

BUKU MONOGRAF

Algoritma Apriori

Penulis

Delima Sitanggang, M.Kom

Mardi Turnip, S.Kom, M.Kom.

Yonata Laia, M.Kom.

Saut Parsaoran Tamba, M.Kom.



Penulis

**Delima Sitanggang, M.Kom
Mardi Turnip, S.Kom, M.Kom.
Yonata Laia, M.Kom.
Saut Parsaoran Tamba, M.Kom.**

Editor

Evta Indra, M.Kom

ISBN

Desain Cover

Saut Parsaoran Tamba

Penerbit

**Unpri Press
Universitas Prima Indonesia**

Redaksi

Jl. Belanga No. 1. Simp. Ayahanda, Medan

Cetakan Pertama

Hak Cipta di Lindungi Undang-undang

**Dilarang memperbanyak karya tulis ini dalam bentuk dan
dengan cara apapun tanpa ijin dari penerbit**

KATA PENGANTAR

Puji dan Syukur penulis panjatkan ke hadirat Allah SWT, karena atas segala berkat dan anugerah-Nya yang diberikan kepada tim penulis, sehingga penulisan buku monograf berhasil diselesaikan tepat pada waktunya.

Buku monograf dengan judul Algoritma apriori atau Apriori algorithm yang berisi tentang menentukan pola frekuensi tertinggi dengan menggunakan pola item-item dalam database yang memiliki nilai support di atas ambang batas tertentu yang disebut dengan istilah minimum support dan mengetahui nilai cofidence.

Penulis menyadari masih banyak kekurangan dalam penulisan monograf ini, untuk itu saran dan kritik sangat diharapkan untuk penyempurnaan.

Penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada semua pihak yang telah memberikan motivasi, semangatnya dan berkontribusi dalam kemampuannya untuk dapat menerbitkan buku monograf, semoga Tuhan Yang Maha Esa membalasnya dengan balasan yang lebih baik.

Penulis

Delima Sitanggang, M.Kom

DAFTAR ISI

KATA PENGANTAR.....	i
DAFTAR ISI.....	ii
BAB I Pengantar Data Mining	1
1.1 Defenisi Data Mining	1
1.2 Tahapan Data Mining.....	2
1.3 Pengelompokan Data Mining.....	5
BAB II Asosiation Rule Mining.....	8
2.1 Pengertian Asosiation Rule	8
2.2.1 Analisa Fola Frekuensi Tinggi.....	9
2.2.2 Pembentukan Aturan Asosiatif.....	10
BAB III Algoritma Apriori.....	11
3.1 Pengertin Algoritma Apriori.....	11
3.2 Kelebihan dan Kekurangan Algoritma Apriori.....	14
BAB IV Penerapan Data Mining Untuk Rekomendasi Paket Pernikahan Menggunakan Metode Algoritma Apriori	16
4.1 Latar Belakang.....	16
4.2 Rumusan Masalah	18
4.3 Batasan Masalah.....	19
4.4.1 Tujuan	19
4.4.2 Manfaat	19
4.5 pendapat para ahli.....	20
4.5.1 Data Mining	20

4.5.2 Algoritma Apriori.....	20
4.6 Cara Kerja	22
4.7 Teknik Pengumpulan Data.....	24
4.7.1 Observasi.....	24
4.7.2 Interview atau Wawancara	24
4.7.3 Studi Literatur	25
4.8 Pre-Processing.....	25
4.8.1 Seleksi Data.....	25
4.8.2 Normalisasi Data	25
4.8.3 Transformasi Data	26
4.9 Indentifikasi Assotion Rule.....	26
4.10 Perhitungan dengan Menggunakan Metode Algoritma Apriori.....	27
4.11 Pembahasan dan Hasil	29
4.11.1 Pembahasan.....	29
4.12 Pengolahan data	31
4.12.1 Normalisasi	31
4.12.2 Seleksi data.....	32
4.12.3 Transformasi data	33
4.13 Implementasi algoritma apriori.....	33
4.14 Pembentukan rule atau aturan asosiasi.....	38
4.15 Impelentasi Apriori pada Python.....	41
4.16 Pembentukan Rule Asosiasi	43
4.17.1 Hasil.....	47
4.17.2 tindakan selanjutnya	48

DAFTAR PUSTAKA..... 49

BAB I PENGANTAR DATA MINING

1.1 Defenisi Data Mining

Secara sederhana data mining adalah penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar. Data mining juga disebut sebagai serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data. Data mining, sering juga disebut sebagai *Knowledge Discovery In Database* (KDD). KDD adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar

Data mining 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. 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. 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. Karakteristik data mining sebagai berikut :

Data mining berhubungan dengan penemuan sesuatu yang tersembunyi dan pola data tertentu yang tidak diketahui sebelumnya. Data mining biasa menggunakan data yang sangat besar. Biasanya data yang besar digunakan untuk membuat hasil lebih dipercaya.

Data mining berguna untuk membuat keputusan yang kritis, terutama dalam strategi. Berdasarkan beberapa pengertian tersebut dapat ditarik kesimpulan bahwa data mining adalah suatu teknik menggali informasi berharga yang terpendam atau tersembunyi pada suatu koleksi data

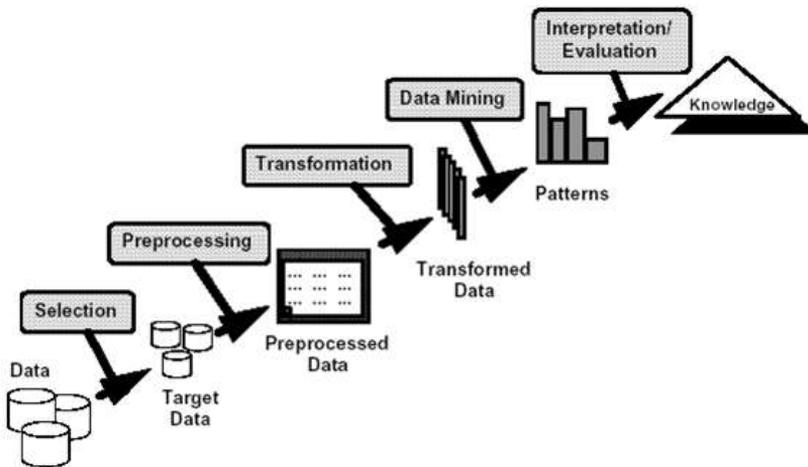
(database) yang sangat besar sehingga ditemukan suatu pola yang menarik yang sebelumnya tidak diketahui. Kata mining sendiri berarti usaha untuk mendapatkan sedikit barang berharga dari sejumlah besar material dasar. Karena itu data mining sebenarnya memiliki akar yang panjang dari bidang ilmu seperti kecerdasan buatan (artificial intelligent), machine learning, statistik dan database. Beberapa metode yang sering disebut-sebut dalam literatur data mining antara lain clustering, lassification, association rules mining, neural network, genetic algorithm dan lain.

1.2 Tahapan Data Mining

Data Mining dapat diartikan sama dengan *Knowledge Discovery In Database* (KDD). Proses yang dilakukan pada KDD sama dengan proses tahapan yang dilakukan pada data mining. Tahapan dimulai dari seleksi data, yaitu dari data sumber ke target. Kemudian tahap *Processsing* digunakan untuk memperbaiki kualitas data dan tranformasi data mining serta tahap interpretasi dan evaluasi yang menghasilkan output

berupa informasi baru yang diharapkan dapat di jadikan informasi lainnya sehingga dapat diolah untuk informasi yang lebih akurat.

Ada beberapa tahapan proses dalam data mining. Diagram dibawah menggambarkan beberapa tahap / proses yang berlangsung dalam data mining. Fase awal dimulai dari data sumber dan berakhir dengan adanya informasi yang dihasilkan dari beberapa tahap, yaitu:



Gambar 1. Fase-Fase dalam Data Miing

Tahapan proses dalam Data Mining dapat dijelaskan sebagai berikut :

1. Seleksi Data

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

2. Pre-processing/ Cleaning (pemilihan data)

Sebelum proses data mining dapat dilaksanakan, perlu dilakukan proses

cleaning pada data yang menjadi fokus KDD.

Proses cleaning mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi).

Juga dilakukan proses enrichment, yaitu proses "memperkaya" data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal.

3. Transformasi

Coding adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses data mining. Proses coding dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data

4. Data mining

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.

5. Interpretasi / Evaluasi

Pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut dengan

interpretation. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya

Proses KDD secara garis besar memang terdiri dari 5 tahap seperti yang telah dijelaskan sebelumnya. Akan tetapi, dalam proses KDD yang sesungguhnya, dapat saja terjadi iterasi atau pengulangan pada tahap tahap tertentu. Pada setiap tahap dalam proses KDD, seorang analis dapat saja kembali ke tahap sebelumnya. Sebagai contoh, pada saat coding atau data mining, analis menyadari proses cleaning belum dilakukan dengan sempurna, atau mungkin saja analis menemukan data atau informasi baru untuk "memperkaya" data yang sudah ada.

1.3 Pengelompokan Data Mining

Berdasarkan fungsi yang dapat dilakukan, *data mining* dibagi menjadi beberapa kelompok:

1. Deskripsi

Dalam proses penelitian terkadang akan mencoba cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data secara sederhana. Sebagai contoh pengumpulan *tweet* yang berhashtag wisata di Kota Malang yang mempunyai kecenderungan untuk berkunjung ke Malang. Penjelasan untuk suatu pola sering kali didapatkan dari deskripsi pada pola dan kecenderungan.

2. Estimasi

Hampir sama dengan klasifikasi, estimasi memiliki pengecuali variabel_target estimasi lebih ke arah numerik dari pada ke arah kategori. Model dibangun dengan *record* lengkap menyediakan nilai dari variabel target sebagai nilai prediksi. Kemudian dalam peninjauan berikutnya estimasi nilai dari variabel target diciptakan berdasarkan nilai variabel prediksi.

3. Prediksi

Prediksi memiliki kesamaan dengan klasifikasi dan

estimasi, perbedaannya terdapat dalam prediksi nilai dari hasil akan ada di masa mendatang.

Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat digunakan (dalam keadaan yang tepat) untuk prediksi.

4. Klasifikasi

Terdapat target variabel kategori dalam klarifikasi. Contohnya penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan rendah, pendapatan sedang, dan pendapatan tinggi.

5. Pengklusteran

Pengklusteran adalah pengelompokan *record*, pengamatan, dan membentuk kelas objek-objek yang memiliki kemiripan. Kluster merupakan kumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan dengan *record-record* dalam kluster lain. Pengklusteran ini beda dengan klasifikasi. Perbedaannya terletak pada ketidak adanya variabel target dalam

pengklusteran. Pengklusteran tidak melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari variabel target.

Namun, algoritma pengklusteran mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan record dalam satu kelompok akan bernilai maksimal, sedangkan dalam kelompok lain akan bernilai minimal pada kemiripan dengan record.

6. Asosiasi

Asosiasi bertugas sebagai penemu atribut yang muncul dalam satu waktu. Dalam dunia bisnis biasa disebut dengan analisis keranjang belanja.

BAB II ASOSIATION RULE MINING

2.1 Pengertian Asosiation Rule

Analisis asosiasi atau *association rule mining* adalah teknik data mining untuk menemukan aturan asosiasi antara suatu kombinasi item. Analisis asosiasi dikenal juga sebagai salah satu teknik data mining yang menjadi dasar dari salah satu teknik data mining lainnya. Bila kita mengambil contoh aturan asosiatif dalam suatu transaksi pembelian barang disuatu minimarket adalah kita dapat mengetahui berapa besar kemungkinan seorang konsumen membeli suatu item bersamaan dengan item lainnya (membeli roti bersama dengan selai). Karena awalnya berasal dari studi tentang database transaksi pelanggan untuk menentukan kebiasaan suatu produk dibeli bersama, maka *association rule* sering juga dinamakan *market basket analysis*

Market basket analysis yaitu analisis yang sering dipakai untuk menganalisa isi keranjang belanja konsumen dalam suatu pasar swalayan. Contoh penerapan dari aturan asosiatif adalah analisa pembelian produk pada sebuah toko alat tulis, pada analisa itu misalkan dapat diketahui berapa besar kemungkinan seorang pelanggan membeli pensil bersamaan dengan membeli

penghapus. Penerapan aturan asosiasi dalam kasus tersebut dapat membantu pemilik toko untuk mengatur penempatan barang, mengatur persediaan atau membuat promosi pemasaran dengan menerapkan diskon untuk kombinasi barang tertentu. Bentuk umum dari *association rule* adalah *Antecedent* dan *Consequent*. Bila kita ambil contoh dalam sebuah transaksi pembelian barang di sebuah minimarket didapat bentuk *association rule* roti selai. Yang artinya bahwa pelanggan yang membeli roti ada kemungkinan pelanggan tersebut juga akan membeli selai, dimana tidak ada batasan dalam jumlah *item-item* pada bagian *antecedent* ataupun *consequent* dalam sebuah *rule*.

2.1.1 Analisis Pola Frekuensi Tinggi

Tahap ini mencari kombinasi item yang memenuhi syarat minimum dari nilai *support* dalam *database*. Nilai *support* (penunjang) yaitu persentase item atau kombinasi item yang ada pada *database*. Nilai *support* sebuah item diperoleh dengan rumus berikut :

$$\text{Support (A)} = \frac{\text{Jumlah Transaksi untuk A}}{\text{Total Transaksi}}$$

Rumus support tersebut menjelaskan bahwa nilai *support* didapat dengan cara membagi jumlah transaksi yang mengandung item A (satu item) dengan jumlah total seluruh transaksi.

Sedangkan nilai *support* dari 2 item diperoleh dari rumus berikut:

$$\text{Support (A, B)} = P (A \cap B)$$

$$\text{Support (A, B)} = \frac{\sum \text{Transaksi Untuk A dan B}}{\sum \text{Transaksi}}$$

Rumus *support* diatas menjelaskan bahwa nilai *support* 2-itemsets didapat dengan cara membagi jumlah transaksi yang mengandung *item* A dan *item* B (item pertama bersamaan dengan item yang lain) dengan jumlah total seluruh transaksi.

2.1.2 Pembentukan Aturan Asosiatif

Setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung nilai *confidence* aturan assosiatif A B. Nilai *confidence* dari aturan A B diperoleh dari rumus sebagai berikut :

$$\text{Confidence} = P (B|A) \frac{\sum \text{Transaksi Untuk A dan B}}{\sum \text{Transaksi}}$$

Rumus diatas menjelaskan bahwa nilai *confidence* diperoleh dengan cara membagi jumlah transaksi yang mengandung *item A* dan *item B* (*item* pertama bersamaan dengan *item* yang lain) dengan jumlah transaksi yang mengandung *item A* (*item* pertama atau *item* yang ada disebelah kiri). Untuk menentukan aturan asosiasi yang akan dipilih maka harus diurutkan berdasarkan *Support* × *Confidence*. Aturan diambil sebanyak n aturan yang memiliki hasil terbesar

BAB III ALGORITMA APRIORI

3.1 Pengertian Algoritma Apriori

Algoritma apriori merupakan algoritma dasar yang diusulkan oleh Agrawal dan Srikant pada tahun 1994 untuk penentuan frequent itemsets untuk aturan asosiasi boolean. Algoritma apriori merupakan salah satu algoritma yang melakukan pencarian *frequent itemset* dengan melakukan teknik *association rule*.

Algoritma apriori digunakan untuk menemukan *association rules* yang memenuhi batas nilai *support* dan *confidence*. Cara kerja apriori adalah dengan menganalisa kumpulan *item* yang diambil atau dipilih secara bersamaan pada beberapa transaksi. penghitung kumpulan pola data yang muncul di dalam database melalui beberapa iterasi atau perulangan. Iterasi i menghitung semua kumpulan data i (kumpulan data yang mengandung elemen i) yang sering muncul.

Setiap iterasi terdiri dari 2 langkah yaitu *candidate generation* (penentuan kandidat) dan *candidate counting and selection* (pemilihan sertapenghitungan kandidat).

Dalam tahap penghitungan, algoritma apriori menghitung nilai *support* terhadap tiap elemen dalam database. Hanya kumpulan data i dengan batas minimal yang telah ditentukan akan dimasukkan ke dalam kandidat barang yang sering muncul.

Pencarian pola kombinasi barang dalam apriori tidak dapat secara langsung menggabungkan elemen i dengan seluruh jenis barang yang diterima. Bila pencarian kombinasi barang dilakukan secara langsung maka waktu yang dibutuhkan untuk mencari kombinasi barang akan lebih lama. Cara yang dilakukan dalam algoritma apriori ini adalah dengan melakukan *self-joining* iterasi sebelumnya.

Proses utama yang dilakukan dalam algoritma apriori untuk menemukan *frequent itemset* yaitu

1. *Join* (Penggabungan)

Pada proses ini dilakukan pengkombinasian *item*

dengan *item* yang lainnya hingga tidak dapat terbentuk kombinasi lagi.

2. *Prune* (Pemangkasan)

Proses ini merupakan hasil dari *item* yang telah dikombinasikan kemudian dipangkas dengan menggunakan minimum *support* yang telah ditentukan *user*.

Langkah pertama pada algoritma apriori yaitu, *support* dari setiap *item* dihitung dengan men-scan database. Setelah *support* dari setiap *item* didapat, *item* yang memiliki *support* lebih besar dari 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* 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.

Untuk selanjutnya iterasi-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 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* tersebut. Ini adalah juga ciri dari algoritma apriori yaitu 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*. Kemudian dihitung *confidence* masing-masing kombinasi *item*. Iterasi berhenti ketika semua *item* telah dihitung sampai tidak ada kombinasi *item* lagi.

3.2 Kelebihan dan Kekurangan Algoritma Apriori

Adapun kelebihan dari algoritma Apriori adalah sebagai berikut:

1. Sederhana dan mudah dipahami di antara algoritma asosiasi

2. Aturan yang dihasilkan intuitif dan mudah dikomunikasikan ke pengguna awam
3. Tidak memerlukan data berlabel karena merupakan algoritma unsupervised
4. Algoritmanya lengkap, sehingga dapat menemukan semua aturan dengan support dan confidence yang ditentukan.

Kelemahan dari algoritma apriori di antaranya:

1. Membutuhkan banyak perhitungan jika itemset sangat besar dan nilai support minimal dipertahankan seminimal mungkin.
2. Perlu melakukan pemindaian penuh seluruh database. Hal ini membuat proses pada algoritma Apriori berjalan sangat lama dan lambat serta menghabiskan banyak sumber daya.

3. Karena perlu melakukan pemindaian penuh, kompleksitas waktu dan ruang dari algoritma apriori adalah $O(2D)$, yang berarti sangat tinggi. Simbol D mewakili lebar horizontal yang ada dalam database

BAB IV

Penerapan Data Mining Untuk Rekomendasi Paket Pernikahan Menggunakan Metode Algoritma Apriori

4.1 Latar Belakang

Pernikahan merupakan tradisi atau adat yang biasanya diikuti oleh pasangan suami istri sebelum memulai sebuah keluarga. Banyak persiapan yang dilakukan sebelum prosesi pernikahan, seperti menyiapkan kartu undangan, menyiapkan timriks pengantin, menyiapkan tempat untuk pernikahan[1].

Wedding Planner adalah perusahaan yang bergerak di bidang jasa, yang menyediakan layanan terpercaya untuk membuat acara pernikahan lebih lancar dan menarik. Layanan ini menawarkan khusus untuk membantu pasangan merencanakan pernikahan, mulai dari tahap perencanaan hingga pelaksanaan rangkaian acara pernikahan[2]. *Wedding planner* pada dasarnya hampir sama dengan *wedding organizer*. Yang membedakannya adalah *wedding planner* penyedia jasa yang bertugas untuk membantu merencanakan segala hal yang berkaitan dengan pernikahan. Mulai dari mencari vendor yang mengurus dekorasi,

biaya, lokasi pesta pernikahan, jadwal acara, dan lain-lain. Sedangkan *wedding organizer* adalah hanya bertugas pada hari pelaksanaan pesta pernikahan. Di zaman yang modern ini banyak calon pengantin yang tidak mau direpotkan dengan berbagai macam persiapan menjelang pernikahan. Oleh karena itu banyak calon pengantin yang membutuhkan jasa *wedding planner*.

Wedding planner akan menawarkan pada calon pengantin berbagai macam paket pernikahan yang ada, yaitu paket *basic*, *premium*, *deluxe*, *ultimate*, dan paket *executive*. Adapun isi paket terdapat pelaminan, *make up wedding*, gaun pengantin, henna, *red carpet*, *photographer*, meja, kursi akad, stand foto, dan lain-lain. Banyak paket yang tersedia membuat calon pengantin bingung untuk memilih paket mana yang akan dipesan. Maka dari itu banyak calon pengantin meminta untuk direkomendasikan paket pernikahan. Dengan adanya rekomendasi bermanfaat bagi calon pengantin terutama bagi calon pengantin yang baru pertama kali memesan. Rekomendasi bermanfaat sebagai referensi bagi calon pengantin yang hendak memesan paket pernikahan. Hasil yang didapatkan dari

rekomendasi tersebut adalah informasi berharga mengenai paket pernikahan apa saja yang sering dipesan berdasarkan histori transaksi *customer-customer* sebelumnya. Untuk mencari rekomendasi paket pernikahan agar lebih jelas dan terarah yaitu dengan cara mengetahui selera calon pengantin, dan *customer* yang dapat dilihat melalui data transaksi pemesanan. Dalam penelitian ini penerapan data mining dibutuhkan untuk mencari jawabannya. Data Mining adalah bidang yang menggabungkan pembelajaran mesin, pengenalan pola, statistik, database, dan visualisasi untuk menemukan informasi dari kumpulan data besar[3]. Data mining juga merupakan proses menemukan pola dan mengekstrak informasi yang berguna dari kumpulan data yang besar, pembelajaran mesin adalah kemampuan komputer untuk menggunakan algoritma yang kompleks dan belajar dari kumpulan data tanpa diprogram secara eksplisit[4]. Ada banyak metode data mining yang saat ini digunakan di berbagai bidang, seperti sistem penyaringan, manajemen analisis risiko, deteksi penipuan, farmasi, *e-commerce*, dan banyak lagi[5]. Data yang diolah

menggunakan metode data mining ini dapat digunakan untuk pengambilan keputusan di masa yang akan datang[6].

Untuk rekomendasi mencari paket pernikahan penelitian ini menggunakan metode algoritma apriori. Algoritma apriori adalah jenis aturan asosiasi data mining. Algoritma ini menemukan kombinasi item dengan nilai frekuensi yang diberikan berdasarkan kriteria yang diinginkan[7]. Algoritma apriori digunakan untuk membantu komputer mempelajari pola aturan asosiasi. Algoritma ini mencari kombinasi item dengan nilai frekuensi tertentu yang memenuhi kriteria atau filter yang diinginkan. Hasil dari algoritma tersebut dapat memberikan pedoman untuk pengambilan keputusan[8].

Dengan menggunakan metode algoritma apriori akan didapatkan rekomendasi paket pernikahan yang sesuai dengan selera calon pengantin atau *customer*.

4.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan di atas, maka dapat ditarik rumusan masalah yaitu: Bagaimana cara menentukan rekomendasi paket pernikahan berdasarkan data

yang didapatkan dari *wedding planner* menggunakan metode algoritma apriori?

4.3 Batasan Masalah

Adapun batasan masalah pada penelitian ini adalah sebagai berikut:

1. Penelitian ini menggunakan sampel data *wedding planner* tahun 2019-2021.
2. Pengujian data menggunakan bahasa pemograman *Python*.
3. Penelitian ini menggunakan metode algoritma apriori.
4. Penelitian ini hanya membahas pemesanan paket pernikahan untuk digedung.

4.4 Tujuan Dan Manfaat Penelitian

4.4.1 Tujuan Penelitian

Adapun tujuan dari penelitian ini yaitu : Untuk mengetahui rekomendasi paket pernikahan berdasarkan data transaksi yang ada metode algoritma apriori serta untuk meningkatkan lagi strategi bisnis dan penjualan pada paket pernikahan lainnya.

4.4.2 Manfaat Penelitian

Adapun manfaat dari penelitian yaitu :

1. Membantu pihak *wedding planner* dalam mengambil keputusan untuk menentukan rekomendasi paket pernikahan
2. Membantu *customer* (calon pengantin) dalam memilih paket wedding terbaik.
3. Dapat menjadi referensi untuk pembaca atau mahasiswa lain agar bisa mengembangkan penelitian ini dan berguna bagi kedepannya.

4.5 Landasan Teori

4.5.1 Data Mining

Data Mining sering juga disebut Knowledge Discovery in Database, adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar. Data mining merupakan proses iterative dan interaktif untuk menemukan pola atau model baru yang lebih (sempurna), bermanfaat dan

dapat dimengerti dalam suatu database yang besar. Hal penting yang terkait dengan data mining.

1. Data mining merupakan suatu proses otomatis terhadap data yang sudah ada.
2. Data yang akan diproses berupa data yang sangat besar.
3. Tujuan data mining adalah untuk mendapatkan hubungan atau pola yang mungkin memberikan indikasi yang bermanfaat.

4.5.2 Algoritma Apriori

Algoritma apriori adalah satu algoritma dasar yang diusulkan oleh Agrawal dan Srikan pada tahun 1994 untuk menemukan frequent itemsets pada aturan asosiasi Boolean. Ide utama pada algoritma apriori adalah : pertama, mencari frequent itemset (himpunan item-item yang memenuhi minimum support.) dari basis data transaksi, kedua - menghilangkan itemset dengan frekuensi yang rendah berdasarkan level minimum support yang telah ditentukan sebelumnya.

Selanjutnya membangun aturan asosiasi dari itemset yang memenuhi nilai minimum confidence dalam basis data

Algoritma apriori atau sering disebut juga dengan analisis asosiasi (association rule mining) adalah teknik data mining untuk menemukan aturan asosiasi antara suatu kombinasi item. Misalkan $I=\{i_1,i_2,i_3,.. i_n\}$ merupakan himpunan items dan D merupakan himpunan database transaksi yang setiap transaksi T merupakan himpunan item yang terkandung dalam transaksi T . Setiap transaksi diasosiasikan dengan identifier yang disebut TID. Aturan yang memenuhi minimum support (min_sup) dan minimum confidence (min_conf) disebut kandidat rule yang kuat. Berdasarkan konvensi, nilai support dan nilai confidence adalah nilai antara 0% sampai 100% sebanding dengan 0 sampai 1,0. Himpunan item disebut dengan itemset. Itemset yang mengandung k items disebut kitemsets. Sebagai contoh himpunan {kopi, gula} merupakan 2-itemset. Kecenderungan kemunculan itemset dalam sejumlah transaksi disebut frequency, support count atau count itemset. Aturan asosiasi dilakukan dengan melalui dua langkah, yaitu :

1. Temukan semua frequent itemset. Hasil dari proses ini adalah sejumlah itemset dengan nilai support lebih besar atau sama dengan minimum support yang diberikan. Nilai support dari itemset $\{A,B\}$ dihitung dengan menggunakan persamaan 1.

$$s = \frac{\text{Frequency}\{A,B\}}{\sum T}$$

2. Menentukan strong rule dari frequent itemset. Hasil dari proses ini adalah semua rule yang memenuhi minimum support dan minimum confidence. Nilai confidence dari suatu kandidat rule $A \diamond B$ dihitung dengan menggunakan persamaan 2.

$$c = \frac{\text{Frequency}\{A,B\}}{\text{Frequency}\{A\}}$$

Untuk membentuk kandidat itemset ada dua proses utama yang dilakukan algoritma

1. Join Step (Penggabungan) Pada proses ini setiap item dikombinasikan dengan item lainnya sampai tidak

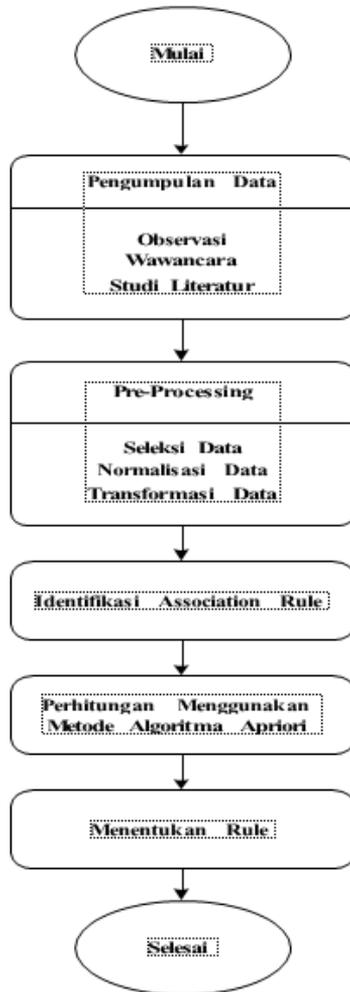
terbentuk kombinasi lagi.

2. Prune Step (Pemangkasan) Pada proses ini, hasil dari item yang dikombinasikan tadi kemudian dipangkas dengan menggunakan minimum support yang telah ditentukan oleh user.

4.6 Metode Penelitian

4.6.1 Prosedur Kerja

Dalam melakukan penelitian ada prosedur kerja yang dapat membantu menyelesaikan penelitian ini dengan baik. Prosedur kerja yang digunakan dalam penelitian ini dapat dilihat pada gambar 2.1 dibawah ini :



Gambar 2.1 Prosedur Kerja

Dari gambar di atas dapat dilihat bahwa ada 5 tahapan

prosedur kerja yang akan dilakukan untuk menyelesaikan penelitian ini meliputi : teknik pengumpulan data melalui observasi, *interview* atau wawancara, dan studi literature, selanjutnya pemrosesan data yaitu dengan cara normalisasi data, seleksi data, dan transformasi data. Tahap selanjutnya melakukan identifikasi penerapan metode *association rule* menggunakan algoritma apriori, tahapan terakhir menentukan rule untuk menemukan jawaban atas permasalahan dari penelitian ini.

4.7 Teknik Pengumpulan Data

Teknik pengumpulan data yang digunakan dalam penelitian adalah sebagai berikut :

4.7.1 Observasi

Pada teknik ini tim peneliti mengamati secara langsung kegiatan yang dilaksanakan pihak *wedding planner* dalam melakukan konsultasi mengenai persiapan acara pernikahan dengan calon pengantin dan survei tempat pelaksanaan

pernikahan.

4.7.2 Interview atau Wawancara

Pada teknik ini tim peneliti melakukan proses tanya jawab (wawancara) dengan pihak - pihak yang bersangkutan langsung dengan objek penelitian. Proses wawancara yang dilakukan oleh tim peneliti membahas seputar masalah penelitian. Contohnya: harga paket pernikahan yang ada di gedung, konsep dekorasi pernikahan, proses transaksi yang dilakukan, dan lain sebagainya.

4.7.3 Studi Literatur

Pada teknik ini tim peneliti membaca, mencatat, dan mempelajari beberapa jurnal, skripsi yang berhubungan dengan objek penelitian.

4.8 Pre-processing

Tahap pra-pemrosesan (*pre-processing*) sangat penting dalam data mining agar mendapatkan hasil yang terbaik. Pra-pemrosesan data membantu algoritma berjalan lebih cepat.

Salah satu langkah pra-pemrosesan adalah normalisasi. Tujuan dari normalisasi data dalam kumpulan data adalah untuk menempatkan data dalam rentang nilai yang sama[11]. Tahap pra-pemrosesan dari penelitian ini dapat dilihat dibawah ini:

4.8.1 Seleksi Data

Tahap seleksi data adalah tahap awal untuk mengubah data mentah yang dikumpulkan dari berbagai sumber menjadi informasi dan bisa digunakan untuk proses selanjutnya. Tahap ini dilakukan agar data yang akan digunakan sesuai dengan kebutuhan.

4.8.2 Normalisasi Data

Normalisasi ialah proses mengukur nilai atribut data sehingga sesuai dengan rentang yang ditentukan[12]. Keuntungan normalisasi adalah sebagai berikut :

1. Normalisasi memastikan bahwa setiap kolom milik tabel yg sah yg ditugaskan & bukan ke tabel lain.
2. Normalisasi menghilangkan penyimpanan fakta yg hiperbola & menyederhanakan akal aplikasi, lantaran

pengembang basis data tidak perlu memikirkan banyak salinan berdasarkan fakta yg sama.

3. Normalisasi memastikan bahwa database hanya berisi satu kolom menggunakan nama & nilai dalam satu waktu[13].

4.8.3 Transformasi Data

Transformation data adalah merupakan proses mentransformasikan data terpilih sebagai akibatnya data tadi sinkron buat proses data mining. Proses transformasi KDD adalah proses kreatif & sangat bergantung dalam jenis atau model informasi yang akan dicari pada database[14].

4.9 Identifikasi Association Rule

Association rule adalah Teknik data mining untuk menemukan aturan asosiatif antara satu kombinasi item. Peraturan Asosiasi memiliki dua tindakan yang relevan, yaitu[15]:

1. *Support* yaitu mengacu pada jumlah nilai penunjang atau persentase campuran media dalam item basis

data.

2. *Confidence* yaitu kuatnya interaksi antar item pada sebuah apriori. *Confidence* dapat dicari sesuai pola frekuensi keluarnya sebuah item ditemukan.

4.10 Perhitungan dengan Menggunakan Metode Algoritma Apriori

Algoritma apriori dibagi menjadi beberapa langkah yang disebut iterasi atau lintasan, yaitu[16]:

1. Pembentukan kandidat *itemset*, kandidat *itemset* terbentuk dari kombinasi $(k-1)$ -*itemset* yang diperoleh dari iterasi sebelumnya.
2. Penghitungan *support* untuk masing-masing kandidat *k-itemset*. *Support* untuk setiap set kandidat *k-item* diperoleh dengan memindai *database* yang bertujuan untuk menghitung jumlah transaksi yang berisi semua *item* dalam kandidat *k-itemset*.
3. Sesuaikan pola frekuensi tinggi. Pola frekuensi tinggi yang megandung *k* item atau *k-itemset* ditentukan dari calon *k-itemset* yang *support*nya lebih besar dari

minimum support.

4. Jika ada pola frekuensi tinggi baru yang ditemukan, maka seluruh proses akan berhenti. Jika tidak, maka k ditambah satu dan kembali ke bagian 1.

Langkah-langkah dari algoritma apriori adalah sebagai berikut[17].

1. Analisa Pola Frekuensi Tinggi

Pada tahap ini mencari kombinasi item yang memenuhi persyaratan minimum dari nilai support dalam database. Nilai support sebuah *item* diperoleh dengan rumus :

$$\text{Support (A)} = \frac{\text{Jumlah transaksi mengandung A}}{\text{Total Transaksi}} \times 100\%$$

[18]

Nilai support dari 2 buah *item* diperoleh dari rumus berikut:

$$\text{Support (A} \cap \text{B)} = \frac{\text{Jumlah transaksi mengandung A dan B}}{\text{Total Transaksi}} \times 100\%$$

2. Pembentukan Aturan Asosiasi

Setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi persyaratan minimum untuk *confidence* dengan menghitung *confidence* aturan asosiatif $A | B$. Nilai *confidence* aturan $A | B$ diperoleh darirumus:

Confidence

$$= P(B|A) \frac{\text{Jumlah transaksi mengandung A dan B}}{\text{Total Transaksi}} \times 100\%$$

Setelah tahap *frequent itemset* selanjutnya dilakukan tahap menentukan *rule* yaitu mencari *confidence* dari *frequent itemset*.

4.11 Pembahasan dan Hasil

4.11.1 Pembahasan

Penelitian ini didasari oleh sudut pandang tentang banyaknya atau tingginya angka jumlah pernikahan yang terjadi pada saat ini. Di era globalisasi saat ini pernikahan mengalami perkembangan yang sangat pesat, banyak hal-hal baru yang ada. Contohnya: paket pernikahan. Paket pernikahan adalah sekumpulan hal-hal yang diperlukan dalam suatu acara pernikahan. Banyaknya paket yang tersedia membuat calon pengantin atau *customer* bingung untuk menentukan paket pernikahan mana yang cocok untuk pernikahannya. Oleh karena dibutuhkan sebuah rekomendasi untuk membantu calon pengantin menentukan paket pernikahannya.

Penelitian dilakukan pada *SM Wedding Decoration*, *SM Wedding Decoration* adalah tempat yang menyediakan jasa untuk mengurus segala hal yang berhubungan dengan pernikahan. Contohnya *wedding decoration*, *wedding organizer*, dan *wedding planner*.

SM Wedding Decoration memiliki beberapa paket pernikahan yang dapat ditawarkan pada *customer*. Paket

pernikahan tersebut terdiri dari paket *basic* sampai paket *executive*. Berikut ini adalah tabel nama-nama paket pernikahan yang ada di *SM Wedding Decoration*.

Tabel 5.1 Nama-nama paket pernikahan

Id Items	Items	Harga
C001	Paket Basic	Rp 7,500,000
C002	Paket Premium	Rp 9,500,000
C003	Paket Deluxe	Rp 13,000,000
C004	Paket Ultimate	Rp 17,000,000
C005	Paket Executive	Rp 23,000,000
C006	Fotografer & Dokumentasi	Rp 3,000,000
C007	Henna Art	Rp 350,000
C008	MUA	Rp 6,000,000

Untuk membuat sebuah rekomendasi paket pernikahan, penelitian ini melakukan pengumpulan data. Dari proses pengumpulan data tersebut, penelitian ini memakai arsip data transaksi pemesanan paket pernikahan dari tanggal 06 Januari 2019 sampai tanggal 5 Desember 2021. Arsip data transaksi

No Transaksi	Hari/Tanggal Acara	Nama Customer	ID Customer	Warna Dekorasi	Paket Pemesanan Gedung	DP-1	DP-2	DP-3	TOTAL	
0	A001	Minggu, 06 Januari 2019	Dina	0001	1. Gold v2, Peach v2, Green	Paket Premium	Rp3.200.000	Rp3.225.000	Rp3.325.000	Rp9.750.000
1	A002	Jumat, 11 Januari 2019	Agus	0002	1. Cinnamon v2, Yellow v3, Beige	Paket Deluxe	Rp3.900.000	Rp4.950.000	Rp4.950.000	Rp13.800.000
2	A003	Sabtu, 12 Januari 2019	Yanto	0003	1. Navy v2, Yellow v3, Blue	Paket Basic, Henna Art	Rp2.300.000	Rp2.747.500	Rp2.747.500	Rp7.800.000
3	A004	Sabtu, 12 Januari 2019	Dani	0004	1. Saki v2, Pale Green v3, White	Paket Deluxe	Rp3.900.000	Rp4.950.000	Rp4.950.000	Rp13.800.000
4	A005	Minggu, 13 Januari 2019	Dhina	0005	1. Black v2, Gold v3, Gray	Paket Premium, Fotografi	Rp3.700.000	Rp4.375.000	Rp4.375.000	Rp12.450.000
...	
292	A293	Senin, 22 November 2021	Nino	0293	1. Navy v2, Yellow v3, Blue	Paket Deluxe, Fotografi	Rp4.800.000	Rp5.900.000	Rp5.900.000	Rp16.600.000
293	A294	Minggu, 28 November 2021	Fitri	0294	1. Cream v2, Brown v3, Ruby	Paket Premium, Henna Art, MUA	Rp4.700.000	Rp5.847.500	Rp5.847.500	Rp16.395.000
294	A295	Selasa, 30 November 2021	Vita	0295	1. Cinnamon v2, Yellow v3, Beige	Paket Deluxe, MUA	Rp5.700.000	Rp6.900.000	Rp6.900.000	Rp19.500.000
295	A296	Kamis, 2 Desember 2021	Fitria	0296	1. Navy v2, Eggplant v2, Gold	Paket Deluxe	Rp3.900.000	Rp4.950.000	Rp4.950.000	Rp13.800.000
296	A297	Minggu, 5 Desember 2021	Andi	0297	1. French Rose v2, Orange v2, Peach	Paket Deluxe, Henna Art	Rp4.000.000	Rp4.875.000	Rp4.875.000	Rp13.750.000

tersebut memiliki jumlah transaksi sebanyak 297 data. Arsip data transaksi dapat dilihat pada gambar 3.1 :

Gambar 5.1 Arsip data transaksi pemesanan paket pernikahan

Berdasarkan gambar diatas dapat diketahui bagaimana arsip data transaksi, didalamnya terdapat kode transaksi, hari/tanggal, nama customer, id customer, warna dekorasi,

nama paket yang dipesan, dan total harga. Arsip data transaksi diatas masih dalam bentuk data mentah yang belum diolah, didalamnya terdapat 2 pemesanan item tapi masih dalam 1 transaksi. Untuk itu diperlukannya normalisasi data agar tidak terdapat grouping data atau data yang sama dalam 1 transaksi.

4.12 Pengolahan data

4.12.1 Normalisasi

Pada langkah awal ini dilakukan normalisasi data, dimana arsip data transaksi pemesanan paket pernikahan yang telah dikumpulkan sebelumnya dinormalisasikan atau dihilangkan *grouping* data atau data yang ada duplikasinya. Jadi dimana maksud pada tahap ini yaitu dalam satu transaksi paket pernikahan dijabarkan apa saja paket pernikahan yang dipesan. Normalisasi data transaksi pemesanan paket pernikahan dapat dilihat pada gambar 3.2:

No Transaksi	Nama Customer	ID Customer	ID Warna	Warna Tams	Hari/Tanggal Acara	ID Items	Items	Harga	Keterangan	
0	A001	Gina	B001	D004	1. Gold in2. Peach in3. Green	Minggu, 06 Januari 2019	C002	Paket Premium	Rp9,500,000	Lunas
1	A002	Agus	B002	D002	1. Cinnamon in2. Yellow in3. Beige	Jumat, 11 Januari 2019	C003	Paket Deluxe	Rp13,000,000	Lunas
2	A003	Yanto	B003	D003	1. Navy in2. Yellow in3. Blue	Sabu, 12 Januari 2019	C001	Paket Basic	Rp7,500,000	Lunas
3	A003	Yanto	B003	D003	1. Navy in2. Yellow in3. Blue	Sabu, 12 Januari 2019	C007	Henna Art	Rp350,000	Lunas
4	A004	David	B004	D001	1. Black in2. Pale Green in3. White	Sabu, 12 Januari 2019	C003	Paket Deluxe	Rp13,000,000	Lunas
...
521	A295	Yola	B295	D002	1. Cinnamon in2. Yellow in3. Beige	Selasa, 30 November 2021	C003	Paket Deluxe	Rp13,000,000	Lunas
522	A295	Yola	B295	D002	1. Cinnamon in2. Yellow in3. Beige	Selasa, 30 November 2021	C008	MUA	Rp6,000,000	Lunas
523	A296	Priskila	B296	D006	1. Navy in2. Eggplant in3. Gold	Kamis, 2 Desember 2021	C003	Paket Deluxe	Rp13,000,000	Lunas
524	A297	Abdi	B297	D006	1. French Rose in2. Orange in3. Peach	Minggu, 8 Desember 2021	C003	Paket Deluxe	Rp13,000,000	Lunas
525	A297	Abdi	B297	D006	1. French Rose in2. Orange in3. Peach	Minggu, 8 Desember 2021	C007	Henna Art	Rp350,000	Lunas

Gambar 5.2 Normalisasi data transaksi

4.12.2 Seleksi data

Setelah dilakukan normalisasi, tahap selanjutnya yaitu seleksi data. Dimana pada tahap ini menghilangkan atau membuang atribut-atribut yang tidak ada kaitannya atau yang tidak diperlukan untuk proses selanjutnya. Seleksi data untuk

No Transaksi	Nama Customer	ID Customer	ID Items	Items	Harga	
0	A001	Gina	B001	C002	Paket Premium	Rp9,500,000
1	A002	Agus	B002	C003	Paket Deluxe	Rp13,000,000
2	A003	Yanto	B003	C001	Paket Basic	Rp7,500,000
3	A003	Yanto	B003	C007	Henna Art	Rp350,000
4	A004	David	B004	C003	Paket Deluxe	Rp13,000,000
...
521	A295	Yola	B295	C003	Paket Deluxe	Rp13,000,000
522	A295	Yola	B295	C008	MUA	Rp6,000,000
523	A296	Priskila	B296	C003	Paket Deluxe	Rp13,000,000
524	A297	Abdi	B297	C003	Paket Deluxe	Rp13,000,000
525	A297	Abdi	B297	C007	Henna Art	Rp350,000

arsip data transaksi paket pernikahan dapat dilihat pada gambar 3.3:

Gambar 5.3 Seleksi data transaksi

Dari gambar 3.3, dapat dilihat beberapa perubahan dari data sebelumnya yaitu pada sebelum seleksi data terdapat no transaksi, hari/tanggal acara, id *customer*, nama *customer*, warna dekorasi, id *items*, *items* dan harga. Setelah dilakukan seleksi data maka perubahan yang dapat dilihat tidak adanya hari/tanggal acara dan warna dekorasi. Dihilangkannya atau dibuang atribut tersebut karena tidak ada kaitannya untuk proses selanjutnya dan mempermudah mengolah data selanjutnya

4.12.3 Transformasi data

Tahap terakhir pada pengolahan data dalam penelitian ini yaitu transformasi data. Dimana pada tahap ini data yang sudah di normalisasikan dan sudah di seleksi, di transformasikan dalam bentuk *binner*. Transformasi data untuk data transaksi paket pernikahan dapat dilihat pada gambar 3.4:

0	NaN	Paket Basic	Paket Premium	Paket Deluxe	Paket Ultimate	Paket Executive	Fotografer & Dokumentasi	Henna Art	MUA
1	1	0	1	0	0	0	0	0	0
2	2	0	0	1	0	0	0	0	0
3	3	1	0	0	0	0	0	1	0
4	4	0	0	1	0	0	0	0	0
--	--	--	--	--	--	--	--	--	--
294	294	0	1	0	0	0	0	1	1
295	295	0	0	1	0	0	0	0	1
296	296	0	0	1	0	0	0	0	0
297	297	0	0	1	0	0	0	1	0
298	JUMLAH = 297	47	85	98	40	27	75	77	77

Gambar 5.4 Transformasi data transaksi

4.13 Implementasi algoritma apriori

Setelah melakukan pengolahan data, tahap selanjutnya data transaksi pemesanan paket pernikahan diimplementasikan ke dalam algoritma apriori. Untuk mendapatkan rekomendasi paket pernikahan, tahap ini menggunakan perhitungan *association rule* dalam menemukan jawabannya atau pembuktiannya.

1. Pembentukan 1 itemset

Langkah awal yang dilakukan yaitu membentuk K_1 atau 1 *itemset*. Sebelum melakukan perhitungan K_1 , tentukan terlebih dahulu nilai minimum *support* dan minimum *confidence*. Dalam

penelitian ini nilai minimum *support* yaitu 3,5% dan minimum *confidence* yaitu 40%. Setelah menentukan minimum *support* dan *confidence*, selanjutnya hitung *support* K_1 setiap *item*nya. Berikut adalah contoh caramenghitung nilai *support* K_1 :

$$\text{Support } (K) = \frac{\sum \text{Transaksi mengandung } K}{\sum \text{Total transaksi}} \times 100\%$$

$$\text{Support } (\text{Paket basic}) = \frac{47}{297} \times 100\%$$

$$\text{Support } (\text{Paket basic}) = 15,8\%$$

Tabel untuk nilai *support* setiap *items* dapat dilihat pada tabel 3.2:

Tabel 5.2 Perhitungan 1 *itemset*

No	Items	Jumlah Transaksi Yang Mengandung Item	Support
1	Paket Basic	47	15.8%
2	Paket Premium	85	28.6%
3	Paket Deluxe	98	32.9%
4	Paket Ultimate	40	13.4%
5	Paket Executive	27	9%
6	Fotografer & Dokumentasi	75	25.2%
7	Henna Art	77	25.9%
8	MUA	77	25.9%

Dari tabel diatas, dapat dilihat nilai *support* untuk setiap *items*. Setelah melakukan perhitungan nilai *support*, setiap *items* memenuhi nilai *support* yang telah ditentukan. Oleh karena itu, perhitungannya dapat dilanjut untuk menghitung kombinasi 2 *items*.

2. Pembentukan 2 itemset

Setelah membentuk K_1 atau 1 *itemset*, langkah selanjutnya membentuk K_2 atau 2 *itemset*. Pada tahap ini yang dilakukan yaitu mengkombinasikan semua *item* menjadi 2 kombinasi, contohnya *item* yang mengandung A di kombinasikan dengan yang mengandung *item* B. Setelah itu hitung nilai *support* untuk setiap kombinasi.

Berikut adalah contoh cara menghitung nilai *support* K_2 :

$$\text{Support}(A, B) = \frac{\sum \text{Transaksi mengandung A dan B}}{\sum \text{Total transaksi}} \times 100\%$$

$$\text{Support}(\text{Paket basic, Fotografer}) = \frac{9}{297} \times 100\%$$

$$\text{Support}(\text{Paket basic, Fotografer}) = 3\%$$

Setelah melakukan perhitungan nilai *support*, Jadi hasil perhitungan 2 *itemset* dapat dilihat pada tabel 3.3:

Tabel 5.3 Hasil Perhitungan 2 *itemset*

No	Items	Jumlah Transaksi Yang Mengandung	Support
			t

		Item	
1	Paket Basic, Henna Art	17	5.7%
2	Paket Premium, Fotografer & Dokumentasi	20	6.7%
3	Paket Premium, Henna Art	20	6.7%
4	Paket Premium, MUA	32	10.7%
5	Paket Deluxe, Fotografer & Dokumentasi	31	10.4%
6	Paket Deluxe, Henna Art	27	9%
7	Paket Deluxe, MUA	32	10.7%
8	Fotografer & Dokumentasi, Henna Art	19	6.3%
9	Fotografer & Dokumentasi, MUA	25	8.4%
10	Henna Art, MUA	23	7.7%

3. Pembentukan 3 itemset

Langkah selanjutnya adalah membentuk K_3 atau 3 *itemset*. Tahap ini hampir sama dengan membentuk K_2 atau 2 *itemset*, hanya saja yang membedakannya pada tahap ini yang dikombinasikan adalah 3 *itemset*. contohnya *item* yang mengandung A dan B di kombinasikan dengan yang mengandung *item* C. Setelah itu hitung nilai *support* untuk setiap kombinasi. Berikut adalah contoh cara menghitung nilai *support* K_3 :

$$\text{Support}(A, B, C) = \frac{\sum \text{Transaksi mengandung } A, B \text{ dan } C}{\sum \text{Total transaksi}} \times 100\%$$

$$\text{Support}(\text{Paket basic, Hemat Art, MUA}) = \frac{2}{297} \times 100\%$$

$$\text{Support}(\text{PPaket basic, Hemat Art, MUA}) = 0,67\%$$

Setelah melakukan perhitungan nilai *support*, jadi hasil perhitungan 3 *itemset* dapat dilihat pada tabel 3.4:

Tabel 5.4 Hasil Perhitungan 3
itemset

No	Items	Jumlah Transaksi Yang Mengandung Item	Support †
1	Paket Deluxe, Fotografer & Dokumentasi, MUA	11	3.7%

Dari tabel diatas, dapat dilihat hasil dari perhitungan 3 kombinasi atau pembentukan 3 *itemset* ditemukan 1 aturan asosiasi. Karena tidak bisa membentuk kombinasi lagi, maka proses perhitungan diberhentikan sampai pembentukan 3 *itemset*.

4.14 Pembentukan *rule* atau aturan asosiasi

Pada tahap ini yang dilakukan yaitu membentuk sebuah *rule* atau aturan asosiasi untuk menghitung nilai *confidence*. Nilai *confidence* memiliki *rule* atau aturan ($A \rightarrow B$) atau jika A maka B. Berikut adalah contoh cara menghitung nilai *confidence* :

$$Confidence (A \rightarrow B) = \frac{Support (A \rightarrow B)}{Support (A)} \times 100\%$$

$$\text{Confidence } (A \rightarrow B) = \frac{0,057}{0,158} \times 100\%$$

$$\text{Confidence } (A \rightarrow B) = 36\%$$

Tabel pembentukan aturan asosiasi dapat dilihat pada tabel 3.5:

Tabel 5.5 Pembentukan aturan asosiasi dari 2 itemset

No	Rule	Support	Confidence
1	Paket Basic → Henna Art	5.7%	36%
2	Henna Art → Paket Basic	5.7%	22%
3	Paket Premium → Fotografer & Dokumentasi	6.7%	23.4%
4	Fotografer & Dokumentasi → Paket Premium	6.7%	26.5%
5	Paket Premium → Henna Art	6.7%	23.4%
6	Henna Art → Paket Premium	6.7%	25.8%
7	Paket Premium → MUA	10.7%	37.4%
8	MUA → Paket Premium	10.7%	41.3%
9	Paket Deluxe → Fotografer & Dokumentasi	10.4%	31.6%
10	Fotografer & Dokumentasi → Paket Deluxe	10.4%	41.2%
11	Paket Deluxe → Henna Art	9%	27.3%
12	Henna Art → Paket Deluxe	9%	34.7%

13	Paket Deluxe → MUA	10.7%	32.5%
14	MUA → Paket Deluxe	10.7%	41.3%
15	Fotografer & Dokumentasi → Henna Art	6.3%	25%
16	Henna Art → Fotografer & Dokumentasi	6.3%	24.3%
17	Fotografer & Dokumentasi → MUA	8.4%	33,3%
18	MUA → Fotografer & Dokumentasi	8.4%	32.4%
19	Henna Art → MUA	7.7%	29.7%
20	MUA → Henna Art	7.7%	29.7%

Berikut ini adalah hasil penjabaran nilai *support* dan nilai *confidence* dari Pembentukan aturan asosiasi 2 itemset :

Tabel 5.6 Pembentukan aturan asosiasi dari 3 itemset

No	Rule	Support	Confidence
1	Paket Deluxe, Fotografer & Dokumentasi → MUA	3.7%	35.5%
2	Paket Deluxe, MUA → Fotografer & Dokumentasi	3.7%	34.5%
3	Fotografer & Dokumentasi, MUA → Paket Deluxe	3.7%	44%
4	Fotografer & Dokumentasi, Paket Deluxe → MUA	3.7%	35.5%
5	MUA, Paket Deluxe → Fotografer & Dokumentasi	3.7%	34.4%
6	MUA, Fotografer & Dokumentasi →	3.7%	44%

	Paket Deluxe		
--	--------------	--	--

Setelah melakukan perhitungan nilai *confidence*, banyak kombinasi yang tidak memenuhi nilai minimum *confidence* yang telah ditentukan. Nilai *confidence* yang telah ditentukan yaitu 40%, oleh karena itu, untuk kombinasi yang tidak memenuhi nilai minimum *confidence* akan dihapus. Jadi hasil perhitungan aturan asosiasi dapat dilihat pada tabel dibawah ini:

Tabel 5.7 Hasil Pembentukan aturan asosiasi

No	Rule	Support	Confidence
1	MUA → Paket Premium	10.7%	41.3%
2	Fotografer & Dokumentasi → Paket Deluxe	10.4%	41.2%
3	MUA → Paket Deluxe	10.7%	41.3%
4	Fotografer & Dokumentasi, MUA → Paket Deluxe	3.7%	44%

Dari perhitungan aturan asosiasi yang telah dilakukan terbentuk 4 aturan asosiasi, rule atau aturan asosiasi yang terbentuk dengan minimum *support* 3.5% dan *confidence* 40% yang dapat dijadikan rekomendasi paket pernikahan dapat dilihat dibawah ini:

1. Jika memesan jasa MUA maka akan memesan juga paket pernikahan yang *Paket Premium*.
2. Jika memesan jasa Fotografer & Dokumentasi maka akan memesan juga paket pernikahan yang *Paket Deluxe*.
3. Jika memesan jasa MUA maka akan memesan juga paket pernikahan yang *Paket Deluxe*.
4. Jika memesan jasa Fotografer & Dokumentasi dan MUA maka akan memesan juga paket pernikahan yang *Paket Deluxe*.

4.15 Impelentasi Apriori pada Python

Pada tahap ini yang dilakukan adalah mengimport *library* *pandas* dan dilanjutkan dengan membaca dataset excel. Hasil pembacaan disajikan pada gambar 3.

No Transaksi	Hari/Tanggal Acara	Nama Customer	ID Customer	Warna Dekorasi	Paket Pemesanan Gedung	DP-1
A001	Minggu, 06 Januari 2019	Gina	B001	1. Gold \n2. Peach \n3. Green	Paket Premium	Rp2,850,000
A002	Jumat, 11 Januari 2019	Agus	B002	1. Cinnamon \n2. Yellow \n3. Beige	Paket Deluxe	Rp3,900,000
A003	Sabtu, 12 Januari 2019	Yanto	B003	1. Navy \n2. Yellow \n3. Blue	Paket Basic , Henna Art	Rp2,355,000
A004	Sabtu, 12 Januari 2019	David	B004	1. Black \n2. Pale Green \n3. White	Paket Deluxe	Rp3,900,000
A005	Minggu, 13 Januari 2019	Elvina	B005	1. Black \n2. Gold \n3. Gray	Paket Premium , Fotografer	Rp3,750,000
...
A293	Senin, 22 November 2021	Nino	B293	1. Navy \n2. Yellow \n3. Blue	Paket Deluxe , Fotografer	Rp4,800,000
A294	Minggu, 28 November 2021	Pingki	B294	1. Green \n2. Brown \n3. Putih	Paket Premium , Henna Art , MUA	Rp4,755,000
A295	Selasa, 30 November 2021	Yola	B295	1. Cinnamon \n2. Yellow \n3. Beige	Paket Deluxe , MUA	Rp5,700,000
A296	Kamis, 2 Desember 2021	Priskila	B296	1. Navy \n2. Eggplant \n3. Gold	Paket Deluxe	Rp3,900,000
A297	Minggu, 5 Desember 2021	Abdi	B297	1. French Rose \n2. Orange \n3. Peach	Paket Deluxe , Henna Art	Rp4,005,000

Gambar 3. Output Read File

Setelah dilakukan *read* dataset maka selanjutnya diterapkan *apriori* dengan *Python* menggunakan nilai minimal *support* sebesar 35% dengan hasil seperti pada gambar 4 dengan menggunakan script berikut ini.

```
frequent_itemsets =
apriori(AllPackage,min_support = 0.035, use_colnames=True)
frequent_itemsets
```

	support	itemsets
0	0.252525	(Fotografer & Dokumentasi)
1	0.259259	(Henna Art)
2	0.259259	(MUA)
3	0.158249	(Paket Basic)
4	0.329986	(Paket Deluxe)
5	0.090909	(Paket Executive)
6	0.288195	(Paket Premium)
7	0.134680	(Paket Ultimate)
8	0.063973	(Henna Art, Fotografer & Dokumentasi)
9	0.084175	(MUA, Fotografer & Dokumentasi)
10	0.104377	(Paket Deluxe, Fotografer & Dokumentasi)
11	0.087340	(Fotografer & Dokumentasi, Paket Premium)
12	0.077441	(Henna Art, MUA)
13	0.057239	(Henna Art, Paket Basic)
14	0.090909	(Henna Art, Paket Deluxe)
15	0.087340	(Henna Art, Paket Premium)
16	0.107744	(Paket Deluxe, MUA)
17	0.107744	(MUA, Paket Premium)
18	0.037037	(MUA, Paket Deluxe, Fotografer & Dokumentasi)

Gambar 4. Output Nilai Support

Langkah terakhir dilakukan penerapan *Apriori* dengan *Python* untuk menentukan minimum *threshold Confidence* dengan nilai minimal sebesar 35% dan hasilnya seperti tersaji pada gambar

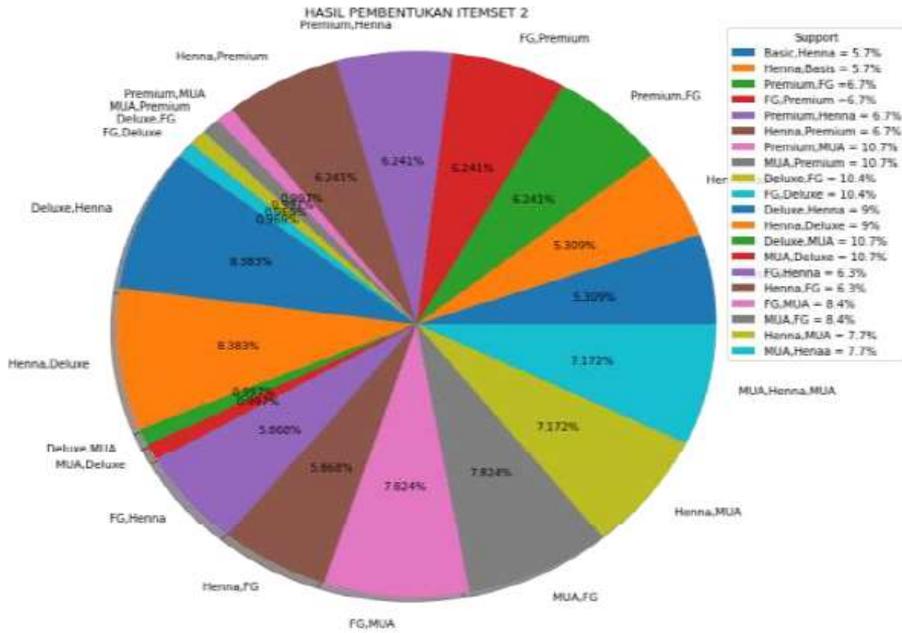
5

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift
3	(Fotografer & Dokumentasi)	(Paket Deluxe)	0.252525	0.329966	0.104377	0.413333	1.252653
11	(MUA)	(Paket Deluxe)	0.259259	0.329966	0.107744	0.415584	1.259475
12	(MUA)	(Paket Premium)	0.259259	0.286195	0.107744	0.415584	1.452101
15	(Fotografer & Dokumentasi, MUA)	(Paket Deluxe)	0.084175	0.329966	0.037037	0.440000	1.333469

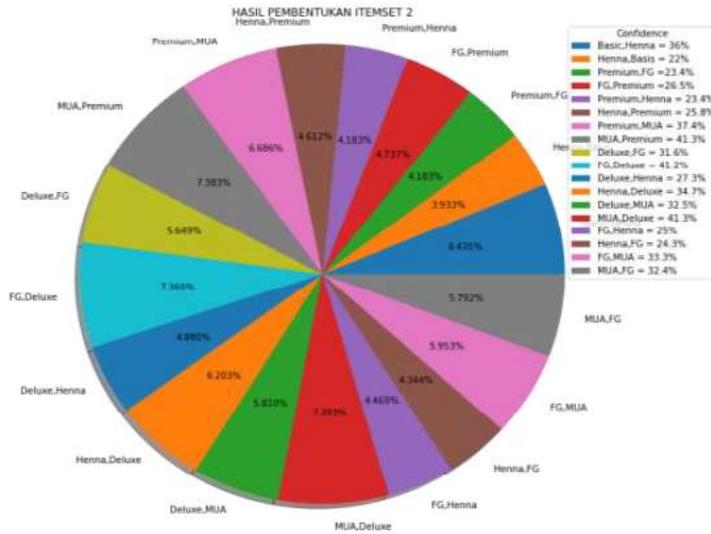
Gambar 5. Output Nilai Confidence

4.16 Pembentukan Rule Asosiasi

Pada tahap ini dibentuk sebuah *rule* atau perintah asosiasi untuk menentukan nilai *confidence*. Gambar dari hasil penjabaran nilai *support* dan nilai *confidence* dari pembentukan aturan asosiasi 2 *itemset* ditunjukkan pada gambar 6 dan gambar 7.

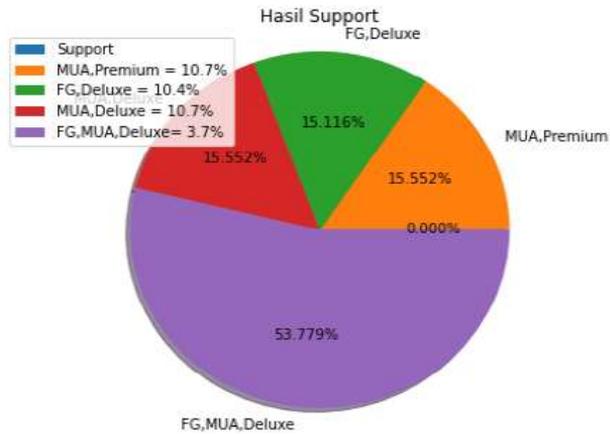


Gambar 6. Nilai support 2 itemset

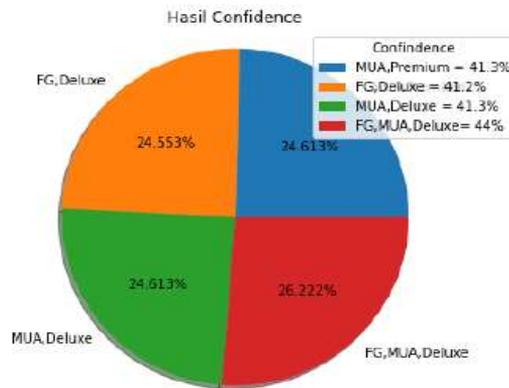


Gambar 7. Nilai Confidence 2 itemset

Selanjutnya dilakukan perhitungan juga untuk melihat nilai *confidence* untuk pembentukan aturan asosiasi 3 itemset.



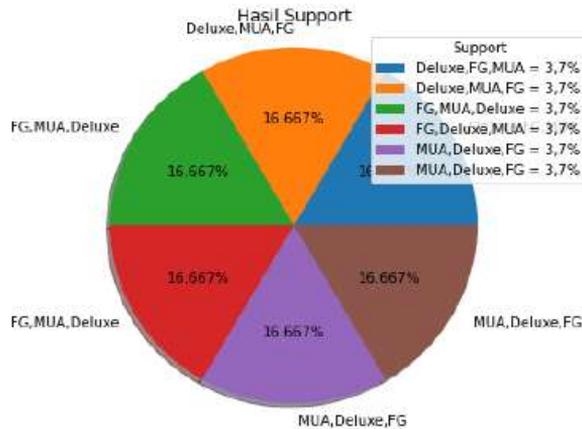
Gambar 8. Nilai *Support* 3 itemset



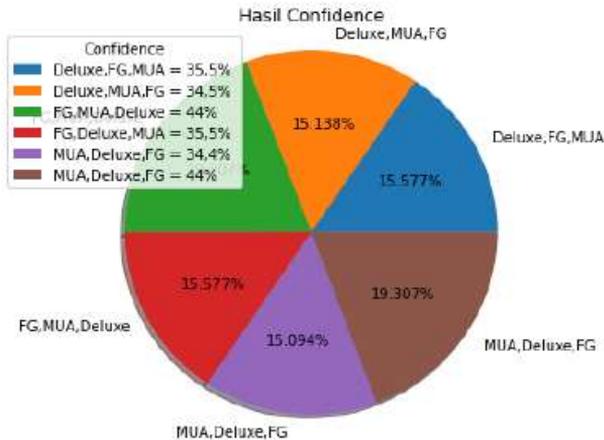
Gambar 9. Nilai *Confidence* 3 itemset

Setelah melakukan perhitungan nilai *confidence*, banyak kombinasi yang tidak memenuhi nilai minimal *confidence* yang telah diterapkan. Nilai *confidence* yang telah diterapkan yaitu

40%, karena itu, untuk kombinasi yang tidak memandai nilai minimal *confidence* akan dihilangkan. Hasil *confidence* untuk pembentukan aturan asosiasi ditunjukkan pada gambar 10.



Gambar 10. Hasil Support Aturan Asosiasi



Gambar 11. Hasil Confidence Aturan Asosiasi

Dari perhitungan dapat ditemukan 4 perintah asosiasi yang telah dihasilkan, dimana terdapat 3 perintah asosiasi yang diperoleh didalam kumpulan *itemset* 2 dan 1 perintah asosiasi yang terbentuk dari kumpulan 3 *itemset*.

Berikut ini rule atau perintah asosiasi yang diterapkan pada rekomendasi paket pernikahan dengan minimal *support* 3.5% dan *confidence* 40%:

1. Jika memesan jasa MUA maka akan memesan juga paket pernikahan yang Paket *Premium*.
2. Jika memesan jasa Fotografer & Dokumentasi maka akan memesan juga paket pernikahan yang Paket *Deluxe*.
3. Jika memesan jasa MUA maka akan memesan juga paket pernikahan yang Paket *Deluxe*.
4. Jika memesan jasa Fotografer & Dokumentasi dan MUA maka akan memesan juga paket pernikahan yang Paket *Deluxe*.

4.17 Kesimpulan Dan Saran

4.17.1 Kesimpulan

Berdasarkan pembahasan yang telah diuraikan sebelumnya dapat dibuat kesimpulan sebagai berikut:

1. Dengan adanya rekomendasi paket pernikahan bermanfaat sebagai referensi bagi *customer* yang hendak memesan paket pernikahan.
2. Hasil yang didapatkan dari rekomendasi tersebut adalah informasi mengenai paket pernikahan apa saja yang sering dipesan berdasarkan histori transaksi *customer* sebelumnya.
3. Dengan menggunakan data mining metode *association rule* untuk mengidentifikasi hubungan antara beberapa item, didapatkan rule yang dapat di jadikan rekomendasi paket pernikahan sesuai dengan selera *customer*.
4. Pemesanan untuk paket Fotografer & Dokumentasi, MUA → Paket *Deluxe* lebih sering dipesan *customer* dan signifikan terhadap data transaksi pemesanan paket sebesar 44%. Untuk paket MUA→Paket *Deluxe*

terhadap data transaksi pemesanan paket sebesar 41.3%. Untuk paket Fotografer & Dokumentasi→Paket *Deluxe* terhadap data transaksi pemesanan paket sebesar 41.2%. Dan untuk paket MUA→Paket *Premium Deluxe* terhadap data transaksi pemesanan paket sebesar 41.3%.

4.17.2 Saran

Berdasarkan kesimpulan penelitian diatas, adapun saran yang dapat disampaikan sebagai berikut:

1. Dengan adanya metode *association rule* membantu pihak *Wedding* dalam hal merekomendasikan paket pernikahan, juga dapat dijadikan referensi dalam meningkatkan paket pernikahan yang lain.
2. Selanjutnya penelitian ini dapat dilanjutkan dengan menggunakan metode lain untuk melihat lebih banyak rekomendasi berdasarkan transaksi setiap bulan, oleh karena itu setiap bulannya rekomendasi paket pernikahan akan berganti.

DAFTAR PUSTAKA

- [1] J. Sofian and H. H. Solihin. 2018. "Pembangunan Aplikasi M-Wedding Sebagai Solusi Usaha Wedding Organizer," *J. Teknol. dan Inf.*, vol. 8, no. 2.
- [2] A. M. Muzaiyin. 2021. "Analisis Sistem Pelayanan Jasa Wedding Organizer Ditinjau Dari Etika Bisnis Islam (Studi Kasus Di UD. Santoso Wedding Organizer)," *Fenomena*, vol. 20, no. 1, 129-144.
- [3] P. Kasih. 2019. "Pemodelan Data Mining Decision Tree Dengan Classification Error Untuk Seleksi Calon Anggota Tim Paduan Suara," *Innov. Res. Informatics*, vol. 1, no. 2, 63-69.
- [4] J. Gladju and B. Sam. 2022. "Smart Agricultural Technology Aplikasi Kerangka Kerja Penambangan Data Dan Pembelajaran Mesin Dalam Akuakultur Dan Perikanan : Tinjauan," vol. 2.
- [5] T. Uçar and A. Karahoca. 2021. "Benchmarking Data Mining Approaches For Traveler Segmentation," *Int. J. Electr. Comput. Eng.*, vol. 11, no. 1, 409-415.
- [6] P. Prahasti. 2018. "Data Mining Dalam Pengelompokan Jenis Dan Jumlah Pembagian Zakat Dengan Menggunakan Metode Clustering K-Means (Studi Kasus: Badan Amil Zakat Kota Bengkulu)," *J. Teknol. Inf.*, vol. 1, no. 2, 211.
- [7] N. Azwanti. 2019. "Combination, Pola, Apriori ANALISIS POLA BELANJA KONSUMEN MENGGUNAKAN

- ALGORITMA APRIORI PADA RAFFA PHOTOCOPY," *J. Teknol. Dan Open Source*, vol. 2, no. 2, 63-73.
- [8] S. Nurajizah. 2019. "Analisa Transaksi Penjualan Obat menggunakan Algoritma Apriori," *INOVTEK Polbeng - Seri Inform.*, vol. 4, no. 1, 35.
- [9] M. Muhammad and Y. Kusnadi, "APLIKASI SISTEM INFORMASI PEMESANAN PAKET PERNIKAHAN BERBASIS WEB (Studi Kasus: SANGGAR LARAS WEDDING BOGOR)," *J. Teknol. Inform. dan Komput.*, vol. 4, no. 2, pp. 19-22, 2018, doi: 10.37012/jtik.v4i2.257.
- [10] E. S. Subrata, S. M. Kuway, and I. D. Ayu Eka Yuliani, "Pemilihan Paket Wedding Menggunakan Metode Electre," *Sisfotenika*, vol. 10, no. 1, p. 103, 2020, doi: 10.30700/jst.v10i1.949.
- [11] Ahmad Harmain, P. Paiman, H. Kurniawan, K. Kusrini, and Dina Maulina. 2022. "Normalisasi Data Untuk Efisiensi K-Means Pada Pengelompokan Wilayah Berpotensi Kebakaran Hutan Dan Lahan Berdasarkan Sebaran Titik Panas," *Tek. Teknol. Inf. dan Multimed.*, vol. 2, no. 2, 83-89.
- [12] D. A. Nasution, H. H. Khotimah, and N. Chamidah. 2019. "Perbandingan Normalisasi Data untuk Klasifikasi Wine Menggunakan Algoritma K-NN," *Comput. Eng. Sci. Syst. J.*, vol. 4, no. 1, 78.
- [13] S. Mulyati, B. A. Sujatmoko, T. I. M. Wira, R. Afif, and R. A. Pratama. 2018. "Normalisasi Database Dan Migrasi

Database Untuk Memudahkan Manajemen Data," *Sebatik*, vol. 22, no. 2, 124-129.

- [14] Desyanti. and F. Sari. 2019. "Penerapan Algoritma Apriori Terhadap Perawatan Tubuh di Kakiku," *SATIN - Sains dan Teknol. Inf.*, vol. 5, no. 1, 51-59.
- [15] A. F. Lestari and M. Hafiz. 2020. "Penerapan Algoritma Apriori Pada Data Penjualan Barbar Warehouse," *INOVTEK Polbeng - Seri Inform.*, vol. 5, no. 1, 96.
- [16] I. Wahyudi, S. Bahri, and P. Handayani. 2019. "Aplikasi Pembelajaran Pengenalan Budaya Indonesia," vol. V, no. 1, 135-138

- [17] I. Ismasari, M. Ramadhan, and W. Hadikristanto. 2020. "Analisis Tingkat Pembelian Konsumen dengan Algoritma Apriori," *J. Ilm. Intech Inf. Technol. J. UMUS*, vol. 2, no. 02.
- [18] Sanjani, H. Fahmi, and A. Sindar. 2019. "Implementasi Data Mining Penjualan Produk Pakaian Dengan Algoritma Apriori," *IJAI (Indonesian Journal of Applied Informatics)*, vol. 4, no. 1, 23-