**NAMA : REFTA SEPDELA**

**NIM : 09031181621009**

**ANALISIS PENERAPAN KNOWLEDGE DISCOVERY IN DATABASES (KDD) TERHADAP PEMBELIAN SEPATU**

1. **PENDAHULUAN**

Knowledge Discovery in Databases (KDD) merupakan sekumpulan proses untuk menggali dan menganalisis sejumlah himpunan data yang besar dan mengekstrak data tersebut sehingga menemukan pengetahuan yang bermanfaat dari data tersebut. KDD berhubungan dengan teknik integrasi dan penemuan ilmiah, interprestasi dan visualisasi dari pola-pola kumpulan data.

Pada proses Knowledge Discovery Database (KDD) di atas terdapat beberapa tahapan yaitu sebagai berikut:

1. Pembersihan Data (*Data Cleaning*)

Digunakan untuk membuang data yang tidak konsisten dan noise.

1. Integrasi Data (*Data Integration*)

Data yang diperlukan untuk data mining tidak hanya berasal dari satu database tetapi juga berasal dari beberapa database atau file teks. Hasil integrasi data sering diwujudkan dalam sebuah data warehouse karena dengan data warehouse, data dikonsolidasikan dengan struktur khusus yang efisien.

1. Pemilihan Data (*Data Selection*)

Mengambil data yang relevan dengan tugas analisis dari database.

1. Transformasi Data (*Data Transformation*)

Pada fase ini yang dilakukan adalah mentransformasi bentuk data yang belum memiliki entitas yang jelas ke dalam bentuk data yang valid atau siap untuk dilakukan proses Data Mining.

1. Penambangan Data (*Data Mining*)

Pada fase ini yang dilakukan adalah menerapkan algoritma atau metode pencarian pengetahuan. Ini adalah langkah penting di mana teknik kecerdasan diterapkan untuk mengekstrak pola informasi yang berpotensi berguna dari data yang dipilih.

1. Evaluasi Pola (*Pattern Evaluation*)

Pada fase ini yang dilakukan adalah menerapkan algoritma atau metode pencarian pengetahuan. Ini adalah langkah penting di mana teknik kecerdasan diterapkan untuk mengekstrak pola informasi yang berpotensi berguna dari data yang dipilih.

1. Presentasi Pengetahuan (*Knowledge Presentation*)

Penggunaan teknik-teknik visualisasi dan representasi untuk menyajikan pengetahuan yang telah diperoleh kepada user.

Data Mining adalah kegiatan untuk menemukan informasi atau pengetahuan yang berguna secara otomatis dari data yang jumlahnya besar. Data Mining merupakan salah satu proses dari keseluruhan proses yang ada pada Knowledge Discovery in Databases (KDD).

Berdasarkan definisi ini terlihat bahwa data mining hanya merupakan salah satu proses dari keseluruhan proses yang ada pada KDD, tetapi merupakan proses yang sangat penting dalam menemukan pola-pola yang berguna dari sejumlah data yang besar (data tersebut bisa disimpan dalam basisdata, Data Warehouse atau media penyimpanan informasi.

1. **DATASET TOKO SEPATU DARI SITUS KDD**

Dataset adalah kumpulan data yang saling berkaitan satu sama lain yang memiliki kesatuan dan bersifat spesifik terhadap suatu kasus, missal dataset transaksi pembelian sepatu pada sebuah toko sepatu. Dataset dapat direprentasikan dalam berbagai bentuk seperti tabel dalam basis data, bentuk matriks, teks, *Comma Separated Value* (CSV), dan sebagainya. *Dataset* dapat dikumpulkan dan dibentuk oleh seseorang, sekelompok bahkan suatu organisasi dan dipublikasikan secara *online* seperti misalnya situs Kaggle (www.kaggle.com) yang berisi berbagai macam *dataset*.

*Dataset* dipergunakan sebagai referensi data yang valid untuk suatu penelitian selanjutnya, misalnya untuk referensi data dalam pembelajaran sistem cerdas seperti sistem pengenalan pola, *machine learning*, dan lain-lain. Selain itu dataset juga digunakan sebagai referensi data dalam pengujian sistem otomatis seperti misalnya pada sistem klasifikasi, klasterisasi, dan sentiment analisis. *Dataset* yang baik memiliki ciri memiliki data yang lengkap, selalu *up to date*, bersifat konsisten dalam representasi datanya, jumlah variabelnya jelas, tidak mengandung *noise*, menarik, dan mudah dipahami.

Pada tulisan kali ini, informasi dataset didapat dari situs Kaggle ([www.kaggle.com](http://www.kaggle.com)) tentang transaksi pada toko sepatu. Dimana dataset ini menjelaskan berbagai informasi dari transaksi yang telah dilakukan pada toko sepatu. Format dataset telah divirtualisasi ke dalam format excel. Dataset tersebut terdiri dari 10001 baris, dan 27 kolom. Kolomnya terdiri dari id, dateAdded, dateUpdated, asins (Amazon Standard Identification Numbers), brand, categories, primaryCategories, colors, dimension, ean (European Article Number), imageURLs, keys, manufacturer, manufacturerNumber, name, price.amountMax, price.amountMin, prices.avalaibility, prices.color, prices,condition, prices.currency, prices.dateAdded, prices.dateSeen, prices.isSale, prices.merchant, prices.offer, prices.returnPolicy, prices.shipping, prices.size, prices.sizes, sourceURLs, upc (Universal Product Code), dan weight.

Di balik transaksi pembelian produk sepatu terdapat informasi berharga yang dapat berguna bagi perusahaan untuk meningkatkan penjualan dengan menyusun strategi penjualan. Untuk itu diperlukan Knowledge Discovery in Databases (KDD) untuk menentukan produk sepatu merk apa yang paling diminati konsumen sehingga dapat menghasilkan informasi tersembunyi yang dapat berguna di masa mendatang bagi perusahaan.

1. **JURNAL PENDUKUNG INFORMASI**
2. **Jurnal 1 “Implementasi Data Mining dengan Algoritma Apriori untuk Menentukan Merek Sepatu yang Diminati pada Mahasiswa Pascasarjana Kelas 14.1A.01 STMIK Nusa Mandiri Jakarta”**.

Maka dapat diperoleh informasi berupa aspek teknologi informasi dalam sistem penjualan produk alas kaki. Sepatu merupakan pakaian yang dibutuhkan untuk aktivitas sehari-hari. Jumlah transaksi penjualan yang terjadi terus menerus menunjukkan bahwa jumlah minat konsumen terhadap sepatu. Hal ini membuat produsen menciptakan berbagai inovasi, seperti merk sepatu. Merk sangat mempengaruhi orang untuk membeli suatu produk. Ada berbagai cara untuk mengantisipasi agar produk selalu menjadi pilihan konsumen. Salah satu caranya yaitu dengan menggunakan teknik penambangan data dengan menggunakan apriori algoritma (asosiasi data mining). Dalam penelitian ini teknik penambangan data apriori algoritma dapat diimplementasikan pada penelusuran pada merk sepatu yang paling diminati. Dalam menganalisa kebutuhan dengan menggunakan algoritma apriori, dapat diketahui pola frekuensi tinggi tentang jenis merek sepatu yang paling banyak digunakan atau paling diminati dengan melakukan kombinasi tiap item dan menggunakan syarat minimum yang telah ditentukan.

Dari penelitian yang telah dilakukan maka didapat pola pembelian merk sepatu yaitu jika membeli Nike maka akan membeli Adidas, jika membeli Converse akan membeli Adidas. Hasil yang diperoleh dari penelitian ini yaitu Nike-Adidas dengan support 25% dan confidence 80% dan Converse-Adidas dengan support 25% dan confidence 66,67%(Rezkiani, 2016).

1. **Jurnal 2 “Penerapan Data Mining Penjualan Sepatu Menggunakan Metode Algoritma Apriori”.**

Maka diperoleh informasi mengenai aktivitas penjualan sepatu setiap hari dimana data penjualan akan meningkat. Data itu tidak hanya berfungsi sebagai arsip bagi perusahaan, data dapat dimanfaatkan dan diolah menjadi informasi yang berguna untuk meningkatkan penjualan produk sepatu. Ketersediaan data penjualan tidak digunakan secara optimal, karena tidak ada sistem pendukung keputusan dan metode yang dapat digunakan untuk merancang strategi bisnis untuk meningkatkan penjualan. Teknik pengolahan data yang digunakan dalam penelitian ini adalah algoritma apriori dan untuk mendapatkan hasil yang lebih baik digunakan alat tanagra versi 1.4. Salah satu tahapan analisis yang menarik perhatian banyak peneliti untuk menghasilkan algoritma apriori yang efisien adalah pola analisis frekuensi suatu asosiasi yang diidentifikasi dengan dua patokan yaitu dukungan dan keyakinan.

Berdasarkan dari calon aturan asosiasi, maka yang memenuhi minimal support 30% dan Minimal Confidence 60% maka dapat dilihat pola pembelian sepatu yaitu Jika membeli sepatu New Balance, maka akan membeli Adidas. Jika membeli sepatu Adidas, maka akan membeli New Balance. Jika membeli sepatu Nike, Maka akan membeli New Balance. Jika membeli sepatu Geox, maka akan membeli New Balance. Jika membeli sepatu Palladium, maka akan membeli New Balance. Jika membeli sepatu Nike, maka akan membeli Adidas. Maka setelah dilakukan penelitian didapat informasi merk sepatu yang paling banyak diminati oleh konsumen adalah New Balance (91,67 %), Adidas (75 %) , Geox (50 %), Nike (41.67 %) dan Palladium (41.67 %)(Sikumbang, 2018).

1. **Jurnal 3 “Sistem Rekomendasi Produk Sepatu dengan Menggunakan Metode Collaborative Filtering”**.

Maka didapatkan informasi mengenai prediksi pola pembelian sepatu dengan merk sepatu yang paling direkomendasi. Penelitian ini dilakukan pada Platinum Shoes. Platinum Shoes merupakan sebuah toko yang bergerak dibidang penjualan sepatu. Banyaknya jumlah produk yang ditawarkan, membuat beberapa pelanggan kesulitan dalam menentukan pilihan mengenai produk apa yang pelanggan pilih dan cocok dengan selera pelanggan. Hal ini yang menjadikan perlunya sebuah sistem rekomendasi yang mampu memberikan rekomendasi produk yang sesuai selera pelanggan untuk memudahkan dalam memilih produk yang akan dibeli.

Penelitian ini menggunakan metode Item-Based Collaborative Filtering, dimana sistem akan mencari kesamaan model pembelian (similarity item) dengan yang lainnya. Selanjutnya, sistem akan mencari rating antara item-item berdasarkan tingkat kemiripan yang ada. Setelah rating antar item didapat, maka rating ini akan digunakan dihitung nilai kemiripan antara item dengan menggunakan pendekatan Adjusted Cosine Similarity. Hasil dari dari perhitungan kemiripan antar item akan digunakan untuk tahap selanjutnya. Tahapan ini memprediksi nilai rating yang belum pernah dilakukan oleh pelanggan terhadap item tertentu. Pendekatan ini menggunakan rumus Weigted Sum yang nilai prediksinya akan dijadikan rekomendasi kepada pelanggan. Dengan menggunakan persamaan item-based filtering produk yang direkomendasikan adalah yang belum pernah di rating oleh pelanggan sebelumnya. Produk yang direkomendasikan adalah yang mempunyai nilai prediksi tertinggi contoh id\_pelanggan 2 mendapatkan rekomendasi produk sepatu vans dan sepatu new balance dengan nilai prediksi 4 dan 3(Kurniawan, 2016).

1. **Jurnal 4 “Analisis Pengaruh Harga, Kualitas Produk, Kualitas Pelayanan, Dan Citra Merek Terhadap Keputusan Pembelian Sepatu Bata”**.

Penelitian ini bertujuan untuk mengetahui pengaruh harga, kualitas produk, kualitas layanan dan citra merek terhadap keputusan pembelian Sepatu Bata di Semarang. Populasi dalam penelitian ini adalah seluruh masyarakat yang membeli Sepatu Bata di Semarang, menggunakan 100 responden. Regresi linear berganda digunakan untuk menganalisis data. Dari hasil yang telah diteliti pengaruh bersama – sama antara variabel harga, kualitas produk, kualitas pelayanan, dan citra merek terhadap keputusan pembelian adalah 77,416 dengan sign. 0,000 < α = 0,05. Nilai sign yang lebih kecil α = 0,05 menyatakan secara bersama - sama variabel harga, kualitas produk, kualitas pelayanan, dan citra merek berpengaruh positif dan sifnifikan terhadap variabel keputusan pembelian Sepatu Bata.

Dari penelitian ini maka didapat pola pembelian sepatu bata dilihat dari beberapa aspek yaitu semakin baik variabel harga ditingkatkan maka semakin naik tingkat keputusan pembelian, semakin tinggi kualitas produk maka keputusan pembelian tentu akan meningkat, semakin tinggi kualitas pelayanan maka semakin tinggi pula keputusan pembelian konsumen, semakin tinggi citra merek maka semakin tinggi pula keputusan pembelian(CIKO, 2016).

1. **Jurnal 5 “Pengaruh Gaya Hidup, Citra Merek, Dan Kualitas Produk Terhadap Keputusan Pembelian Sepatu Sneakers Merek Converse (Studi Pada Mahasiswa Universitas Diponegoro Semarang)”**.

Penelitian ini diperoleh informasi berupa perilaku konsumen yang bervariasi dan susah untuk diprediksi menyebabkan dunia bisnis mengalami perubahan yang begitu cepat. Konsumen menuntut suatu produk yang sesuai dengan selera, gaya hidup dan kebutuhan mereka. Hal ini menyebabkan perusahaan-perusahaan mengalami persaingan yang begitu ketat karena tuntutan keinginan konsumen yang semakin kompleks. Salah satu produk yang selalu mengikuti selera konsumen adalah sepatu. Sepatu tak lagi hanya berfungsi sebagai alas kaki untuk melindungi kaki dari benda-benda tajam, namun sepatu juga sebagai produk fashion sehingga desain, bahan dan coraknya pun menjadi sangat variatif. Pada kurun tahun 2009-2013 penjualan Sepatu sneakers merek Converse mengalami fluktuasi, penjualannya dan selalu berada di bawah Adidas, Nike, dan Bata. Tujuan penelitian ini adalah untuk mengetahui pengaruh gaya hidup, citra merek dan kualitas produk terhadap keputusan pembelian sepatu sneakers merek Converse.

Berdasarkan analisis hasil penelitian membuktikan bahwa gaya hidup memiliki pengaruh terhadap keputusan pembelian sepatu sneakers merek Converse sebesar 35,4% dan pengaruhnya tergolong sedang. Namun terdapat beberapa indikator yang masih dianggap kurang seperti, sepatu sneakers merek Converse kurang menunjang aktivitas sehari-hari karena desainnya yang monoton kurang menarik untuk beraktivitas. Citra merek sneakers Converse dikatakan baik karena merek Converse adalah merek yang pertama kali disebutkan responden untuk sepatu jenis sneakers. Hasil analisis penelitian membuktikan bahwa citra merek memiliki pengaruh terhadap keputusan pembelian sepatu sneakers merek Converse sebesar 42% dan pengaruhnya tergolong kuat. Kualitas produk dikatakan baik karena sneakers Converse yang awet dan nyaman dipakai. Hasil analisis penelitian menyatakan bahwa kualitas produk memiliki pengaruh terhadap keputusan pembelian sepatu sneakers merek Converse sebesar 51,3% dan pengaruhnya tergolong kuat(Marchiani, Hidayat, & Dewi, 2015).

1. **KESIMPULAN**

Dari beberapa contoh kasus jurnal diatas maka dapat disimpulkan ada beberapa pola yang ditemukan pada pembelian sepatu dilihat dari berbagai aspek yaitu apabila membeli sepatu Nike maka akan membeli Adidas. Jika membeli Converse akan membeli Adidas. Jika membeli sepatu New Balance, maka akan membeli Adidas. Jika membeli sepatu Adidas, maka akan membeli New Balance. Jika membeli sepatu Nike, Maka akan membeli New Balance. Jika membeli sepatu Geox, maka akan membeli New Balance. Jika membeli sepatu Palladium, maka akan membeli New Balance. Selain itu juga didapat pola seperti kualitas sepatu yang bagus, harga yang sesuai, merk terkenal, design yang bagus, dan nyaman dipakai maka konsumen akan mencari produk sesuai dengan pola yang inginkan. Dari pola-pola yang telah didapat maka perusahaan dapat mencari strategi penjualan untuk mencari merk sepatu yang paling diminati oleh konsumen sehingga perusahaan dapat tau prediksi penjualan untuk kedepannya yang dapat meningkatkan penjualan terhadap permintaan sepatu yang paling diminati.

**DAFTAR PUSTAKA**

CIKO, A. (2016). Analisis Pengaruh Harga, Kualitas Produk, Kualitas Pelayanan dan Citra merek Terhadap Keputusan Pembelian Sepatu Bata (Studi Pada Konsumen Sepatu Bata Semarang). *Skripsi, Fakultas Ekonomi & Bisnis*.

Kurniawan, A. (2016). Sistem Rekomendasi Produk Sepatu Dengan Menggunakan Metode Collaborative Filtering. *Universitas Jenderal Ahmad Yani, Cimahi*.

Marchiani, N. D., Hidayat, W., & Dewi, R. S. (2015). Pengaruh Gaya Hidup, Citra Merek, Dan Kualitas Produk Terhadap Keputusan Pembelian Sepatu Sneakers Merek Converse (Studi Pada Mahasiswa Universitas Diponegoro Semarang). *Jurnal Ilmu Administrasi Bisnis*, *4*(3), 324–332.

Rezkiani, R. (2016). IMPLEMENTASI DATA MINING DENGAN ALGORITMA APRIORI UNTUK MENENTUKAN MEREK SEPATU YANG DIMINATI PADA MAHASISWA PASCASARJANA KELAS 14.1 A. 01 STMIK NUSA MANDIRIJAKARTA. *Konferensi Nasional Ilmu Pengetahuan Dan Teknologi*, *2*(1), 49–INF.

Sikumbang, E. D. (2018). Penerapan Data Mining Penjualan Sepatu Menggunakan Metode Algoritma Apriori. *Jurnal Teknik Komputer*, *4*(1), 156–161.