# BAB II LANDASAN TEORI

## 2.1. Data Mining

Data mining adalah gabungan dari beberapa disiplin ilmu yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, database, dan visualisasi untuk penanganan permasalahan pengambilan informasi dari basis data yang besar [8]. Data Mining Data mining merupakan salah satu teknik penggolongan data dengan tujuan mencari kolerasi antar data yang tidak diketahui oleh pengguna dan menyajikannya dalam hasil yang mudah dipahami dan kolerasi data tercatat dapat digunakan sebagai dasar pengambilan keputusan. Penambangan data dapat terbagi dari beberapa kelompok seperti Deskripsi, Evaluasi, Prediksi, Klasifikasi, Pengelompokan dan Pemetaan berdasarkan tugas yang akan dilakukan.

Data mining adalah suatu proses dimana sejumlah besar data digali dan dianalisis untuk mendapatkan suatu yang valid, baru dan bermanfaat, alhasil bisa menemukan pola atau formula dalam data tersebut [9]. Penambangan data secara umum dapat diklasifikasikan menjadi dua kategori utama, penambangan deskriptif dan penambangan prediktif. Deskriptif Mining adalah proses menemukan informasi penting dari database. Sedangkan prediktif adalah Proses penentuan polaritas data dengan menentapkan beberapa variabel untuk memperkirakan variabel lain di masa mendatang. Data Mining ialah bagian integral dari pengetahuan pada basis data atau sering disebut dengan Knowledge Discovery in Databases (KDD), yang berarti progres umum transformasi data mentah menjadi

model yang berguna, yaitu informasi yang dibutuhkan pengguna sebagai informasi. Dibawah ini ialah proses terjadinya dalam sebuah KDD: 1. Data selection Data selection berasal dari sekumpulan data oprasional yang harus dilakukan sebelum tahap ekstraksi data pada KDD. Hasil Pemilihan data yang akan digunakan dalam tahap data mining ditampilkan dalam satu halaman berdasarkan data operasional. 2. Preprocessing Sebelum tahap selanjutnya dilakukan, Diperlukan metode pembersihan untuk data yang menjadi fokus KDD. Dalam tahapan ini mencangkup penghapusan duplikat data, pengecekan data yang tidak konsisten serta koreksi kesalahan dalam data seperti kekeliruan dalam cetak. 3. Transformation Coding merupakan transformasi data yang dipilih agar cocok untuk langkah penambangan data. Coding KDD adalah langkah kreatif yang sangat bergantung oleh jenis data atau model yang diambil dari setdata. 4. Data mining Dalam langkah ini ialah digunakannya teknik atau metode dengan tujuan menemukan pola atau hasil yang menarik. Teknik, metode, atau algoritma penambangan data yang bervariasi, menentukan metode atau algoritma yang tepat sangat bergantung dalam tujuan serta keseluruhan tahapan dalam KDD. 5. Interpretation/Evaluation Model data yang diperoleh dari langkah data mining dapat memberikan hasil yang dapat mudah dimengerti daripada yang berkepentingan. Tujuan dari bagian ini adalah untuk menentukan apakah kebijakan atau informasi yang diperoleh bertentangan dengan informasi.

#### 2.2. Algoritma K-Means

K-Means adalah salah satu algoritma untuk penambangan data dan digunakan untuk data cluster atau grup berdasarkan kesamaan properti. Algoritma ini berpusat

di sekitar banyak kelompok awal dan menghitung penghapusan setiap data di tengah cluster menggunakan metode seperti jarak Euclidean [20]. Data kemudian dirangkum ke dalam cluster berikutnya, dan pusat cluster berulang kali diperbarui sampai departemen grup tidak berubah secara signifikan. Keuntungan utama K-means adalah kemampuannya untuk menangani sejumlah besar data dengan waktu komputasi yang relatif cepat. Namun, kelemahannya terletak pada sensitivitasnya untuk memilih jumlah kelompok awal yang dapat memengaruhi hasil kelompok [21].

Saat menganalisis pengadaan produk dan layanan yang digunakan oleh K-means, ini dapat membantu Anda mengelompokkan data berdasarkan pola spesifik seperti: Pengelompokan ini memungkinkan wawasan lebih lanjut di depot sehubungan dengan proses perencanaan dan tren pengadaan yang mendorong alokasi anggaran yang lebih efisien [22]. Selain itu, hasil pengelompokan dapat digunakan sebagai dasar untuk menentukan strategi pengadaan yang lebih akurat.

# 2.3. Algoritma Naïve Bayes

Algoritma Naïve Bayes adalah salah satu metode klasifikasi yang didasarkan pada Teorema Bayes, yang menghitung probabilitas bersyarat dari suatu kategori berdasarkan fitur-fitur yang ada dalam dataset. Algoritma ini bekerja dengan mengasumsikan bahwa setiap fitur bersifat independen satu sama lain, meskipun dalam kenyataannya asumsi ini tidak selalu berlaku. Namun, dengan kesederhanaannya, Naïve Bayes tetap menjadi salah satu metode yang efisien dalam menangani data berukuran besar serta mampu memberikan hasil klasifikasi yang cukup akurat. Algoritma ini sering digunakan dalam berbagai bidang, seperti

analisis teks, deteksi spam, klasifikasi dokumen, hingga sistem rekomendasi karena kemampuannya dalam menangani data dengan berbagai macam fitur secara cepat.

Pemanfaatan Naïve Bayes dalam analisis pengadaan barang dan jasa dapat membantu dalam mengklasifikasikan jenis barang atau jasa yang lebih sering dibutuhkan serta mengidentifikasi pola pengadaan yang terjadi dalam kurun waktu tertentu. Dengan menghitung probabilitas berdasarkan data historis pengadaan, metode ini mampu memberikan prediksi mengenai kategori produk yang memiliki tingkat permintaan tinggi maupun rendah. Selain itu, algoritma ini juga dapat membantu dalam pemilihan penyedia barang atau jasa yang lebih sesuai dengan kebutuhan berdasarkan rekam jejak transaksi sebelumnya. Dengan demikian, penerapan Naïve Bayes dapat meningkatkan efektivitas pengambilan keputusan serta membantu optimalisasi dalam pengelolaan anggaran pengadaan barang dan jasa.

### 2.4. Pengadaan Barang dan Jasa

Pengadaan barang dan jasa merupakan salah satu aspek penting dalam administrasi pemerintahan yang bertujuan untuk memenuhi kebutuhan operasional serta mendukung pelaksanaan program kerja [3]. Proses ini mencakup berbagai tahapan, mulai dari perencanaan, pemilihan penyedia, hingga evaluasi hasil pengadaan. Efisiensi dan transparansi dalam pengadaan sangat krusial untuk memastikan penggunaan anggaran yang tepat dan menghindari potensi penyimpangan [4]. Namun, tantangan seperti pemilihan penyedia yang kurang optimal, keterlambatan dalam proses pengadaan, serta kurangnya analisis berbasis data sering kali menjadi hambatan dalam implementasi pengadaan yang efektif.

Oleh karena itu, diperlukan pendekatan yang lebih sistematis dan berbasis teknologi untuk meningkatkan akurasi dalam pengambilan keputusan terkait pengadaan barang dan jasa.

Dalam konteks penelitian ini, pengadaan barang dan jasa di Kantor Bupati Labuhanbatu dianalisis menggunakan metode data mining untuk mengidentifikasi pola dan tren dalam proses pengadaan [5]. Metode K-Means digunakan untuk mengelompokkan data berdasarkan karakteristik tertentu, seperti jenis barang/jasa, waktu pengadaan, dan penyedia yang sering digunakan. Sementara itu, metode Naïve Bayes diterapkan untuk melakukan klasifikasi terhadap data pengadaan guna memprediksi barang atau jasa yang memiliki potensi lebih diminati serta membantu dalam perencanaan pengadaan yang lebih efisien. Dengan menerapkan kedua metode ini, diharapkan dapat tercipta sistem pengadaan yang lebih akurat, efisien, serta mampu mendukung transparansi dalam pengelolaan anggaran di Kantor Bupati Labuhanbatu.

# 2.5. Knowledge Discovery in Database (KDD)

Knowledge Discovery in Database (KDD) adalah metode dan cara mendapatkan sebuah informasi melalui basis data yang telah tersedia [6]. Berikut tahap penyelesainnya, yaitu :

- Data Selection adalah pengambilan data dari sebuah dataset yang nantinya akan diolah.
- 2. Data Preprocessing adalah tahap untuk melakukan sebuah proses awal dalam pengolahan data. Pada tahap ini data yang akan diolah bertujuan

untuk menghindarkan dari data yang mengganggu (noise) atau data yang tidak konsisten.

- 3. Transformation adalah tahap membuat data yang akan diproses menjadi sesuai dengan model ataupun algoritma yang ingin digunakan dalam tahap pemrosesan data [7].
- 4. Data Mining Proses pencarian dan penggalian pengetahuan hingga diperoleh suatu model yang dapat digunakan untuk menjadi sebuah informasi penting dan berguna.
- 5. Evaluation adalah Merepresentasikan hasil model yang telah diperoleh serta menguji akurasi dan kesesuaian terhadap data-data yang bersinggungan.

# 2.6. Database dan Data Processing

Basis data adalah sistem terstruktur untuk mengelola, menyimpan dan mengatur data, memungkinkan data diakses dan diproses secara efisien [10]. Di dunia teknologi informasi, basis data memainkan peran penting dalam menyimpan berbagai data, dari data transaksional hingga informasi pelanggan dan data operasional organisasi. Proses manajemen basis data melibatkan memori, pemrosesan, dan manipulasi data menggunakan Sistem Manajemen Database (DBMS) seperti MySQL, PostgreSQL, dan MongoDB. Pemrosesan data adalah sejumlah langkah yang dilakukan untuk mengubah data mentah menjadi informasi yang berguna. Proses ini mencakup level seperti pengumpulan, pembersihan, konversi, dan analisis tampilan data dalam format yang mudah dipahami [11]. Pemrosesan data yang optimal memungkinkan organisasi untuk mengeksplorasi

pengetahuan yang lebih dalam daripada data untuk mendukung proses pengambilan keputusan yang lebih akurat.

Dalam konteks penelitian ini, database berfungsi sebagai tempat untuk menyimpan data historis untuk pengadaan barang dan jasa di kantor Bupati Labuhuhanbatu, termasuk informasi tentang jenis barang/jasa, harga, volume pengadaan, dan jenis penyedia jenis penyedia . . Data yang disimpan melalui proses data yang diproses untuk memungkinkan data yang digunakan dalam analisis menjadi bersih, terstruktur, dan diproses lebih lanjut. Dengan menggunakan metode penambangan data, terutama algoritma K-Means untuk pengelompokan dan bayes naif, hasil pemrosesan data digunakan untuk mengidentifikasi pola pengadaan dan memprediksi produk yang paling efisien dalam proses pengadaan. Oleh karena itu, pemrosesan basis data terstruktur dari kunci utama untuk generasi analisis yang akurat dan pemrosesan data yang efektif, dan batasan keputusan yang lebih efisien dalam sistem pengadaan barang dan jasa.

#### 2.7. Visualization

Visualization adalah teknik yang digunakan untuk menyajikan data dalam bentuk grafis atau visual agar lebih mudah dipahami dan dianalisis [12]. Dengan menggunakan grafik, diagram, atau peta interaktif, visualisasi dapat membantu mengidentifikasi pola, tren, dan hubungan dalam kumpulan data yang kompleks. Representasi visual ini memungkinkan pengambilan keputusan yang lebih cepat dan akurat karena informasi yang disajikan menjadi lebih intuitif dibandingkan dengan tabel atau angka mentah [13]. Dalam dunia analisis data, visualization

memainkan peran penting dalam menyampaikan hasil penelitian dengan cara yang lebih menarik dan mudah diinterpretasikan oleh berbagai pemangku kepentingan.

Penerapan visualisasi dalam analisis pengadaan barang dan jasa dapat membantu dalam memahami distribusi anggaran, pola transaksi, serta efisiensi pemilihan penyedia. Misalnya, dengan menggunakan diagram batang atau heatmap, dapat terlihat kapan periode pengadaan paling sibuk atau penyedia mana yang paling sering digunakan [14]. Grafik klasifikasi dari metode Naïve Bayes dapat menunjukkan probabilitas suatu barang atau jasa menjadi pilihan utama berdasarkan data historis. Sementara itu, hasil klasterisasi K-Means dapat divisualisasikan dalam bentuk scatter plot untuk menunjukkan kelompok-kelompok pola pengadaan yang serupa. Dengan pendekatan ini, pengelolaan data pengadaan menjadi lebih transparan, memudahkan pemangku kebijakan dalam mengoptimalkan strategi pengadaan yang lebih efektif dan efisien.

# 2.8. Statistik

Statistik merupakan cabang ilmu yang berfokus pada pengumpulan, analisis, interpretasi, dan penyajian data dalam bentuk yang lebih terstruktur dan bermakna [15]. Dengan menggunakan berbagai metode seperti deskriptif dan inferensial, statistik memungkinkan seseorang untuk memahami pola dalam data, mengidentifikasi tren, serta membuat prediksi berdasarkan informasi yang tersedia. Statistik deskriptif digunakan untuk menggambarkan data melalui ukuran seperti mean, median, dan standar deviasi, sedangkan statistik inferensial memungkinkan pengambilan kesimpulan yang lebih luas dengan mengandalkan teknik seperti uji hipotesis dan regresi [16]. Dalam berbagai bidang, statistik menjadi alat yang sangat

penting untuk mendukung pengambilan keputusan berbasis data yang lebih akurat dan sistematis.

Dalam analisis pengadaan barang dan jasa, statistik memainkan peran penting dalam mengidentifikasi pola pembelian, mengelompokkan data berdasarkan karakteristik tertentu, serta melakukan prediksi terhadap tren pengadaan di masa mendatang. Dengan menerapkan metode statistik, dapat diperoleh informasi mengenai kecenderungan belanja instansi, efisiensi anggaran, serta preferensi terhadap jenis barang atau jasa tertentu. Selain itu, statistik juga membantu dalam menilai performa penyedia barang dan jasa berdasarkan data historis, sehingga proses pengadaan dapat dilakukan dengan lebih optimal. Analisis statistik yang akurat memungkinkan perencanaan yang lebih baik, mengurangi risiko pemborosan anggaran, serta meningkatkan transparansi dalam setiap tahap pengadaan.

### 2.9. Pattern Recognition

Pattern recognition adalah cabang ilmu dalam kecerdasan buatan dan pembelajaran mesin yang berfokus pada identifikasi pola atau struktur dalam data [17]. Metode ini digunakan untuk mengklasifikasikan atau mengenali pola berdasarkan karakteristik yang ada, baik dalam bentuk teks, gambar, suara, maupun data numerik. Teknik pattern recognition bekerja dengan cara menganalisis sekumpulan data, mengidentifikasi hubungan antar elemen, serta mengelompokkan atau memprediksi berdasarkan pola yang ditemukan [18]. Beberapa metode yang sering digunakan dalam pattern recognition meliputi klasifikasi, klasterisasi, dan

deteksi anomali, yang diterapkan dalam berbagai bidang seperti pengolahan citra, analisis teks, hingga sistem rekomendasi.

Dalam sistem pengadaan barang dan jasa, pattern recognition dapat dimanfaatkan untuk mengidentifikasi tren pengadaan, menentukan kategori barang yang paling sering dibeli, serta menganalisis perilaku penyedia barang dan jasa. Teknik ini memungkinkan analisis berbasis data untuk membantu pengambilan keputusan yang lebih akurat, mengurangi pemborosan anggaran, serta meningkatkan efisiensi proses pengadaan [19]. Dengan mengombinasikan metode klasifikasi seperti Naïve Bayes dan klasterisasi seperti K-Means, pola dalam data pengadaan dapat dikenali dengan lebih jelas, sehingga mempermudah dalam menentukan strategi optimal dalam pemilihan penyedia dan perencanaan waktu pengadaan.

#### 2.10. Model Klaster

Model klaster merupakan salah satu pendekatan dalam data mining yang digunakan untuk mengelompokkan data berdasarkan karakteristik yang serupa [23]. Metode ini bekerja dengan cara membagi sekumpulan data ke dalam beberapa kelompok (klaster) tanpa adanya label awal, sehingga setiap klaster yang terbentuk memiliki pola yang berbeda dari klaster lainnya. Salah satu algoritma yang sering digunakan dalam model klaster adalah K-Means, yang mengelompokkan data dengan mencari titik pusat dari masing-masing klaster dan memperbaruinya secara iteratif hingga mencapai konvergensi [24]. Pendekatan ini sangat efektif dalam menemukan pola tersembunyi dalam data yang besar, membantu analisis tren, serta

meningkatkan pemahaman terhadap hubungan antar data yang mungkin sulit diidentifikasi secara manual.

Dalam proses pengadaan barang dan jasa, model klaster dapat digunakan untuk mengelompokkan berbagai jenis barang atau jasa berdasarkan karakteristik seperti harga, volume pembelian, serta frekuensi pengadaan. Pengelompokan ini dapat membantu dalam merancang strategi pengadaan yang lebih efisien, seperti menentukan prioritas dalam pemilihan penyedia atau mengidentifikasi kelompok barang yang sering dibutuhkan secara bersamaan. Dengan adanya analisis berbasis klaster, pengambil kebijakan dapat membuat keputusan yang lebih tepat, mengoptimalkan perencanaan anggaran, serta meningkatkan efisiensi dalam proses pengadaan tanpa harus bergantung sepenuhnya pada metode konvensional yang sering kali memakan waktu lebih lama.

#### 2.11. Model Klasifikasi

Model klasifikasi merupakan salah satu pendekatan dalam data mining yang bertujuan untuk mengelompokkan data ke dalam kategori tertentu berdasarkan pola yang telah dipelajari dari data historis. Salah satu model klasifikasi yang populer adalah Naïve Bayes, yang bekerja berdasarkan Teorema Bayes dengan asumsi bahwa setiap fitur dalam data bersifat independen satu sama lain. Meskipun asumsi ini jarang sepenuhnya terpenuhi dalam data nyata, algoritma Naïve Bayes tetap efektif dalam banyak kasus karena kesederhanaannya dan kemampuannya dalam menangani data berukuran besar. Model ini menghitung probabilitas suatu data masuk ke dalam kategori tertentu berdasarkan distribusi probabilitas fitur-fiturnya, sehingga dapat memberikan prediksi yang cepat dan cukup akurat dalam berbagai

aplikasi, termasuk dalam pengelolaan data bisnis, analisis teks, dan sistem rekomendasi.

Penerapan model klasifikasi Naïve Bayes dalam pengadaan barang dan jasa memungkinkan identifikasi pola pengadaan yang lebih sistematis dan berbasis data. Dengan menggunakan data historis, model ini dapat memprediksi kategori barang atau jasa yang memiliki tingkat permintaan tinggi, membantu dalam pemilihan penyedia yang lebih sesuai, serta memperkirakan waktu pengadaan yang optimal. Proses ini dapat meningkatkan efisiensi perencanaan anggaran dan mengurangi ketidaktepatan dalam pengambilan keputusan. Selain itu, karena Naïve Bayes dapat menangani berbagai variabel dengan cepat, sistem pengadaan dapat lebih adaptif terhadap perubahan tren dan kebutuhan yang dinamis dalam pemerintahan.

## 2.12. Alat Bantu Aplikasi Orange

Orange adalah salah satu perangkat lunak open-source yang digunakan dalam analisis data dan data mining. Aplikasi ini memiliki antarmuka berbasis visual yang memudahkan pengguna dalam mengolah data tanpa perlu menulis banyak kode pemrograman [25]. Dengan fitur drag-and-drop, pengguna dapat menyusun berbagai komponen analisis seperti preprocessing data, klasifikasi, klasterisasi, visualisasi, dan evaluasi model. Orange juga mendukung berbagai metode pembelajaran mesin, termasuk Naïve Bayes dan K-Means, sehingga memungkinkan analisis yang lebih mendalam terhadap pola dan tren dalam dataset. Selain itu, Orange dapat diintegrasikan dengan berbagai format data serta mendukung penggunaan bahasa pemrograman Python bagi pengguna yang ingin mengembangkan analisis lebih lanjut [26].

Penggunaan Orange dalam analisis pengadaan barang dan jasa dapat memberikan kemudahan dalam mengelompokkan data berdasarkan karakteristik tertentu serta mengklasifikasikan pola pengadaan yang terjadi. Dengan metode K-Means, data dapat diklasterkan berdasarkan kesamaan fitur seperti jenis barang, harga, atau waktu pengadaan, sehingga lebih mudah untuk melihat pola pembelian yang sering terjadi. Sementara itu, metode Naïve Bayes dapat digunakan untuk memprediksi kategori barang atau penyedia yang paling sesuai berdasarkan data historis. Kombinasi kedua metode ini dalam Orange memungkinkan pemrosesan data yang lebih cepat dan efisien, membantu pengambilan keputusan yang lebih tepat dalam proses pengadaan barang dan jasa.

# 2.13. Penelitian Terdahulu

Penelitian terdahulu merupakan Penelitian yang sudah dilakukan sebelumnya yang berhubungan dengan Penelitian ini, hal ini dilakukan guna untuk membuktikan bahwa metode K-Means bagus dalam melakukan klastering data.

Tabel 2. 1. Penelitian Terdahulu

Penelitian	1
Judul	Application of the K-Means Clustering Agorithm to Group Train Passengers in Labuhanbatu
Nama Penulis	Indri Cahaya Indah1)*, Mila Nirmala Sari2), Muhammad Halmi Dar3)
Tahun	2023
Hasil	Metode K-Means digunakan untuk melakukan clustering data penumpang kereta berdasarkan kesamaan data mereka. K-Means dipilih karena mampu mengelompokkan data secara efektif sesuai dengan pola yang diinginkan. Proses clustering dilakukan dengan memasukkan metode ini ke dalam pola widget di aplikasi Orange, yang kemudian mengolah data dan menghasilkan tiga kelompok atau cluster. Hasil clustering ini

	mengelompokkan penumpang ke dalam tiga kategori, termasuk kelas bisnis dan eksekutif, sehingga dapat memberikan gambaran yang lebih jelas mengenai karakteristik masingmasing kelompok [27].
Penelitian	2
Judul	Data Mining Sales of Skin Care Products Using the K-Means Method
Nama Penulis	Dasril Aldo <sup>1</sup>
Tahun	2023
Hasil	Metode K-Means dalam data mining digunakan untuk mengelompokkan data penjualan produk kecantikan dan perawatan kulit guna menyesuaikan stok dengan permintaan pasar. Dalam penelitian ini, produk dikelompokkan menjadi dua kategori, yaitu produk terlaris dan tidak terjual. Dengan menggunakan 30 sampel data, hasil analisis menunjukkan 18 produk tidak terjual dengan baik dan 12 produk masih kurang laku. Perhitungan manual dan perhitungan menggunakan aplikasi RapidMiner menghasilkan kesamaan 100%, membuktikan bahwa algoritma K-Means dapat menjadi solusi efektif dalam mengelola stok produk berdasarkan pola penjualan [28].
Penelitian	3
Judul	Clustering Analysis of Tweets About COVID19 Using the K-Means Algorithm
Nama Penulis	Andi1)*, Carles Juliandy2), David3)
Tahun	2023
Hasil	Penelitian ini menggunakan metode K-Means untuk mengelompokkan tweet tentang COVID-19 guna mempermudah pengguna dalam membaca dan memilah informasi. Data yang digunakan berasal dari dataset Omicron-Covid-19 Variant Tweets di Kaggle serta hasil scraping dengan Bright Data, dengan total 4.103 data. Penentuan jumlah klaster optimal menggunakan metode Elbow menunjukkan bahwa k = 5 adalah jumlah klaster terbaik. Hasil pengelompokan menunjukkan bahwa klaster terbesar adalah klaster 4 dengan 1.185 tweet, diikuti klaster 1 dengan 1.047 tweet, klaster 2 dengan 757 tweet, klaster 3 dengan 744 tweet, dan klaster terkecil adalah klaster 5 dengan 370 tweet [29].
Penelitian	4

Judul	Comparison of the K-Means Algorithm and C4.5 Against Sales Data
Nama Penulis	Abdi Dharma1)*, Eko Bambang Wijaya2), Daniel Heyneker3), Jeff Vanness4)
Tahun	2023
Hasil	Metode K-Means digunakan dalam penelitian ini sebagai teknik non-hierarkis untuk mengelompokkan data berdasarkan jarak minimum ke pusat klaster (centroid). Dengan menentukan jumlah klaster sendiri, peneliti dapat mengelompokkan data secara lebih fleksibel. Evaluasi kualitas klaster dilakukan menggunakan Silhouette Coefficient dan Elbow Method. Hasil penelitian menunjukkan perbedaan signifikan sebelum dan sesudah penerapan metode ini, di mana K-Means mampu mengidentifikasi pola data dengan lebih baik, khususnya dalam menentukan barang dengan tingkat penjualan tertinggi [30].