

Penerapan Top-K High-Utility Itemset Mining Untuk Menentukan Kombinasi Produk Dengan Nilai Penjualan Tertinggi Sebagai Rekomendasi Bundling

Novi Wulandari^{*1}, Amat Damuri², Abdul Rohim³, Iwan Mulyana⁴

^{*1,2}Manajemen Informatika STMIK Al Muslim, Bekasi

^{3,4}Ilmu Komputer STMIK Al Muslim, Bekasi

e-mail: ^{*1}wulandari.novi1@gmail.com, ²amatdamuri@gmail.com, ³abdul.rohim.pranata18@gmail.com,

⁴iwan.mulyana@almuslim.ac.id

Abstrak

Strategi bundling memerlukan informasi mengenai kombinasi produk yang tidak hanya sering dibeli bersama, tetapi juga menghasilkan nilai penjualan tinggi. Penelitian ini bertujuan menerapkan Top-K High-Utility Itemset Mining untuk menentukan kombinasi produk bernilai penjualan tertinggi sebagai rekomendasi bundling. Data yang digunakan adalah Online Retail II dari UCI Machine Learning Repository. Tahapan penelitian mencakup pembersihan data, penghapusan transaksi pembatalan dan nilai nonpositif, agregasi produk pada setiap faktur, perhitungan utility berdasarkan kuantitas dan harga satuan, pembentukan basis data utility, serta penambangan pola menggunakan utility-list. Proses penambangan menggunakan nilai K sebesar 10 dengan panjang kombinasi dua sampai empat produk. Setelah preprocessing, diperoleh 36.302 transaksi valid untuk dianalisis. Hasil penelitian menunjukkan bahwa sembilan pola teratas merupakan kombinasi dua produk dan satu pola merupakan kombinasi tiga produk. Jumbo Bag Strawberry dan Jumbo Bag Red White Spotty menghasilkan utility tertinggi sebesar £115.525,29 dengan support 3,2753%. Sementara itu, Jumbo Bag Pink with White Spots dan Jumbo Bag Red White Spotty memiliki support tertinggi sebesar 3,9337% dengan utility £109.344,48. Temuan ini menunjukkan bahwa kombinasi paling sering dibeli tidak selalu menghasilkan nilai penjualan tertinggi. Metode tersebut dapat menghasilkan rekomendasi bundling dengan mempertimbangkan frekuensi, kuantitas, dan nilai penjualan produk secara bersamaan.

Abstract

Product bundling strategies require information about product combinations that are not only frequently purchased together but also generate high sales value. This study aims to apply Top-K High-Utility Itemset Mining to identify product combinations with the highest sales value as bundling recommendations. The Online Retail II dataset from the UCI Machine Learning Repository was used. The research stages included data cleaning, removal of cancelled transactions and non-positive values, aggregation of products within each invoice, utility calculation based on quantity and unit price, utility database construction, and pattern mining using utility lists. The mining process used a K value of 10, with combination lengths ranging from two to four products. After preprocessing, 36,302 valid transactions were obtained for analysis. The results showed that nine of the top patterns consisted of two products, while one pattern consisted of three products. The combination of Jumbo Bag Strawberry and Jumbo Bag Red White Spotty generated the highest utility of £115,525.29, with a support of 3.2753%. Meanwhile, Jumbo Bag Pink with White Spots and Jumbo Bag Red White Spotty achieved the highest support of 3.9337%, with a utility of £109,344.48. These findings indicate that the most frequently purchased combination does not necessarily generate the highest sales value. This method can produce bundling recommendations by simultaneously considering product frequency, quantity, and sales value.

Article Info

Kata Kunci:

bundling produk;
high-utility itemset mining;
nilai penjualan;
Top-K

Keywords:

product bundling;
high-utility itemset mining;
sales value;
Top-K.

I. PENDAHULUAN

Perkembangan perdagangan elektronik menyebabkan perusahaan ritel menghasilkan data transaksi dalam jumlah besar. Data tersebut tidak hanya berfungsi sebagai catatan penjualan, tetapi juga dapat diolah menjadi pengetahuan mengenai kecenderungan pembelian konsumen. Pengetahuan mengenai produk yang dibeli secara bersamaan dapat dimanfaatkan untuk menyusun promosi, mengatur penempatan produk, merencanakan persediaan, dan membentuk paket bundling yang lebih relevan.

Market Basket Analysis merupakan salah satu pendekatan data mining yang digunakan untuk mengidentifikasi hubungan antarproduk dalam transaksi. Sejumlah penelitian di Indonesia telah menggunakan algoritma Apriori dan FP-Growth untuk menemukan frequent itemset dan aturan asosiasi. Artsitella et al. (2021) menunjukkan bahwa pola asosiasi dapat mendukung penentuan tata letak, katalog, dan voucher belanja, sedangkan Ridho et al. (2024) menggunakan FP-Growth untuk memperoleh kombinasi produk yang sering dibeli bersama. Pendekatan tersebut bermanfaat untuk memahami frekuensi pembelian, tetapi tidak secara langsung mempertimbangkan perbedaan kuantitas dan nilai ekonomi setiap produk.

Frequent itemset mining menilai pola terutama berdasarkan support. Akibatnya, produk berharga rendah yang sering muncul dapat dianggap lebih penting daripada kombinasi yang sedikit lebih jarang tetapi menghasilkan nilai penjualan jauh lebih tinggi. Dalam konteks bundling, keterbatasan ini dapat menyebabkan rekomendasi hanya berfokus pada popularitas dan mengabaikan kontribusi ekonomi. *High-Utility Itemset Mining* (HUIM) dikembangkan untuk mengatasi permasalahan tersebut dengan memperhitungkan utility internal, seperti kuantitas pembelian, dan utility eksternal, seperti harga atau keuntungan produk.

HUIM konvensional membutuhkan minimum utility threshold yang harus ditentukan pengguna. Penetapan nilai yang terlalu rendah dapat menghasilkan sangat banyak pola dan meningkatkan biaya komputasi, sedangkan nilai yang terlalu tinggi berisiko menghilangkan pola yang relevan. *Top-K High-Utility Itemset Mining* mengubah kebutuhan

tersebut menjadi parameter K, yaitu jumlah pola berutility tertinggi yang ingin diperoleh. Tseng et al. (2016) memperkenalkan kerangka Top-K HUIM serta algoritma TKO yang menambang pola dalam satu fase menggunakan utility-list dan peningkatan ambang utility secara dinamis.

Penelitian ini menggunakan dataset Online Retail II yang berisi transaksi nyata perusahaan ritel daring berbasis di Inggris selama dua tahun. Dataset memiliki 1.067.371 baris, mencakup informasi faktur, kode dan deskripsi produk, kuantitas, tanggal, harga satuan, pelanggan, dan negara (Chen, 2012). Karakteristik tersebut memungkinkan utility didefinisikan sebagai nilai penjualan kotor, yaitu kuantitas dikalikan harga satuan.

Berdasarkan permasalahan tersebut, penelitian menerapkan *Top-K High-Utility Itemset Mining* untuk menentukan kombinasi produk dengan nilai penjualan tertinggi. Pola yang ditemukan kemudian dilengkapi dengan support dan rata-rata utility agar rekomendasi bundling tidak hanya bernilai tinggi secara agregat, tetapi juga dapat dinilai dari frekuensi kemunculannya.

II. METODE PENELITIAN

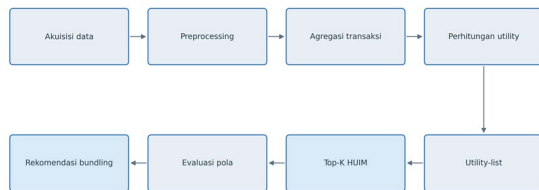
Desain dan Sumber Data

Penelitian ini merupakan penelitian kuantitatif terapan dengan pendekatan eksperimen komputasional. Tujuannya adalah menerapkan *Top-K High-Utility Itemset Mining* untuk menemukan kombinasi produk bernilai penjualan tinggi sebagai kandidat bundling. Analisis menggunakan data sekunder dan tidak melibatkan intervensi langsung terhadap pelanggan.

Data berasal dari Online Retail II pada *UCI Machine Learning Repository*. Dataset memuat 1.067.371 baris transaksi sebuah perusahaan ritel daring berbasis di Inggris selama 1 Desember 2009 sampai 9 Desember 2011 (Chen, 2012). Atribut yang digunakan adalah *InvoiceNo*, *StockCode*, *Description*, *Quantity*, *InvoiceDate*, *UnitPrice*, *CustomerID*, dan *Country*. Seluruh lembar kerja pada berkas *online_retail_II.xlsx* dibaca dan digabungkan. Unit analisis penelitian adalah *invoice*, sedangkan

item yang dianalisis adalah *StockCode* di dalam setiap invoice.

Tahapan penelitian meliputi akuisisi data, preprocessing, transformasi menjadi basis transaksi utility, penambangan *Top-K*, evaluasi pola, dan penyusunan rekomendasi bundling. Alur penelitian ditunjukkan pada Gambar 3.1.



Gambar 1. Alur penelitian Top-K High-Utility

Preprocessing Data

Preprocessing dilakukan agar setiap transaksi memenuhi persyaratan *utility mining* dan hanya memuat produk yang dapat dibundel. Nama kolom dari seluruh lembar terlebih dahulu diseragamkan karena dataset menggunakan variasi Invoice atau *InvoiceNo* serta *Price* atau *UnitPrice*. Baris dengan nilai kosong pada *InvoiceNo*, *StockCode*, *Description*, *Quantity*, atau *UnitPrice* dihapus, kemudian *Quantity* dan *UnitPrice* dikonversi menjadi numerik dan *InvoiceDate* menjadi format tanggal.

Invoice yang diawali huruf C dihapus karena menunjukkan transaksi pembatalan. Baris dengan $Quantity \leq 0$ atau $UnitPrice \leq 0$ juga dihapus. Kode dan deskripsi operasional yang tidak merepresentasikan produk, seperti POST, DOT, BANK CHARGES, *postage*, *carriage*, *adjustment*, dan *discount*, dikeluarkan agar pola akhir hanya berisi barang yang dapat ditawarkan sebagai bundling.

Produk yang sama dalam invoice yang sama diagregasi berdasarkan pasangan *InvoiceNo*–*StockCode*. *Utility* baris dihitung sebelum agregasi untuk mengakomodasi perubahan harga. Transaksi yang memiliki kurang dari dua produk unik dihapus karena tidak dapat membentuk bundling. Jumlah transaksi setelah *preprocessing* dilaporkan pada bagian hasil agar pemisahan antara prosedur dan temuan tetap jelas.

Perhitungan Utility dan Pembentukan Transaksi

Utility pada penelitian ini didefinisikan sebagai nilai penjualan kotor. *Utility internal* adalah *Quantity*, sedangkan *utility eksternal* adalah *UnitPrice*. Dataset tidak menyediakan harga pokok, sehingga *utility tidak* ditafsirkan sebagai laba atau keuntungan.

$$u(i, T) = q(i, T)xp(i, T) \quad (1)$$

Pada Persamaan (1), $u(i, T)$ adalah *utility* item i pada transaksi T , $q(i, T)$ adalah kuantitas, dan $p(i, T)$ adalah harga satuan. Untuk menghindari permasalahan presisi bilangan desimal, nilai penjualan dikalikan 100 dan disimpan sebagai bilangan bulat *pence*. *Utility* suatu itemset dihitung dengan menjumlahkan *utility* seluruh item anggotanya pada semua transaksi yang memuat itemset tersebut.

$$u(X) = \sum (T \supseteq X) \sum (i \in X) u(i, T) \quad (2)$$

Setiap *StockCode* dipetakan menjadi *ItemID*. Satu *utility transaction* terdiri atas daftar *ItemID*, *transaction utility*, dan *utility* masing-masing item. *Transaction utility* merupakan jumlah *utility* seluruh item pada suatu invoice. Struktur ini digunakan untuk membentuk *utility-list* dan menjalankan pencarian *Top-K*.

Penerapan Top-K High-Utility Itemset Mining

Mining menggunakan pendekatan satu fase berbasis *utility-list* yang mengikuti prinsip TKO (Tseng et al., 2016). Tahap awal menghitung *Transaction-Weighted Utility* (TWU) setiap item dan membentuk urutan global. *Utility-list* menyimpan *transaction identifier*, itemset *utility* (*iutil*), dan *remaining utility* (*rutil*). Jumlah *iutil* merupakan *utility* aktual pola, sedangkan jumlah *iutil* dan *rutil* digunakan sebagai batas atas perluasan pola.

$$UB(X) = \sum iutil(X) + \sum rutil(X) \quad (3)$$

Ambang *utility internal* atau border dinaikkan secara dinamis ketika pola ber*utility* lebih tinggi ditemukan. Cabang pencarian dipangkas apabila batas atasnya lebih kecil daripada border. *Utility-list* kandidat dibentuk melalui *intersection* berdasarkan *transaction identifier*. Lower bound awal diperoleh secara aman dari *utility* pasangan produk pada maksimum 2.000 transaksi bernilai tinggi. Strategi

tersebut mempercepat kenaikan border tanpa melebihi-lebihkan utility pada keseluruhan basis data.

Nilai K ditetapkan sebesar 10 agar hasil cukup ringkas untuk diinterpretasikan dan diprioritaskan sebagai kandidat bundling. Panjang minimum ditetapkan dua item karena bundling harus memuat lebih dari satu produk, sedangkan panjang maksimum empat item digunakan untuk membatasi ruang pencarian serta menjaga keterterapan paket. Apabila beberapa pola mempunyai utility yang sama tepat pada ambang peringkat ke-K, seluruh pola tersebut dipertahankan sehingga jumlah keluaran dapat melebihi K.

Tabel 1. Konfigurasi penambangan Top-K HUIM

Parameter	Nilai	Keterangan
K	10	Jumlah pola utama yang dicari
Panjang minimum	2	Bundling minimal dua produk
Panjang maksimum	4	Membatasi ruang pencarian dan kompleksitas paket
Skala mata uang	100	Konversi pound sterling menjadi pence
Seed transactions	2	Pembentukan lower bound awal
Penanganan tie	Dipertahankan	Seluruh pola pada ambang ke-K disertakan

Evaluasi dan Validasi Hasil

Pola diurutkan menurun berdasarkan utility. Setiap pola kemudian dievaluasi menggunakan support count, support relatif, rata-rata utility per transaksi, dan kontribusi terhadap keseluruhan nilai penjualan. Support count menunjukkan jumlah invoice yang memuat seluruh produk dalam pola, sedangkan support relatif dihitung menggunakan Persamaan (4).

$$support(X) = \frac{|\{T \in D : X \subseteq T\}|}{|D|} \quad (4)$$

$$average\ utility(X) = \frac{u(X)}{support\ count(X)} \quad (5)$$

Utility digunakan sebagai dasar pemeringkatan, sedangkan support digunakan untuk menilai seberapa luas pola terjadi. Rata-rata utility membantu membedakan pola berfrekuensi tinggi dari pola yang lebih jarang tetapi bernilai besar per transaksi. Penjumlahan utility beberapa itemset tidak

ditafsirkan sebagai pendapatan unik karena satu transaksi dapat berkontribusi pada lebih dari satu pola.

Kebenaran implementasi divalidasi dengan membandingkan keluaran miner terhadap enumerasi brute-force pada basis transaksi berukuran kecil. Pengujian memeriksa kesamaan keanggotaan pola, nilai *utility*, hasil tanpa *lower bound* awal, dan pola yang tie pada ambang K. Pengujian tambahan dilakukan pada basis transaksi acak untuk memastikan bahwa strategi *pruning* tidak menghilangkan pola Top-K yang seharusnya ditemukan.

Implementasi dijalankan menggunakan Python 3 pada Google Colab. Pustaka utama yang digunakan adalah pandas dan NumPy untuk pengolahan data, openpyxl untuk membaca workbook, serta Matplotlib dan Seaborn untuk visualisasi. Informasi lingkungan komputasi dan parameter eksekusi dicatat untuk mendukung reproduktibilitas.

III. HASIL DAN PEMBAHASAN

Hasil Pengolahan Data

Setelah preprocessing, basis analisis terdiri atas 36.302 transaksi. Jumlah tersebut diperoleh secara konsisten dari perbandingan SupportCount dan Support pada seluruh hasil Top-10. Utility dihitung sebagai nilai penjualan kotor; karena data tidak memiliki harga pokok, hasil tidak ditafsirkan sebagai laba.

Hasil Top-K High-Utility Itemset Mining

Proses penambangan dengan $K = 10$ menghasilkan sembilan kombinasi dua produk dan satu kombinasi tiga produk. Tidak munculnya kombinasi empat produk pada Top-10 tidak berarti pola tersebut tidak ada, melainkan utility-nya belum termasuk sepuluh tertinggi.

Tabel 2. Hasil Top-10 kombinasi produk berdasarkan utility

Rank	Kombinasi produk	Len	Support	Support (%)	Utility (£)
1	Jumbo Bag Strawberry Jumbo Bag Red White Spotty	2	1.189	32753%	115.525,29
2	Jumbo Bag Pink with White Spots Jumbo Bag Red White Spotty	2	1.428	39337%	109.344,48
3	Paper Chain Kit Vintage Christmas Paper Chain Kit 50's Christmas	2	839	23112%	101.376,91
4	Red Hanging Heart T-Light Holder White Hanging Heart T-Light Holder	2	1.242	34213%	96.009,91
5	Party Bunting Regency Cakestand 3 Tier	2	631	17382%	93.868,74
6	Regency Cakestand 3 Tier White Hanging Heart T-Light Holder	2	766	21101%	93.762,65
7	Roses Regency Teacup and Saucer Regency Cakestand 3 Tier	2	744	20495%	90.257,31
8	Pack of 72 Retro Spot Cake Cases Regency Cakestand 3 Tier	2	689	18980%	89.035,94
9	Jumbo Bag Baroque Black White Jumbo Bag Red White Spotty	2	1.113	30659%	87.509,45
10	Green Regency Teacup and Saucer Roses Regency Teacup and Saucer Regency Cakestand 3 Tier	3	578	15922%	86.505,72

Kombinasi *Jumbo Bag Strawberry* dan *Jumbo Bag Red White Spotty* menempati peringkat pertama dengan utility £115.525,29 dan muncul pada 1.189 transaksi. Peringkat kedua memiliki support tertinggi, yaitu 1.428 transaksi, tetapi utility lebih rendah sebesar £109.344,48. Perbedaan ini membuktikan bahwa pola paling sering tidak selalu menghasilkan nilai penjualan terbesar.

Kombinasi *Paper Chain Kit Vintage Christmas* dan *Paper Chain Kit 50's Christmas* berada pada peringkat ketiga dengan utility £101.376,91. Pola tersebut memiliki karakter musiman, sehingga penerapan bundling lebih relevan menjelang periode Natal daripada sebagai promosi sepanjang tahun.

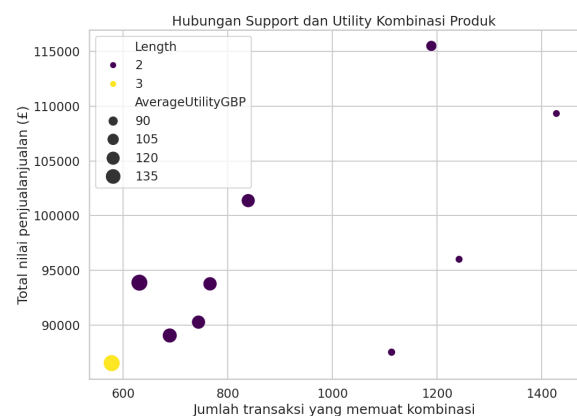
Analisis Kombinasi Berdasarkan Utility



Gambar 2. Top-K kombinasi produk berdasarkan utility Itemset Mining

Nilai utility Top-10 berada pada rentang £86.505,72 sampai £115.525,29. Produk *Jumbo Bag Red White Spotty* muncul dalam tiga kombinasi Top-10 dan dapat diposisikan sebagai *anchor product* untuk kategori tas. *Regency Cakestand 3 Tier* muncul pada lima kombinasi, menunjukkan potensinya sebagai pusat bundling untuk kategori dekorasi, pesta, dan penyajian.

Hubungan Support dan Utility



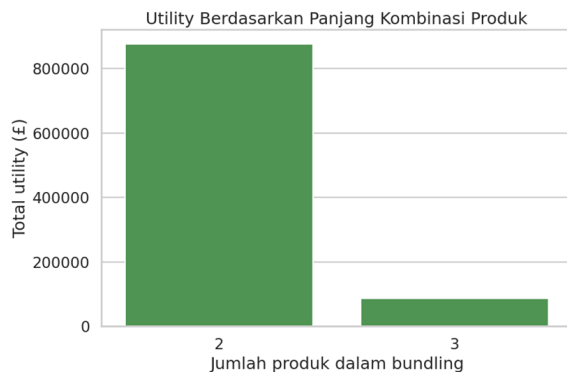
Gambar 3. Hubungan support dan utility kombinasi produk

Scatter plot menunjukkan kecenderungan hubungan positif antara support dan utility, tetapi tidak linier sempurna. Korelasi Pearson deskriptif pada sepuluh pola adalah sekitar 0,653. Kombinasi

Party Bunting dan Regency Cakestand 3 Tier hanya muncul pada 631 transaksi, tetapi rata-rata utility-nya mencapai £148,76. Sebaliknya, pola dengan support tertinggi memiliki rata-rata utility £76,57.

Kombinasi tiga produk pada peringkat kesepuluh mempunyai support terendah, yaitu 578 transaksi, tetapi rata-rata utility tertinggi sebesar £149,66. Temuan ini menunjukkan pentingnya membaca support dan utility secara bersamaan. Pola dengan frekuensi lebih rendah dapat tetap menarik sebagai paket premium.

Utility Berdasarkan Panjang Kombinasi



Gambar 4. Utility berdasarkan panjang kombinasi produk

Sembilan kombinasi dua produk menghasilkan penjumlahan utility sebesar £876.690,68, sedangkan satu kombinasi tiga produk menghasilkan £86.505,72. Berdasarkan ringkasan pola Top-10, kombinasi dua produk menyumbang 91,02% dan kombinasi tiga produk 8,98%. Dominasi tersebut mendukung penerapan paket sederhana yang mudah dipahami konsumen dan lebih mudah dikelola dari sisi stok.

Penjumlahan utility pada grafik tidak boleh ditafsirkan sebagai pendapatan unik karena satu transaksi dapat memuat lebih dari satu itemset. Grafik hanya menggambarkan komposisi utility pola Top-10 menurut panjangnya.

Rekomendasi Bundling

Seluruh pola berstatus Recommended karena support count jauh melampaui batas minimum dua transaksi. Akan tetapi, ambang tersebut terlalu rendah untuk membedakan prioritas bisnis. Implementasi nyata perlu menambahkan aturan mengenai margin, stok, diskon, musim, dan kesesuaian fungsi produk.

Implikasi dan Keterbatasan

Hasil membuktikan bahwa Top-K HUIM dapat memberikan rekomendasi yang lebih berorientasi nilai daripada frequent itemset mining. Parameter K memudahkan pengguna menentukan jumlah pola yang dibutuhkan tanpa mencoba berbagai minimum utility. Namun, utility penelitian hanya mewakili nilai penjualan kotor. Kombinasi berutility tinggi belum tentu berlaba tinggi apabila harga pokok, biaya promosi, dan biaya persediaan diperhitungkan.

Pola yang ditemukan merupakan asosiasi ko-transaksi dan tidak membuktikan bahwa promosi bundling akan menyebabkan peningkatan penjualan. Rekomendasi sebaiknya diuji melalui A/B testing dengan membandingkan konversi, average order value, dan margin kelompok promosi terhadap kelompok kontrol. Analisis juga menggabungkan dua periode sehingga perubahan perilaku dari waktu ke waktu belum dinilai

Tabel 3. Prioritas rekomendasi bundling

Prioritas	Kombinasi	Dasar	Strategi
1	Jumbo Bag Strawberry + Jumbo Bag Red White Spotty	Utility tertinggi	Paket dua motif/diskon produk kedua
2	Jumbo Bag Pink with White Spots + Jumbo Bag Red White Spotty	Support tertinggi	Bundling varian warna
3	Dua Paper Chain Kit Christmas	Utility tinggi dan tema seragam	Promosi musiman Natal
4	Red + White Hanging Heart T-Light Holder	Support tinggi dan komplementer	Paket dekorasi dua warna
5	Party Bunting + Regency Cakestand 3 Tier	Rata-rata utility tinggi	Paket pesta bernilai tinggi
6	Dua Regency Teacup + Regency Cakestand	Rata-rata utility tertinggi	Paket premium penyajian teh

IV. KESIMPULAN

Penerapan Top-K High-Utility Itemset Mining pada 36.302 transaksi hasil preprocessing dataset Online Retail II berhasil menghasilkan sepuluh kombinasi produk dengan nilai penjualan tertinggi. Sembilan pola terdiri atas dua produk dan satu pola terdiri atas tiga produk. Kombinasi Jumbo Bag Strawberry dan Jumbo Bag Red White Spotty menghasilkan utility tertinggi sebesar £115.525,29 dengan support 3,2753%. Sementara itu, kombinasi Jumbo Bag Pink with White Spots dan Jumbo Bag Red White Spotty memiliki support tertinggi sebesar 3,9337%, tetapi utility-nya lebih rendah, yaitu £109.344,48.

Hasil tersebut menunjukkan bahwa kombinasi yang paling sering dibeli tidak selalu memberikan nilai penjualan tertinggi. Top-K High-Utility Itemset Mining dapat menghasilkan rekomendasi bundling dengan mempertimbangkan frekuensi, kuantitas, dan harga produk secara bersamaan. Produk Jumbo Bag Red White Spotty dan Regency Cakestand 3 Tier muncul pada beberapa pola Top-10 sehingga berpotensi digunakan sebagai produk utama dalam penyusunan paket bundling.

Utility dalam penelitian ini merepresentasikan nilai penjualan kotor karena dataset tidak menyediakan harga pokok dan margin keuntungan. Penelitian selanjutnya dapat memasukkan informasi biaya, ketersediaan stok, dan karakteristik pelanggan, serta menguji efektivitas rekomendasi melalui A/B testing.

V. REFERENSI

- Agrawal, R., & Srikant, R. (1994). Fast algorithms for mining association rules in large databases. In *Proceedings of the 20th International Conference on Very Large Data Bases* (pp. 487–499).
- Artsitella, C. R., Apriliani, A. R., & Ashari, S. (2021). Penerapan association rules–market basket analysis untuk mencari frequent itemset dengan algoritma FP-Growth. *Jurnal Al-Azhar Indonesia Seri Sains dan Teknologi*, 6(2). <https://doi.org/10.36722/sst.v6i2.661>
- Chen, D. (2012). Online Retail II [Data set]. *UCI Machine Learning Repository*. <https://doi.org/10.24432/C5CG6D>
- Fournier-Viger, P., Wu, C.-W., Zida, S., & Tseng, V. S. (2014). FHM: Faster high-utility itemset mining using estimated utility co-occurrence pruning. In *Foundations of Intelligent Systems* (pp. 83–92). Springer. https://doi.org/10.1007/978-3-319-08326-1_9
- Krishnamoorthy, S. (2019). Mining top-k high utility itemsets with effective threshold raising strategies. *Expert Systems with Applications*, 117, 148–165. <https://doi.org/10.1016/j.eswa.2018.09.051>
- Kumar, R., & Singh, K. (2024). Top-k high utility itemset mining: Current status and future

directions. *The Knowledge Engineering Review*, 39, e5.
<https://doi.org/10.1017/S0269888924000055>

Liu, M., & Qu, J. (2012). Mining high utility itemsets without candidate generation. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management* (pp. 55–64).
<https://doi.org/10.1145/2396761.2396773>

Ridho, A. N., Masa, A. P. A., & Widagdo, P. P. (2024). Implementasi market basket analysis pada data penjualan CV. XYZ menggunakan algoritma FP-Growth. *Jurnal Informatika dan Teknik Elektro Terapan*, 12(3).
<https://doi.org/10.23960/jitet.v12i3.4541>

Saputra, J. P. B., Rahayu, S. A., & Hariguna, T. (2023). Market basket analysis using FP-Growth algorithm to design marketing strategy by determining consumer purchasing patterns. *Journal of Applied Data Sciences*, 4(1).
<https://doi.org/10.47738/jads.v4i1.83>

Tseng, V. S., Wu, C.-W., Fournier-Viger, P., & Yu, P. S. (2016). Efficient algorithms for mining top-k high utility itemsets. *IEEE Transactions on Knowledge and Data Engineering*, 28(1), 54–67.
<https://doi.org/10.1109/TKDE.2015.2458860>

Wulandari, U. M., Suseno, A. T., & Kholilurrahman, M. (2025). Market basket analysis using FP-Growth and Apriori on distro store sales transaction. *MATICS: Jurnal Ilmu Komputer dan Teknologi Informasi*, 17(1).
<https://doi.org/10.18860/mat.v17i1.28820>