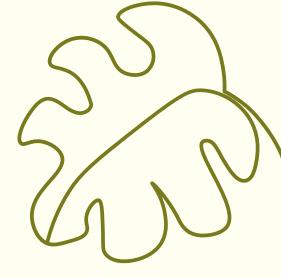
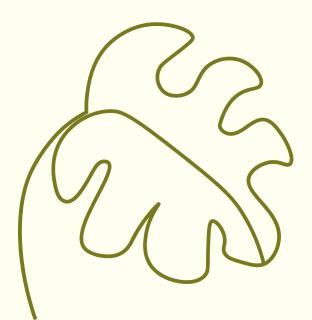


PENDAHULUAN





Sistem rekomendasi adalah sistem yang memprediksi preferensi pengguna terhadap suatu item berdasarkan profil pengguna. Sistem ini dapat meningkatkan penjualan, kepuasan, dan pengambilan keputusan. Sistem rekomendasi dapat dibedakan menjadi berbasis kolaboratif, berbasis konten, dan hibrida. Metode K-Means Clustering adalah algoritma yang mengelompokkan data berdasarkan kesamaan karakteristik. Metode ini dapat digunakan untuk memberikan rekomendasi produk berdasarkan kategorisasi pelanggan. Metode ini mudah dan cepat, tetapi memerlukan pemilihan jumlah kluster secara manual.









Dataset Memiliki 541909 baris dengan 8 kolom

	InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice	Customerl	Country	
	536365	85123A	WHITE HANGING HEART T-LIGHT HOLDER	6	12/1/2010 8:26	2.55	17850	United Kingo	dom
	536365	71053	WHITE METAL LANTERN	6	12/1/2010 8:26	3.39	17850	United Kingo	dom
	536365	84406B	CREAM CUPID HEARTS COAT HANGER	8	12/1/2010 8:26	2.75	17850	United Kingo	dom
	536365	84029G	KNITTED UNION FLAG HOT WATER BOTTLE	6	12/1/2010 8:26	3.39	17850	United Kingo	dom
	536365	84029E	RED WOOLLY HOTTIE WHITE HEART.	6	12/1/2010 8:26	3.39	17850	United Kingo	dom
	536365	22752	SET 7 BABUSHKA NESTING BOXES	2	12/1/2010 8:26	7.65	17850	United Kingo	dom
	536365	21730	GLASS STAR FROSTED T-LIGHT HOLDER	6	12/1/2010 8:26	4.25	17850	United Kingo	dom
A	536366	22633	HAND WARMER UNION JACK	6	12/1/2010 8:28	1.85	17850	United Kingo	dom
Ш	536366	22632	HAND WARMER RED POLKA DOT	6	12/1/2010 8:28	1.85	17850	United Kingo	dom
HL			,				,	-	-



METODE

• Data Cleansing: isnull() - dropna():





- R (Recency): Number of days since last purchase/ interaksi pelanggan terakhir. dengan mengextract Invoice Date untuk mengambil max days
- F (Frequency): Number of tracsactions / jumlah transaksi per pelanggan dengan mengelompokkan Invoice No berdasarkan Customer ID
- M (Monetary): Total amount of transactions (revenue contributed) / total transaksi dengan mengggunakan kolom Quantity dikali Unit Price.
- 2.Rescaling/ normalisasi pada data RFM



K-Means:

- Menentukan Nilai optimal K menggunakan Elbow Method
- Mengkluster dengan k = 3 pada data RFM
- Metode evaluasi kualitas klaster untuk ukuran sejauh mana suatu objek dalam satu klaster mirip dengan objek dalam klaster lain. menggunakan Silhouette value. Hasilnya 0.5084896296141937 (Medium Structure)



Langkah-langkah:

- Mencari ID produk target berdasarkan nama produk.
- Mengambil kluster dari produk target.
- Memfilter produk dalam kluster yang sama.
- Menghapus produk target dari rekomendasi (jika ada).
- Menghitung skor kesamaan kosinus antara produk target dan produk dalam kluster.
- Mensortir produk berdasarkan skor kesamaan dan mengambil jumlah rekomendasi yang diinginkan.





THANK YOU





