4d
23	DUs, FTd, 22:45:00, 0s, 0d
24	SMs, DUd, 18:30:00, 0s, 4d
25	DUs, SMd, 04:30:00, 1s, 4d
26	DUs, SMd, 23:00:00, 1s, 4d
27	DUs, SMd, 23:00:00, 0s, 4d
28	SMs, DUd, 18:30:00, 1s, 4d
29	SMs, DUd, 05:30:00, 0s, 4d
30	DUs, SMd, 04:30:00, 0s, 4d
31	FTs, DUd, 22:45:00, 0s, 4d
32	DUs, FTd, 22:45:00, 1s, 0d
33	FTs, DUd, 07:45:00, 0s, 0d
34	SMs, DUd, 18:30:00, 0s, 0d
35	DUs, SMd, 23:00:00, 0s, 4d
35	FTs, DUd, 07:45:00, 0s, 0d
37	DUs, FTd, 18:45:00, 0s, 0d

b. Melakukan pengujian menggunakan aplikasi



Gambar 4.4 Hasil Perhitungan Aplikasi Weka

Hasil perhitungan pada **Error! Reference source not found.** sama dengan perhitungan manual, yang membedakan adalah aplikasi weka akan melakukan pencarian rule untuk setiap *itemset* yang terbentuk, mulai dari *itemset* L_1-L_4.

4.7 Pengujian Testing

Pengujian testing merupakan tahapan untuk melakukan serangkaian test terhadap sistem yang telah dibuat, dengan tujuan untuk mengetahui apakah sistem tersebut telah mencapai tujuan yang diinginkan ataukah belum. Dalam penelitian ini, metode pengujian yang dilakukan adalah dengan menggunakan pengujian *eksperimental*, pengujian ini dilakukan dengan cara membandingkan aturan asosiasi yang dihasilkan oleh sistem dengan kondisi sebenarnya. Adapun data yang akan digunakan pada pengujian adalah sebagai berikut :

1. Data latih yang akan digunakan adalah data periode bulan September 2016 hingga Desember 2016 dengan total *record* penjadwalan sebanyak 6.505 *record.*
2. Data uji yang akan digunakan untuk membandingkan hasil dari sistem dengan keadaan sebenarnya, yaitu data periode bulan Januari 2017 hingga Februari 2017, dengan total *record* penjadwalan sebanyak 2.618 *record.*
3. Perhitungan keakurasian estimasi keterlambatan dengan menghitung total data prediksi yang dihasilkan *data mining* dibagi dengan data pada keadaan nyata yang memiliki kesamaan dengan prediksi.
4. Perhitungan keakurasian jumlah penumpang untuk setiap jadwal keberangkatan dengan menghitung total data prediksi yang dihasilkan *data mining* dibagi dengan data pada keadaan nyata yang memiliki kesamaan dengan prediksi.
5. Perhitungan tidak ditemukannya prediksi untuk data nyata atau *data not found* dengan cara membagi data yang tidak ditemukan dengan jumlah data penjadwalan nyata yang terjadi, hal ini dilakukan untuk mendapatkan *minimum support* dan *minimum confidence* yang sesuai.

Skenario pengujian yang digunakan pada penelitian ini adalah dengan mengubah nilai *minimum support, minimum confidence,* dan maksimal waktu pemrosesan (*data stream mining)*. Ketentuan yang digunakan pada penelitian ini adalah sebagai berikut :

a. *Minimun support* yang digunakan adalah, rata-rata *itemset*, 1%, 3%, 5%, 10%, 15% dan 20% dari rata-rata *itemset.*
b. *Minimum confidence* yang digunakan adalah 0%, 10%, 20%, 30%, 40%, 50%, 60%, dan 70%.
c. Waktu pemrosesan yang digunakan dalam melakukan *data mining* adalah1 detik kemudian diikuti dengan kelipatan 5 detik hingga 1 menit. Untuk melihat batas rata-rata waktu yang dibutuhkan untuk mendapatkan informasi.

4.5.1 Pengujian Nilai Support Rata-rata

Pengujian ini melakukan perhitungan rata-rata *itemset* data latih. Berdasarkan hasil perhitungan rata-rata *itemset,* diperoleh nilai *minimum support* sebanyak 516 transaksi atau kemunculan. Setelah melakukan pengujian dengan aplikasi tidak ditemukan *rulebase*.

4.5.2 Pengujian Nilai Support 1%

Pengujian ini menghitung nilai 1% dari rata-rata *itemset* yang telah didapat sebelumnya yaitu 516. Pada perhitungan diperoleh *minimum support* sebanyak 5 transaksi atau setara dengan minimal 5 kemunculan untuk setiap kombinasi.

Setelah dilakukan pengujian menggunakan pembatasan waktu proses diketahui bahwa hasil untuk waktu proses 1 detik hingga 20 detik memiliki hasil yang sama untuk akurasi *delay* keberangkatan memiliki rata-rata nilai *confidence* 51,95% dengan akurasi sebesar 50,83% dan untuk akurasi penumpang memiliki rata-rata nilai *confidence* 20,08% dengan akurasi sebesar 35,99%. Persentase tidak adanya informasi untuk data uji sebesar 48,19%.

Pengujian untuk waktu proses 25 detik hingga 1 menit memiliki hasil yang sama untuk akurasi *delay* keberangkatan memiliki rata-rata nilai *confidence* 44,49% dengan akurasi sebesar 49,58% dan untuk akurasi penumpang memiliki rata-rata nilai *confidence* 26,96% dengan akurasi sebesar 35,28%.Persentase tidak adanya informasi untuk data uji sebesar 44,58%. Hasil pengujian untuk nilai *minimum support* 1% dapat dilihat pada Tabel 4.22.

Tabel 4.22 Pengujian Nilai Support 1%

Skenario	Waktu	Minimum Confidence	Akurasi Delay	Akurasi Penumpang	Data not found
1	1 Detik	0%	42,43%	25,69%	31,33%
		10%	42,43%	30,10%	31,33%
		20%	44,17%	35,63%	33,73%
		30%	46,76%	43,83%	38,55%
		40%	49,58%	0%	44,58%
		50%	50,24%	0%	46,99%
		60%	58,84%	0%	56,63%
		70%	78,64%	0%	65,06%
2	25 Detik	0%	36,15%	26,29%	31,33%
		10%	42,43%	27,81%	31,33%
		20%	44,17%	32,18%	33,73%
		30%	46,78%	36,56%	38,55%
		40%	49,58%	39,06%	44,58%
		50%	50,24%	46,01%	46,99%
		60%	58,84%	52,63%	56,63%
		70%	78,64%	0%	65,06%

4.5.3 Pengujian Nilai Support 3%

Pengujian ini menghitung nilai 3% dari rata-rata *itemset* yang telah didapat sebelumnya yaitu 516. Pada perhitungan diperoleh *minimum support* sebanyak 15 transaksi atau setara dengan minimal 15 kemunculan untuk setiap kombinasi.

Setelah dilakukan pengujian menggunakan pembatasan waktu proses diketahui bahwa hasil untuk waktu proses 1 detik hingga 20 detik memiliki hasil yang sama untuk akurasi *delay* keberangkatan memiliki rata-rata nilai *confidence* 52,64% dengan akurasi sebesar 52,82% dan untuk akurasi penumpang memiliki rata-rata nilai *confidence* 22,45% dengan akurasi sebesar 39,19%. Persentase tidak adanya informasi untuk data uji sebesar 53,01%.

Pengujian untuk waktu proses 25 detik hingga 1 menit memiliki hasil yang sama untuk akurasi *delay* keberangkatan memiliki rata-rata nilai *confidence* 46,65% dengan akurasi sebesar 50,87% dan untuk akurasi penumpang memiliki rata-rata nilai *confidence* 31,39% dengan akurasi sebesar 37,14%.Persentase tidak adanya informasi untuk data uji sebesar 38,55%. Hasil pengujian untuk nilai *minimum support* 3% dapat dilihat pada Tabel 4.23.

Tabel 4.23 Pengujian Nilai Support 3%

Skenario	Waktu	Minimum Confidence	Akurasi Delay	Akurasi Penumpang	Data not found
1	1 Detik	0%	43,12%	32,52%	33,73%
		10%	43,12%	35,52%	33,73%
		20%	43,92%	35,80%	34,94%
		30%	46,42%	43,83%	39,76%
		40%	49,39%	0%	46,99%
		50%	50,07%	0%	49,40%
		60%	58,58%	0%	59,04%
		70%	78,28%	0%	67,47%
2	25 Detik	0%	46,55%	29,85%	30,12%
		10%	46,55%	29,85%	30,12%
		20%	47,31%	32,19%	30,12%
		30%	49,25%	36,49%	30,12%
		40%	52,32%	38,95%	31,33%
		50%	51,94%	46,01%	42,17%
		60%	59,92%	52,63%	54,22%
		70%	78,34%	0%	65,06%

4.5.4 Pengujian Nilai Support 5%

Pengujian ini menghitung nilai 5% dari rata-rata *itemset* yang telah didapat sebelumnya yaitu 516. Pada perhitungan diperoleh *minimum support* sebanyak 26 transaksi atau setara dengan minimal 26 kemunculan untuk setiap kombinasi.

Setelah dilakukan pengujian menggunakan pembatasan waktu proses diketahui bahwa hasil untuk waktu proses 1 detik hingga 20 detik memiliki hasil yang sama untuk akurasi *delay* keberangkatan memiliki rata-rata nilai *confidence* 54,65% dengan akurasi sebesar 52,82% dan untuk akurasi penumpang memiliki rata-rata nilai *confidence* 30,30% dengan akurasi sebesar 42,52%. Persentase tidak adanya informasi untuk data uji sebesar 54,22%.

Pengujian untuk waktu proses 25 detik hingga 1 menit memiliki hasil yang sama untuk akurasi *delay* keberangkatan memiliki rata-rata nilai *confidence* 47,85% dengan akurasi sebesar 51,31% dan untuk akurasi penumpang memiliki rata-rata nilai *confidence* 30,30% dengan akurasi sebesar 42,52%.Persentase tidak adanya informasi untuk data uji sebesar 40,96%. Hasil pengujian untuk nilai *minimum support* 5% dapat dilihat pada Tabel 4.24.

Tabel 4.24 Pengujian Nilai Support 5%

Skenario	Waktu	Minimum Confidence	Akurasi Delay	Akurasi Penumpang	Data not found
1	1 Detik	0%	44,40%	37,41%	37,35%
		10%	44,40%	37,41%	37,35%
		20%	44,40%	37,41%	37,35%
		30%	46,42%	43,90%	40,96%
		40%	49,39%	0%	48,19%
		50%	50,07%	0%	50,6%
		60%	58,58%	0%	60,24%
		70%	78,28%	0%	68,67%
2	25 Detik	0%	47,39%	37,41%	31,33%
		10%	47,39%	37,41%	31,33%
		20%	47,39%	37,41%	31,33%
		30%	49,06%	43,90%	31,33%
		40%	52,09%	0%	32,53%
		50%	51,94%	0%	43,37%
		60%	59,92%	0%	55,42%
		70%	78,34%	0%	66,27%

4.5.5 Pengujian Nilai Support 10%

Pengujian ini menghitung nilai 10% dari rata-rata *itemset* yang telah didapat sebelumnya yaitu 516. Pada perhitungan diperoleh *minimum support* sebanyak 52 transaksi atau setara dengan minimal 52 kemunculan untuk setiap kombinasi.

Setelah dilakukan pengujian menggunakan pembatasan waktu proses diketahui bahwa hasil untuk waktu proses 1 detik hingga 20 detik memiliki hasil yang sama untuk akurasi *delay* keberangkatan memiliki rata-rata nilai *confidence* 61,74% dengan akurasi sebesar 61,31% dan untuk prediksi penumpang dengan

minimum support 10% tidak ditemukan kombinasi. Persentase tidak adanya informasi untuk data uji sebesar 66,27%.

Pengujian untuk waktu proses 25 detik hingga 1 menit memiliki hasil yang sama untuk akurasi *delay* keberangkatan memiliki rata-rata nilai *confidence* 59,60% dengan akurasi sebesar 60,45% dan untuk akurasi penumpang memiliki rata-rata nilai *confidence* 57,50% dengan akurasi sebesar 51,46%.Persentase tidak adanya informasi untuk data uji sebesar 60,24%. Hasil pengujian untuk nilai *minimum support* 10% dapat dilihat pada Tabel 4.25.

Tabel 4.25 Pengujian Nilai Support 10%

Skenario	Waktu	Minimum Confidence	Akurasi Delay	Akurasi Penumpang	Data not found
1	1 Detik	0%	49,73%	0%	54,22%
		10%	49,73%	0%	52,22%
		20%	49,73%	0%	52,22%
		30%	49,73%	0%	52,22%
		40%	49,73%	0%	52,22%
		50%	50,42%	0%	56,63%
		60%	58,58%	0%	65,06%
		70%	78,28%	0%	73,49%
2	25 Detik	0%	52,30%	47,82%	43,37%
		10%	52,30%	47,82%	43,37%
		20%	52,30%	47,82%	43,37%
		30%	52,30%	47,82%	43,37%
		40%	52,30%	47,82%	43,37%
		50%	52,40%	47,82%	49,4%
		60%	60,45%	52,63%	60,24%
		70%	78,34%	0%	69,88%

4.5.6 Pengujian Nilai Support 15%

Pengujian ini menghitung nilai 15% dari rata-rata *itemset* yang telah didapat sebelumnya yaitu 516. Pada perhitungan diperoleh *minimum support* sebanyak 77 transaksi atau setara dengan minimal 77 kemunculan untuk setiap kombinasi.

Setelah dilakukan pengujian menggunakan pembatasan waktu proses diketahui bahwa hasil untuk waktu proses 1 detik hingga 20 detik memiliki hasil yang sama untuk akurasi *delay* keberangkatan memiliki rata-rata nilai *confidence* 75,73% dengan akurasi sebesar 83,96% dan untuk prediksi penumpang dengan

minimum support 15% tidak ditemukan kombinasi. Persentase tidak adanya informasi untuk data uji sebesar 90,36%.

Pengujian untuk waktu proses 25 detik hingga 1 menit memiliki hasil yang sama untuk akurasi *delay* keberangkatan memiliki rata-rata nilai *confidence* 75,73% dengan akurasi sebesar 83,96% dan untuk prediksi penumpang dengan minimum support 15% tidak ditemukan kombinasi.Persentase tidak adanya informasi untuk data uji sebesar 90,36%. Hasil pengujian untuk nilai *minimum support* 5% dapat dilihat pada Tabel 4.26.

Tabel 4.26 Pengujian Nilai Support 15%

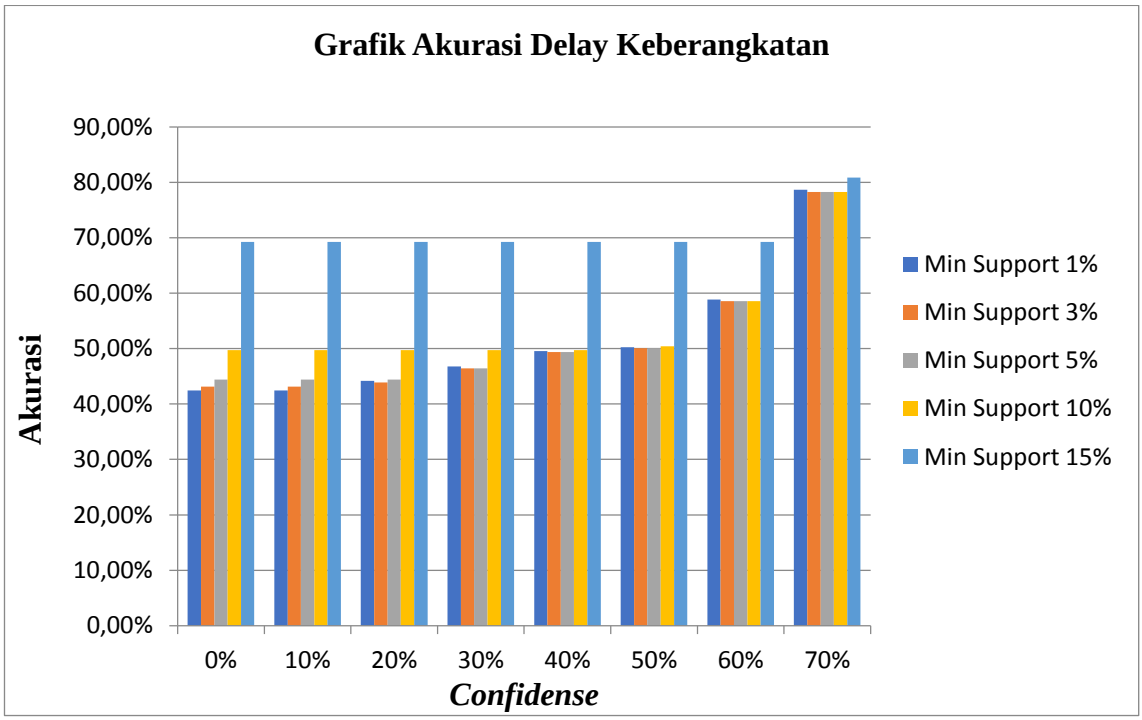
Skenario	Waktu	Minimum Confidence	Akurasi Delay	Akurasi Penumpang	Data not found
1	1 Detik	0%	69,25%	0%	86,75%
		10%	69,25%	0%	86,75%
		20%	69,25%	0%	86,75%
		30%	69,25%	0%	86,75%
		40%	69,25%	0%	86,75%
		50%	69,25%	0%	86,75%
		60%	69,25%	0%	86,75%
		70%	80,86%	0%	90,36%
2	25 Detik	0%	69,25%	0%	86,75%
		10%	69,25%	0%	86,75%
		20%	69,25%	0%	86,75%
		30%	69,25%	0%	86,75%
		40%	69,25%	0%	86,75%
		50%	69,25%	0%	86,75%
		60%	69,25%	0%	86,75%
		70%	80,86%	0%	90,36%

4.5.7 Pengujian Nilai Support 20%

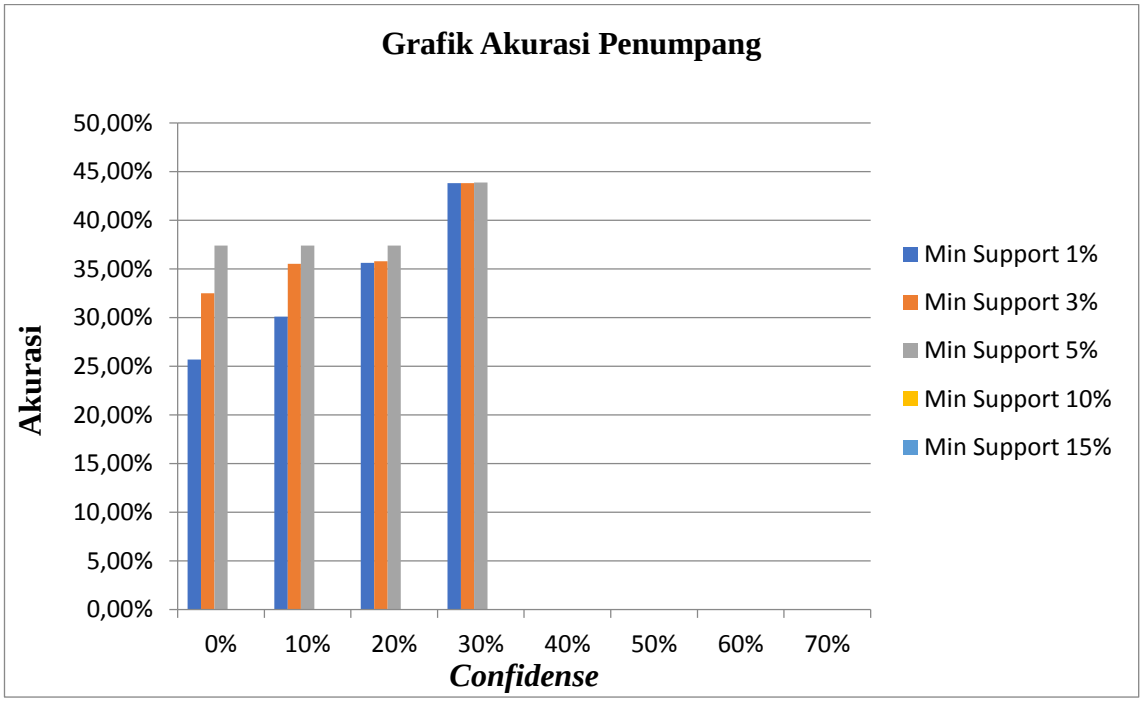
Pada pengujian dengan nilai *minimum support* 20% ditemukan hasil yang sama dengan pengujian dengan nilai *minimum support* 15%. Sehingga pengujian dihentikan dengan nilai *minimum support* 15% dari rata-rata *minimum support* yang telah didapat atau dengan minimal terdapat 77 kemunculan untuk setiap kombinasi yang terbentuk..

Berdasalkan hasil pengujian yang telah dilakukan, dapat dilihat rata-rata *rulebase* ditemukan untuk batasan waktu satu detik hanya dapat ditemukan

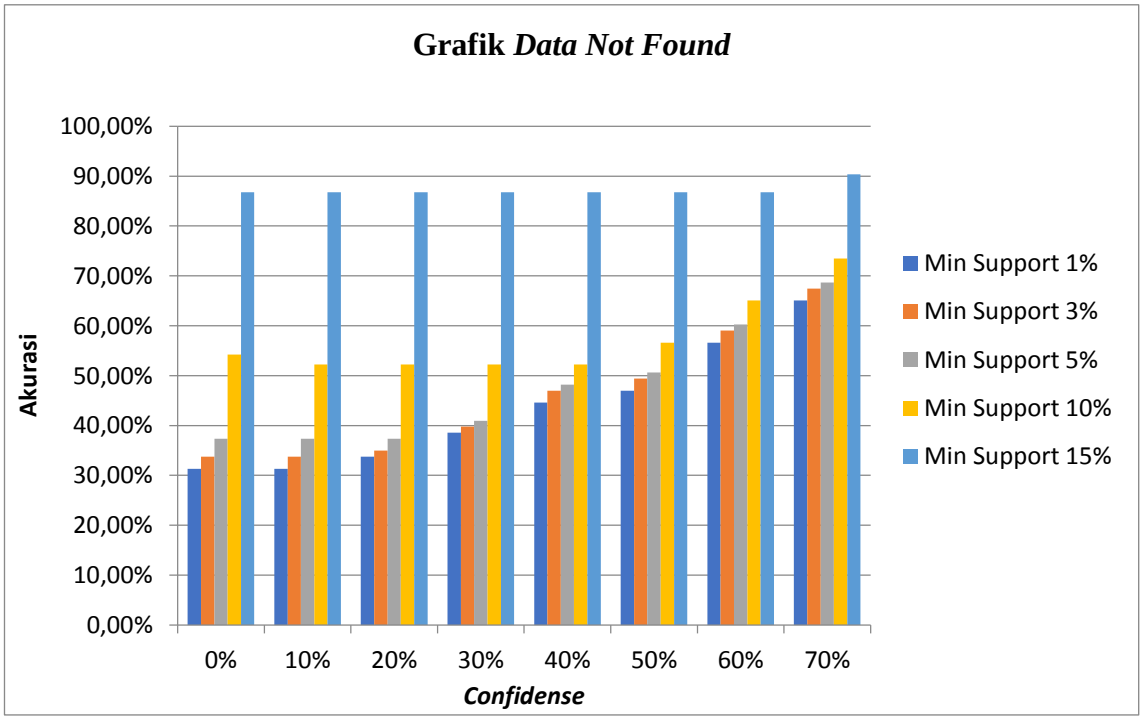
dengan *minimum confidence* yang ditentukkan maksimal 30%, khususnya untuk aturan keakuratan jumlah penumpang. Oleh sebab itu pada penelitian ini disarankan untuk menggunakan *minimum confidence* tidak lebih dari 30% dan menggunakan nilai *minimum support* 5%. Untuk grafik pengujian akurasi delay keberangkatan dengan batasan waktu satu detik dapat dilihat pada Gambar 4.5, untuk pengujian akurasi jumlah penumpang dapat dilihat pada Gambar 4.6 dan grafik pengujian tidak ditemukannya informasi dapat dilihat pada Gambar 4.7.



Gambar 4.5 Grafik Akurasi *Delay* Keberangkatan



Gambar 4.6 Grafik Akurasi Penumpang

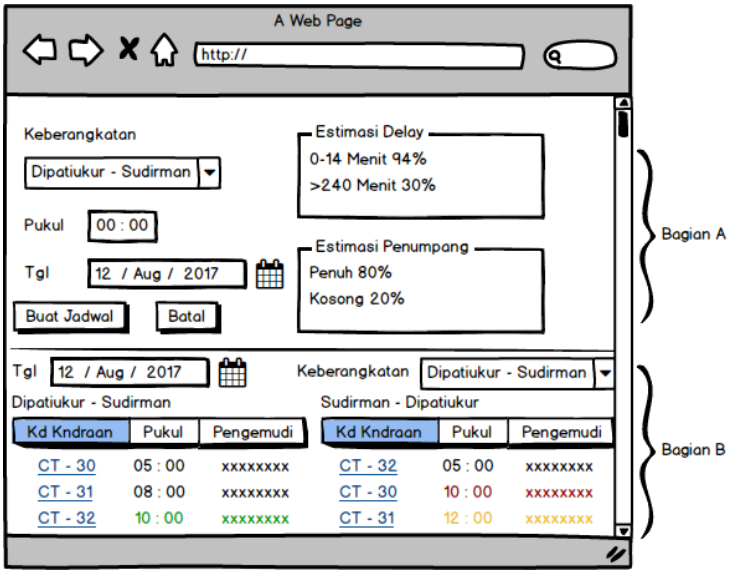


Gambar 4.7 Grafik *Data Not Found*

Berdasarkan grafik di atas dapat dilihat bahwa semakin tinggi nilai *minimum support* dan *minimum confidence* yang ditentukkan, maka tingkat *data not found* akan semakin tinggi, hal ini disebabkan karena berkurangnya ragam *itemset* dan juga aturan pada *rulebase.* Dengan semakin tingginya *minimum confidence* juga mengakibatkan tingkat kesamaan atau keakuratan untuk delay keberangkatan dan keakurasian prediksi penumpang semakin tinggi, karena hanya terpilih kombinasi dengan persentase tinggi yang hampir mendekati data nyata yang akan terjadi.

4.8 Prediksi *Delay*

Prediksi *delay* keberangkatan merupakan hasil analisa yang telah dilakukan menggunakan algoritma *apriori* berdasarkan *minimum support* dan *minimum support* yang telah ditentukkan sebelumnya. Adapun penempatan prediksi pada aplikasi dapat dilihat pada Gambar 4.8.



Gambar 4.8 Tampilan Prediksi

Terdapat dua bagian yang akan digunakan sebagai alat bantu dalam proses penjadwalan yang terdiri dari :

Bagian A

Pada bagian a, proses prediksi dilakukan sebelum dilakukannya proses penjadwalan. Pada bagian ini pengguna akan diberikan prediksi mengenai estimasi keterlambatan dan estimasi jumlah penumpang untuk setiap rute menggunakan algoritma *apriori*. Prediksi estimasi akan muncul dengan *trigger* pengguna memilih rute tujuan dan waktu keberangkatan. Pengguna akan di berikan peringatan saat menekan tombol buat jadwal apabila prediksi keterlambatan melebihi 50%. Dengan diberikannya peringatan ini akan membuat pengguna lebih awas saat melakukan penjadwalan terhadap kendaraan untuk menghindari keterlambatan pada keberangkatan selanjutnya.

Bagian B

Pada bagian b, proses prediksi memperlihatkan estimasi keterlambatan menggunakan warna dengan ketentuan :

1. Warna hijau apabila prediksi keterlambatan $\leq$ 30%
2. Warna kuning apabila prediksi keterlambatan $\geq$ 31% hingga $\leq$ 70%
3. Warna merah apabila prediksi keterlambatan $\geq$ 71%

Pemberian warna pada penjadwalan kendaraan memberikan informasi agar penjadwalan selanjutnya tidak mengalami keterlambatan dengan melakukan persiapan pergantian terhadap kendaraan yang memiliki kemungkinan terlambat besar.