

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1 Tinjauan Teori**

Dalam sebuah penelitian membutuhkan teori-teori yang akan menjadi landasan teoritis dan menjadi pedoman dalam melaksanakan penelitian.

##### **2.1.1. Tinjauan Teori Umum**

###### **2.1.1.1. Pengertian Data Mining**

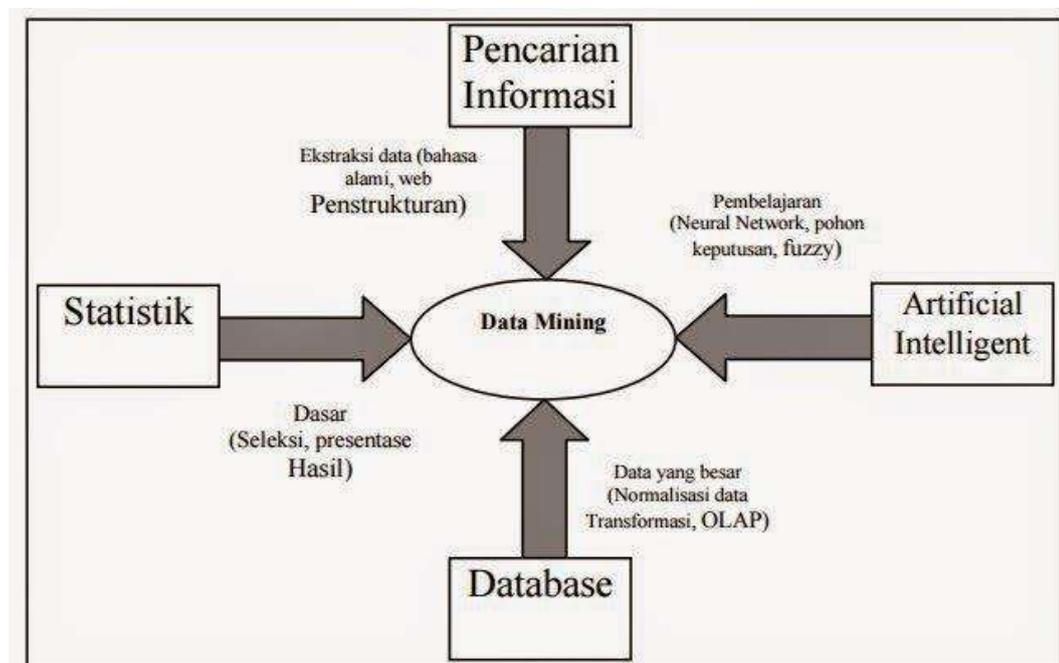
Proses ekstraksi informasi dari kumpulan data melalui penggunaan algoritma dan teknik yang melibatkan bidang ilmu statistik, mesin pembelajaran dan sistem manajemen database. Data mining digunakan untuk ekstraksi informasi penting yang tersembunyi dari dataset yang besar. Adanya data mining akan didapatkan suatu permata berupa pengetahuan didalam kumpulan data-data yang banyak jumlahnya. (Lestari & Hafiz, 2020).

Data mining merupakan metode untuk melihat melalui kumpulan data besar yang disimpan dalam penyimpanan untuk menemukan tautan, pola, dan tren yang signifikan dalam penyimpanan memerlukan penggunaan teknik pengenalan pola seperti statistik dan matematika. (Ependi & Akbar, 2019). Data Mining adalah sesuatu ilmu yang menguraikan hasil temuan berupa pengetahuan pada sekumpulan informasi untuk menemukan pola dan aturan yang berarti. Pola tersebut didapatkan dari relasi database, data transaksi (Haidar, 2021).

Data Mining ialah sebuah istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database (Ria dkk., 2019). Data mining menjadi alat yang semakin penting untuk mengubah data tersebut menjadi informasi. Hal

ini sering digunakan dalam berbagai praktek profil, seperti pemasaran, pengawasan, penipuan deteksi dan penemuan ilmiah (Wardah & Fitriana, 2021).

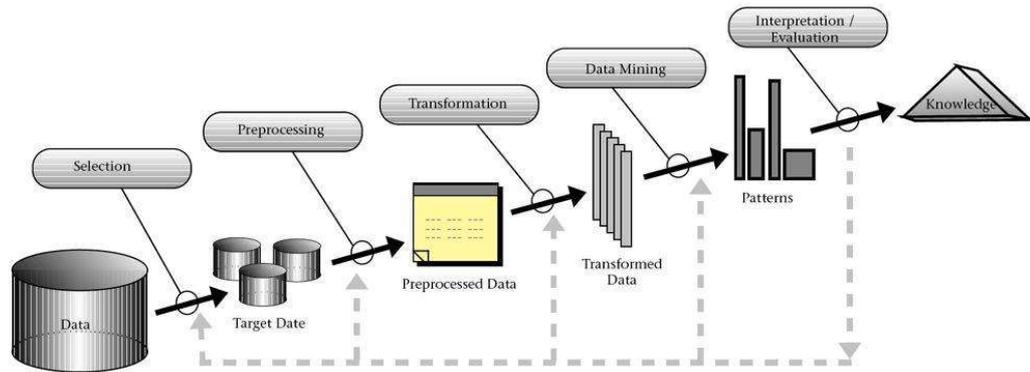
Dapat disimpulkan bahwa data mining merupakan disiplin ilmu untuk menganalisa data dan menggali informasi atau pola pengetahuan baru yang tidak diketahui sebelumnya dan akan bermanfaat dimasa yang akan datang. Dari beberapa pengertian diatas gambaran bidang ilmu data mining dapat dilihat pada gambar 2. 1 berikut:



**Gambar 2. 1** Bidang Ilmu Data Mining

Kusrini & Lutfi dalam (Ulfa, 2018) terdapat istilah lain yang memiliki makna yang sama dengan data mining yaitu *knowledge discovery in database* (KDD). Istilah data mining dan KDD seringkali digunakan secara bergantian untuk menjelaskan suatu proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Sebenarnya kedua istilah tersebut memiliki konsep yang

berbeda tetapi berkaitan satu sama lain. Dan salah satu tahapan keseluruhan proses KDD adalah data mining.



**Gambar 2.2** Tahapan dalam KDD

Dalam (Rumahorbo & Arnomo, 2020) terdapat 4 langkah data mining yaitu:

1. Menjelaskan permasalahan yang ingin diuraikan.
2. Mengumpulkan data yang akan digunakan untuk penelitian.
3. Validasi data
4. Memilih aplikasi yang sesuai untuk menguji data dan memberikan hasil yang tepat.

#### 2.1.1.2. Pengelompokan Data Mining

Berdasarkan fungsionalitasnya, tugas-tugas data mining bisa dikelompokkan ke dalam kelompok berikut ini (Femy Mulya dkk., 2019):

1. Klustering

Mengelompokkan obyek ke dalam beberapa kelompok berdasarkan kemiripan antar obyek, dimana dalam satu klaster harus berisi obyek yang saling mirip dan antar klaster obyek saling tidak mirip.

## 2. Klasifikasi

Melakukan pengelompokan obyek berdasarkan kelompok yang sudah ada. Berbeda dengan klustering, klasifikasi ini memerlukan data pelatihan yang sudah diberi label kelompok/kelas.

## 3. Regresi / Estimasi

Regresi pada dasarnya mirip dengan klasifikasi, yakni memerlukan data pelatihan yang sudah diberi label. Bedanya, output klasifikasi adalah nilai diskrit, sedangkan output dari regresi adalah kontinyu, regresi ini mencari model hubungan antara atribut prediktor dan atribut dependen, dimana atribut dependennya juga berupa nilai kontinyu.

## 4. Asosiasi

Asosiasi bertugas menemukan atribut-atribut yang terjadi bersamaan. Asosiasi mencoba untuk menemukan aturan untuk mengkuantifikasi hubungan antara dua atau lebih atribut. Aturan asosiasi berbentuk *If antecedent, then consequent* dan *confidence* yang berhubungan dengan aturan.

### 2.1.1.3. Association Rule Mining

Association Rules merupakan salah satu metode data mining yang bertujuan untuk menemukan semua aturan asosiasif dari beberapa himpunan item. Tahapan analisis asosiasi yang perlu diperhatikan dalam Association Rules adalah frequent patterns yang bertujuan untuk mengetahui seberapa sering munculnya kombinasi antar item (Rhomadhona dkk., 2021): Analisis asosiasi atau association rule mining adalah teknik data mining untuk menemukan aturan asosiasi antara suatu

kombinasi item. *Interestingness measure* yang dapat digunakan dalam data mining adalah sebagai berikut.

1. *Support* adalah suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu item atau itemset dari keseluruhan transaksi.
2. *Confidence* adalah suatu ukuran yang menunjukkan hubungan antar dua item secara *conditional* (berdasarkan suatu kondisi tertentu). (Ria dkk., 2019):

Analisis asosiasi berguna untuk menemukan hubungan penting yang tersembunyi diantara set data yang sangat besar. Hubungan yang sudah ada tersebut direpresentasikan dalam bentuk aturan asosiasi (*association rules*) atau aturan item yang sering muncul

#### **2.1.1.4. Aturan Asosiasi**

Pencarian pola asosiasi berawal dari pengolahan data transaksi penjualan buku, kemudian dicari hubungan antar buku yang dibeli. Proses pencarian asosiasi ini menggunakan bantuan algoritma apriori yang merupakan algoritma yang digunakan untuk menghasilkan *association rule* dengan pola “*if then*” yang berfungsi untuk membentuk kombinasi item yang mungkin, kemudian diuji apakah kombinasi tersebut memenuhi parameter *support* dan *confidence* minimum yang merupakan nilai ambang yang diberikan oleh user. Dengan algoritma apriori tersebut akan menghasilkan pola kombinasi item dan rules sebagai ilmu pengetahuan dan informasi penting dari data transaksi penjualan.

Untuk memperoleh ketentuan asosiatif dibutuhkan pencarian ketentuan yang mempunyai pola frekuensi besar (PFT). PFT dicari dengan cara mencari

ketentuan yang penuh nilai support minimum. Nilai *support* (penunjang) merupakan persentase item ataupun campuran item yang terdapat pada totalitas informasi (Iswandi dkk., 2020).

Untuk menemukan aturan asosiasi seperti yang diharapkan maka harus menemukan nilai dari support yang telah ditentukan. Support tersebut merupakan jumlah item pada setiap transaksi yang ada didalam database. Untuk dapat menemukan nilai support kita dapat mencari semua aturan yang jumlah support minsup. Dalam hal ini dapat digunakan sebagai cara untuk menemukan sebuah nilai *confidence*. Nilai *confidence* ditentukan dari nilai support suatu aturan dalam sebuah transaksi.

Jika itemset pada setiap transaksi tidak sering muncul (*infrequent*), maka kandidat yang tidak sesuai dengan nilai support minsup tersebut harus segera dipangkas tanpa harus menghitung nilai *confidence*-nya. Strategi umum digunakan oleh banyak algoritma penggalian aturan asosiasi adalah memecahkan masalah kedalam dua pekerjaan utama, yaitu (Suyanto, 2017);

1. *Frequent itemset generation*

Tujuannya adalah mencari semua itemset yang memenuhi ambang batas minsup. Itemset itu disebut itemset frekuen (itemset yang sering muncul).

2. *Rules generation*

Tujuannya adalah mengekstrak aturan dengan *confidence* tinggi dari itemset frekuent yang ditemukan dalam langkah sebelumnya. Aturan ini kemudian disebut aturan yang kuat (*strong rules*).

### 2.1.1.5. Algoritma Apriori

Algoritma apriori termasuk jenis aturan asosiasi pada data mining yang dikembangkan pertama kali oleh R. Agrawal dan R. Srikant pada tahun 1994. Algoritma ini didasarkan pada fakta bahwa apriori menggunakan pengetahuan sebelumnya dari suatu itemset dengan frekuensi kemunculan yang sering atau disebut frequent itemset. Apriori menggunakan pendekatan iteratif dimana kitemset digunakan untuk mengeksplorasi (k+1)- itemset berikutnya. Prinsip metode apriori adalah jika suatu itemset sering muncul (frequent), maka semua subset dari itemset tersebut juga harus sering muncul dalam suatu database (Rhomadhona dkk., 2021):

Algoritma apriori adalah algoritma pengambilan data dengan aturan asosiatif (*association rule*) untuk menentukan hubungan asosiatif suatu kombinasi item. Analisis pola frekuensi tinggi dengan algoritma apriori mencari kombinasi item yang memenuhi syarat minimum dari nilai support dalam basis data. Nilai support sebuah item diperoleh dengan menggunakan rumus berikut

$$\text{Support (A)} = \frac{\text{Jumlah transaksi mengandung A}}{\text{Total transaksi}} \times 100\%$$

Nilai Support dari 2item diperoleh dengan menggunakan rumus:

$$\text{Support (A, B)} = P(A \cap B)$$

$$\text{Support (A, B)} = \frac{\sum \text{transaksi mengandung A dan B}}{\sum \text{transaksi}} \times 100\%$$

**Rumus 2.1** Perhitungan Support

**Sumber:** (Haidar, 2021)

Pembentukan aturan asosiasi setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan asosiatif A U B. Nilai *confidence* dari aturan A U B diperoleh dengan rumus berikut:

$$Confidence = P(B|A) = \frac{\sum \text{transaksi mengandung A dan B}}{\sum \text{transaksi mengandung A}} \times 100\%$$

**Rumus 2.2** Pembentukan Aturan Asosiasi

**Sumber:** (Haidar, 2021)

Untuk menentukan aturan asosiasi yang akan dipilih maka harus diurutkan berdasarkan  $Support \times Confidence$ . Aturan diambil sebanyak n aturan yang memiliki hasil terbesar.

### 2.1.2. Tinjauan Teori Khusus

#### 2.1.2.1. Tanagra

Menurut (Oktariani & Utnasari, 2020) Tanagra merupakan program open source yang setiap orang yang menggunakannya dapat mengakses dengan langsung pada kode source program dan melakukan penambahan algoritma-algoritma yang sesuai izin dari perangkat lunak. Adapun tujuan tanagra yaitu:

1. Tanagra bertujuan untuk memberikan kemudahan untuk peneliti dan mahasiswa dalam mengakses perangkat lunak data mining.
2. Tanagra juga bertujuan dalam memberikan usulan kepada peneliti arsitektur agar dapat menambahkan metode data mining pada data mereka dan dijadikan sebagai tolak ukur kinerja mereka.
3. Tujuan ketiga ditunjukkan kepada pengembang pemula, yang dimana dalam menyebarkan metodologi yang digunakan dalam membuat software dengan

jenis yang sama. Mereka juga bisa memperoleh suatu keuntungan dalam mengakses gratis ke kode sumber ini, yang dimana digunakan untuk mengetahui bagaimana perangkat lunak ini dibuat.

#### **2.1.2.2. Penjualan**

Penjualan yaitu syarat dalam berlangsungnya sebuah usaha, dengan melakukan proses penjualan usaha tersebut akan mendapatkan sebuah keuntungan. Semakin tinggi atau semakin banyak proses jual beli yang terjadi keuntungan yang didapatkan akan berbanding lurus. Pasar pokok penjualan ialah suatu ilmu dan seni untuk memberikan pengaruh pada pribadi pola penjualan untuk mengajak dan menarik perhatian orang lain untuk membeli barang yang ditawarkan.

Dalam (Oktariani, 2020) adapun beberapa faktor yang memberikan pengaruh terhadap penjualan sebagai berikut :

1. **Kondisi Pasar**

Pasar merupakan tempat terjadinya penjualan antar pihak penjual dan pembeli. Faktor kondisi pasar yaitu jenis kelompok pasar pembeli, segmen pasar, daya beli, frekuensi pembelian, dan kebutuhannya sehingga penjual dapat menghasilkan pemasukan yang maksimal.

2. **Kondisi dan Kemampuan Penjual**

Transaksi jual beli pada prinsipnya melibatkan dua pihak yaitu penjual sebagai pihak pertama dan pembeli sebagai pihak kedua. Karena itu, penjual atau pihak pertama harus bisa meyakinkan pembelinya atau pihak kedua agar bisa mencapai sasaran penjualan yang diinginkan.

### 3. Modal

Pada awalnya penjual membutuhkan modal untuk dapat memperlihatkan kepada pembeli produk yang diperjualkan melalui sarana usaha seperti biaya transportasi, biaya mempromosikan barang, biaya lainnya.

### 4. Kondisi Organisasi Perusahaan

Pada perusahaan besar penjualan akan ditangani oleh bagian marketing perusahaan atau pihak yang ahli dalam bidang penjualan. Lain halnya dengan perusahaan kecil atau usaha-usaha kecil yang dimana proses penjualan ditangani oleh pihak yang juga mengurus banyak hal dalam perusahaan atau usaha tersebut.

#### **2.1.2.3. Tas**

Menurut Gunarto dan Sugiyono di dalam (Leona, 2019) tas adalah tempat untuk menyimpan sesuatu baik alat-alat ataupun barang-barang yang diperlukan, misalnya sebagai tempat surat, buku, pakaian, dan sebagainya. Tas digunakan oleh berbagai kalangan, mulai dari anak sekolah, mahasiswa, dosen, karyawan, hingga pemilik perusahaan. Sehingga dapat dikatakan bahwa semua kalangan membutuhkan tas dan penggunaan tas tidak dibatasi oleh usia maupun pekerjaan. Tas mempunyai berbagai macam bentuk dan material yang disesuaikan dengan kegunaannya. Selain dari segi fungsionalitasnya, tas pada zaman sekarang juga digunakan oleh perempuan metropolitan sebagai barang pelengkap padu padan busana untuk menyelaraskan gaya dan tampil modern.

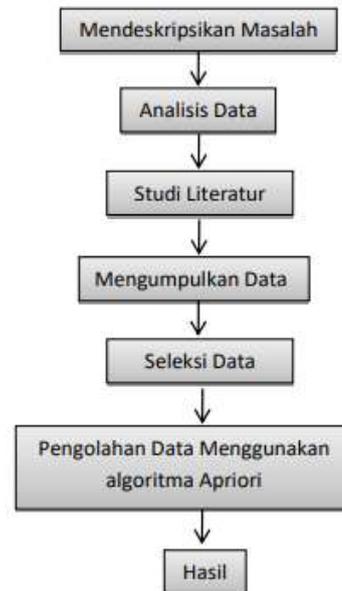
Tas adalah wadah tertutup yang dapat dibawa bepergian materi untuk membuat tas antara lain adalah kertas, plastik, kulit, kain, dan lain-lain. Biasanya

digunakan untuk membawa pakaian, buku, dan lain-lain. Tas yang dapat digendong di punggung disebut ransel, sedangkan tas yang besar untuk memuat pakaian disebut koper (dari bahasa Belanda *koffer*). adapula tas yang hanya berbentuk kotak yang biasanya dipergunakan oleh kaum wanita untuk membawa peralatan kecantikannya, biasanya disebut dengan tas kecantikan atau beauty case. Sekarang ini, banyak sekali model tas yang tersedia dan setiap model punya gaya serta fungsi tersendiri (Junayni, 2020). Tas dapat dibuat dari berbagai macam bahan antara lain dari bahan logam, kulit, plastik, kayu, bahkan dari bahan kain. Seiring berkembangnya teknologi tas pada jaman sekarang lebih banyak menggunakan bahan sintetis. Bahan sintetis merupakan bahan yang terbuat dari proses kimiawi, seperti tas berbahan kulit sintetis, kanvas sintetis, togo, dan clemence (Septyani&Musdalifah, 2019).

Berbagai kegiatan atau acara membutuhkan satu jenis model tas. Tas juga jadi barang koleksi, baik secara khusus oleh para pencinta fashion maupun sekedar untuk memiliki banyak opsi ketika dibutuhkan saja. Berikut berbagai jenis model tas :

1. Berdasarkan Pengguna, antara lain tas wanita, tas pria dan tas anak-anak.
2. Berdasarkan bentuk dan ukuran, antara lain tas selempang (*sling bag*), ransel, tas pesta, *beauty case*, tas laptop, tas hobo, ibag dan koper.
3. Berdasarkan bahan, antara lain tas kertas, tas plastik, tas rajut, tas *spunbound*, tas berbahan alami dan tas daur ulang.

## 2.2. Kerangka Pemikiran



**Gambar 2.3** Kerangka Pemikiran

**Sumber :** (Peneliti, 2022)

Penjelasan dari kerangka pemikiran pada *flowchart* di atas: Adapun desain penelitian pada penelitian ini dijelaskan sebagai berikut:

1. Mendeskripsikan Masalah

Mendeskripsikan masalah dalam penelitian dengan menentukan dan mendefinisikan batasan masalah yang akan diteliti, sehingga dapat memudahkan dan membantu dalam mendapatkan solusi terbaik dari masalah tersebut.

2. Analisis Masalah

Analisis masalah merupakan langkah dalam memahami sebuah masalah yang ditentukan ruang lingkup dan batasan masalahnya, maka masalah yang dapat dipahami dengan baik.

3. Studi Literatur

Untuk mencapai tujuan, maka peneliti mempelajari literatur yang berhubungan dengan konsep dari penerapan Association rule dan Algoritma apriori. Sumber literatur dapat diperoleh didalam buku, maupun jurnal dan situs-situs yang dapat menjadi penunjang lainnya untuk digunakan saat menentukan teknik yang tepat untuk menyelesaikan masalah yang diteliti.

#### 4. Mengumpulkan Data

Dalam mengumpulkan data yang dibutuhkan dalam penelitian ini yaitu dilakukan dengan cara observasi yaitu pengamatan secara langsung di PT. Tasindo Total Inproducts sehingga dapat diketahui permasalahan yang sedang terjadi secara langsung. Kemudian melakukan interview yang bertujuan untuk mendapatkan informasi maupun data yang peneliti butuhkan.

#### 5. Seleksi Data

Data yang telah didapatkan pada proses pengumpulan data selanjutnya dilakukan proses seleksi supaya data yang telah dipilih tidak ada yang duplikat, menghilangkan kesalahan pada data (tipografi) dan menghilangkan data yang kosong, pada tahap ini data dipilih untuk digunakan pada proses mining.

#### 6. Implementasi Algoritma

Adapun tahapan dalam langkah-langkah pengimplementasian algoritma apriori yaitu:

- a. Analisis pola frekuensi tinggi.
- b. Pembentukan *Association rule*.

## 7. Pengujian Hasil

Pada tahap ini, peneliti melakukan pengujian data pada penjualan PT Tasindo Total Inproducts menggunakan Software open source Tanagra, sehingga mendapatkan suatu hubungan dari data tersebut.

### 2.3. Hipotesis Penelitian

Hipotesis penelitian berdasarkan kerangka pemikiran diatas adalah sebagai berikut:

1. Diduga proses analisa data mining pada transaksi data penjualan dengan algoritma apriori untuk mengetahui pola pembelian konsumen.
2. Diduga nilai *support* dan *confidence* yang terbentuk sehingga dapat menentukan pola pembelian konsumen

### 2.4. Penelitian Terdahulu

Melihat masalah dan judul penelitian yang akan diteliti, maka diperlukan adanya pemaparan tentang penelitian terdahulu guna mengungkapkan fenomena yang sama dalam sudut pandang yang berbeda sehingga diharapkan dapat memperkaya pengetahuan. Secara ringkas, penelitian terdahulu yang mendukung penelitian yang akan dilakukan dapat dilihat di bawah ini:

**Tabel 2. 1** Hasil Penelitian Terdahulu

No.	Nama Peneliti	Judul Penelitian	ISSN	Hasil
1	Cep Adiwihardja, Murni Cahyati, & Reynida Hilma (2018)	Implementasi Data Mining Penjualan Tas Pada Toko Fabella Shop Menggunakan Algoritma Apriori	2088-0154	Dari hasil penelitian telah diketahui bahwasannya konsumen yang membeli tas merek burberry maka akan membeli tas merek chanel menghasilkan nilai support 33,33% dan confidence 66,67%
2	Agus Salim, & Mochammad Nizar (2020)	<i>Application of Apriori Algorithm Method in Sales Analysis of Mountain Bag Brands in Post Stores 1</i>	2549-6255	Bedasarkan Analisa yang dilakukan penulis, bahwa carrier yang paling diminati oleh pendaki adalah Avtech, Consina, Cozmed. Dari hasil tersebut bisa digunakan oleh toko pos 1 untuk mempersiapkan stok merek produk tas gunung yang banyak dibeli oleh pembeli dan menambah persediaan merek.
3	Errina Azizah (2019)	Aplikasi Penjualan Pada Toko Tas Dan Sepatu Tiga Puluh Berbasis Website Menggunakan Framework Codeigniter	Skripsi	Hasil dari penelitian adalah membuat aplikasi penjualan pada Toko Tas Tiga Puluh berbasis website. Metode pengembangan sistem yang digunakan adalah waterfall, yaitu analisa, desain, pengkodean, pengujian, dan pemeliharaan.

4	Muhammad Iqbal Islami, & Adhi Prasetio (2018)	Analisis Dan Prediksi Penjualan Pada Marketplace Berdasarkan Pendekatan Klasifikasi Dengan Metode Pohon Keputusan (Studi Kasus Pada Data Katalog Dan Penjualan Tas Pria & Wanita Di Tokopedia Secara Nasional)	2355-9357	Diketahui semakin tinggi sebuah rating produk tas pria & wanita semakin tinggi juga rata – rata pengunjung yang tertarik untuk membuka detail produk. Jenis produk tas pria & wanita yang paling tinggi penjualan dan viewernya adalah jenis produk Ransel.
5	Bagas Praditya Yusuf Putra, Sari Wulandari & Bobby Hera Sagita (2019)	Perancangan Program Komunikasi Pemasaran Tas Pada Ukm Levaya Menggunakan Metode Benchmarking	2355-2085	Berdasarkan hasil penelitian ini diperoleh rancangan program komunikasi pemasaran yang baru untuk UKM Levaya pada produk tas wanitanya.

**Sumber:** (Peneliti, 2022)