

BAB I PENDAHULUAN

1.1 Latar Belakang

Kebutuhan masyarakat yang semakin meningkat dewasa ini menyebabkan meningkatnya pula usaha ritel. Hal ini terlihat dari semakin bertambahnya jumlah minimarket yang dapat dilihat di kota – kota besar maupun kecil. Berdasarkan data Aprindo diperkirakan jumlah minimarket di Indonesia mencapai 13.000 toko dan masih ditargetkan untuk meningkat setiap tahunnya [SIL-12].

Setiap minimarket ini menangani transaksi setiap harinya, sehingga akan tersimpan banyak data transaksi. Data transaksi ini dapat dimanfaatkan untuk mencari pola belanja konsumen. Pola belanja ini dapat berbentuk item – item yang cenderung muncul bersama dalam sebuah transaksi. Selanjutnya pola ini dapat dimanfaatkan untuk strategi pemasaran, contohnya menempatkan item – item yang sering dibeli bersamaan di area yang berdekatan, merancang tampilan item di katalog, atau memberi promo diskon untuk item tertentu yang dibeli oleh pelanggan [HAN-06]. Pencarian pola tersebut merupakan salah satu metode dalam data mining.

Salah satu metode dalam *data mining* yang dapat digunakan untuk mencari pola pembelian konsumen adalah *Sequential Pattern Mining*. *Sequential pattern* sendiri adalah pola yang menggambarkan urutan waktu terjadinya peristiwa. Pola tersebut dapat ditemukan apabila data yang disimpan relatif besar dan peristiwa yang berurutan terjadi beberapa kali [HAN-06]. Contohnya seorang pelanggan yang tercatat sebagai member suatu minimarket, melakukan transaksi belanja berulang kali pada minimarket tersebut, sehingga dapat ditemukan pola belanja dari pelanggan ini.

Karena *Sequential Pattern Mining* ini berkaitan dengan urutan atau waktu terjadinya peristiwa, sehingga pola hasil dari proses ini juga dapat digunakan untuk promosi yang berkaitan dengan urutan pembelian atau waktu. Contohnya apabila seorang pelanggan membeli kamera maka pelanggan itu akan mendapatkan diskon apabila membeli lensa kamera dalam jangka waktu 1 bulan

setelah pembelian kamera tersebut. Hal ini diharapkan dapat menarik pelanggan untuk membeli barang yang dipromosikan.

Sebelum dikembangkan algoritma untuk mencari *sequential pattern*, untuk mencari pola belanja konsumen digunakan algoritma untuk mencari *frequent pattern*. Berbeda dengan *sequential pattern*, *frequent pattern* hanya mencari hubungan antar item yang sering muncul bersamaan tanpa mempertimbangkan urutan atau waktu pembelian item. Diharapkan juga dengan diperhitungkannya urutan atau waktu pembelian item dalam *sequential pattern mining*, pola yang dihasilkan akan lebih bervariasi untuk mendapatkan hasil yang lebih maksimal. Dalam *Sequential Pattern Mining* untuk menemukan pola agar mendapat informasi yang berguna harus dicari *frequent sequences* atau urutan peristiwa tertentu yang sering muncul. Algoritma yang digunakan sebelumnya untuk mencari *frequent sequence* ini seperti *Generalized Sequential Pattern* (GSP) masih membutuhkan pencarian *database* lengkap untuk mencari *frequent sequence* yang semakin panjang, sehingga membutuhkan waktu dan proses yang lama [ZAK-01]

Algoritma *Sequential Pattern Discovery using Equivalent classes* (SPADE) adalah algoritma salah satu algoritma dalam *Sequential Pattern Mining* yang dapat digunakan untuk mengatasi kekurangan algoritma sebelumnya dimana pencarian *database* secara lengkap harus dilakukan berkali – kali. SPADE menggunakan *id-list* vertikal untuk memudahkan pencarian dalam *database*. SPADE dapat mencari *frequent sequences* dengan beberapa kali pencarian *database* saja. Berdasarkan hasil penelitian Kumar [KUM-12], SPADE menghasilkan performa yang jauh lebih baik dalam hal waktu komputasi dibandingkan algoritma – algoritma pencari *frequent sequences* sebelumnya seperti AprioriAll dan GSP. Selain itu, berdasarkan penelitian Wiwin [WIW-11] yang menerapkan algoritma SPADE untuk simulasi prediksi banjir, algoritma ini menghasilkan akurasi hingga 85% dalam hasil uji menggunakan data histori banjir dengan data klimatologi. Sehingga diharapkan algoritma SPADE ini akan cocok diterapkan pada data transaksi penjualan, karena jumlah data transaksi penjualan yang besar. Pada penelitian ini akan mencoba menggunakan algoritma SPADE untuk mencari *sequential pattern* pada data transaksi penjualan. Berdasarkan latar

belakang tersebut maka skripsi ini berjudul “**Sequential Pattern Mining Pada Data Transaksi Penjualan Menggunakan Algoritma Sequential Pattern Discovery using Equivalent classes (SPADE)**”.

1.2 Rumusan Masalah

Rumusan masalah dalam skripsi ini adalah :

1. Bagaimana merancang dan mengimplementasikan SPADE pada data transaksi penjualan?
2. Bagaimana keakuratan rule yang didapatkan dari *sequential pattern* dari data transaksi yang dihasilkan SPADE?

1.3 Batasan Masalah

Batasan masalah yang diangkat dalam skripsi ini adalah :

1. Dataset yang digunakan adalah data transaksi penjualan dari suatu toko elektronik yang diambil dari www.customers-dna.com.
2. Data yang digunakan untuk mencari *sequential pattern* dari dataset hanya id konsumen, timestamp dan item yang dibeli.

1.4 Tujuan

Tujuan yang ingin dicapai dalam skripsi ini adalah :

1. Menerapkan SPADE pada dataset transaksi penjualan untuk mencari *sequential pattern*.
2. Menguji keakuratan rule yang didapatkan dari *sequential pattern* yang dihasilkan dari dataset transaksi penjualan.

1.5 Manfaat

Manfaat hasil skripsi ini diharapkan akan menghasilkan perangkat lunak yang dapat mencari *sequential pattern* dari suatu data transaksi sehingga ke depannya dapat dimanfaatkan untuk membantu pemasaran, contohnya untuk diskon suatu item apabila membeli item tertentu dalam jangka waktu yang ditentukan.

1.6 Sistematika Penulisan

BAB I : PENDAHULUAN

Bab ini membahas mengenai latar belakang, rumusan masalah, batasan masalah, tujuan, manfaat dan sistematika penulisan skripsi.

BAB II : DASAR TEORI

Bab ini menjelaskan uraian dasar teori terkait yang menjadi dasar rujukan pada penelitian ini.

BAB III : METODOLOGI DAN PERANCANGAN

Bab ini berisi penjelasan metode yang digunakan dalam penelitian dan bagaimana kerangka dasar solusi yang diusulkan.

BAB IV : HASIL DAN PEMBAHASAN

Bab ini berisi uraian mengenai implementasi dari rancangan solusi yang diajukan dan pembahasan mengenai hasil pengujian.

BAB V : PENUTUP

Bab ini berisi kesimpulan dari penelitian yang telah dilakukan, dan saran untuk pengembangan penelitian selanjutnya.

