

BAB III

METODOLOGI PENELITIAN

3.1 Subjek dan Objek Penelitian

Subjek penelitian yang digunakan oleh peneliti adalah pola perilaku pelanggan dalam melakukan pembelian barang. Objek yang menjadi tujuan utama penelitian ini dilakukan untuk mengamati pola pada data transaksi yang terdapat dalam data penjualan *e-commerce* yang diperoleh secara daring.

3.2 Alat dan Bahan Penelitian

Alat dan bahan yang dibutuhkan dan digunakan dalam pembuatan penelitian penempatan produk ini antara lain:

3.2.1 Perangkat Keras

Perangkat keras yang digunakan dalam penelitian ini adalah sebagai berikut :

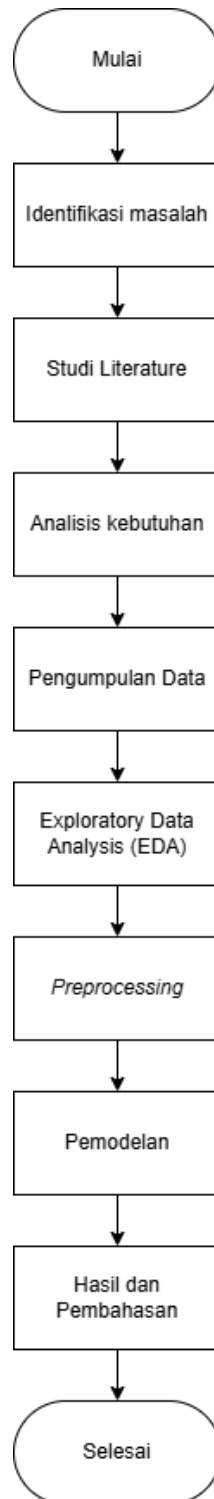
1. Prosesor AMD Ryzen 7 5800H (16 CPUs), ~3,2GHz
2. Ram 16 GB GDDR4
3. Storage SSD 512 GB
4. Keyboard
5. Mouse

3.2.2 Perangkat Lunak

Perangkat Lunak yang digunakan dalam pengembangan penelitian ini adalah sebagai berikut :

1. Sistem Operasi Windows 11 Home 64-Bit
2. Python 3.11
3. Google Collaboratory
4. Jupyter Notebook
5. Microsoft Edge

3.3 Diagram alir penelitian



Gambar 3. 1 Diagram Alir Penelitian

3.3.1 Identifikasi Masalah

Pada Tahap ini akan melakukan analisa terhadap pola pembelian konsumen, dengan mengidentifikasi produk yang sering muncul dan dibeli secara bersamaan, dan segmentasi pasar dengan cara menganalisa kota dengan penjualan terbesar dan kategori3 pada dataset dengan penjualan terbesar, yang kemudian divisualiasikan untuk mempermudah memahami informasi yang disajikan, serta optimasi tata letak produk berdasarkan hasil analisa aturan asosiasi dengan tiga kategori nilai yaitu *support* yang merupakan persentasi kemunculan item secara bersamaan pada *datasest*, *confidence* atau nilai kepastian untuk melakukan analisa persentase frekuensi kemunculan item B(*consequent*) muncul dalam data transaksi yang mengandung item A(*antecedents*), *lift ratio* yang digunakan sebagai parameter untuk menentukan apakah aturan yang dihasilkan oleh aturan asosiasi valid atau tidak, dengan ketentuan dimana jika nilai pada *lift ratio* lebih besar dari satu maka A (*antecedent*) dan b (*consequent*) muncul lebih sering dari yang diharapkan, dan kemunculan A memiliki efek positif terhadap kemunculan B, sedangkan jika nilai pada *lift ratio* lebih kecil dari satu maka A dan B muncul lebih jarang dari yang diharapkan, Kemunculan aturan pada item A tidak memiliki efek positif terhadap kemunculan item B, dan jika nilai pada *lift ratio* sama dengan satu kemunculan A dan B hampir selalu muncul secara bersamaan, tetapi kemunculan aturan A tidak memiliki pengaruh positif terhadap kemunculan aturan B.

3.3.2 Studi Literatur

Bagian ini merupakan tahapan awal dimana penulis melakukan kegiatan membaca untuk mengembangkan konsep dan permasalahan terkait analisis keranjang belanja menggunakan algoritma eclat yang diambil dari jurnal. Dengan melakukan penelitian kepustakaan, penulis telah memperoleh dasar-dasar dalam melakukan dan menulis penelitian.

3.3.3 Analisi Kebutuhan

Pada tahap ini dilakukan proses analisis mengenai kebutuhan penelitian dan pengembangan system untuk melakukan penelitian ini, kebutuhan penulisan dan kebutuhan data merupakan kebutuhan penting pada saat membuat system.

Pentingnya Analisa kebutuhan pada saat membuat suatu system untuk mengantisipasi kekurangan data yang dapat menghambat fungsi system dengan baik. Fase ini penting untuk memaksimalkan proses perancangan dan pengembangan sistem

3.3.4 Pengumpulan Data

Pada penelitian yang dilakukan data yang digunakan merupakan data sekunder yang diperoleh menggunakan data penjualan *e-commerce* Tokopedia pada bulan April dan Juni pada tahun 2023. data pada *e-commerce* Tokopedia yang digunakan memiliki 34185 baris dan 23 kolom. Berikut adalah beberapa *variable* pada data Tokopedia bulan April-Juni “LinkProduk”, “Hyperlink”, “NamaProduk”, “LinkToko”, “Hyperlink2”, “NamaToko”, “Harga”, “HargaAtas”, “OmsetMinimum”, “JumlahTerjual”, “Kategori1”, “Kategori2”, “Kategori3”, “KotaPengiriman”, “PenjualanSebelumnya”, “KenaikanPenjualan”, “PersentaseKenaikanPenjualan”, “ReveiwSebelumnya”, “KenaikanReviewPenjualan”, dan “PersentaseKenaikanReview” data ini selanjutnya harus melalui proses *preprocessing* sebelum bisa diimplementasikan pada model.

NamaToko	Harga	HargaAtas	OmsetMinimum	JumlahTerjual	Kategori1	Kategori2	Kategori3	KotaPengiriman	PenjualanSebelumnya	KenaikanPenjualan	PersentaseKenaikanPenjualan	ReviewSebelumnya
Sgameshop	102000	102000	497760000	4880	Gaming	Aksesoris Game Console	Adapter Game Console	Jakarta Barat	4878.0	2.0	0.041000	2023.0
Butikgames	200000	200000	285200000	1426	Gaming	Aksesoris Game Console	Adapter Game Console	Jakarta Pusat	1400.0	26.0	1.857143	840.0
Timur Game Shop	28000	28000	66500000	2375	Gaming	Aksesoris Game Console	Adapter Game Console	Bekasi	2306.0	69.0	2.992194	922.0
cello game	13000	13000	25610000	1970	Gaming	Aksesoris Game Console	Adapter Game Console	Tangerang Selatan	1876.0	94.0	5.010661	570.0
Timur Game Shop	8500	8500	35810500	4213	Gaming	Aksesoris Game Console	Adapter Game Console	Bekasi	4059.0	154.0	3.794038	768.0

Gambar 3. 2 Data Tokopedia Bulan April-Juni

3.3.5 Exploratory Data Analysis

Pada bagian ini EDA dilakukan untuk memahami dan mengidentifikasi data untuk kebutuhan *preprocessing* yang lebih spesifik. Kemudian EDA dilakukan untuk mencari informasi pada *dataset* untuk kepentingan penelitian.

3.3.6 Preprocessing

Pada bagian *preprocessing* sebelum data dapat diimplementasikan pada model, data harus melalui beberapa tahapan seperti *data cleansing* yang bertujuan untuk membersihkan data dari NaN (Not a Number) dan nilai yang hilang seperti *Missing Value*, kemudian melakukan *data transformation* yang bertujuan untuk mengubah data menjadi format yang dapat diproses oleh algoritma. Seperti yang bisa dilihat pada Gambar 3.3 sampai 3.5 yang merupakan hasil dari proses *data cleansing* dan *data transformation*.

```
[5 rows x 23 columns]
##### NA #####
LinkProduk          0
Hyperlink           0
NamaProduk          0
LinkToko             0
Hyperlink2          0
NamaToko            0
Harga               0
HargaAtas           0
OmsetMinimum        0
JumlahTerjual       0
JumlahReview        0
JumlahRating        0
ValidOrder          1
Kategori1           18
Kategori2           18
Kategori3           288
KotaPengiriman      0
PenjualanSebelumnya 7882
KenaikanPenjualan   7882
PersentaseKenaikanPenjualan 7883
ReviewSebelumnya    7882
KenaikanReviewPenjualan 7882
PersentaseKenaikanReview 7886
dtype: int64
```

Gambar 3. 3 Data sebelum melalui tahap data cleansing

Pada gambar 3.3 dapat dilihat jumlah kolom yang memiliki nilai yang bukan angka atau nilai yang kosong diberbagai kolom sangat banyak, maka diperlukan proses *Data cleansing* untuk membersihkan data sehingga data yang digunakan akurat dan sesuai dengan kebutuhan.

```
[5 rows x 23 columns]
##### NA #####
LinkProduk          0
Hyperlink           0
NamaProduk          0
LinkToko            0
Hyperlink2          0
NamaToko            0
Harga               0
HargaAtas           0
OmsetMinimum        0
JumlahTerjual       0
JumlahReview        0
JumlahRating        0
ValidOrder          0
Kategori1           0
Kategori2           0
Kategori3           0
KotaPengiriman     0
PenjualanSebelumnya 0
KenaikanPenjualan   0
PersentaseKenaikanPenjualan 0
ReviewSebelumnya   0
KenaikanReviewPenjualan 0
PersentaseKenaikanReview 0
dtype: int64
```

Gambar 3. 4 Data setelah melalui tahap preprocessing

Pada gambar 3.4 dengan menggunakan perintah *dropna* data yang kosong atau tidak memiliki nilai akan dihilangkan, dan mengganti tipe data pada kolom “Harga” sebagai *numeric* atau *integer* dan mengatur nilai pada kolom “Harga”, dan “KenaikanPenjualan” harus lebih besar dari nol sehingga kolom yang tidak bernilai tidak digunakan.

	Motherboard	3D Puzzle	AC Portable	AC Reflektor	AC Standing	Abaya	Abon	Abs Roller	Access Control Door	Action Camera	...	Yoga Mat	Yogurt	Yukata Pria	Yukata Wanita	benang dan tusuk gigi	eyeshadow	gasing	iOS
0	False	False	False	False	True	True	False	False	False	True	...	False	False	True	False	False	False	False	False
1	False	False	False	False	False	False	False	False	False	False	...	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False	False	...	False	False	False	False	False	False	True	False
3	False	False	False	False	False	True	False	False	False	False	...	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False	False	...	False	False	False	False	False	False	False	False
5	True	False	False	False	True	True	False	False	True	True	...	False	False	False	False	False	True	False	False
6	False	False	False	False	False	False	False	False	True	True	...	False	True	False	False	False	False	False	True
7	False	False	False	False	False	False	False	False	False	False	...	False	False	False	False	False	False	False	False
8	True	False	True	True	True	False	False	False	False	False	...	False	False	False	False	False	False	False	False
9	False	False	False	False	True	False	False	False	False	False	...	False	False	False	False	False	False	False	False

Gambar 3. 5 Hasil dari data transformation

Pada gambar 3.5 *dataset* diubah menjadi bentuk yang dapat diolah oleh algoritma menggunakan *library TransactionEncoder* kemudian menyimpan nilainya kedalam vairabel *onehot* dengan nilai *binary* yang berisi satu atau nol, atau true atau false.

3.3.7 Pemodelan

Pada tahap ini data yang telah diimplementasikan pada sistem selanjutnya digunakan untuk desain tata letak produk pada *e-commerce* berdasarkan pola pembelian dan perilaku pelanggan.

3.3.7.1 *Market Basket Analysis*

Pada bagian ini *Market Basket Analysis* bertujuan untuk menemukan barang, produk atau jasa yang sering dibeli pada satu transaksi yang sama, untuk mencari produk, barang atau jasa yang paling sering dibeli pada satu transaksi yang sama dibutuhkan algoritma asosiasi. Pada penelitian ini algoritma asosiasi yang digunakan adalah algoritma Eclat.

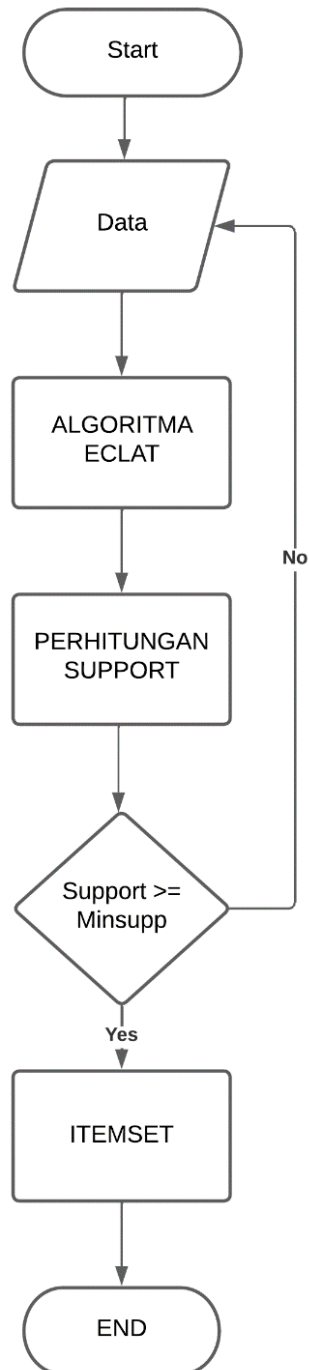
3.3.7.1 Algoritma Eclat

Pada tahap pemodelan algoritma Eclat mencari nilai *support* untuk menemukan kemunculan *item* yang paling sering muncul pada *dataset* kemudian hasil yang ditemukan disimpan pada vairabel *frequent_itemsets*, *frequent_itemsets* digunakan untuk mencari hubungan antar *item* pada *dataset* menggunakan aturan asosiasi yang kemudian menghasilkan tiga kategori nilai yaitu *support*, *confidence*, dan *lift ratio*

3.3.8 Hasil dan Pembahasan

Pada bagian ini peneliti memaparkan hasil dari visualisasi penjualan pada kolom “KotaPengiriman” untuk melihat kota yang memiliki jumlah penjualan terbesar pada *dataset* visualiasasi dilakukan dengan menggunakan *scatter plot* atau *px.scatter* pada bahasa pemrograman *python*, kemudian visualisasi pada kolom “Kategori3” yang bertujuan untuk melihat barang apa saja yang memiliki penjualan terbesar pada *dataset*. Selanjutnya membahas hasil dari visualisasi yang berisi pola atau tren barang pada tiap kota dan barang yang ada di *dataset*. Hasil dari pemrosesan algoritma kemudian ditampilkan dalam bentuk tabel, dan divisualisasikan menggunakan *network graph*. *Rules* atau aturan yang ditemukan kemudian diimplementasikan pada UI yang bertujuan sebagai strategi penjualan pada *e-commerce* Tokopedia.

3.4 Diagram alir algoritma *Eclat*



Gambar 3. 6 Diagram alir algoritma *Eclat*

3.4.1 Diagram alir *Eclat*

Pada tahap ini data yang sudah dipersiapkan kemudian akan diproses oleh algoritma *Eclat*, selanjutnya *Eclat* akan melakukan perhitungan *support* jika *item* atau produk tidak lebih besar atau sama dengan ambang batas pada nilai *support* yang sudah ditetapkan maka *item* akan di *drop* atau dikembalikan pada *dataset* dan tidak digunakan, sebaliknya jika *item* sesuai dengan nilai ambang batas yang sudah ditetapkan maka akan di simpan pada *itemsets*.

3.4.3 Perhitungan nilai *support*

Perhitungan nilai *support* dilakukan dengan rumus berikut, yang mana jumlah transaksi yang terdapat A (*antedecent*) didapatkan dengan cara melihat kemunculan item A pada data transaksi, dan jumlah total transaksi didapatkan dengan menghitung jumlah baris yang digunakan pada data transaksi

$$Support (A) = \frac{\text{Jumlah transaksi yang terdapat A}}{\text{Jumlah Total transaksi}} \quad (3.4)$$

Sehingga untuk menemukan nilai *support* pada data transaksi dilakukan dengan rumus seperti pada rumus 3.1 dengan membagi jumlah transaksi yang terdapat A dan jumlah total transaksi.

Tabel 3. 1 Perhitungan nilai *support*

NO	<i>Item</i>	Jumlah transaksi yang terdapat A	<i>Jumlah Total Transaksi</i>	<i>Support (%)</i>
1	Joystick Game	1195	23901	5
2	Casing & Silikon Game Console	956	23901	4
3	CD PC dan Laptop Gaming	956	23901	4
4	Gunting Kuku	717	23901	3
5	Playstation	956	23901	4

Pada tabel 3.1 hasil perhitungan berdasarkan rumus disajikan dalam bentuk tabel yang berisi Jumlah transaksi yang mengandung A(*antedecent*) dan jumlah total transaksi, serta nilai *support* yang dihasilkan dari perhitungan.