



PENEMUAN POLA KETERKAITAN PRODUK DALAM DATA PENJUALAN RETAIL ONLINE MENGUNAKAN ALGORITMA APRIORI

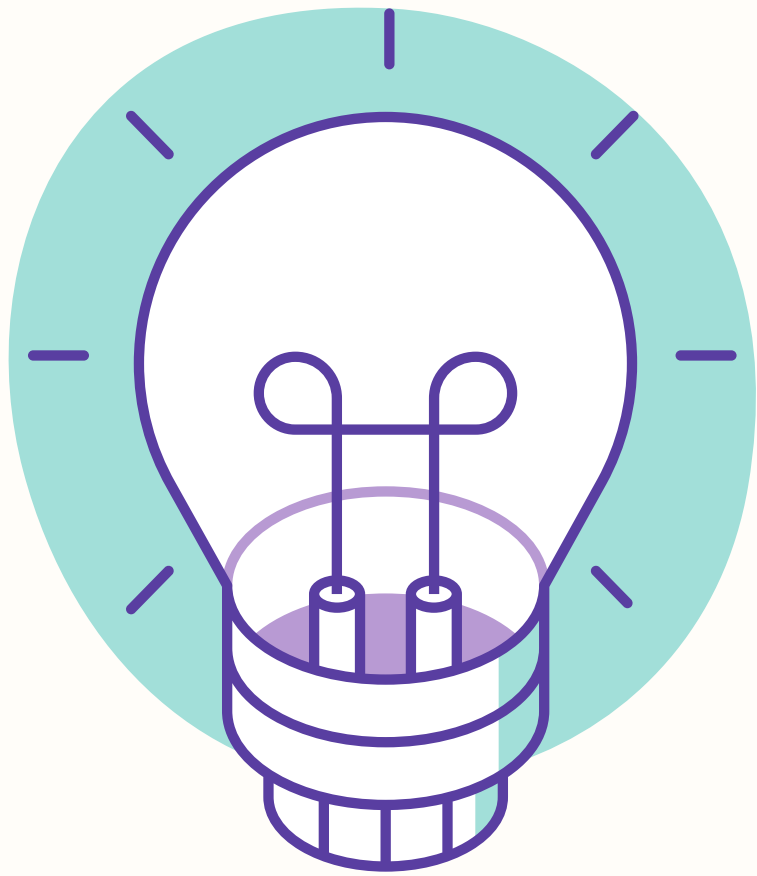


Bagas Maulana Akbar (21081010307)



Latar Belakang

Perkembangan e-commerce yang pesat telah menghasilkan data transaksi dalam jumlah besar. Data ini menyimpan informasi penting tentang perilaku konsumen yang dapat dimanfaatkan untuk strategi pemasaran. Namun, metode manual tidak efektif untuk analisis data yang begitu besar. Oleh karena itu, algoritma Apriori digunakan untuk mengidentifikasi pola keterkaitan produk guna mendukung keputusan pemasaran berbasis data.



Research Gap

Penelitian sebelumnya menunjukkan bahwa algoritma Apriori banyak digunakan dalam retail offline. Namun, belum banyak yang menerapkannya secara mendalam pada data retail online. Penelitian ini mengisi kekosongan tersebut dengan memanfaatkan dataset retail online untuk menemukan pola keterkaitan produk dengan pendekatan sistematis.



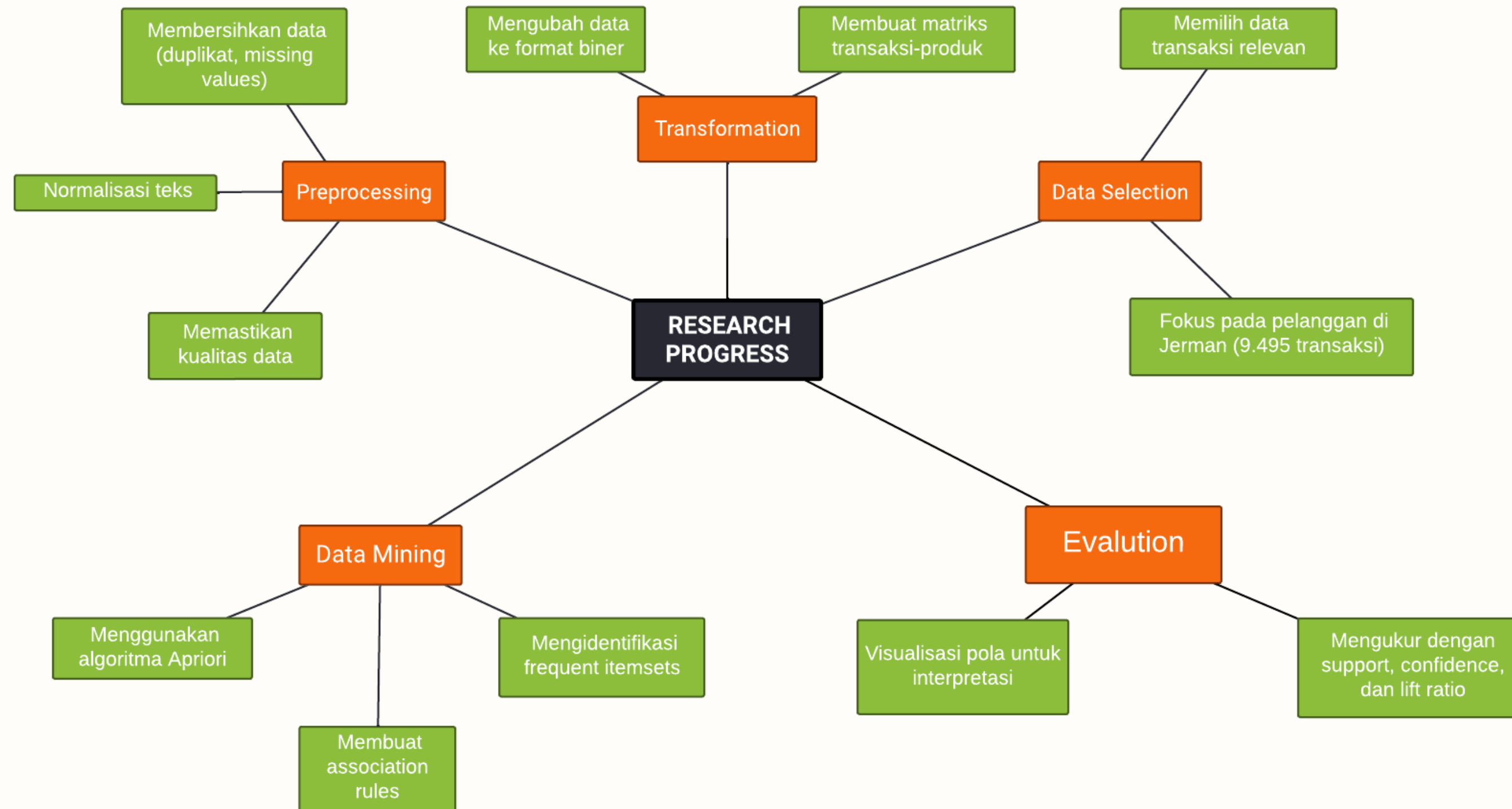


Tujuan Penelitian

Penelitian ini bertujuan untuk mengaplikasikan algoritma Apriori dalam menganalisis data retail online, menemukan pola produk yang sering dibeli bersama, dan memberikan rekomendasi berbasis data untuk meningkatkan efektivitas pemasaran.



Mind Mapping





Metodologi

- Dataset: Data transaksi retail online dari UCI Machine Learning Repository.

- Parameter Evaluasi:

Support: $\geq 5\%$



Confidence: $\geq 50\%$

Lift Ratio: > 1

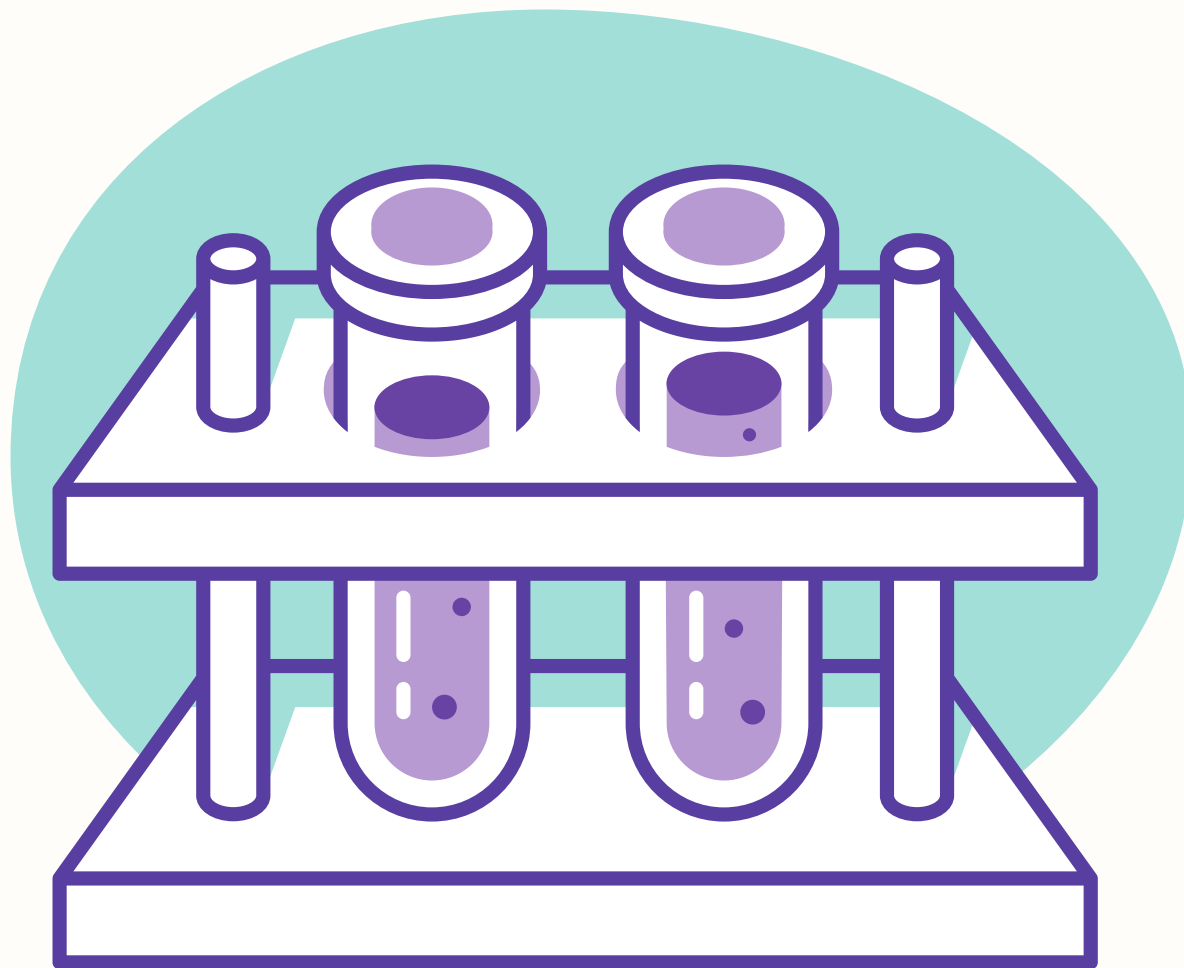
- Tools:

Python dengan pustaka seperti mlxtend, pandas, numpy, dan matplotlib.

Penelitian ini menggunakan dataset dari UCI Machine Learning Repository yang berisi 541.910 transaksi. Analisis dilakukan dengan algoritma Apriori menggunakan parameter evaluasi seperti support, confidence, dan lift ratio. Proses penelitian menggunakan Python dan pustaka pendukung untuk analisis data dan visualisasi.



Research Progress




- **Seleksi Data:** Fokus pada transaksi pelanggan di Jerman (9.495 transaksi).
- **Preprocessing:** Normalisasi teks, penghapusan data duplikat, dan missing values.
- **Transformasi:** Representasi data ke dalam format biner.
- **Data Mining:** Implementasi algoritma Apriori untuk menemukan frequent itemsets.





Testing (Evaluation)

Metrics

- Support: Frekuensi kemunculan kombinasi produk.
 - Confidence: Probabilitas produk A muncul bersamaan dengan produk B.
 - Lift Ratio: Mengukur kekuatan hubungan antarproduk dibandingkan kemunculan secara independen.
- 


Hasil Analisis

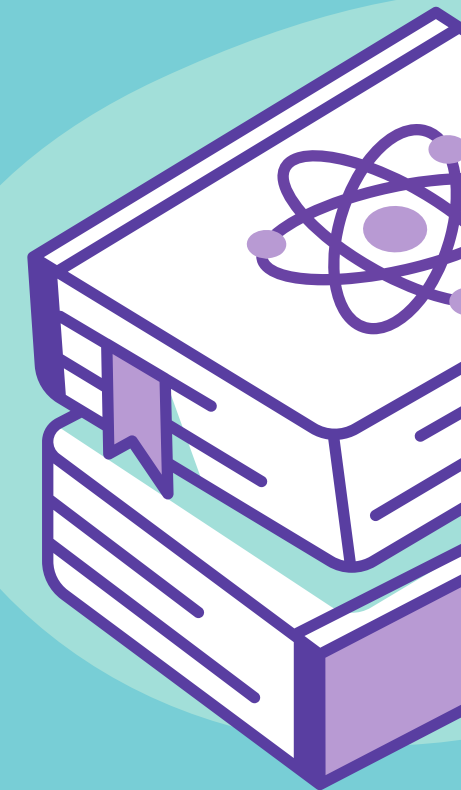
Nama Itemset	Apriori		
	Support	Confidence	Lift
round snack boxes set of 4	0.103806	0.833333	4.300595
fruits => round snack boxes set of 4 woodland			
spaceboy lunch box => round snack boxes set of 4 woodland	0.055363	0.680851	3.513678
plasters in tin circus parade => plasters in tin woodland animals	0.053633	0.584906	5.366277
round snack boxes set of 4 woodland => plasters in tin woodland animals	0.058824	0.539683	2.785147
round snack boxes set of 4 woodland => round snack boxes set of 4 fruits	0.103806	0.535714	4.300595
round snack boxes set of 4 woodland => woodland charlotte bag	0.079585	0.5	0.050173



Kesimpulan

Penelitian ini membuktikan bahwa algoritma Apriori adalah metode yang efektif untuk menemukan pola keterkaitan produk dalam data retail online. Rekomendasi yang diberikan meliputi implementasi promosi silang, bundling produk, dan sistem rekomendasi berbasis data untuk meningkatkan strategi pemasaran.





THANK YOU

