

BAB II

LANDASAN TEORI

2.1 Kajian Pustaka

Penelitian ini menggunakan algoritma *fp-growth* yang diusulkan oleh Han pada tahun 2004 [3]. Algoritma analisis keranjang belanja berbasis pohon (*tree*) ini sangat efektif karena hanya memerlukan dua kali *scan* basis data untuk menyelesaikan tugas penambangan yang ada [3].

Pada penelitian terdahulu algoritma *fp-growth* mampu menyelesaikan beberapa permasalahan. Penelitian yang dilakukan oleh Nelisa dan Halim pada tahun 2021 yang berjudul "Identifikasi Pola Penjualan Kategori Barang dalam Menjaga Stabilitas Stok Barang Menggunakan Algoritma *FP-Growth*" dengan menggunakan algoritma ini *mini market* dapat menyampaikan informasi lebih cepat dan akurat dengan berpacu pada pola yang sudah ditemukan dan dihasilkan oleh *rule* dengan nilai *confidence* 70% dan *minimum support* 30%, sehingga berguna untuk menjaga stabilitas stok produk dan hasil penjualan menjadi lebih meningkat dan terkontrol pada *mini market* Ulfamart [5].

Dalam penelitian yang dilakukan oleh Ikhwan et al. pada tahun 2015 dengan judul "Penerapan Data Mining dengan Algoritma *FP-Growth* untuk Mendukung Strategi Promosi Pendidikan (Studi Kasus Kampus STMIK Triguna Dharma)" yang bertujuan untuk menemukan strategi pemasaran dalam mempromosikan perguruan tinggi yang efektif dengan menggunakan *rule* dari data yang ada [6].

Selanjutnya, pada penelitian yang dilakukan oleh Zia Ghassani et al. pada tahun 2021 yang berjudul “*Market Basket Analysis Menggunakan Algoritma FP-Growth dalam Menentukan Cross-selling*” dilakukan pada November 2019 di koperasi KAOCHEM Sinergi Mandiri, peneliti menggunakan *tool RapidMiner Studio* dengan menggunakan algoritma *fp-growth* dapat ditentukan nilai *minimum confidence*-nya 50% dan *minimum support*-nya 3%, hal tersebut menghasilkan 3 *strong rules* yang dapat diterapkan untuk meningkatkan hasil penjualan dengan teknik pemasaran *cross-selling* [7].

Penelitian yang dilakukan oleh Wahyu Romadhon pada tahun 2018 dengan judul “*Data Mining Asosiasi dalam Menentukan Cross-selling Produk Citramart Menggunakan Algoritma FP-Growth*” tujuan dari penelitian tersebut yaitu menemukan pola pembelian belanja mahasiswa dan karyawan Universitas Amikom Yogyakarta, menemukan aturan *rule*, menemukan keterkaitan barang untuk dijadikan strategi *marketing*, dengan cara mengolah data transaksi yang ada menggunakan algoritma *fp-growth* yang diolah menggunakan *tool RapidMiner* [8].

Dalam penelitian selanjutnya dilakukan oleh Melati dan Wahyuni pada tahun 2020 yang berjudul “*Association Rule dalam Menentukan Cross-Selling Produk Menggunakan Algoritma FP-Growth*” bertujuan untuk menemukan keterkaitan barang dari pola pembelian *customer* yang ada dengan menggunakan algoritma *fp-growth*, serta mengolahnya dengan *tool Weka 8.3* karena *tool* ini memiliki tingkat akurasi *tools* yang lebih baik daripada *RapidMiner* [9]. Adapun untuk melihat hasil penelitian secara lebih detail terdapat pada Tabel 2.1.

Tabel 2. 1 Pemetaan Penelitian Terdahulu

No.	Nama, Tahun	Judul Penelitian	Persamaan	Perbedaan
1.	Zia Ghassani et al. (2021)	<i>Market Basket Analysis</i> Menggunakan Algoritma <i>FP-Growth</i> dalam Menentukan <i>Cross-selling</i>	Algoritma yang digunakan <i>fp-growth</i>	Menggunakan tool <i>RapidMiner Studio</i> , menentukan <i>cross-selling</i>
2.	Nelisa dan Halim (2021)	Identifikasi Pola Penjualan Kategori Barang dalam Menjaga Stabilitas Stok Barang Menggunakan Algoritma <i>FP-Growth</i>	Algoritma yang digunakan <i>fp-growth</i>	Menggunakan tool <i>RapidMiner</i>
3.	Melati dan Wahyuni (2020)	<i>Association Rule</i> dalam Menentukan <i>Cross-Selling</i> Produk Menggunakan Algoritma <i>FP-Growth</i>	Algoritma yang digunakan <i>fp-growth</i>	Menggunakan tool <i>Weka 8.3</i> , menentukan <i>cross-selling</i> produk
4.	Wahyu Romadhon (2018)	<i>Data Mining</i> Asosiasi dalam Menentukan <i>Cross-selling</i> Produk Citramart Menggunakan Algoritma <i>FP-Growth</i>	Algoritma yang digunakan <i>fp-growth</i>	Menggunakan tool <i>RapidMiner</i> , hasilnya berupa aplikasi web

No.	Nama, Tahun	Judul Penelitian	Persamaan	Perbedaan
5.	Ikhwan et al. (2015)	Penerapan Data Mining dengan Algoritma <i>FP-Growth</i> untuk Mendukung Strategi Promosi Pendidikan (Studi Kasus Kampus STMIK Triguna Dharma)	Algoritma yang digunakan <i>fp-growth</i>	Menggunakan <i>tool RapidMiner</i> 5.2

2.2 Dasar Teori

Pembahasan pada dasar teori ini, peneliti menjelaskan tentang semua teori-teori dasar yang terkait dengan penelitian ini.

2.2.1 Data Mining

Data mining adalah kumpulan teknik untuk penemuan otomatis yang efisien dari pola yang sebelumnya tidak diketahui, valid, baru, berguna dan dapat dimengerti dalam *database* besar. Pola harus dapat ditindaklanjuti sehingga dapat digunakan untuk pengambilan keputusan perusahaan [10].

Data mining sendiri dibagi menjadi enam metode untuk melakukan tugasnya, yaitu [6]:

a. Deskripsi (*Description*)

Pada metode ini analisis atau peneliti secara sederhana mencari cara untuk menggambarkan pola dan kecenderungan pada data yang ada.

b. Estimasi (*Estimation*)

Metode estimasi memiliki variabel target estimasi yang lebih kearah numerik daripada kearah kategori.

c. Prediksi (*Prediction*)

Pada metode ini hampir sama dengan estimasi dan klasifikasi akan tetapi berbeda dalam prediksi nilai dari hasil yang akan ada di masa mendatang.

d. Klasifikasi (*Classification*)

Dalam metode klasifikasi ini hampir sama dengan estimasi akan tetapi metode klasifikasi terdapat variabel kategori.

e. Pengelompokan (*Clustering*)

Metode ini merupakan pengelompokan *record*, pengamatan maupun memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan.

f. Asosiasi (*Association*)

Pada metode asosiasi yaitu dengan cara menemukan *attribute* yang muncul dalam satu waktu yang sama atau sering disebut sebagai keranjang belanja.

2.2.2 Association Rule

Association rule atau sering dikenal sebagai '*Market Basket Analysis*' adalah teknik yang sangat efektif untuk menemukan asosiasi penjualan *item X* dengan *item Y*. Dengan kata sederhana, *Market Basket Analysis* terdiri dari

pemeriksaan barang-barang dalam keranjang pembeli yang di-check out untuk melihat jenis barang apa saja yang dibeli secara bersamaan [10].

Association rule adalah proses *data mining* yang berguna untuk menghasilkan seluruh aturan hubungan asosiasi yang memenuhi syarat untuk nilai *minimum support* dan *confidence* pada kumpulan data transaksi yang ada [4].

Persamaan 2.1 merupakan formula untuk menentukan nilai *support* dan **persamaan 2.2** untuk menentukan nilai *confidence* [11].

$$\text{Support}(A, B) = P(A \cap B) = \frac{\text{Jumlah transaksi yang mengandung } A \& B}{\text{Total Transaksi}}$$

Persamaan 2.1.

$$\begin{aligned} \text{Confidence}(A \rightarrow B) &= P(A|B) \\ &= \frac{\text{Jumlah transaksi yang mengandung } A \& B}{\text{Jumlah transaksi yang mengandung } A} \end{aligned}$$

Persamaan 2.2.

2.2.3 Lift Ratio

Suatu transaksi dapat dinilai valid jika nilai *lift* lebih dari satu nilai, itu berarti dalam transaksi tersebut produk A dan B dibeli secara bersamaan [12]. *Lift ratio* mengukur seberapa penting *rule* yang terbentuk berdasarkan nilai *support* dan *confidence* yang ada. *Lift ratio* sendiri digunakan untuk mengukur tingkat keakurasian dari *fp-growth* [11]. Ketika nilai *lift ratio* <1 artinya menandakan kolerasi negatif atau tidak kuat, dan apabila nilainya >1 itu artinya menandakan

korelasi positif atau kuat [8]. **Persamaan 2.3** merupakan rumus untuk menghitung *lift ratio*:

$$Lift(A, B) = \frac{Support(A \cap B)}{Support(A) \times Support(B)}$$

Persamaan 2.3.

2.2.4 Cross-selling

Cross-selling merupakan salah satu teknik *marketing* untuk meningkatkan pembelian produk yang sebelumnya hanya membeli satu produk menjadi beberapa produk dengan cara menawarkan produk yang berkaitan [2].

2.2.5 Algoritma FP-Growth

Algoritma *fp-growth* merupakan algoritma yang digunakan untuk menentukan himpunan data yang paling sering muncul atau sering disebut dengan *frequent itemset* dalam sebuah kumpulan data. Algoritma ini menggunakan struktur penyimpanan data yang dimampatkan yaitu *fp-tree* [11]. Algoritma ini memiliki 3 tahap pengerjaan, yaitu [11]:

a. Tahap Pembangkitan *Conditional Pattern Base*

Conditional pattern base adalah *subdatabase* yang berisikan *prefix path* atau lintasan *prefix*, dan *suffix pattern* atau pola akhiran. Pembangkitan ini didapatkan melalui *FP-Tree* yang telah dibangun sebelumnya.

b. Tahap Pembangkitan *Conditional FP-Tree*

Di tahap ini, *support count* dari setiap item pada setiap tahap pembangkitan sebelumnya dijumlahkan, lalu setiap item yang memiliki jumlah *support count* lebih besar sama dengan *minimum support count* akan dibangkitkan dengan *Conditional FP-Tree*.

c. Tahap Pencarian *Frequent Itemset*

Apabila *Conditional FP-Tree* merupakan lintasan tunggal, maka didapatkan *frequent itemset* dengan melakukan kombinasi item untuk setiap *Conditional FP-Tree*.

Ketiga tahap tersebut merupakan tahap yang akan dilakukan untuk mendapatkan *frequent itemset*.

2.2.6 Weka 3.9.5



Gambar 2. 1 Icon Weka

Weka adalah perangkat lunak *open-source* dibawah GNU (*General Public Licence System*) dengan bahasa pemrograman *Java*. Dikembangkan oleh *Machine Learning Group*, Universitas Waikato, di Selandia Baru. *Weka* adalah perangkat

lunak *data mining* dan merupakan seperangkat algoritma pembelajaran mesin yang dapat diterapkan ke *dataset* secara langsung. *Weka* memiliki *tools* untuk pra-pemrosesan data, klasifikasi, regresi, pengelompokan, aturan asosiasi, dan visualisasi [10]. Pada penelitian kali ini *tool Weka* yang digunakan yaitu *Weka* versi 3.9.5.

