

ASSOCIATION RULE MINING UNTUK MENEMUKAN POLA PEMILIK UMKM DENGAN ALGORITMA FP-GROWTH

Tutik Khotimah^{1*}, Darsin², Endang Supriyati³, Anastasya Latubessy⁴, Arief Susanto⁵

¹³⁴⁵ Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muria Kudus
Gondangmanis, PO Box 53, Bae, Kudus 59352

² Program Studi Manajemen, Fakultas Ekonomika dan Bisnis, Universitas Pandanaran
Jl. Banjarsari No 1 Tembalang, Semarang 50275

*Email: tutik.khotimah@umk.ac.id

Abstrak

Penelitian ini bertujuan untuk mengetahui pola pemilik UMKM di Kecamatan Banyumanik Kabupaten Semarang. Data yang digunakan sebanyak 449 record. Atribut yang digunakan meliputi jenis kelamin dan tingkat pendidikan. Algoritma yang digunakan untuk Association Rule Mining adalah algoritma FP-Growth. Dari penelitian ini dihasilkan 13 rule dengan nilai support terkecil 0,025. Dari ke-13 rule, dibagi menjadi 4 kelompok berdasarkan nilai support dan confidence. Kelompok pertama, dengan nilai support dan confidence yang tinggi, ada 1 aturan, yaitu aturan jika SMA maka Perempuan. Kelompok kedua, dengan nilai support tinggi sedangkan nilai confidence rendah, ada 1 aturan, yaitu aturan jika Perempuan maka SMA. Kelompok ketiga, dengan nilai support rendah sedangkan confidence tinggi, ada 5 aturan, yaitu aturan jika Sarjana maka Perempuan, jika SMP maka Perempuan, jika Tidak Sekolah maka Perempuan, jika Diploma maka Perempuan, dan jika SD maka Perempuan. Kelompok keempat, dengan nilai support dan confidence rendah, ada 6 aturan, yaitu aturan jika Perempuan maka Sarjana, jika Laki-laki maka Sarjana, jika SMP maka Laki-laki, jika SMA maka Laki-laki, jika Sarjana maka Laki-laki, dan jika Laki-laki maka SMA.

Kata kunci: Association Rule Mining; FP-Growth; UMKM

1. PENDAHULUAN

Keberhasilan Usaha Mikro Kecil dan Menengah (UMKM) tidak terlepas dari peran seorang pemilik. Utari dan Dewi (2014) meneliti pengaruh tingkat pendidikan pemilik terhadap pendapatan UMKM dan hasilnya adalah positif dan signifikan. Sementara itu, Nainggolan (2016) menyatakan bahwa pendidikan pemilik tidak memberi pengaruh terhadap penghasilan UMKM, tetapi gender memberikan pengaruh yang signifikan terhadap penghasilan UMKM.

Berdasarkan data yang diperoleh dari Dinas Koperasi dan UMKM Semarang tahun 2017, Kecamatan Banyumanik Kabupaten Semarang memiliki sekitar 449 UMKM. Sebagian besar UMKM tersebut bergerak di bidang *handycraft* dan olahan pangan. Dilihat dari jenis kelamin pemiliknya, 128 orang pemilik berjenis kelamin laki-laki dan 321 orang berjenis kelamin perempuan. Sedangkan dilihat dari tingkat pendidikan, ada 14 orang pemilik yang tidak bersekolah, 28 orang lulusan SD/MI, 44 orang lulusan SMP/MTs, 235 lulusan SMA/SMK/MA, 40 lulusan Diploma, 82 lulusan Sarjana, dan 6 orang lulusan lainnya. Penelitian ini bertujuan untuk mengetahui pola pemilik UMKM dengan melakukan *Association Rule Mining* terhadap jenis kelamin dan tingkat pendidikan pemilik UMKM yang ada di Kecamatan Banyumanik, Semarang.

Association Rule Mining adalah salah satu metode yang ada di *Data Mining* yang digunakan untuk menemukan hubungan antar variabel dalam *database* yang besar (Gorunescu, 2011). Pane (2013) menggunakan metode ini untuk mengetahui pola penjualan produk elektronik dengan algoritma *Apriori*. Algoritma yang sama digunakan oleh Santoso (2003) dengan mengimplementasikannya pada sebuah perangkat lunak dari bahasa pemrograman Visual C++. *Apriori* adalah salah satu algoritma yang digunakan dalam *Association Rule Mining*. Dasar dari algoritma ini adalah menemukan *frequent itemsets* (Han & Kamber, 2007). Kelemahan dari algoritma ini adalah adanya *candidate generation*.

Kelemahan ini dapat ditangani oleh algoritma *Frequent Pattern-Growth (FP-Growth)* yang dapat menemukan *frequent itemsets* tanpa menggunakan *candidate generation* (Han & Kamber, 2007). *FP-Growth* menerapkan struktur data *tree* yang disebut *Frequent Pattern Tree (FP-Tree)* (Witten dkk, 2011). Erwin (2009) membandingkan kecepatan proses antara algoritma *Apriori* dan algoritma *FP-Growth*. Kesimpulan yang diperoleh adalah algoritma *FP-Growth* lebih cepat dibandingkan algoritma *Apriori*. Pada penelitian ini, digunakan algoritma *FP-Growth* untuk menemukan pola pemilik UMKM di Kecamatan Banyumanik.

2. METODE PENELITIAN

2.1. Pengumpulan Data

Data yang digunakan adalah data UMKM yang diperoleh dari Dinas Koperasi dan UMKM Semarang tahun 2017.

2.2. Penyeleksian Data

Data yang dipilih adalah data UMKM yang berada di Kecamatan Banyumanik Kabupaten Semarang. Sedangkan atribut yang digunakan dalam penelitian ini adalah kelamin dan pendidikan pemilik UMKM. Tabel 1 menunjukkan seleksi data UMKM.

Tabel 1. Seleksi data UMKM

Record ke-	ID	Kelamin	Pendidikan
1	10	L	Akademi
2	17	P	
3	24	P	
4	82	P	SMP/MTs
5	84	P	SMA/SMK/MA
6	97	P	Sarjana St
7	101	P	Akademi
8	121	P	SMA/SMK/MA
9	124	L	SMA/SMK/MA
10	142	P	Akademi
⋮	⋮	⋮	⋮
449	6879	L	SMP/MTs

2.3. Preprocessing

Pada tahap ini dilakukan pembersihan data terhadap data yang kurang lengkap. Pada atribut Pendidikan, ada beberapa *record* yang tidak memiliki *value*. *Record-record* tersebut dibersihkan dengan memberikan nilai “Tidak Sekolah”.

2.4. Data Mining

Pada tahap ini, digunakan algoritma *FP-Growth*. Proses *mining* data menggunakan aplikasi Rapid Miner.

2.5. Interpretation

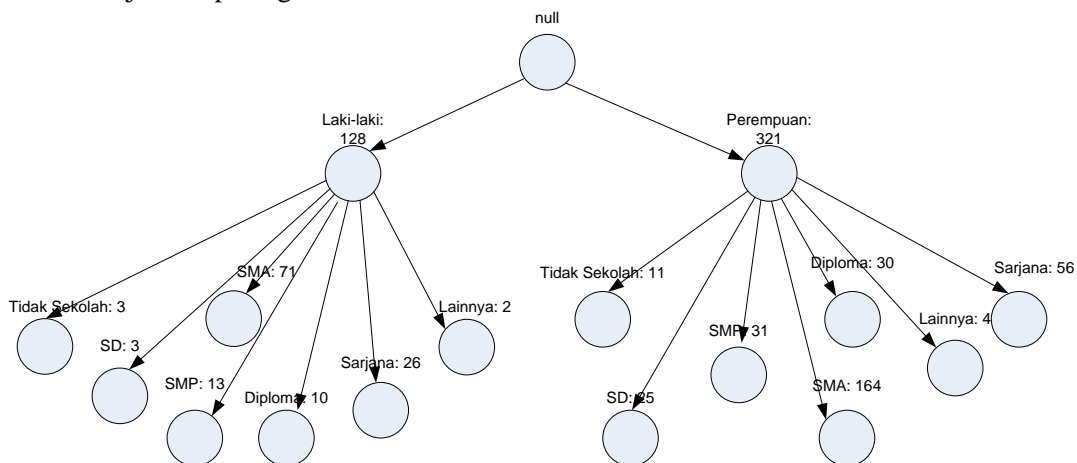
Pada tahap ini, dilakukan pengelompokan terhadap aturan yang dihasilkan dari algoritma *FP-Growth*. Pengelompokan dilakukan berdasarkan nilai *support* dan *confidence* dari masing-masing aturan. Untuk mempermudah pengelompokan, nilai *support* dan *confidence* di-*scalling* terlebih dahulu.

3. HASIL DAN PEMBAHASAN

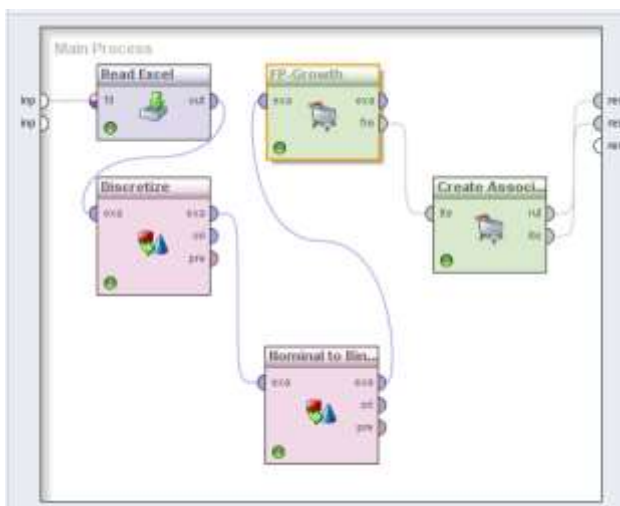
3.1. Implementasi

Algoritma *FP-Growth* dimulai dengan menghitung berapa kali pasangan atribut-nilai yang terjadi dalam *dataset*. Setelah itu dibentuk struktur data *tree* (Witten dkk, 2011).

Pembentukan *FP-Tree* ditunjukkan pada gambar 1. Sedangkan proses *mining* data di Rapid Miner ditunjukkan pada gambar 2.



Gambar 1. Pembentukan *FP-Tree*



Gambar 2. Proses di Rapid Miner

Dengan nilai minimal *support* 0,025 dihasilkan 13 aturan. Aturan-aturan tersebut ditunjukkan pada gambar 3. Sedangkan nilai *support* dan *confidence* dari masing-masing aturan ditunjukkan pada gambar 4.

```

AssociationRules

Association Rules
[Perempuan] --> [Sarjana] (confidence: 0.164)
[Laki-laki] --> [Sarjana] (confidence: 0.208)
[SMP] --> [Laki-laki] (confidence: 0.256)
[SMA] --> [Laki-laki] (confidence: 0.280)
[Sarjana] --> [Laki-laki] (confidence: 0.310)
[Perempuan] --> [SMA] (confidence: 0.508)
[Laki-laki] --> [SMA] (confidence: 0.557)
[Sarjana] --> [Perempuan] (confidence: 0.690)
[SMA] --> [Perempuan] (confidence: 0.720)
[SMP] --> [Perempuan] (confidence: 0.744)
[Tidak Sekolah] --> [Perempuan] (confidence: 0.786)
[Diploma] --> [Perempuan] (confidence: 0.806)
[SD] --> [Perempuan] (confidence: 0.893)
    
```

Gambar 3. Aturan yang dihasilkan

No.	Premises	Conclusion	Support	Confidence
1	Perempuan	Sarjana	0.121	0.164
2	Laki-laki	Sarjana	0.054	0.208
3	SMP	Laki-laki	0.025	0.256
4	SMA	Laki-laki	0.146	0.280
5	Sarjana	Laki-laki	0.054	0.310
6	Perempuan	SMA	0.375	0.508
7	Laki-laki	SMA	0.146	0.557
8	Sarjana	Perempuan	0.121	0.690
9	SMA	Perempuan	0.375	0.720
10	SMP	Perempuan	0.072	0.744
11	Tidak Sekolah	Perempuan	0.027	0.786
12	Diploma	Perempuan	0.072	0.806
13	SD	Perempuan	0.062	0.893

Gambar 4. Nilai *support* dan nilai *confidence*

3.2. Analisis Hasil

Dari ke-13 aturan yang dihasilkan, dikelompokkan berdasarkan nilai *support* dan nilai *confidence*. Sebelum dikelompokkan, nilai *support* dan nilai *confidence* di-*scalling* terlebih dahulu menggunakan persamaan 1 dengan nilai batas atas (BA) = 0,9 dan nilai batas bawah (BB) = 0,1.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} * (BA - BB) + BB \quad (1)$$

Tabel 2 menunjukkan nilai *support* dan nilai *confidence* sebelum dan sesudah dilakukan proses *scalling*

Tabel 2. Nilai *support* & *confidence* sebelum dan sesudah *scalling*

No	Sebelum Scalling		Sesudah Scalling	
	Support	Confidence	Support	Confidence
1	0,121	0,164	0,320	0,100
2	0,054	0,208	0,168	0,148
3	0,025	0,256	0,100	0,202
4	0,146	0,280	0,376	0,227
5	0,054	0,310	0,168	0,260
6	0,375	0,508	0,900	0,478
7	0,146	0,557	0,376	0,531
8	0,121	0,690	0,320	0,678
9	0,375	0,720	0,900	0,711
10	0,072	0,744	0,207	0,736
11	0,027	0,786	0,106	0,782
12	0,072	0,806	0,207	0,804
13	0,062	0,893	0,185	0,900

Hasil *scalling* dikelompokkan menjadi 4. Kelompok pertama memiliki nilai *support* dan nilai *confidence* yang tinggi. Kelompok kedua memiliki nilai *support* tinggi dan nilai *confidence* rendah. Kelompok ketiga memiliki nilai *support* rendah dan nilai *confidence* tinggi. Kelompok keempat memiliki nilai *support* rendah dan nilai *confidence* rendah. Transformasi nilai *support* dan nilai *confidence* yang sudah di-*scalling* ke masing-masing kelompok ditunjukkan pada tabel 3.

Tabel 3. Transformasi nilai support & confidence ke pengelompokan

Nilai Scalling	Kelompok			
	1	2	3	4
Support	Tinggi 0,6 - 0,9	Tinggi 0,6 - 0,9	Rendah 0,1 - 0,5	Rendah 0,1 - 0,5
Confidence	Tinggi 0,6 - 0,9	Rendah 0,1 - 0,5	Tinggi 0,6 - 0,9	Rendah 0,1 - 0,5

Secara lengkap, nilai *support* dan nilai *confidence* sebelum dan sesudah dilakukan proses scalling dan hasil pengelompokannya dapat dilihat pada tabel 3.

Tabel 4. Pengelompokan rule berdasarkan nilai support & confidence

No	Premises	Coclusion	Hasil Scalling		Kelompok
			Support	Confidence	
1	Perempuan	Sarjana	0,320	0,100	4
2	Laki-laki	Sarjana	0,168	0,148	4
3	SMP	Laki-laki	0,100	0,202	4
4	SMA	Laki-laki	0,376	0,227	4
5	Sarjana	Laki-laki	0,168	0,260	4
6	Perempuan	SMA	0,900	0,478	2
7	Laki-laki	SMA	0,376	0,531	3
8	Sarjana	Perempuan	0,320	0,678	3
9	SMA	Perempuan	0,900	0,711	1
10	SMP	Perempuan	0,207	0,736	3
11	Tidak Sekolah	Perempuan	0,106	0,782	3
12	Diploma	Perempuan	0,207	0,804	3
13	SD	Perempuan	0,185	0,900	3

Penerapan *Association Rule Mining* dengan *FP-Growth* untuk mengetahui pola pemilik UMKM menghasilkan 13 aturan yang dikelompokkan menjadi 4 kelompok sebagai berikut:

- (1) Kelompok 1 terdiri dari 1 aturan, yaitu aturan jika SMA maka Perempuan dengan nilai *support* 0,375 dan *confidence* 0,72.
- (2) Kelompok 2 terdiri dari 1 aturan, yaitu aturan jika Perempuan maka SMA dengan nilai *support* 0,375 dan *confidence* 0,5.
- (3) Kelompok 3 terdiri dari 5 aturan, yaitu aturan jika Sarjana maka Perempuan, jika SMP maka Perempuan, jika Tidak Sekolah maka Perempuan, jika Diploma maka Perempuan, dan jika SD maka Perempuan. Kelompok ini memiliki nilai *support* kurang dari 0,25 sedangkan *confidence* lebih dari 0,6.
- (4) Kelompok 4 terdiri dari 6 aturan, yaitu aturan jika Perempuan maka Sarjana, jika Laki-laki maka Sarjana, jika SMP maka Laki-laki, jika SMA maka Laki-laki, jika Sarjana maka Laki-laki, dan jika Laki-laki maka SMA. Kelompok ini memiliki nilai *support* kurang dari 0,25 dan *confidence* kurang dari 0,6.

4. KESIMPULAN

Penerapan *Association Rule Mining* dengan *FP-Growth* untuk mengetahui pola pemilik UMKM menghasilkan 13 aturan yang dikelompokkan menjadi 4 kelompok. Dari keempat kelompok tersebut, kelompok pertama adalah kelompok yang memiliki nilai *support* dan nilai *confidence* yang tinggi dibanding kelompok lain. Kelompok ini terdiri dari satu aturan yaitu jika SMA maka Perempuan.

DAFTAR PUSTAKA

- Erwin, (2009), Analisis Market Basket dengan Algoritma Apriori dan FP-Growth, Jurnal Generic, Vol 4, No 2, pp. 26-30
- Gorunescu, F., (2011), Data Mining: Concepts, Models and Techniques, Intelligent Systems Reference Library Vol 12, Springer, London

- Han, J., and Kamber, M., (2007), *Data Mining: Concept and Techniques*, 2nd Ed., Morgan Kaufmann, New York
- Nainggolan, R., (2016), Gender, Tingkat Pendidikan dan Lama Usaha sebagai Determinan Penghasilan UMKM Kota Surabaya, *Kinerja*, Vol 20, No 1, pp. 1-12.
- Pane, D.K., (2013), Implementasi Data Mining pada Penjualan Produk Elektronik dengan Algoritma Apriori (Studi Kasus: Kreditplus), *Pelita Informatika Budi Darma*, Vol 4, No 3, pp 25-29
- Santoso, L.W., (2003), Pembuatan Perangkat Lunak Data Mining untuk Penggalan Kaidah Asosiasi Menggunakan Metode Apriori, *Jurnal Informatika*, Vol 6, No 2, pp. 49-56
- Utari, T., dan Dewi, P.M., (2014), Pengaruh Modal, Tingkat Pendidikan dan Teknologi Terhadap Pendapatan Usaha Mikro Kecil dan Menengah (UMKM) di Kawasan Imam Bonjol Denpasar Barat, *E-Jurnal Ekonomi Pembangunan Universitas Udayana*, Vol 3, No. 12, pp.576-585
- Witten, I.H., Frank, E., and Hall, M.A., (2011), *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd Ed, London