

# Penerapan *Data Mining* Pada Penjualan Tiket Pesawat Menggunakan Algoritma Apriori (Studi Kasus: PT. Sinar Jagad Indonesia)

Larita Ayu Wahyuni S.<sup>1</sup>

**Abstrak**— Semakin banyaknya jumlah travel yang berkembang saat ini membuat para pengelola ingin menunjukkan strategi pemasaran yang lebih baik. Banyak cara di mana travel untuk meningkatkan penjualan, dan juga banyak strategi yang digunakan. Satu hal tentang strategi yang digunakan oleh manajemen adalah aplikasi perangkat lunak dengan sistem informasi. Sebuah era yang terus berkembang teknologi informasi setiap kali maka perlu adanya inovasi baru. Agar tidak terkesan monoton dalam proses menjual tiket pesawat dan diharapkan dapat memberikan manfaat bagi mereka yang menggunakan, karena dapat meningkatkan penjualan dengan memberikan rekomendasi kepada pembeli. Untuk itu dalam penulisan ini di kembangkan aplikasi analisis association untuk mengekstraksi dan menginterpretasi pola kecenderungan penjualan tiket pesawat yang sering dijual secara bersamaan dari data transaksi menggunakan algoritma apriori. Algoritma apriori ini akan membentuk *frequentitemset* banyak yang telah ditentukan sebelumnya berdasarkan dua parameter, *support* dan *confidence*, untuk menemukan aturan asosiasi antara suatu kombinasi item. Proses yang dilakukan diawali dengan persiapan data melalui preprocessing data kemudian ditransformasi kedalam bentuk yang dapat diolah pada proses selanjutnya yaitu join dan purne hingga pembentukan *association rules*. Dengan algoritma ini didapatkan kumpulan tiket pesawat yang sering terjual yaitu Garuda, Lion dan Trigana. Asosiasi ini didapat dengan nilai *support* minimal 50% dan nilai *confidence* 70%.

**Kata Kunci :** *Data Mining*, Algoritma Apriori, *Association Rules*, Penjualan Tiket Pesawat.

**Abstract**—The increasing number of travel that is developing now makes managers want to show a better marketing strategy. There are many ways in which travel increases sales, and also many strategies are used. One thing about the strategy used by management is software applications with information systems. An era that continues to develop information technology every time it is necessary to have new innovations. So as not to seem monotonous in the process of selling plane tickets and is expected to provide benefits for those who use, because it can increase sales by providing recommendations to buyers. For this reason, in the writing of this application, association analysis was developed to extract and interpret patterns of tendency of airline ticket sales which are often sold simultaneously from transaction data using a priori algorithm. This a priori algorithm will form as many pre-determined frequent itemset based on two parameters, support and confidence, to find the association rules between a combination of items. The process begins with the preparation of data through preprocessing data and then transformed into a form that

can be processed in the next process that is join and retire until the formation of association rules. With this algorithm, we get a collection of airline tickets that are often sold, namely Garuda, Lion and Trigana. This association is obtained with a support value of at least 50% and a confidence value of 70%.

**Keywords:** *Data Mining*, *Apriori Algorithms*, *Association Rules*, *Airplane Ticket Sales*.

## I. PENDAHULUAN

### A. Latar Belakang

---

<sup>1</sup> Mahasiswa Jurusan Sistem Informasi Fakultas Ilmu Komputer Universitas Darwan Ali, Jl. Batu Berlian Kab. Kotawaringin Timur 74322 INDONESIA (tlp:053-121 5227; fax: 053-121 5227; e-mail: info@unda.ac.id)

PT. Sinar Jagad Indonesia merupakan bergerak dalam bidang travel yang didirikan pada tahun 2012 yang beralamat di JL.Kapt. Mulyono Perum Setia Griya Minimalize Blok.B No.13. Namun sejatinya Semakin banyak jumlah travel di Indonesia yang berkembang saat ini khususnya daerah Jakarta, para pengelola PT Sinar Jagad Indonesia ingin menunjukkan strategi pemasaran yang lebih baik, Maka dari itu para pengelola harus mencermati pola-pola pembelian tiket pesawat yang di pesan oleh konsumen. Seperti pola penjualan tiket sehari-hari di PT. Sinar Jagad Indonesia yang penulis amati ketika sedang melihat data tiket pesawat disana, Selain itu Penulis menemukan banyak kekurangan yang terjadi pada PT. Sinar Jagad Indonesia tersebut diantaranya data semua penjualan tercampur menjadi satu dengan data peminjaman mobil, tiket kereta api, penginapan hotel, Hal ini tentu saja sangat disayangkan disebabkan karena data-data tersebut bisa diolah untuk jadi bahan promosi bagi perusahaan, Dengan adanya kegiatan penjualan setiap hari data semakin lama akan semakin bertambah banyak, Data tersebut tidak hanya berfungsi sebagai arsip bagi perusahaan, data tersebut dapat dimanfaatkan dan diolah menjadi informasi yang berguna untuk peningkatan penjualan tiket pesawat dan promosi, Hal ini disebabkan oleh karena PT. Jagad Indonesia tidak memanfaatkan data transaksi penjualan yang ada dan biasanya data transaksi penjualan tersebut hanya sebagai arsip. Sehingga terjadi penumpukan data yang tidak diketahui apa manfaatnya.,

diperlukan algoritma apriori. Algoritma apriori pada saat ini telah diimplementasikan ke berbagai bidang, salah satunya adalah dibidang bisnis atau perdagangan dan bidang pendidikan, dibidang bisnis misalnya implementasi *datamining* algoritma apriori untuk sistem penjualan tujuannya untuk membantu para pebisnis meningkatkan penjualan produk. Sedangkan dibidang pendidikan misalnya implementasi *datamining* untuk menemukan pola hubungan tingkat kelulusan mahasiswa dengan data induk mahasiswa. Adanya kegiatan operasional sehari-hari data semakin lama akan semakin bertambah banyak. Jika dibiarkan saja, maka data-data transaksi tersebut hanya menjadi sampah yang tidak berarti. Dengan adanya dukungan perkembangan teknologi, semakin berkembang pula kemampuan dalam mengumpulkan dan mengolah data. Ketatnya persaingan pada sektor penjualan, membuat pelaku usaha harus pintar dalam menganalisis pasar. Selain itu, ketersediaan barang yang dijual pun menjadi salah satu faktor yang harus dianalisis supaya kebutuhan pelanggan dapat terpenuhi. Untuk mendukung hal itu, data penjualan sebelumnya dapat dimanfaatkan untuk menganalisis pasar dan kebutuhan pelanggan [1].

Teknologi informasi yang semakin lama semakin maju dan setiap saat selalu ada perkembangan sangatlah mempunyai peranan penting dalam segala aspek kehidupan, salah satu aspek yang tidak bisa lepas dari teknologi informasi adalah aspek perekonomian terutama dalam sistem penjualan produk. Perlu sebuah kreativitas dan inovasi dari produsen agar penjualan produknya bisa ditingkatkan, apalagi melihat masyarakat sekarang yang mempunyai tingkat konsumtif tinggi terhadap barang-barang baru. Ada bermacam-macam cara untuk mensiasati agar produk yang kita jual bisa meningkat dan diminati para konsumen. Adanya tuntutan seperti di atas maka memunculkan ide-ide baru dalam dunia teknologi informasi, dengan cara membuat aplikasi yang sekiranya bisa membantu para produsen untuk meningkatkan penjualan produk. Salah satu caranya adalah memanfaatkan teknik *datamining* dalam hal ini menggunakan algoritma apriori (asosiasi *datamining*) [2].

Semakin banyak jumlah travel yang berkembang saat ini membuat para pengelola ingin menunjukkan strategi pemasaran yang lebih baik. Semakin banyak jumlah travel yang berkembang saat ini membuat para pengelola ingin menunjukkan strategi pemasaran yang lebih baik. Untuk itu maka para pengelola harus mencermati pola pola pembelian yang dilakukan oleh konsumen. Seperti pola penjualan tiket pesawat sehari-hari. Untuk itu maka para pengelola harus mencermati pola-pola pembelian yang dilakukan oleh konsumen. Seperti pola penjualan tiket pesawat sehari-hari di PT. Sinar Jagad Indonesia yang penulis amati ketika sedang memesan tiket pesawat di sana. Penulis menemukan banyak kekurangan yang terjadi pada PT. Sinar Jagad Indonesia tersebut di antaranya permasalahan yang sering timbul antara lain, sering sekali pemesanan tiket pesawat yang diinginkan konsumen tidak ada atau habis karena mereka tidak mengamati transaksi yang ada. Hal ini tentu sangat mengecewakan konsumen

yang hendak memesan tiket pesawat, karena persediaan tiket pesawat di PT. Sinar Jagad Indonesia tidak terkontrol baik.

Dengan adanya kegiatan penjualan setiap hari, data semakin lama akan semakin bertambah banyak. Data tersebut tidak hanya berfungsi sebagai arsip bagi perusahaan, data tersebut dapat dimanfaatkan dan diolah menjadi informasi yang berguna untuk peningkatan penjualan tiket pesawat dan promosi. Hal ini disebabkan oleh karena PT. Sinar Jagad Indonesia tidak memanfaatkan data transaksi penjualan yang ada dan biasanya data transaksi penjualan tersebut hanya sebagai arsip. Sehingga terjadi penumpukan data yang tidak diketahui apa manfaatnya. Pada dasarnya kumpulan data tersebut memiliki informasi-informasi yang bermanfaat, yang bisa digunakan untuk mengambil suatu keputusan dan untuk memperoleh pengetahuan yang baru (*knowledge*) tentang pola penjualan tiket pesawat. Oleh karena itu penulis mencoba untuk menyelesaikan permasalahan yang ada di PT. Sinar Jagad Indonesia dengan salah satu teknik yang digunakan dalam pengolahan data tersebut dengan menggunakan Algoritma Apriori. Algoritma apriori termasuk jenis aturan asosiasi pada data mining, Algoritma Apriori yang bertujuan untuk menemukan frequent item sets dijalankan pada sekumpulan data. Analisis Apriori didefinisikan suatu proses untuk menemukan semua aturan apriori yang memenuhi syarat minimum untuk support dan syarat minimum untuk confidence

## B. Perumusan Masalah

Berdasarkan latar belakang yang telah diuraikan di atas, maka dapat dirumuskan sebuah permasalahan yang dapat dijadikan sebagai acuan yaitu :

1. Bagaimana menerapkan algoritma apriori pada penjualan tiket pesawat ?
2. Bagaimana memperoleh presentasi penjualan tiket pesawat yang paling banyak terjual ?

## C. Tujuan dan Manfaat Penelitian

Adapun yang menjadi tujuan dalam penelitian ini adalah untuk mengetahui sejauh mana algoritma apriori dapat membantu pengembangan strategi pemasaran dalam penjualan tiket pesawat pada PT. Sinar Jagad Indonesia.

Adapun manfaat dari penelitian ini adalah membantu perusahaan untuk mengetahui penjualan tiket pesawat yang paling banyak terjual.

## D. Batasan Masalah

Batasan masalah dari penulisan ini adalah sebagai berikut :

1. Tiket yang dijual di PT. Sinar Jagad Indonesia di antaranya tiket pesawat Garuda, tiket pesawat Loin Air, tiket pesawat Garuda, tiket pesawat Trigana Air, tiket pesawat Citilink dan tiket pesawat Lion.

## II. LANDASAN TEORI

### A. Data Mining

*Datamining* merupakan suatu proses pendukung pengambil keputusan di mana kita mencari pola informasi dalam data. Pencarian ini dapat dilakukan oleh pengguna, misalnya dengan menggunakan *query* atau dapat dibantu dengan suatu aplikasi yang secara otomatis mencari pola informasi pada basis data. Pencarian ini disebut *discovery* [3]. *Data mining* merupakan bidang dari beberapa bidang keilmuan yang menyatukan teknik dari pembelajaran mesin, pengolahan pola, statistik, database, dan visualisasi untuk penanganan permasalahan pengambilan informasi dari database yang besar. Hubungan yang dicari dalam data mining dapat berupa hubungan antara dua atau lebih dalam satu dimensi. Misalnya dalam dimensi produk dapat melihat keterkaitan pembelian suatu produk dengan produk yang lain. Selain itu, hubungan juga dapat dilihat antara dua atau lebih atribut dan dua atau lebih objek. Algoritma *apriori* termasuk jenis aturan asosiasi pada datamining. Algoritma *Apriori* yang bertujuan untuk menemukan frequent item sets dijalankan pada sekumpulan data. Data mining bukanlah suatu bidang yang sama sekali baru. Perkembangan data mining (DM) yang pesat tidak dapat lepas dari perkembangan teknologi informasi yang memungkinkan data dalam jumlah besar terakumulasi. Sebagai contoh, toko swalayan merekam setiap penjualan barang dengan memakai alat POS (point of sales). Database data penjualan tsb. bisa mencapai beberapa GB setiap harinya untuk sebuah jaringan toko swalayan berskala nasional. Perkembangan internet juga punya andil cukup besar dalam akumulasi data. Tetapi pertumbuhan yang pesat dari akumulasi data itu telah menciptakan kondisi yang sering disebut sebagai “rich of data but poor of information” karena data yang terkumpul itu tidak dapat digunakan untuk aplikasi yang berguna. Tidak jarang kumpulan data itu dibiarkan begitu saja seakan-akan “kuburan data” (data tombs). Salah satu kesulitan untuk mendefinisikan data mining adalah kenyataan bahwa data mining mewariskan banyak aspek dan teknik bidang-bidang ilmu yang sudah mapan terlebih duhulu, *Data mining*, sering juga disebut *knowledge discovery in database*(KDD), adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar. Keluaran dari data mining bisa dipakai untuk memperbaikin pengambilan keputusan dimasa depan. *datamining* adalah kegiatan menemukan pola yang menarik dari data dalam jumlah besar, data dapat disimpan dalam database, data warehouse, atau penyimpanan informasi lainnya. Data mining berkaitan dengan bidang ilmu-ilmu lain, seperti database system, data warehousing, statistik, machine learning, information retrieval, dan komputasi tingkat tinggi[4]. Selain itu, data mining didukung oleh ilmu lain seperti neural network, pengenalan pola, spatial data analysis, image database, signal processing. Data mining didefinisikan sebagai

proses menemukan pola-pola dalam data. Definisi publik dari data mining merupakan metode pencarian pola-pola pengetahuan yang tersembunyi yang tidak diketahui sebelumnya dari suatu sekumpulan data yang sangat besar dalam database, data warehouse, atau media penyimpanan lainnya. Data mining digunakan untuk menggali nilai tambah berupa informasi yang tidak diketahui secara manual dari suatu basis data. Informasi diperoleh dengan cara mengekstrasi dan mengenali pola yang penting atau menarik dari data yang terdapat dalam basis data. Proses ini otomatis atau seringnya semiotomatis. Pola yang ditemukan harus penuh arti dan pola tersebut memberikan keuntungan, biasanya keuntungan secara ekonomi. Data yang dibutuhkan dalam jumlah besar Tentunya tidak terlepas dari pembahasan Data Mining erat kaitan dengan klasifikasi. Klasifikasi bertujuan untuk mengklasifikasikan item data menjadi satu dari beberapa kelas standar. Sebagai contoh, suatu program email dapat mengklasifikasikan email yang sah dengan email spam. Beberapa algoritma klasifikasi antara lain pohon keputusan, nearest neighbor, nave bayes, neural networks dan support vector machines, Hasil dari data mining sering kali diintegrasikan dengan decision support system (DSS). Sebagai contoh, dalam aplikasi bisnis informasi yang dihasilkan oleh data mining dapat diintegrasikan dengan tool manajemen kampanye produk sehingga promosi pemasaran yang efektif yang dilaksanakan dan dapat diuji. Integrasi demikian memerlukan langkah postprocessing yang menjamin bahwa hanya hasil yang valid dan berguna yang akan digabungkan dengan DSS. Salah satu pekerjaan dan postprocessing adalah visualisasi yang memungkinkan analyst untuk mengeksplor data dan hasil data mining, dari berbagai sudut pandang. Ukuran-ukuran statistik dan metode pengujian hipotesis dapat digunakan selama postprocessing untuk membuang hasil data mining yang palsu. Secara khusus, data mining menggunakan ide-ide seperti (1) pengambilan contoh, estimasi, dan pengujian hipotesis, dari statistika dan (2) algoritma pencarian, teknik pemodelan, dan teori pembelajaran dari kecerdasan buatan, pengenalan pola, dan machine learning. Data mining juga telah megadopsi ide-ide dari area lain meliputi optimisasi, evolutionary computing, teori informasi, pemrosesan sinyal, visualisasi dan information retrieval. Berbagai Sejumlah area lain juga memberikan peran pendukung dalam data mining, seperti sistem basis data yang dibutuhkan untuk menyediakan tempat penyimpanan yang efisien, indexing dan pemrosesan kueri tentunya.

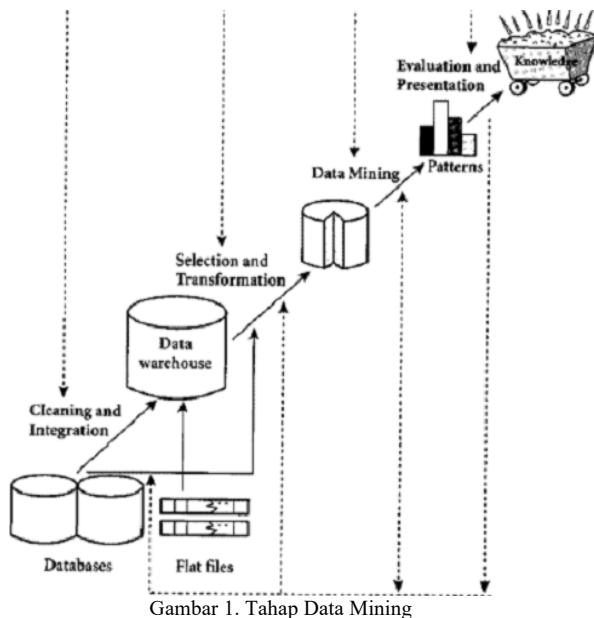
#### A.1 Pengelompokan Data Mining

Dalam Data Mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan,[5] yaitu:

- a. Deskripsi, Pada pengelompokan *datamining* deskripsi ini, peneliti dan analisis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Deskripsi dari pola kecenderungan sering memberikan kemungkinan

- penjelasan untuk suatu pola atau kecenderungan. Deskripsi bertujuan untuk mengidentifikasi pola yang muncul secara berulang pada suatu data dan mengubah pola tersebut menjadi aturan dan kriteria yang dapat mudah dimengerti oleh para ahli pada domain aplikasinya. Aturan yang dihasilkan harus mudah dimengerti agar dapat dengan efektif meningkatkan tingkat pengetahuan (knowledge) pada sistem. Tugas deskriptif merupakan tugas data mining yang sering dibutuhkan pada teknik postprocessing untuk melakukan validasi dan menjelaskan hasil dari proses data mining. Postprocessing merupakan proses yang digunakan untuk memastikan hanya hasil yang valid dan berguna yang dapat digunakan oleh pihak yang berkepentingan.
- b. Estimasi, Pengelompokan *datamining* ini yaitu Estimasi. Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik dari pada ke arah kategori. Model dibangun menggunakan 6 baris data (record) lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi, selain itu sejatinya Estimasi hampir sama dengan prediksi, kecuali variabel target estimasi lebih ke arah numerik dari pada ke arah kategori. Model dibangun menggunakan record lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi. Sebagai contoh, akan dilakukan estimasi tekanan darah sistolik pada pasien rumah sakit berdasarkan umur pasien, jenis kelamin, berat badan, dan level sodium darah. Hubungan antara tekanan darah sistolik dan nilai variabel prediksi dalam proses pembelajaran akan menghasilkan model estimasi.
- c. Prediksi, selanjutnya yaitu Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang. Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi, selain itu juga menurut pemahamannya Prediksi memiliki kemiripan dengan klasifikasi, akan tetapi data diklasifikasikan berdasarkan perilaku atau nilai yang diperkirakan pada masa yang akan datang. Contoh dari tugas prediksi misalnya untuk memprediksi adanya pengurangan jumlah pelanggan dalam waktu dekat dan prediksi harga saham dalam tiga bulan yang akan datang.
- d. Klasifikasi, Dalam klasifikasi, terdapat target variabel kategori, Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah. Klasifikasi juga merupakan suatu proses untuk menyatakan suatu objek ke salah satu kategori yang sudah didefinisikan sebelumnya. Proses pembelajaran fungsi target (model klasifikasi) yang memetakan setiap sekumpulan atribut x (input) ke salah satu klas y yang didefinisikan sebelumnya. Inputnya adalah sekumpulan record (training set) dari setiap record tersebut terdiri atas sekumpulan atribut.
- e. Pengklasteran (Clustering) Pengklasteran merupakan pengelompokan record, pengamatan, atau memperhatikan dan membentuk kelas obyek-obyek yang memiliki kemiripan. Klaster adalah kumpulan record yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan record dalam klaster yang lain. Berbeda dengan klasifikasi, pada pengklasteran tidak ada variabel target. Pengklasteran tidak melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari variabel target, akan tetapi, algoritma pengklasteran mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan (homogen), yang mana kemiripan record dalam satu kelompok akan bernilai maksimal, sedangkan kemiripan dengan record dalam kelompok lain akan bernilai minimal, selain itu pada menurut pemahamannya sejatinya juga diartikan bahwa Clustering merupakan pengelompokan data tanpa berdasarkan kelas data tertentu ke dalam kelas objek yang sama. Sebuah kluster adalah kumpulan record yang memiliki kemiripan suatu dengan yang lainnya dan memiliki ketidakmiripan dengan record dalam kluster lain. Tujuannya adalah untuk menghasilkan pengelompokan objek yang mirip satu sama lain dalam kelompok-kelompok. Semakin besar kemiripan objek dalam suatu cluster dan semakin besar perbedaan tiap cluster maka kualitas analisis cluster semakin baik.
- f. Asosiasi, Tugas asosiasi dalam Data Mining adalah untuk menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja (Market Basket Analysis), ataupun Dalam dunia bisnis lebih umum disebut analisis keranjang belanja (market basket analysis) tidak hanya itu Asosiasi juga diartikan bahwa Tugas asosiasi dalam data mining adalah menemukan atribut yang muncul dalam suatu waktu. Tugas asosiasi berusaha untuk mengungkap aturan untuk mengukur hubungan antara dua atau lebih atribut.

## A.2 Tahap Data Mining



Gambar 1. Tahap Data Mining

Tahap-tahap tersebut bersifat interaktif di mana pemakai terlibat langsung atau dengan perantaraan knowledge base. [6] Tahapan-tahapan tersebut, diantaranya : 1. Pembersihan data (untuk membuang data yang tidak konsisten dan noise)

Pada umumnya data yang diperoleh, baik dari database suatu perusahaan maupun hasil eksperimen, memiliki isian-isian yang tidak sempurna seperti data yang hilang, data yang tidak valid atau juga hanya sekedar salah ketik. Selain itu, ada juga atribut-atribut data yang tidak relevan dengan hipotesa data mining yang kita miliki. Data-data yang tidak relevan itu juga lebih baik dibuang karena keberadaannya bisa mengurangi mutu atau akurasi dari hasil data mining nantinya. Garbage in garbage out (hanya sampah yang akan dihasilkan bila yang dimasukkan juga sampah) merupakan istilah yang sering dipakai untuk menggambarkan tahap ini. Pembersihan data juga akan mempengaruhi performasi dari sistem data mining karena data yang ditangani akan berkurang jumlah dan kompleksitasnya.

2. Integrasi data (penggabungan data dari beberapa sumber)

Tidak jarang data yang diperlukan untuk data mining tidak hanya berasal dari satu database tetapi juga berasal dari beberapa database atau file teks. Integrasi data dilakukan pada atribut-atribut yang mengidentifikasi entitas-entitas yang unik seperti atribut nama, jenis produk, nomor pelanggan dsb. Integrasi data perlu dilakukan secara cermat karena kesalahan pada integrasi data bisa menghasilkan hasil yang menyimpang dan bahkan menyesatkan pengambilan aksi nantinya. Sebagai contoh bila integrasi data berdasarkan jenis produk ternyata menggabungkan produk dari kategori yang berbeda maka akan didapatkan korelasi antar produk yang sebenarnya tidak ada. Dalam integrasi data ini juga perlu dilakukan transformasi dan pembersihan data karena seringkali data

dari dua database berbeda tidak sama cara penulisannya atau bahkan data yang ada di satu database ternyata tidak ada di database lainnya.

3. Transformasi data (data diubah menjadi bentuk yang sesuai untuk di-mining)

Beberapa teknik data mining membutuhkan format data yang khusus sebelum bisa diaplikasikan. Sebagai contoh beberapa teknik standar seperti analisis asosiasi dan klustering hanya bisa menerima input data kategorikal. Karenanya data berupa angka numerik yang berlanjut perlu dibagi-bagi menjadi beberapa interval. Proses ini sering disebut binning. Disini juga dilakukan pemilihan data yang diperlukan oleh teknik data mining yang dipakai. Transformasi dan pemilihan data ini juga menentukan kualitas dari hasil data mining nantinya karena ada beberapa karakteristik dari teknik-teknik data mining tertentu yang tergantung pada tahapan ini.

4. Aplikasi teknik data mining.

Aplikasi teknik data mining sendiri hanya merupakan salah satu bagian dari proses data mining. Ada beberapa teknik data mining yang sudah umum dipakai. Kita akan membahas lebih jauh mengenai teknik-teknik yang ada di seksi berikutnya. Perlu diperhatikan bahwa ada kalanya teknik-teknik data mining umum yang tersedia di pasar tidak mencukupi untuk melaksanakan data mining di bidang tertentu atau untuk data tertentu. Sebagai contoh akhir-akhir ini dikembangkan berbagai teknik data mining baru untuk penerapan di bidang bioinformatika seperti analisa hasil microarray untuk mengidentifikasi DNA dan fungsi-fungsinya.

5. Evaluasi pola yang ditemukan (untuk menemukan yang menarik/bernilai)

Dalam tahap ini hasil dari teknik data mining berupa pola-pola yang khas maupun model prediksi dievaluasi untuk menilai apakah hipotesa yang ada memang tercapai. Bila ternyata hasil yang diperoleh tidak sesuai hipotesa ada beberapa alternatif yang dapat diambil seperti : menjadikannya umpan balik untuk memperbaiki proses data mining, mencoba teknik data mining lain yang lebih sesuai, atau menerima hasil ini sebagai suatu hasil yang di luar dugaan yang mungkin bermanfaat. Ada beberapa teknik data mining yang menghasilkan hasil analisa berjumlah besar seperti analisis asosiasi. Visualisasi hasil analisa akan sangat membantu untuk memudahkan pemahaman dari hasil data mining.

6. Presentasi pola yang ditemukan untuk menghasilkan aksi

Tahap terakhir dari proses data mining adalah bagaimana memformulasikan keputusan atau aksi dari hasil analisa yang didapat. Ada kalanya hal ini harus melibatkan orang-orang yang tidak memahami data mining. Karenanya presentasi hasil data mining dalam bentuk pengetahuan yang bisa dipahami semua orang adalah satu tahapan yang diperlukan dalam proses data mining. Dalam presentasi ini, visualisasi juga bisa membantu mengkomunikasikan hasil data min.

## B. Algoritma Apriori

Algoritma Apriori adalah suatu algoritma dasar yang diusulkan oleh Agrawal & Srikant pada tahun 1994 untuk penentuan *frequentitemsets* untuk aturan asosiasi *boolean*. Algoritma apriori termasuk jenis aturan asosiasi pada *datamining*. Aturan yang menyatakan asosiasi antara beberapa atribut sering disebut *affinity analysis* atau *marketbasket analysis*. Analisis asosiasi atau *associationrulemining* adalah teknik *datamining* untuk menemukan aturan suatu kombinasi *item*. Salah satu tahap analisis asosiasi yang menarik perhatian banyak peneliti untuk menghasilkan algoritma yang efisien adalah analisis pola frekuensi tinggi (*frequentpatternmining*). Penting tidaknya suatu asosiasi dapat diketahui dengan dua tolak ukur, yaitu *support* dan *confidence*. *Support* (nilai penunjang) adalah persentase kombinasi item tersebut dalam *database*, sedangkan *confidence* (nilai kepastian) adalah kuatnya hubungan antar-item dalam aturan asosiasi [7].

Algoritma apriori merupakan salah satu algoritma klasik data mining. Algoritma apriori digunakan agar komputer dapat mempelajari aturan asosiasi, mencari pola hubungan antar satu atau lebih item dalam suatu dataset. Algoritma apriori banyak digunakan pada data transaksi atau biasa disebut market basket, misalnya sebuah swalayan memiliki market basket, dengan adanya algoritma apriori, pemilik swalayan dapat mengetahui pola pembelian seorang konsumen, jika seorang konsumen membeli item A, B, punya kemungkinan 50% dia akan membeli item C, pola ini sangat signifikan dengan adanya data transaksi selama ini. Penting tidaknya suatu aturan assosiatif dapat diketahui dengan dua parameter, *support* (nilai penunjang) yaitu persentase kombinasi item tersebut dalam *database* dan *confidence* (nilai kepastian) yaitu kuatnya hubungan antar item dalam aturan assosiatif. Aturan tersebut cukup signifikan karena mewakili 40% dari catatan transaksi selama ini.” Analisis asosiasi didefinisikan suatu proses untuk menemukan semua aturan assosiatif yang memenuhi syarat minimum untuk *support* (minimum *support*) dan syarat minimum untuk *confidence* (minimum *confidence*). Tetapi di lain pihak Apriori memiliki kelemahan karena harus melakukan scan *database* setiap kali iterasi, sehingga waktu yang diperlukan bertambah dengan makin banyak iterasi. Masalah ini yang dipecahkan oleh algoritma-algoritma baru seperti FP-growth.

Algoritma Apriori dibagi menjadi beberapa tahap yang disebut iterasi atau pass. Tiap iterasi menghasilkan pola frekuensi tinggi dengan panjang yang sama dimulai dari pass pertama yang menghasilkan pola frekuensi tinggi dengan panjang satu. Di iterasi pertama ini, *support* dari setiap item dihitung dengan men-scan *database*. Setelah *support* dari setiap item didapat, item yang memiliki *support* diatas minimum *support* dipilih sebagai pola frekuensi tinggi dengan panjang 1 atau sering disingkat 1-itemset. Singkatan k-itemset berarti satu set yang terdiri dari k item.

Iterasi kedua menghasilkan 2-itemset yang tiap set-nya memiliki dua item. Pertama dibuat kandidat 2-itemset dari kombinasi semua 1-itemset. Lalu untuk tiap kandidat 2-itemset ini dihitung *support*-nya dengan men-scan *database*. *Support* disini artinya jumlah transaksi

dalam *database* yang mengandung kedua item dalam kandidat 2-itemset. Setelah *support* dari semua kandidat 2-itemset didapatkan, kandidat 2-itemset yang memenuhi syarat minimum *support* dapat ditetapkan sebagai 2-itemset yang juga merupakan pola frekuensi tinggi dengan panjang 2. Selain itu tentunya Untuk selanjutnya pada iterasi ke-k dapat dibagi lagi menjadi beberapa bagian :

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

2. Penghitungan support dari tiap kandidat k-itemset. *Support* dari tiap kandidat k-itemset didapat dengan men-scan *database* untuk menghitung jumlah transaksi yang memuat semua item di dalam kandidat k-itemset tsb. Ini adalah juga ciri dari algoritme Apriori dimana diperlukan penghitungan dengan scan seluruh *database* sebanyak k-itemset terpanjang.

3. Tetapkan pola frekuensi tinggi. Pola frekuensi tinggi yang memuat k item atau k-itemset ditetapkan dari kandidat k-itemset yang *support*-nya lebih besar dari minimum *support*.

4. Bila tidak didapat pola frekuensi tinggi baru maka seluruh proses dihentikan. Bila tidak, maka k ditambah satu dan kembali ke bagian 1.

## C. Analisis Pola Frekuensi Tinggi dengan Algoritma Apriori

Tahap ini mencari kombinasi *item* yang memenuhi syarat minimum dari nilai *support* dalam basis data. Nilai *support* sebuah *item* diperoleh dengan menggunakan rumus berikut : [8]

$$support(A) = \frac{\text{Jumlah transaksi mengandung } A}{\text{Total transaksi}}$$

Sementara, nilai *support* dari 2 *item* diperoleh dengan menggunakan rumus :

$$support(A, B) = \frac{\sum \text{transaksi mengandung } A \text{ dan } B}{\sum \text{transaksi}}$$

Frequent itemset menunjukkan itemset yang memiliki frekuensi kemunculan lebih dari nilai minimum yang ditentukan ( $\emptyset$ ). Misalkan  $\emptyset = 2$ , maka semua *itemsets* yang frekuensi kemunculannya lebih dari atau sama dengan 2 kali disebut *frequent*. Himpunan dari *frequent-itemset* dilambangkan dengan *Fk*.

## D. Pembentukan Aturan Asosiasi

Setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung

*confidence* aturan asosiatif  $A \rightarrow B$ . Nilai *confidence* dari aturan  $A \rightarrow B$  diperoleh dengan rumus berikut : [9]

$$\text{confidence } P(B|A) = \frac{\sum \text{transaksi mengandung } A \text{ dan } B}{\sum \text{transaksi mengandung } A}$$

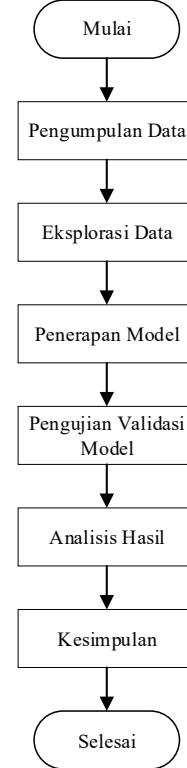
Untuk menentukan aturan asosiasi yang akan dipilih maka harus diurutkan berdasarkan *Support*  $\times$  *Confidence*. Aturan diambil sebanyak  $n$  aturan yang memiliki hasil terbesar.

### III. METODOLOGI PENELITIAN

#### A. Metodologi Penelitian

Metodologi penelitian dilakukan dengan mengumpulkan data. Proses pengumpulan data dilakukan dengan dua cara yaitu studi kepustakaan dan studi lapangan. Studi kepustakaan dilakukan untuk memahami penerapan *data mining* seperti mengumpulkan teori-teori yang berhubungan dengan *datamining*. Studi kepustakaan adalah kegiatan untuk menghimpun informasi yang relevan dengan topik atau masalah yang menjadi obyek penelitian. Informasi tersebut dapat diperoleh dari buku-buku, karya ilmiah, tesis, disertasi, ensiklopedia, internet, dan sumber-sumber lain. Dengan melakukan studi kepustakaan, peneliti dapat memanfaatkan semua informasi dan pemikiran-pemikiran yang relevan dengan penelitiannya. Peranan studi kepustakaan sebelum penelitian sangat penting sebab dengan melakukan kegiatan ini hubungan antara masalah, penelitian-penelitian yang relevan dan teori akan menjadi lebih jelas[10]. Selain itu penelitian akan lebih ditunjang, baik oleh teori-teori yang sudah ada maupun oleh bukti nyata, yaitu hasil-hasil penelitian, kesimpulan dan saran. Studi kepustakaan adalah tugas yang terus menerus dilakukan selama kegiatan penelitian. Sebuah penelitian akan menghasilkan suatu karya ilmiah, karena itu haruslah mampu memberi sumbangan kepada kemajuan ilmu pengetahuan. Pemeriksaan yang teliti perlu dilakukan, dari mulai memilih judul, agar jangan sampai terjadi duplikasi terhadap masalah yang sudah diteliti oleh orang lain. Meskipun masalah yang sama sekali baru (original) sangat jarang, namun studi atau hasil penelitian yang terdahulu tidak harus ditiru seutuhnya, kecuali teknik-teknik yang dipergunakan terbukti tidak tepat atau hasil penelitian dan kesimpulannya meragukan, atau telah diketemukan informasi baru yang dapat memberikan pemecahan lain. Selanjutnya yaitu Studi Lapangan, merupakan penelitian yang dilakukan dengan cara melakukan pengamatan langsung pada objek yang diteliti untuk memperoleh data primer. Tentunya studi lapangan dilakukan mendapatkan data penelitian, sejatinya juga dalam pemahaman mengenai Studi Lapangan ini, langsung turun ke objek dan subjek penelitian untuk meneliti keadaan sebenarnya, langsung merasakan dampak ataupun melihat langsung keadaan yang terjadi seperti halnya yang dilakukan pada penelitian ini. Penelitian kepada agen penjual tiket pesawat Data yang diperoleh akan dianalisis

untuk mengetahui kebutuhan dalam proses asosiasi data dalam penyelesaian masalah. Berikut merupakan langkah-langkah yang digunakan dalam penelitian ini.



Gambar 2. Langkah-Langkah Penelitian

#### B. Studi awal

Tahap ini bertujuan dalam merumuskan latar belakang, tujuan dan permasalahan yang akan dibahas adapun beberapa hal yang dilakukan pada tahapan ini adalah [11]:

a. Mengidentifikasi masalah

Pada penelitian ini diawali dengan memperoleh dan menentukan topic yang akan diselesaikan dan dibahas dalam penelitian ini. Setelah peneliti memperoleh dan menentukan topic penelitian, maka langkah awal yang harus dilakukan adalah mengidentifikasi permasalahan yang akan dipelajari. Maksud identifikasi adalah sebagai penegasan batas – batas permasalahan, agar pembahasan penelitian tidak keluar dari tujuannya.

b. Menentukan ruang lingkup masalah

Selanjutnya dalam penelitian ini sejatinya dilakukan adalah menentukan ruang lingkup masalah. Langkah yang peneliti lakukan setelah merancang penelitian adalah merumuskan pertanyaan ataupun dengan kata lain yaitu menentukan ruang lingkup masalah. Nantinya, rumusan pertanyaan ini akan digunakan sebagai acuan peneliti dalam mengumpulkan data. Ketika merumuskan pertanyaan, peneliti harus menggunakan bahasa yang mudah dipahami agar pembaca penelitian mengerti inti penelitian ini.

Banyak hal yang tentunya harus diperhatikan ketika merumuskan pertanyaan merupakan topik, subjek, judul, rumusan masalah, dan variabel, dimana pada

pembahasannya Topik merupakan area masalah yang akan dikaji dalam penelitian. Subjek merujuk kepada orang, tempat, maupun benda yang diamati selama penelitian. Peneliti melakukan beberapa strategi dalam pemilihan judul penelitian ini, yakni Judul penelitian sebaiknya mampu menjelaskan subjek serta topik yang digunakan. Penelitian juga memerlukan variabel, yaitu faktor-faktor yang diduga berperan dalam peristiwa atau gejala yang akan diteliti.

c. Mempelajari beberapa literatur atau mengumpulkan data

Sesuai dengan Metode Penelitian yang peneliti lakukan yaitu dengan studi pustaka dan studi Lapangan, peneliti melakukan Pengumpulan data dengan dilakukan melalui instrument pengumpulan data, observasi, maupun melalui dokumentasi. Data yang dikumpulkan berupa data primer dan sekunder sejatinya Data primer diperoleh dari sumber langsung berupa interview atau observasi. Data sekunder diperoleh dari sumber tidak langsung berupa data dokumentasi dan arsip – arsip resmi.

d. Analisa atau mengolah Data

Kegiatan mengolah data bertujuan untuk memperoleh kesimpulan hasil analisis, kemudian dalam Hasil analisis data yang menjadi dasar penolakan atau penerimaan hipotesis dan harus dibaha dalam pemecahan masalah. Apakah hasil analisi data menyimpulkan bahwa hipotesis harus diterima atau ditolak, diskusi tentang hal tersebut akan menentukan bobot penelitian yang dilakukan. Untuk mempermudah dan mempercepat proses maka pengolahan data dapat dilakukan dengan menggunakan computer, dalam hal ini peneliti menggunakan hasil penelitian langsung dengan studi pustaka dan studi lapangan. Data yang telah dikumpulkan kemudian diolah dan dianalisis agar menjadi informasi yang relevan dengan masalah penelitian dan dapat dipahami oleh pembaca. Metode penelitian sosial mengenal beberapa cara mengolah dan menganalisis data, tiga di antaranya adalah dengan teknik tabel frekuensi, statistik sederhana, dan tabulasi silang.

Teknik tabel frekuensi mencakup penyajian data dan pengelompokan data ke dalam suatu daftar atau tabel dan kelas interval dari hasil penelitian. Statistik sederhana digunakan untuk merangkum data dan menampilkannya dalam bentuk yang dapat dimengerti oleh setiap orang, biasanya sebagai tabel dan grafik. Sementara itu, tabulasi silang adalah salah satu analisis korelasional yang digunakan untuk melihat hubungan antarvariabel (minimal 2 variabel) yang menggunakan data nominal ataupun ordinal.

#### E. Merumuskan dan Menyajikan Hasil Penelitian

Langkah terakhir dari metode penelitian sosial adalah menyajikan hasil penelitian. Proses ini dilakukan setelah semua data diolah dan dianalisis, sehingga peneliti dapat merumuskan dan menyajikan hasil penelitian secara garis besar. Umumnya, bagian ini menjabarkan tiga poin, yaitu hasil penelitian, simpulan, dan saran.

Hasil penelitian adalah hasil akhir yang didapat peneliti dan merupakan jawaban dari rumusan masalah yang telah dikemukakan. Simpulan adalah inti dari

seluruh informasi yang diperoleh selama proses penelitian, dan saran memuat harapan peneliti dari apa yang telah diperoleh serta kekurangan yang telah dilakukan sehingga penelitian selanjutnya bisa menjadikannya evaluasi.

### C. Pengolahan Data dengan Data Mining

Pada tahap pengolahan data terlebih dahulu melakukan identifikasi masalah yang ada dan sering dihadapi oleh pihak perusahaan, untuk kemudian mendeskripsikan masalah-masalah tersebut untuk diperoleh solusinya. Tahap selanjutnya dilakukan analisa masalah menggunakan teknik *datamining* dengan algoritma Apriori untuk mendapatkan hasil sebagai tujuan yang akan dicapai. Seperti pemahamannya Data Mining adalah suatu proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan pembelajaran mesin untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai big data. Istilah data mining memiliki hakikat sebagai suatu disiplin ilmu yang tujuannya adalah untuk menemukan, menggali, atau mendapat pengetahuan dari data atau informasi yang telah dimiliki, yang dihubungkan dengan Algoritma Apriori, dimana Algoritma Apriori dapat diartikan bahwa Algoritma apriori merupakan salah satu algoritma klasik data mining. Algoritma apriori digunakan agar komputer dapat mempelajari aturan asosiasi, mencari pola hubungan antar satu atau lebih item dalam suatu dataset. Algoritma apriori banyak digunakan pada data transaksi atau biasa disebut market basket sehingga dapat dihubungkan yang membuat permasalahan dalam penelitian ini dapat diselesaikan dengan baik tentunya.

## IV. HASIL DAN PEMBAHASAN

### A. Implementasi Algoritma Apriori

Untuk melakukan proses pengolahan data transaksi penjualan tiket pesawat dengan menggunakan data transaksi penjualan yang jumlahnya semakin banyak, maka akan dilakukan Analisis Asosiasi algoritma apriori dengan membuat nilai *support* dengan menggunakan data transaksi penjualan seperti berikut :

TABEL I  
TRANSAKSI

Transaksi	Item yang dijual			
1	Garuda	Lion	Citilink	Trigana
2	Garuda	Citilink		Trigana
3	Garuda			
4	Citilink	Lion		
5	Trigana	Lion	Garuda	
6	Lion	Citilink	Garuda	
7	Citilink	Garuda		
8	Garuda	Trigana	Lion	
9	Lion	Trigana	Citilink	
10	Garuda	Trigana	Lion	

## B. Representasi Data Transaksi

Data pada transaksi tabel di atas direpresentasikan ke dalam bentuk tabel berikut :

TABEL II  
REPRESENTASI DATA TRANSAKSI

Transaksi	Item
1	Garuda
2	Citilink
3	Lion
4	Trigana

## C. Tabulasi Data Transaksi

Data transaksi pada tabel 3 dibentuk tabel tabular yang akan mempermudah dalam mengetahui beberapa banyak item yang ada dibeli dalam setiap transaksi seperti pada tabel berikut :

TABEL III  
FORMAT TABULAR DATA TRANSAKSI

Transaksi	Garuda	Citilink	Lion	Trigana
1	1	1	1	1
2	1	1	0	1
3	1	0	0	0
4	0	1	1	0
5	1	0	1	1
6	1	1	1	0
7	1	1	0	0
8	1	0	1	1
9	0	1	1	1
10	1	0	1	1
Jumlah	8	6	7	6

## D. Pembentukan Itemset

Berikut ini adalah penyelesaian berdasarkan kasus yang sudah disediakan pada tabel 3. Proses pembentukan C1 atau disebut dengan 1 itemset dengan jumlah dengan rumus sebagai berikut :

$$Support A = \frac{\text{Jumlah Transaksi Mengandung A}}{\text{Total Transaksi}}$$

$$1. Support (Garuda) = \frac{8}{10} = 0,8$$

$$2. Support (Citilink) = \frac{6}{10} = 0,6$$

$$3. Support (Lion) = \frac{7}{10} = 0,7$$

$$4. Support (Trigana) = \frac{6}{10} = 0,6$$

Hasil dari perhitungan di atas diubah ke dalam bentuk persentase untuk mengetahui *confidence* dari masing-masing item penjualan, hasil tersebut dapat kita lihat pada tabel berikut :

TABEL IV  
SUPPORT SETIAP ITEM

Transaksi	Item	Jumlah	Confidence
1	Garuda	8	80%
2	Citilink	6	60%
3	Lion	7	70%
4	Trigana	6	60%

Dalam proses pembentukan itemset pada tabel 4 hasil pembentukan 1 itemset akan dilakukan kombinasi 2 itemset.

## Kombinasi 2 Itemset

Proses pembentukan C2 atau disebut dengan 2 itemset dengan rumus :

$$support(A, B) = \frac{\sum \text{transaksi mengandung } A \text{ dan } B}{\sum \text{transaksi}}$$

TABEL V  
KOMBINASI 2 ITEMSET

Nama Itemset	Jumlah	Support	
Garuda	Citilink	4	40,0%
Garuda	Lion	4	40,0%
Garuda	Trigana	5	50,0%
Citilink	Lion	4	40,0%
Citilink	Trigana	3	30,0%
Lion	Trigana	5	50,0%

Dari tabel tersebut di atas, ditetapkan nilai  $\emptyset = 5$  sehingga didapat C2 = (Garuda, Trigana) & (Lion, Trigana). Pada tabel ini dapat dipahami dan dianalisis bahwa Garuda dengan Support 40% bertahan dengan Citilink, Lion, walaupun pada Jumlahnya berbeda beda seperti halnya Garuda Citilink dengan Jumlah 4, Sedangkan Garuda dengan Lion Jumlah 4 berbeda halnya dengan Garuda dan Trigana dengan Jumlah 5 dengan persentase support nya yakni 50% dibandingkan dengan dua itemset lainnya, namun sejatinya persentase support diantar beberapa itemset tersebut berbeda karena jumlahnya pun berbeda.

## Kombinasi 3 Itemset

Kombinasi dari itemset pada C2 bisa digabung menjadi 3 itemset. Proses pembentukan C3 atau disebut dengan 3 itemset dapat dilihat dari tabel di bawah ini :

TABEL VI  
KOMBINASI 3 ITEMSET

Nama Itemset		Jumlah
Garuda	Trigana	Lion

Dengan demikian, C3 = (Garuda, Trigana, Lion) karena hanya kombinasi itulah yang memiliki frekuensi  $\Rightarrow \emptyset$ .

## E. Pembentukan Aturan Asosiasi

Setelah pola frekuensi tinggi ditemukan kemudian dicari aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* dengan menghitung *confidence* aturan asosiatif  $A \rightarrow B$ , nilai *confidence* dari aturan  $A \rightarrow B$  dapat dirumuskan sebagai berikut :

$$confidence P(B|A) = \frac{\sum \text{transaksi mengandung } A \text{ dan } B}{\sum \text{transaksi mengandung } A}$$

1.  $Confidence P(Garuda|Trigana) = \frac{4}{5} = 0,8$
2.  $Confidence P(Garuda|Loin) = \frac{2}{4} = 0,5$
3.  $Confidence P(Trigana|Citilink) = \frac{2}{3} = 0,66666667$

Dari C3 yang telah ditemukan, bisa dilihat besarnya nilai *support* dan *confidence* dari aturan asosiasi seperti tabel di bawah ini :

TABEL VII  
ATURAN ASOSIASI DARI C3

Rule	Confidence	
Jika beli Garuda, Trigana maka beli Lion	4/5	80,00%
Jika beli Citilink, Lion maka beli Garuda	2/4	50,00%
Jika beli Trigana, Citilink maka beli Garuda	2/3	66,67%

Berdasarkan tabel 7 nilai *confidence* minimal adalah 70% sehingga aturan yang bisa terbentuk adalah aturan dengan *antecedent* berikut :

- Jika beli Garuda, Trigana maka beli Lion

Sementara itu aturan asosiasi untuk C2 bisa dilihat pada tabel berikut :

TABEL VIII  
ATURAN ASOSIASI DARI C2

Rule	Confidence	
Jika beli Garuda maka beli Citilink	4/8	50,00%
Jika beli Garuda maka beli Lion	5/8	62,50%
Jika beli Garuda maka beli Trigana	5/8	62,50%
Jika beli Citilink maka beli Lion	4/6	66,67%
Jika beli Citilink maka beli Trigana	3/6	50,00%
Jika beli Lion maka beli Trigana	5/7	71/43%

Sementara itu aturan asosiasi final secara berurut dengan nilai *confidence* minimal adalah 70% dapat dilihat pada tabel berikut.

TABEL IX  
ATURAN ASOSIASI FINAL

Rule	Confidence	
Jika beli Garuda, Trigana maka beli Lion	4/5	80,00%
Jika beli Lion maka beli Trigana	5/7	71,43%

Hasil dan analisis sistem yang mencari pola pembelian tiket pesawat berdasarkan data transaksi penjualan tiket, dijelaskan bahwa tiket yang sering dibeli konsumen bersamaan adalah :

1. Jika konsumen membeli tiket Garuda, Trigana maka konsumen juga akan membeli tiket Lion dengan tingkat *confidence* 80%.
2. Jika konsumen membeli tiket Lion maka konsumen juga akan membeli tiket Trigana dengan tingkat *confidence* 71,43%.

Dengan mengetahui hasil tersebut, maka pihak travel agent dapat menyusun strategi dalam penentuan promo maskapai untuk menyusun strategi pemasaran pesawat lainnya.

## V. KESIMPULAN

Berdasarkan hasil analisis dan pengujian hipotesis sebagaimana yang telah disajikan pada pembahasan yang ada pada bab sebelumnya, maka dapat diambil beberapa simpulan. Simpulan tersebut dipaparkan sebagai berikut:

1. Penerapan algoritma apriori pada penjualan tiket pesawat efektif untuk melihat pola pembelian konsumen dimana pada saat pengolahan data dapat kita ketahui tiket yang sering di pesan oleh pembeli. Dimana disini dapat di lihat tiket yang sering di beli yaitu tiket Garuda, Trigana dan Lion sehingga pihak pemilik travel dapat mengatur strategi pemasarannya, seperti mengadakan promo untuk meningkatkan penjualan tiket pesawat untuk maskapai lainnya yang sehingga semua maskapai tiket dapat terjual dengan banyak. Maka pihak travel agent dapat menyusun strategi dalam penentuan promo maskapai untuk menyusun strategi pemasaran pesawat lainnya.
2. Penggunaan data mining algoritma apriori untuk menemukan pola kombinasi *itemset* dan *association rules*, dimana dengan penjualannya itu nilai *support* dan *confidence* tertinggi adalah Lion dan Trigana serta Garuda dan Trigana dengan nilai *support* 50% dan *confidence* 70%. Dengan rincian sebagai berikut, jika membeli tiket Lion maka akan membeli tiket Trigana dengan *support* 50% dan *confidence* sebesar 70%, jika membeli tiket Garuda, Trigana maka membeli tiket lion dengan nilai *support* 50% dan *confidence* sebesar 70%.

## DAFTAR PUSTAKA

- [1] E. Buulolo, "IMPLEMENTASI ALGORITMA APRIORI PADA SISTEM PERSEDIAAN OBAT (STUDI KASUS : APOTIK RUMAH SAKIT ESTOMIHI MEDAN)," pp. 71–83, 2013.
- [2] S. R. Siregar, "IMPLEMENTASI DATA MINING PADA PENJUALAN TIKET PESAWAT MENGGUNAKAN ALGORITMA APRIORI (Studi Kasus : Jumbo Travel Medan)," *Pelita Inform. Budi Darma*, Vol. Vii, Nomor 1, Juli 2014, vol. VII, pp. 152–156, 2014.
- [3] H. Santoso, I. P. Hariyadi, and Prayitno, "Data Mining Analisa Pola PembelianProduk," *Tek. Inform.*, no. 1, pp. 19–24, 2016.
- [4] G. Gunadi and D. I. Sensuse, "PenerapanMetode Data Mining Market Basket Analysis Terhadap Data PenjualanProdukBukuDenganMenggunakanAlgoritmaApriori Dan Frequent Pattern Growth ( FP-GROWTH )," vol. 4, no. 1, 2012.
- [5] Kusrini & Luthfi, E. Taufiq. 2009, "Algorima Data Mining", Penerbit Andi, Yogyakarta.
- [6] Frontline Solvers. (2012),"XL Miner Data Mining add-in for Excel." Diakses 17 Agustus 2020,darihttp://www.solver.com/xlminerdatamining.
- [7] A. Nursikuaguset *et al.*, "Implementasi algoritma apriori untuk analisis penjualan dengan berbasis web," vol. 7, no. 2, pp. 701–706, 2016.
- [8] M. Badrul, P. Studi, and S. Informasi, "Algoritma asosiasi dengan algoritma apriori untuk analisa data penjualan," no. 2, pp. 121–129, 2016.

- [9] Pramudiono, I.(2003),"Pengantar Data Mining".Diakses 17 Agustus 2020, dari <http://ikc.depsos.go.id/umum/iko-datamining.php>
- [10] Sandjaja, B. dan Albertus Heriyanto. 2011."Panduan Penelitian." Jakarta: PrestasiPustakaraya
- [11] Arikunto, Suharasimi. 2010. "Prosedur Penelitian Suatu Pendekatan Praktik." Jakarta :RinekaCipta

