

PERBANDINGAN ALGORITMA APRIORI DAN ALGORITMA FP-GROWTH UNTUK MENGETAHUI POLA PENGGUNAAN TRANSPORTASI ONLINE

Desti Fitriati¹, Musi Hardiyanto²

¹² Teknik Informatika Universitas Pancasila Jakarta

Jalan Raya Lenteng Agung Timur No.56-80, Srengseng Sawah, Jagakarsa, Jakarta Selatan

³Email: desti.fitriati@univpancasila.ac.id

Abstrak

Kemajuan teknologi saat ini sudah tidak dapat dicegah. Kemajuan teknologi mempermudah pekerjaan manusia dalam berbagai bidang termasuk dibidang transportasi. Kemajuan teknologi dibidang transportasi dapat membantu masyarakat mencapai tempat tujuan dengan cepat dan menghemat waktu tempuh. Kemajuan teknologi dibidang transportasi juga terjadi di Indonesia sekarang. Masyarakat lebih cenderung memilih transportasi online dibandingkan transportasi yang sudah ada. Tentunya masyarakat memilih transportasi online juga didasari alasan kenapa mereka memilih transportasi online. Berdasarkan permasalahan diatas, tujuan dari penelitian ini adalah ingin mengetahui faktor-faktor apa saja yang menjadi alasan masyarakat memilih transportasi online dengan metode algoritma apriori dan algoritma fpgrowth. Tentunya dengan diketahuinya faktor-faktor tersebut, tentunya dapat meningkatkan pelayanan yang diberikan oleh perusahaan transportasi online kepada konsumennya. Tentunya dengan pelayanan yang terus ditingkatkan, dapat meningkatkan laba perusahaan transportasi online.

Kata kunci: Apriori; Data mining; FP-Growth; Laba; Pola Penggunaan Transportasi; Transportasi Online

1. PENDAHULUAN

1.1. Latar Belakang

Perkembangan teknologi yang sangat pesat di era globalisasi saat ini telah memberikan banyak manfaat dalam kemajuan diberbagai aspek sosial. Penggunaan teknologi oleh manusia dalam membantu menyelesaikan pekerjaan merupakan hal yang menjadi keharusan dalam kehidupan. Perkembangan teknologi ini juga harus diikuti dengan perkembangan pada Sumber Daya Manusia (SDM).

Kemajuan teknologi sudah masuk dalam semua aspek kehidupan manusia, seperti pendidikan, kesehatan, makanan, dan juga transportasi. Kemajuan teknologi dibidang transportasi membuat orang lebih mudah untuk melakukan aktifitas seperti berangkat kerja, pergi ke sekolah, jalan – jalan bersama keluarga dan lain – lain. Kemajuan teknologi dibidang transportasi memberi dampak munculnya perusahaan yang bergerak di bidak transportasi berlandaskan teknologi. Dimana perusahaan tersebut memberikan layanan kepada pelanggannya seperti cepat untuk sampai lokasi penjemputan, dapat mengurangi waktu tempuh perjalanan, dan harga yang terjangkau kepada masyarakat.

Faktor lain yang menentukan orang ingin menggunakan transportasi online yaitu cuaca, dan jumlah orang yang ingin pergi secara bersamaan. Dalam kondisi cuaca hujan dan pergi secara berkelompok biasanya orang lebih memilih transportasi online yang menggunakan mobil, karena dinilai dapat mengangkut semua orang dan terhindar dari hujan yang dapat menimbulkan penyakit.

Berdasarkan latar belakang tersebut, tujuan dari penelitian ini adalah ingin mengetahui faktor – faktor apa saja yang mempengaruhi orang memilih transportasi online sehingga dengan penggunaan teknologi informasi dapat meningkatkan pelayanan bagi konsumen, driver, dan laba perusahaan. Dan mengetahui metode manakah yang memiliki nilai akurasi lebih baik.

1.2. Rumusan Masalah

Rumusan Masalahnya adalah bagaimana membentuk aturan asosiasi dari data kuesioner yang ada dengan algoritma apriori dan algoritma fpgrowth dan metode manakah yang memiliki nilai lebih baik dari metode algoritma apriori dan algoritma fpgrowth.

1.3. Tujuan dan Manfaat Penelitian

Tujuan dari penelitian ini yaitu dapat meningkatkan laba perusahaan berdasarkan hasil aturan asosiasi yang terbentuk dari faktor-faktor penentu masyarakat memilih transportasi online dan mengetahui metode manakah yang memiliki nilai akurasi lebih baik. Dan manfaatnya tentu tertuju kepada pemilik perusahaan transportasi online dengan mudah dapat meningkatkan pelayanan kepada konsumen berdasarkan dari aturan asosiasi yang sudah terbentuk.

1.4. Metodologi Penelitian

Metode penelitian yang digunakan dalam penerapan data mining adalah metode *Waterfall*. Pengembangan metode *Waterfall* sendiri melalui beberapa tahapan yaitu *Analisis, Desain, Coding, Testing*.

2. LANDASAN TEORI

2.1. Algoritma apriori

Algoritma apriori adalah salah satu algoritma yang melakukan pencarian frekuensi itemset dengan menggunakan teknik *association rule*. Algoritma apriori menggunakan pengetahuan frekuensi atribut yang telah diketahui sebelumnya untuk memproses informasi selanjutnya. Pada algoritma apriori, menentukan kandidat yang mungkin muncul dengan cara memperhatikan *minimum support* dan *minimum confidence*.

2.2. Tahapan proses algoritma apriori

- a) Tentukan *minimum support*.
- b) Iterasi 1 : hitung item-item dari *support* (transaksi yang memuat seluruh item) dengan men-scan database untuk 1-itemset, setelah 1-itemset didapatkan, dari 1-itemset apakah diatas *minimum support*, apabila telah memenuhi minimum support, 1-itemset tersebut akan menjadi pola frekuensi tinggi.

$$Support(A) = \frac{\text{Jumlah Transaksi Mengandung A}}{\text{Total Transaksi}} \quad (1)$$

- c) Iterasi 2 : untuk mendapatkan 2-itemset, harus dilakukan kombinasi dari 1 itemset, kemudian scan database lagi untuk hitung item-item yang memuat *support*. Itemset yang memenuhi *minimum support* akan dipilih sebagai pola frekuensi tinggi dari kandidat.

$$Support(A \cup B) = \frac{\text{Jumlah Transaksi Mengandung A dan B}}{\text{Total Transaksi}} \quad (2)$$

- d) Tetapkan nilai k-itemset dari support yang telah memenuhi minimum support dari k-itemset.
- e) Lakukan proses untuk iterasi selanjutnya hingga tidak ada lagi k-itemset yang memenuhi minimum support.
- f) Hitung nilai *confidence* dan buat tabel aturan asosiasi.

$$Confidence P(B|A) = \frac{\text{Jumlah Transaksi Mengandung A dan B}}{\text{Total Transaksi Mengandung A}} \quad (3)$$

2.3. Algoritma Fp-growth

Frequent Pattern Growth (FP-Growth) adalah salah satu alternatif algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (frequent itemset) dalam sebuah kumpulan data. Dalam Algoritma *Frequent Pattern Growth (FP-Growth)* menyimpan informasi mengenai frequent itemset dalam bentuk struktur prefix-tree atau sering di sebut FP-Tree. Tidak melakukan candidate generation dalam proses pencarian frequent itemset, sehingga dapat mengurangi scan database secara berulang dalam proses mining dan dapat berlangsung lebih cepat.

2.4. Tahapan Proses Algoritma Fpgrowth

- Penyiapan Dataset.
- Pencarian itemset yang sering muncul.
- Urutkan itemset berdasarkan jumlah terbesar hingga jumlah terkecil.
- Buat fptree berdasarkan itemset yang sudah diurutkan.
- Buat conditional pattern base.
- Buat conditional fptree .
- Buat frequent pattern.
- Hitung nilai support dari frequent pattern yang sudah terbentuk.

$$Support (A \cup B) = \frac{\text{Jumlah Transaksi Mengandung A dan B}}{\text{Total Transaksi}} \quad (4)$$

- Hitung nilai *confidence* dari *frequent pattern* yang sudah terbentuk.

$$Confidence P(B|A) = \frac{\text{Jumlah Transaksi Mengandung A dan B}}{\text{Total Transaksi Mengandung A}} \quad (5)$$

- Buat aturan asosiasi berdasarkan *minimum support* dan *minimum confidence* yang sudah ditentukan.

3. HASIL DAN PEMBAHASAN

3.1. Kumpulan Kategori

Berikut ini adalah penjelasan dari kumpulan kategori :

Tabel 1. Data kategori

| No | Kategori | Keterangan |
|----|----------|------------------------------------|
| 1 | A1 | Remaja Awal (12 - 16) |
| 2 | A2 | Remaja Akhir (17 - 25) |
| 3 | A3 | Dewasa Awal (26 - 35) |
| 4 | A4 | Dewasa Akhir (36 - 45) |
| 5 | B1 | Perempuan |
| 6 | B2 | Laki - Laki |
| 7 | C1 | Pelajar (SMP - SMA DAN MAHASISWA) |
| 8 | C2 | Ibu Rumah Tangga |
| 9 | C3 | Wiraswasta |
| 10 | C4 | Pns |
| 11 | C5 | Kary.Swasta |
| 12 | C6 | Bumn (Peg. Bank) |
| 13 | C7 | Tenaga Pengajar (Guru, Dosen) |
| 14 | C8 | Dokter |
| 15 | C9 | Tentara Nasional (Tni, Polisi) |
| 16 | D1 | Motor |
| 17 | D2 | Mobil |
| 18 | E1 | Manual(Motor) |
| 19 | E2 | Matic(Motor) |
| 20 | E3 | Avanza |
| 21 | E4 | Xenia |

| No | Kategori | Keterangan |
|----|----------|--|
| 22 | E5 | Mobilio |
| 23 | F1 | Gojek |
| 24 | F2 | Grab |
| 25 | F3 | Uber |
| 26 | G1 | Sendiri |
| 27 | G2 | 1 Orang |
| 28 | G3 | 2 Orang |
| 29 | G4 | 3 Orang |
| 30 | G5 | >4 Orang |
| 31 | H1 | Harga |
| 32 | H2 | Cuaca |
| 33 | H3 | Waktu Tempuh |
| 34 | H4 | Promo |
| 35 | H5 | Lainnya |
| 36 | I1 | 0 - 15 Menit |
| 37 | I2 | 16 - 30 Menit |
| 38 | I3 | 30 - 45 Menit |
| 39 | I4 | 45 - 1 Jam |
| 40 | I5 | >1 Jam |
| 41 | J1 | Pakai Promo (Ya) |
| 42 | J2 | Pakai Promo (Tidak) |
| 43 | K1 | 2triwhid25000 |
| 44 | K2 | Gratisoba |
| 45 | K3 | Maretasik |
| 46 | K4 | Jabarpasti |
| 47 | K5 | Picodiid30k |
| 48 | K6 | Tidak Pakai Promo |
| 49 | L1 | Motor |
| 50 | L2 | Mobil |
| 51 | M1 | Harga |
| 52 | M2 | Kenyamanan |
| 53 | M3 | Promo |
| 54 | M4 | Waktu Tempuh |
| 55 | N1 | Jas Hujan |
| 56 | N2 | Masker |
| 57 | N3 | Penutup Kepala |
| 58 | N4 | Permen |
| 59 | N5 | Ac Dan Musik |
| 60 | N6 | Tidak Dapat Fasilitas |
| 61 | R1 | Sikap Driver (Ramah,Sopan) |
| 62 | R2 | Sikap Berkendara (Keselamatan Berkendara, Keselamatan Berkendara, Nyaman Berkendara, Selamat Ditujuan) |
| 63 | R3 | Fasilitas |
| 64 | R4 | Tidak Ada Kriteria |
| 65 | S1 | Cash |
| 66 | S2 | Electronic Payment |
| 67 | Z1 | 1 (Jumlah Aplikasi Transportasi Online) |
| 68 | Z2 | 2 (Jumlah Aplikasi Transportasi Online) |
| 69 | Z3 | 3 (Jumlah Aplikasi Transportasi Online) |
| 70 | ZA1 | Sendiri |
| 71 | ZA2 | Orang Lain |
| 72 | Q1 | History Driver(Ya) |
| 73 | Q2 | History Driver (Tidak) |

3.2. Data

Data yang diolah merupakan data kuesioner dari pengguna transportasi online baik itu yang menggunakan mobil atau motor yang berjumlah 500 kuesioner. Dari 500 kuesioner

tersebut dibagi menjadi 300 data kuesioner sebagai data training dan 200 data sebagai data testing.

Berikut ini adalah contoh data kuesioner yang telah diisi oleh pengguna transportasi online yang berisi kumpulan alasan masyarakat memilih transportasi online.

Tabel 2. Data kuesioner

| No | Data |
|----|---|
| 1 | A3, B2, C3, D2, E3, F1, G1, H1, I4, J2, K6, L2, M1, N5, Q1, R2, S1, Z2, ZA1 |
| 2 | A2, B1, C1, D1, E2, F3, G1, H1, I1, J2, K6, L1, M1, N3, Q2, R1, S1, Z3, ZA1 |
| 3 | A2, B1, C2, D1, E2, F2, G1, H1, I2, J2, K6, L1, M1, N3, Q2, R4, S1, Z1, ZA1 |
| 4 | A2, B1, C2, D2, E4, F3, G2, H3, I4, J2, K6, L2, M4, N4, Q2, R3, S2, Z3, ZA1 |
| 5 | A1, B1, C1, D1, E2, F1, G1, H1, I2, J2, K6, L1, M1, N1, Q2, R1, S1, Z3, ZA1 |

3.3. Proses Data Training

Pada proses testing menghasilkan aturan asosiasi yang diambil hanya kombinasi itemset terbesar saja:

3.3.1 Algoritma Apriori

a) Kombinasi 2 itemset

Tabel 3. Kombinasi 2 itemset

| No | Rule | Support | Confidence | SxC | Lift Ratio |
|----|--|---------|------------|-------|------------|
| 1 | Jika masyarakat memilih transportasi online dengan alasan ZA1 maka masyarakat memilih dengan alasan Q2 | 91% | 96.81% | 88.1% | 1,61 |

b) Kombinasi 3 itemset

Tabel 4. Kombinasi 3 itemset

| No | Rule | Support | Confidence | SxC | Lift Ratio |
|----|---|---------|------------|--------|------------|
| 1 | Jika masyarakat memilih transportasi online dengan alasan G1, Q2 maka masyarakat memilih dengan alasan D1 | 84% | 98.44% | 82.69% | 1,64 |

c) Kombinasi 4 itemset

Tabel 5. Kombinasi 4 itemset

| No | Rule | Support | Confidence | SxC | Lift Ratio |
|----|--|---------|------------|--------|------------|
| 1 | Jika masyarakat memilih transportasi online dengan alasan G1, Q2, ZA1 maka masyarakat memilih dengan alasan D1 | 79% | 98.34% | 77.69% | 1,64 |

d) Kombinasi 5 itemset

Tabel 6. Kombinasi 5 itemset

| No | Rule | Support | Confidence | SxC | Lift Ratio |
|----|--|---------|------------|--------|------------|
| 1 | Jika masyarakat memilih transportasi online dengan alasan D1, G1, L1, ZA1 maka masyarakat memilih dengan alasan Q2 | 67% | 97.1% | 65.06% | 1,62 |

e) Kombinasi 6 itemset

Tabel 7. Kombinasi 6 itemset

| No | Rule | Support | Confidence | SxC | Lift Ratio |
|----|--|---------|------------|--------|------------|
| 1 | Jika masyarakat memilih transportasi online dengan alasan G1, L1, Q2, ZA1, K6 maka masyarakat memilih dengan alasan D1 | 50.67% | 99.35% | 50.34% | 1,66 |

3.3.2 Algoritma Fp-growth**Tabel 8. Hasil asosiasi Fpgrowth**

| No | Rule | Support | Confidence | SxC | Lift Ratio |
|----|--|---------|------------|-------|------------|
| 1 | Jika masyarakat memilih transportasi online dengan alasan ZA1 maka masyarakat memilih dengan alasan Q2 | 91% | 96.81% | 88.1% | 1,61 |

3.4 Proses Data Testing**3.4.1 Algoritma Apriori**

a) Kombinasi 2 itemset

Tabel 9. Kombinasi 2 itemset

| No | Rule | Jumlah | Akurasi |
|----|--|--------|---------|
| 1 | Jika masyarakat memilih transportasi online dengan alasan ZA1 maka masyarakat memilih dengan alasan Q2 | 81 | 40,5% |

b) Kombinasi 3 itemset

Tabel 10. Kombinasi 3 itemset

| No | Rule | Jumlah | Akurasi |
|----|---|--------|---------|
| 1 | Jika masyarakat memilih transportasi online dengan alasan G1, Q2 maka masyarakat memilih dengan alasan D1 | 162 | 81% |

c) Kombinasi 4 itemset

Tabel 11. Kombinasi 4 itemset

| No | Rule | Jumlah | Akurasi |
|----|--|--------|---------|
| 1 | Jika masyarakat memilih transportasi online dengan alasan G1, Q2, ZA1 maka masyarakat memilih dengan alasan D1 | 67 | 33,5% |

d) Kombinasi 5 itemset

Tabel 12. Kombinasi 5 itemset

| No | Rule | Jumlah | Akurasi |
|----|--|--------|---------|
| 1 | Jika masyarakat memilih transportasi online dengan alasan D1, G1, L1, ZA1 maka masyarakat memilih dengan alasan Q2 | 61 | 30,5% |

e) Kombinasi 6 itemset

Tabel 13. Kombinasi 6 itemset

| No | Rule | Jumlah | Akurasi |
|----|--|--------|---------|
| 1 | Jika masyarakat memilih transportasi online dengan alasan G1, L1, Q2, ZA1, K6 maka masyarakat memilih dengan alasan D1 | 44 | 22% |

3.4.2 Algoritma Fpgrowth

Tabel 14. Hasil asosiasi Fp-growth

| No | Rule | Jumlah | Akurasi |
|----|--|--------|---------|
| 1 | Jika masyarakat memilih transportasi online dengan alasan ZA1 maka masyarakat memilih dengan alasan Q2 | 81 | 40,5% |

4. PENUTUPAN

4.1. Kesimpulan

Berdasarkan hasil pengolahan data kuesioner menggunakan algoritma apriori dan algoritma fpgrowth didapatkan hasil akurasi yang sama dengan kombinasi itemset pada algoritma apriori jika masyarakat memilih transportasi online dengan alasan “transportasi motor (**D1**)” maka masyarakat memilih dengan alasan “tidak pernah memiliki history driver” dengan akurasi sebesar 86%. Sedangkan hasil uji menggunakan algoritma fpgrowth didapatkan kombinasi itemset jika masyarakat memilih transportasi online dengan alasan “menaiki transportasi online dengan sendirian (**G1**)” maka masyarakat memilih dengan alasan “transportasi motor (**D1**)” dengan akurasi sebesar 86%. Jadi pengolahan data mining menggunakan algoritma apriori dan algoritma fpgrowth memiliki tingkat akurasi yang baik untuk meningkatkan pelayanan perusahaan transportasi online kepada masyarakat dan tentunya dapat meningkatkan laba perusahaan dengan pelayanan yang terus ditingkatkan.

4.2. Saran

Dengan sudah terbentuknya hasil aturan asosiasi, diharapkan penyedia jasa transportasi *online* dapat meningkatkan pelayanan supaya masyarakat semakin nyaman menggunakan jasa transportasi *online*. Tujuan lainnya tentu dapat meningkatkan laba perusahaan dengan semakin banyaknya masyarakat yang menggunakan jasa transportasi *online* akibat pelayanan yang terus ditingkatkan.

DAFTAR PUSTAKA

- Dr. Reyna Usma, dkk. 2014. *Klasifikasi Baku Jabatan Indonesia*. Jakarta: Kementerian Ketenagakerjaan dan Badan Pusat Statistik.
- Fauzy, Mohamad, Saleh Kemas Rahmat dan Asror Ibnu. 2016. *Penerapan Metode Association Rule Menggunakan Algoritma Apriori Pada Simulasi Prediksi Hujan Wilayah Kota Bandung*. Bandung: Universitas Telkom.
- Fitria, Rizki. Nengsih Warnia dan Qudsi Dini Hidayatul. 2017. *Implementasi Algoritma Fpgrowth Dalam Penentuan Pola Hubungan Kecelakaan Lalu Lintas*. Riau: Politeknik Caltex.
- Fitriati, Desti. 2016. *Implementasi Data Mining Untuk Menentukan Kombinasi Media Promosi Barang Berdasarkan Perilaku Pembelian Pelanggan Menggunakan Algoritma Apriori*. Jakarta: Universitas Pancasila.
- Hadi, Diki Alfarabi. 2016. *Pengertian dan Cara Menggunakan Codeigniter*. <https://www.malasngoding.com/pengertian-dan-cara-menggunakan-codeigniter>. Diakses pada 28 April 2018.
- Larasati, Dyah Pramesti, Nasrun Muhammad, S.si, MT dan Ahamad Umar Ali, ST., MT. *Analisis Dan Implementasi Algoritma Fpgrowth Pada Aplikasi Smart Untuk Menentukan Market Basket Analysis Pada Usaha Retail*. Bandung: Universitas Telkom.
- Miraldi, Novta Rama, Rachmat Antonius dan Susanto Budi. 2014. *Implementasi Algoritma Fpgrowth Untuk Sistem Rekomendasi Buku Di Perpustakaan UKDW*. Yogyakarta: Universitas Kristen Duta Kencana.
- Maulida, Linda. 2011. *Kategori Umur Menurut Depkes RI (2009)*. <https://id.scribd.com/doc/149375854/penggolongan-usia>. Diakses pada 1 Juni 2018.
- Nita, Pria Utari dan Fajriya Hakim. 2015. *Penerapan Metode Association Rule Menggunakan Algoritma Apriori Untuk Analisa Pola Data Kecelakaan Pesawat Dari Tahun 1967 – 2014 Di Indonesia*. Yogyakarta: Universitas Islam Indonesia.
- Prayitno, Deden. 2015. *Penerapan Association Rule Mining Pada Data Bursa Efek Indonesia Dengan Algoritma Apriori Untuk Memilih Saham Unggulanggulan*. Jakarta: Perbanas Institute.
- Ririanti. 2014. *Implementasi Algoritma Fpgrowth Pada Aplikasi Prediksi Persediaan Sepeda Motor (Studi Kasus PT. Pillar Deli Labumas)*. Medan: STMIK Budi Darma.
- Santoso, Heroe, Hariyadi I Putu dan Prayitno. 2016. *Data Mining Analisa Pola Pembelian Produk Dengan Menggunakan Metode Algoritma Apriori*. Yogyakarta: STMIK Amikom.
- Tampubolon, Kennedy, Saragih Hoga dan Reza Bobby. 2013. *Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Alat – Alat Kesehatan*. Jakarta: Universitas Bakrie.
- Yanto, Robi dan Khoiriah Riri. 2015. *Implementasi Data Mining Dengan Metode Algoritma Apriori Dalam Menentukan Pola Pembelian Obat*. Lubuklingau: STMIK Bina Nusantara