

SYSTEMIC: Information System and Informatics Journal

ISSN: 2460-8092, 2548-6551 (e)

Vol 6 No 2 – Desember 2020

Pemanfaatan Algoritma FP-Growth Untuk Menentukan Strategi Penjualan Pada Kedai Kopi Teras Garden

Adrian Marvel Ugrasena¹, Achmad Zakki Falani²^{1,2}) Universitas Narotama Surabayaadrian.marvel@mhs.fasilkom.narotama.ac.id¹, achmad.zakki@nanrotama.ac.id²

Kata Kunci

Association rule, fp-growth, pola pembelian.

Abstrak

Data mining merupakan proses pengambilan informasi dari sekumpulan data. Pada penelitian ini akan mencoba menerapkan data mining pada data transaksi di Kedai Kopi Teras Garden untuk strategi penjualan dengan cara membuat rekomendasi item mana yang cocok dan sesuai untuk dijual secara sistem paket atau dijual bersamaan. Penelitian ini menggunakan metode asosiasi rule dengan algoritma fp-growth untuk menemukan pola pembelian customer pada Kedai Kopi Teras Garden. Penelitian ini nantinya akan menghasilkan beberapa rule untuk digunakan sebagai rekomendasi penjualan sesuai dengan data yang sudah di kumpulkan dan diproses dengan metode association rule. Data akan dibagi menjadi 2 yaitu data pada musim kemarau dan musim hujan karena ada perubahan pola pembelian customer sesuai musim yang sedang berlangsung.

Keywords

Association rule, fp-growth, purchase pattern.

Abstract

Data mining is a process extracting data from the dataset. In this paper will try to apply data mining on data transaction in Kedai Kopi Teras Garden for sales strategy by creating recommendation item that suit to be sold by package system or the other words by to be sold simultaneously. This paper use association rule method with fp-growth algorithm to find customer purchase pattern on Kedai Kopi Teras Garden. The output of this paper will create some rules for recommendation item that can be sold as a package according to data that already collected and processed with association rule method. Data will be divided by 2, dry season data and rainy season data because there's customer pattern change accordingly with the season.

1. Pendahuluan

Perumahan *The Quality Garden* mempunyai sebuah kedai kopi yang berada didepan pintu masuk yang bernama Kedai Kopi Teras Garden. Kedai ini berlokasi di Gamping Wetan, Gamping, Kecamatan Krian, Kabupaten Sidoarjo, Jawa Timur. Banyak muda mudi yang menghabiskan waktu bersama dikedai ini sambil menikmati makanan dan minuman yang dijual dikedai ini. Kedai ini juga menyediakan *live music* setiap hari sabtu malam untuk meramaikan suasana Kedai Kopi Teras Garden. Pihak manajemen dari Kedai Kopi Teras Garden ingin menjual makanan dan minuman dengan sistem paket yang efektif atau dijual secara bersamaan dengan tujuan meningkatkan penjualan. Tetapi pihak manajemen belum memahami bagaimana cara menemukan kombinasi item mana yang cocok atau sesuai untuk dijual secara bersamaan. Tetapi pihak manajemen belum memahami bagaimana cara menemukan kombinasi item mana yang cocok atau sesuai untuk dijual

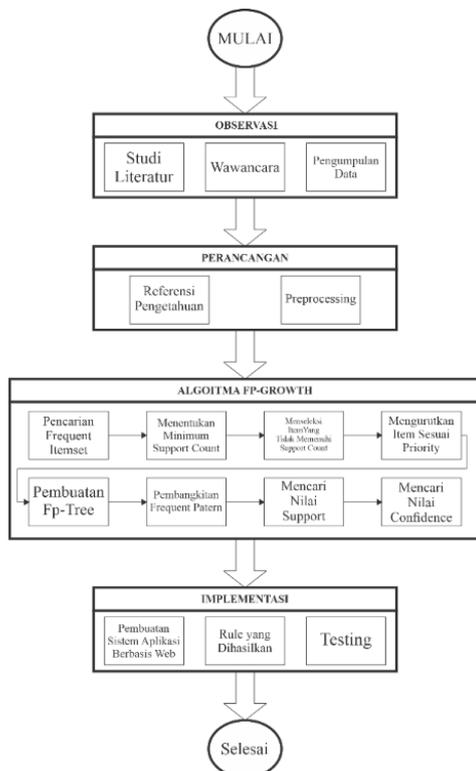
secara bersamaan. Pihak Manajemen perlu untuk merubah sistem yang tradisional ke sistem yang lebih professional[1].

Data mining biasanya digunakan untuk memproses data yang berukuran sangat besar untuk mengambil informasi dan pengetahuan dari data tersebut [2]. Data mining bisa juga digunakan untuk menemukan pola pembelian dari *customer* dalam membeli beberapa kombinasi barang dengan menggunakan metode tertentu. Data mining sering disebut juga *knowledge discovery in database* (KDD) adalah sebuah proses yang meliputi pengumpulan data, penggunaan data untuk menentukan pola atau hubungan dalam sebuah data yang besar [3]. Metode yang digunakan untuk menemukan pola pembelian dari *customer* yaitu *association rule mining*. Ada beberapa algoritma yang biasanya digunakan dalam metode ini, salah satunya adalah *fp-growth* (*frequent pattern growth*).

Dalam beberapa tahun yang lalu, *frequent pattern mining* digunakan secara signifikan untuk menemukan pengetahuan yang menarik yang tersembunyi dari sekumpulan data. Banyak algoritma yang telah dikembangkan untuk mempercepat proses pengambilan informasi [4]. Algoritma *fp-growth* adalah pengembangan dari algoritma apriori dalam menemukan kombinasi item yang tepat menurut data yang sudah di kumpulkan sebelumnya. Algoritma apriori yang diciptakan oleh Agrawal et al pada 1990 membutuhkan untuk membaca database beberapa kali dan menghasilkan jumlah kandidat kombinasi yang besar, Han et al menciptakan *fp-growth* yang mempunyai struktur data berupa *fp-tree* untuk memperbaiki kinerja algoritma apriori [5]. *Fp-growth* akan memakan banyak waktu dalam pembuatan *fp-tree* dan membangkitkan *frequent pattern*. Pada *association rule* dapat ditentukan dengan 2 nilai parameter yaitu *support* dan nilai *confidence* [6].

Data yang digunakan didapat dari pihak manajemen Kedai Kopi Teras Garden. Karena biasanya pola pembelian *customer* berubah sesuai dengan cuaca yang sedang berlangsung, maka data akan dibagi menjadi 2 yaitu data musim kemarau dan musim hujan. Sebagai contoh di musim kemarau, *customer* membeli lebih banyak minuman dingin dari pada minuman hangat. Begitu juga sebaliknya pada musim dingin *customer* membeli lebih banyak minuman hangat dari pada minuman dingin.

2. Metode penelitian



Gambar 1 Flowchart

Penelitian ini menggunakan metode *association rule* dengan algoritma *frequent pattern growth* untuk mencari kombinasi dari makanan dan minuman yang di jual di Kedai Kopi Teras Garden. Penelitian ini memilih tahapan atau proses yang bisa dilihat pada Gambar 1 untuk menggambarkan bagaimana jalannya penelitian ini.

2.1 Association Rule

Penelitian ini bertujuan untuk menemukan pola pembelian dari *customer* yang nanti digunakan untuk menemukan kombinasi item yang cocok atau sesuai dijual secara bersamaan di Kedai Kopi Teras Garden.

1. Observasi

Pada tahap ini, peneliti mengamati strategi penjualan apa yang sudah ada pada Kedai Kopi Teras Garden. Studi literature dilakukan untuk mengumpulkan jurnal yang digunakan sebagai referensi atau rujukan untuk penelitian ini. Wawancara dilakukan kepada pihak manajemen tentang bagaimana penjualan yang sedang berjalan, menu apa yang paling sering terjual dan menu mana yang biasanya dibeli *customer* secara bersamaan. Pengumpulan data juga dilakukan untuk menganalisa pola pembelian *customer*. Pengumpulan data dilakukan dengan cara meminta langsung kepada pihak manajemen, yang diperlukan untuk penelitian ini hanya data penjualan.

2. Perancangan

Dibutuhkan perancangan terlebih dahulu sebelum masuk ke pembuatan aplikasi untuk menganalisa pola pembelian *customer*. Dilakukan pengumpulan referensi pengetahuan berupa pengetahuan tentang bahasa pemrograman python beserta library *fp-growth* yang digunakan dalam penelitian ini. Selanjutnya tahap *preprocessing*, pada tahap ini data di olah terlebih dahulu sebelum dimasukan ke dalam proses analisa untuk hasil yang lebih optimal. Tahap *preprocessing* yang dilakukan pada penelitian ini adalah menghapus item yang sama pada satu pesanan atau transaksi untuk menghindari hasil rekomendasi item yang sama seperti jika membeli *ice tea* maka akan membeli *ice tea*. Hasil rekomendasi ini akan kurang relevan untuk dijadikan hasil penelitian.

3. Algoritma FP-Growth

Data yang telah melalui tahap *preprocessing*, akan dilakukan analisa pada tahap ini. Adapun beberapa langkah dari algoritma ini sebagai berikut.

a. Pencarian *frequent itemset*

Pada tahap pencarian *frequent itemset* ini, data yang sudah melalui tahap *preprocessing* sebelumnya akan dimasukan kedalam tabel yang mempunyai atribut nomer transaksi dan pesanan seperti pada Tabel 1.

Tabel 1 Daftar pesanan

No	Pesanan
1	Ice Tea,IceCapuccino,French Fries
2	Ice Tea,Lemon Tea,Kebab,French Fries
3	Ice Capuccino,French Fries,Donat
4	Ice Tea,Spagethi,French Fries
5	Ice Choco Oreo,Ice Choco Delight,French Fries
6	Hot Vietnam Drip,Pisang Coklat Keju
7	Lemon Tea,Hot Vietnam Drip,Jamur Crispy,Onion Ring
8	Hot Capuccino,Ice Coffe Latte,Donat
9	Ice Tea,Ice Choco Delight,Ice Choco Oreo,Onion Ring
10	Lemon Tea,ice Coffe Latte,French Fries,Pisang Coklat Keju

b. Menentukan *Minimum Support Count*

Perlu menentukan *minimum support count* untuk menyeleksi item yang kurang laku atau jarang dibeli untuk memudahkan tahap berikutnya. Pada tahap ini peneliti akan menentukan *minimum support count* sebesar 20%. Jadi item yang memiliki support count dibawah 20% akan di buang atau tidak akan digunakan untuk perhitungan kedepannya. Untuk menentukan support count digunakan persamaan(1), jumlah transaksi yang mengandung item A dibagi dengan total transaksi. Untuk mencari nilai support dari 2 item digunakan persamaan (2), Jumlah transaksi yang mengandung item A dan B dibagi dengan total transaksi.

$$Support(A) = \frac{Jumlah\ transaksi\ yang\ mengandung\ item\ A}{Total\ transaksi} (1)$$

$$Support(A \cap B) = \frac{Jumlah\ transaksi\ yang\ mengandung\ item\ A\ dan\ B}{Total\ transaksi} (2)$$

c. Menyeleksi Item Yang Tidak Memenuhi *Support Count*

Asumsi minimum support adalah 20% dari total transaksi pada penelitian ini, jadi minimum *support count* adalah $(20\% \times 10) = 2$. Karena minimum support adalah 2, maka item yang memiliki nilai *support* kurang dari 2 akan dihilangkan.

d. Mengurutkan Item Sesuai *Priority*

Pada tahap ini item akan di urutkan sesuai *priority* atau mengurutkan item dari frekuensi yang tertinggi ke frekuensi yang terendah. Seperti pada tabel 3 *french fries* mempunyai frekuensi tertinggi sebesar 6, makan *french fries* akan diletakan pada nomor 1.

Kedua adalah *ice tea* yang mempunyai frekuensi sebesar 4 maka diletakan nomor 2. Tahap ini akan dilakukan pada semua item yang memenuhi support count. Dan akan dikembalikan ke tabel data transaksi dengan sesuai urutan dari frekuensi yang terbesar ke frekuensi yang terkecil seperti pada tabel 4.

Tabel 2 *Support count*

Item	Frekuensi
Ice Tea	4
French Fries	6
Lemon Tea	3
Kebab	1
Ice Capuccino	2
Ice Choco Delight	2
Donat	2
Hot Capuccino	1
Spagethi	1
Ice Choco Oreo	1
Hot Vietnam Drip	2
Pisang Coklat Keju	1
Jamur Crispy	1
Ice Coffe Latte	1
Onion Ring	2
Ice Coffe Latte	1

Tabel 3 Yang sudah diseleksi dan di urutkan

No	Pesanan
1	French Fries,Ice Tea,IceCapuccino
2	French Fries,Ice Tea,Lemon Tea
3	French Fries,Ice Capuccino,Donat
4	French Fries,Ice Tea
5	French Fries,Ice Choco Delight
6	Hot Vietnam Drip
7	Lemon Tea,Hot Vietnam Drip,Onion Ring
8	Hot Capuccino,Donat
9	Ice Tea,Ice Choco Delight,Onion Ring
10	French Fries,Lemon Tea

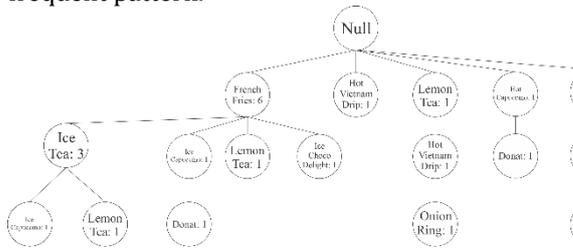
Tabel 4 Pesanan yang sudah di urutkan

Item	Frekuensi
French Fries	6
Ice Tea	4
Lemon Tea	3
Ice Capuccino	2
Ice Choco Delight	2
Donat	2
Hot Vietnam Drip	2
Onion Ring	2

e. Pembuatan *FP-Tree*

Setelah diurutkan sesuai *priority*, tahap berikutnya adalah membuat *fp-tree*. Sebuah *fp-tree* akan selalu dimulai dengan node null. Di dalam *fp tree* juga terdapat support count yang menandakan item melewati transaksi yang sama. Dan di isi mulai dari transaksi nomor 1. French Fries, Ice Tea dan Ice Capuccino diletakan dalam satu jalur kebawah. Transaksi nomor 2 French Fries, Ice Tea dan Lemon Tea. Tidak perlu membuat jalur baru untuk French Fries karena sudah ada sebelumnya, tinggal menambahkan support count pada French Fries sebelumnya. Ice Tea juga tidak membentuk jalur baru setelah French Fries, tinggal menambahkan support count pada Ice Tea sebelumnya. Lalu akan membuat jalur baru

setelah Ice tea karena Belum ada Lemon Tea setelah French Fries dan Ice Tea. Akan selalu membuat jalur baru jika belum ada yang melewati item tersebut dan jika sudah ada, tinggal menambahkan pada support count seperti pada gambar 2. Semua pesanan pada data transaksi yang sudah diurutkan semuanya akan ditulis pada fp tree ini untuk diproses ke tahap selanjutnya. Setelah semua pesanan pada data transaksi ditulis kedalam fp tree, maka akan dilanjutkan pada tahap pembangkitan frequent pattern.



f. Pembangkitan Frequent Pattern

Pada langkah ini semua *suffix* atau akhir dari semua jalur yang terdapat pada fp tree akan di tulis pada sebuah tabel beserta jalur yang telah di lalui.

Tabel 5 frequent pattern

Suffix	Frequent Itemset
Ice Capuccino	{Ice Capuccino, Ice Tea, French Fries}
Lemon Tea	{Lemont Tea, Ice Tea, French Fries}, {Lemon Tea, French Fries}
Hot Capuccino	{Hot Capuccino, Ice Tea, French Fries}
Ice Tea	{Ice Tea, French Fries}
Donat	{Donat, Ice Capuccino, French Fries}, {Donat, Hot Capuccino}
Ice Choco Delig	{Ice Choco Delight, French Fries}
Hot Vietnam Dr	{Hot Vietnam Drip}
Onion Ring	{Onion Ring, Hot Vietnam Drip, Lemon Tea}, {Onion Ring, Ice Choco Delight, Ice Tea}

Dari tabel ini didapatkan beberapa pattern, selanjutnya akan dicari nilai support dan nilai confidence untuk setiap pattern. Untuk mencari nilai confidence digunakan persamaan (3), Jumlah transaksi yang mengandung item A dan B dibagi dengan jumlah transaksi yang mengandung item A.

$$Confidence P(A|B) = \frac{\text{Jumlah transaksi yang mengandung A dan B}}{\text{Jumlah transaksi yang mengandung A}} \quad (3)$$

Jika semua sudah tertulis pada tabel maka akan memunculkan pattern seperti berikut.

- Pattern 1: {Ice Capuccino, Ice Tea, French Fries} support = 0.1 & confidence = 1
- Pattern 2: {Lemon Tea, Ice Tea, French Fries} support = 0,1 & confidence = 1
- Pattern 3: {Lemon Tea, French Fries} support = 0,2 confidence = 0,66
- Pattern 4: {Hot Capuccino, Ice Tea, French Fries} support = 0,1 & confidence = 1
- Pattern 5: {Ice Tea, French Fries} support = 0,3 & confidence = 0,75

- Pattern 6: {Donat, Ice Capuccino, French Fries} support = 0,1 & confidence = 1
- Pattern 7: {Donat, Hot Capuccino} support = 0,1 & confidence = 1
- Pattern 8: {Ice Choco Delight, French Fries} support = 0,1 & confidence = 1
- Pattern 9: {Hot Vietnam Drip} support = 0,1 & confidence = 1
- Pattern 10: {Onion Ring, Hot Vietnam Drip, Lemon Tea} support = 0,1 & confidence = 1
- Pattern 11: {Onion Ring, Ice Choco Delight, Ice Tea} support = 0,1 & confidence = 1

Dari perhitungan nilai support dan nilai confidence yang sudah dilakukan, didapatkan hasil 2 rule dengan nilai terbesar, yaitu.

- a. Rule 1, jika membeli Ice Tea maka akan membeli French Fries dengan nilai support mencapai 30% dan nilai confidence mencapai 75%.
- b. Rule 2, Jika Membeli Lemon Tea maka akan membeli French Fries dengan nilai support mencapai 20% dan nilai confidence mencapai 66%.

Dua rule ini akan dijadikan rekomendasi kombinasi item yang akan dijual secara bersamaan atau sistem paket karena 2 rule mempunyai kombinasi item yang cukup kuat.

3. Hasil dan Diskusi

Pada bagian ini akan menjelaskan hasil dan evaluasi dari metode yang digunakan dalam penelitian ini. Penelitian ini menggunakan jupyter lab untuk menganalisa data. Penelitian ini juga menggunakan Bahasa pemrograman python dan menggunakan library fp growth dari python.

3.1 Dataset

Data yang digunakan dalam metode ini adalah data yang berformat csv yang telah melalui tahap *preprocessing* sebelumnya untuk mendapatkan hasil yang optimal. Data didapat dari pihak manajemen Kedai Kopi Teras Garden langsung. Data akan dibagi menjadi 2 yaitu data musim kemarau dan musim hujan untuk menghasilkan hasil rekomendasi yang sesuai untuk musim kemarau dan musim hujan. Data pada musim kemarau sebanyak 5727 transaksi, sedangkan data pada musim hujan sebanyak 4371. Data musim kemarau diambil dalam periode 6 bulan mulai dari bulan Maret 2019 sampai dengan bulan Agustus 2019. Data musim hujan diambil dalam periode 6 bulan juga dari bulan September 2019 sampai dengan bulan Februari 2020. Data pada musim hujan lebih sedikit dikarenakan ada penurunan jumlah *customer* pada musim hujan. Kedai Kopi Teras Garden mempunyai pengunjung yang lebih sedikit pada musim hujan daripada musim kemarau. Perbedaan musin juga mempengaruhi hasil yang didapat karena pada musim kemarau pengunjung Kedai Kopi Teras

Garden cenderung untuk membeli minuman dingin. Begitu juga sebaliknya, pada musim hujan pengunjung lebih banyak memesan minuman hangat.

3.2 Testing

Pengujian pada penelitian ini menggunakan jupyter lab untuk analisa data yang didapat dari manajemen Kedai Kopi Teras Garden. Untuk hasil yang maksimal, maka diperlukan tahap *preprocessing* terlebih dahulu. Tahap pertama untuk *preprocessing* adalah menghilangkan item yang sama pada setiap transaksi. Hal ini perlu dilakukan untuk menghindari hasil output dengan item yang sama contoh jika memesan ice tea maka akan membeli ice tea. Hasil ini kurang relevan untuk dijadikan rekomendasi penjualan. Tahap kedua adalah menghapus *cell* kosong pada file csv untuk bisa diproses lebih lanjut. Tahap berikutnya adalah mengubah file csv ke dalam struktur data berbentuk list dalam pemrograman python.

Data pada musim kemarau dan musim hujan akan ditentukan *minimum support count* sebesar 20% dan *minimum confidence* sebesar 0,5. Berdasarkan data yang di dapat dari 6 bulan musim kemarau dan musim hujan didapatkan 1 rule setiap musim yang sesuai dengan *minimum support* dan *minimum confidence* yang sudah ditentukan pada penelitian ini. Pada musim kemarau didapatkan rule French Fries dan Ice Tea dengan tingkat *confidence* sebesar 0,61. Dapat diartikan jika Membeli French Fries maka akan membeli Ice Tea.

```
[11]: {'French Fries',): ('Ice Tea',), 0.6151639344262295}
```

Gambar 2 hasil dari data musim kemarau

```
[28]: {'Donat',): ('Hot Cappuccino',), 0.5837696335078534}
```

Gambar 3 hasil dari data musim hujan

Sedangkan pada musim hujan didapatkan rule Donat dan Hot Capuccino dengan tingkat *confidence* sebesar 0, 58. Dapat di artikan Jika membeli Donat maka akan membeli Hot Capuccino. Hasil kedua analisa tersebut dilakukan pada Jupyter lab dengan menggunakan Bahasa pemrograman python dan menggunakan library fp growth.

4. Kesimpulan

Berdasarkan hasil uji coba yang dilakukan pada penelitian ini didapatkan 2 rule yang terdiri dari 1 rule pada musim kemarau dan 1 rule pada musim hujan yang memenuhi syarat *minimum support count* sebesar 20% dan *minimum confidence* sebesar 0,5 untuk digunakan sebagai rekomendasi

penjualan dengan sistem paket atau dijual secara bersamaan. Rule yang dihasilkan sebagai berikut. Pada musim kemarau, jika *customer* membeli French Fries maka *customer* juga akan membeli Ice Tea dengan tingkat *confidence* sebesar 0,61. Sedangkan pada musim Hujan, jika *customer* membeli Donat maka *customer* juga akan membeli Hot Capuccino dengan tingkat *confidence* sebesar 0,58. Data yang di dapat dari Kedai Kopi Teras Garden pada musim hujan lebih sedikit karena terdapat penurunan jumlah pengunjung saat musim hujan.

5. DAFTAR PUSTAKA

- [1] B. A. Arasy Alimudin, Soebandi, Ahmad Zakki Falani, Eman Setiawan, "The Use of Fuzzy Logic in Developing Competitiveness Strategy Toward Small and Medium Enterprises (Groceries)," *Adv. Sci. Lett.*, vol. 23, no. 12, 2017.
- [2] M. Narvekar and S. Fatma, "An optimized algorithm for association rule mining using FP tree," *Procedia - Procedia Comput. Sci.*, vol. 45, pp. 101-110, 2015, doi: 10.1016/j.procs.2015.03.097.
- [3] A. Ikhwan, "A NOVELTY OF DATA MINING FOR FP-GROWTH ALGORITHM," vol. 9, no. 7, pp. 1660-1669, 2018.
- [4] K. Gadia and K. Bhowmick, "Parallel text mining in multicore systems using FP-Tree algorithm," *Procedia - Procedia Comput. Sci.*, vol. 45, pp. 111-117, 2015, doi: 10.1016/j.procs.2015.03.100.
- [5] Y. Li and S. Yin, "Mining Algorithm for Weighted FP-Growth Frequent Item Sets based on Ordered FP-Tree," vol. 5, no. October, pp. 154-158, 2019.
- [6] E. Regar and W. Silfianti, "Visit Patterns Analysis of Foreign Tourist in Indonesian Territory Using Frequent Pattern Growth (FP - Growth) Algorithm," vol. 7, no. 9, pp. 212-217, 2018.