

Nama : Muhammad nur sopin

Nim : G.231.22.0087

Matkul : Data Mining

Kelas : B3

Berikut adalah penjelasan tentang langkah-langkah melakukan Market Basket Analysis (Analisis Keranjang Pasar) menggunakan Python berdasarkan artikel di pbpython.com:

Persiapan Data

1. Memuat Data: Dataset yang berisi catatan transaksi dimuat ke dalam DataFrame Pandas.
2. Membersihkan Data: Transaksi yang diidentifikasi sebagai pengembalian (faktur yang mengandung 'C') dihapus untuk memastikan hanya transaksi pembelian yang dianalisis.

Transformasi Data

3. Memfilter Berdasarkan Negara: Analisis dibatasi hanya pada transaksi dari Inggris.
4. Mengelompokkan Transaksi: Data dikelompokkan berdasarkan 'InvoiceNo' dan 'Description', dan kuantitas setiap item per transaksi dijumlahkan.
5. Hot Encoding: Data yang dikelompokkan diubah sehingga setiap baris mewakili satu transaksi dengan kolom untuk setiap item, diisi dengan kuantitas yang sesuai. Nilai yang hilang diisi dengan nol.

Encoding Kuantitas

6. Binary Encoding: Kuantitas dikonversi ke nilai biner (1 jika item dibeli, 0 jika tidak).

Penerapan Algoritma Apriori

7. Menghasilkan Frequent Itemsets: Menggunakan fungsi apriori dari library mlxtend, itemset yang sering muncul dihasilkan dengan ambang batas minimum support yang ditentukan.
8. Menghasilkan Aturan Asosiasi: Fungsi association_rules dari mlxtend digunakan untuk menghasilkan aturan dari itemset yang sering muncul. Parameter seperti metric (misalnya, 'lift') dan min_threshold (tingkat kepercayaan minimum) ditentukan untuk menyaring aturan.

Pendekatan Manual

9. Metode Alternatif: Artikel ini juga membahas pendekatan manual untuk Market Basket Analysis, yang melibatkan menghasilkan pasangan item dan menghitung

support serta confidence tanpa bergantung pada mlxtend. Metode ini berguna saat bekerja dengan dataset besar di mana operasi pivot/unstack mungkin tidak efisien.

Analisis dan Interpretasi

10. Menafsirkan Hasil: Aturan yang dihasilkan dianalisis untuk memahami hubungan antara item. Insight ini dapat digunakan untuk cross-selling, sistem rekomendasi, atau manajemen inventaris.

Proses ini membantu mengidentifikasi produk mana yang sering dibeli bersama, memungkinkan bisnis membuat keputusan berbasis data untuk meningkatkan strategi penjualan.