

Product Recommendation: Anticipating Sales Loss

"Leveraging Hybrid Product Recommendation to Anticipate Sales Loss in E-Commerce"

Team Thalassa



MENTOR
Muhammad
Anwar Sanusi



FASILITATOR
Jasmine Azalia



Adliyan Rizky
Pratama



Rofik



Gustian



Muslim
Musthofa Febri



Siti Salwa
Nurhadiva

Outline

01 **Problem**

02 **Goals**

03 **Framework**

04 **Data Understanding**

05 **Data Preparation**

06 **EDA**

07 **Model Evaluation**

08 **Insight and Recommendation**

Problem



Trend

Perusahaan e-commerce memiliki **tren penjualan yang meningkat setiap tahunnya**.



Problem

Tahun 2022 perusahaan mengalami penurunan penjualan yang ekstrim.



Churn

Perusahaan **banyak kehilangan pelanggan** yang menjadi sumber daya bisnis.



Cost

Berdasarkan literatur pada penelitian sebelumnya, usaha dan **biaya menggaet pelanggan baru 5-10% lebih besar** dibanding mempertahankan pelanggan lama.



Goals

Mempertahankan pelanggan untuk mengatasi penurunan penjualan.

Goals

01

**Meningkatkan
retensi
pelanggan.**

02

**Meningkatkan
penjualan
produk
unggulan.**

03

**Mengantisipasi
penurunan
penjualan
dengan
implementasi
model produk
rekomendasi.**

Framework

01 Pemahaman Pelanggan

Pemahaman preferensi pelanggan untuk implementasi model Product Recommendation.



02 Peningkatan Penjualan

Dengan memahami preferensi pelanggan, juga **berpengaruh pada peningkatan angka penjualan.**



Data Understanding

Dipilih **3 dataset** yang relevan dengan konteks rekomendasi sistem.

Dataset Customers

- customer_id,
- first_name,
- last_name,
- username,
- email,
- gender,
- birthdate,
- device_type,
- device_id,
- device_version,
- home_location_lat,
- home_location_long,
- home_location,
- home_country,
- first_join_date

Dataset Product

- id,
- gender,
- masterCategory,
- subCategory,
- articleType,
- baseColour,
- season,
- year,
- usage,
- productDisplayName

Dataset Transaction

- created_at,
- customer_id,
- booking_id,
- session_id,
- product_metadata,
- payment_method,
- payment_status,
- promo_amount,
- promo_code,
- shipment_fee,
- shipment_date_limit,
- shipment_location_lat,
- shipment_long,
- total_amount

Data Preparation

Seleksi fitur-fitur dari tiap dataset yang sudah dipilih untuk melakukan modeling.

Dataset Customers

- customer_id,
- gender,
- birthdate,
- home_location,
- first_join_date

Dataset Product

- id,
- gender,
- masterCategory,
- subCategory,
- articleType,
- baseColour,
- season,
- usage,
- productDisplayName

Dataset Transaction

- created_at,
- customer_id,
- product_metadata,
 - product_id,
 - quantity,
 - item_price

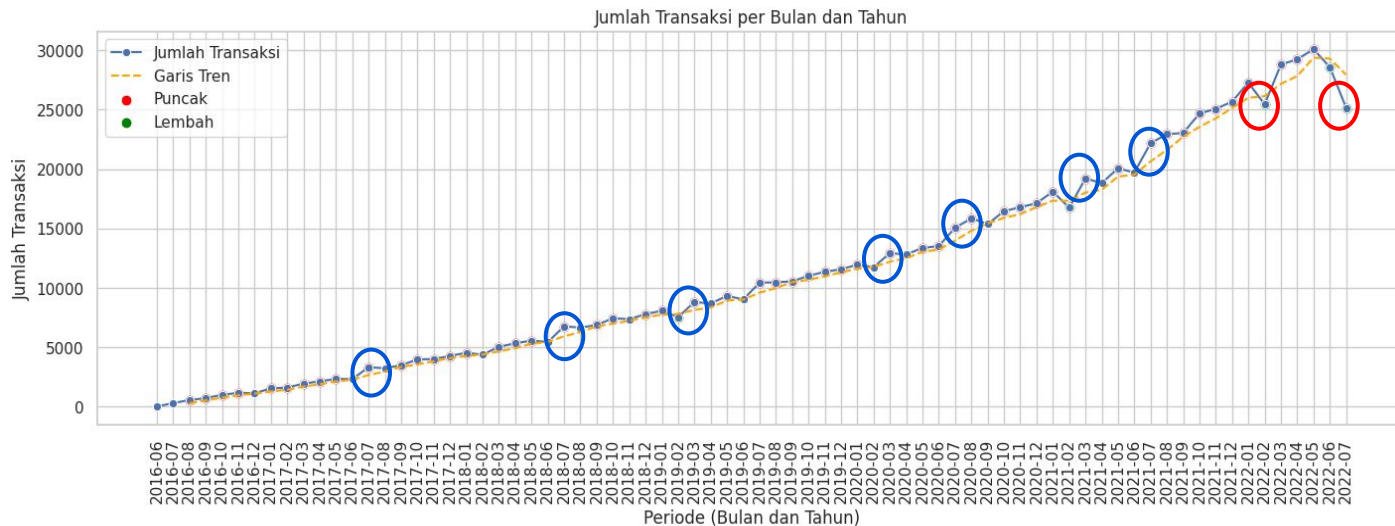
Data Preprocessing

1. **Inputodus** pada NaN di fitur-fitur pada dataset product.
2. **Konversi tipe data** yang masih salah.

Feature Engineering

1. Pembuatan **fitur age** (perhitungan dari birthdate)
2. **Lama pelanggan** bergabung (dari first join date)
3. **Revenue** (dari item price dan quantity)
4. Fitur **churn** dari terakhir transaksi (last created_at)

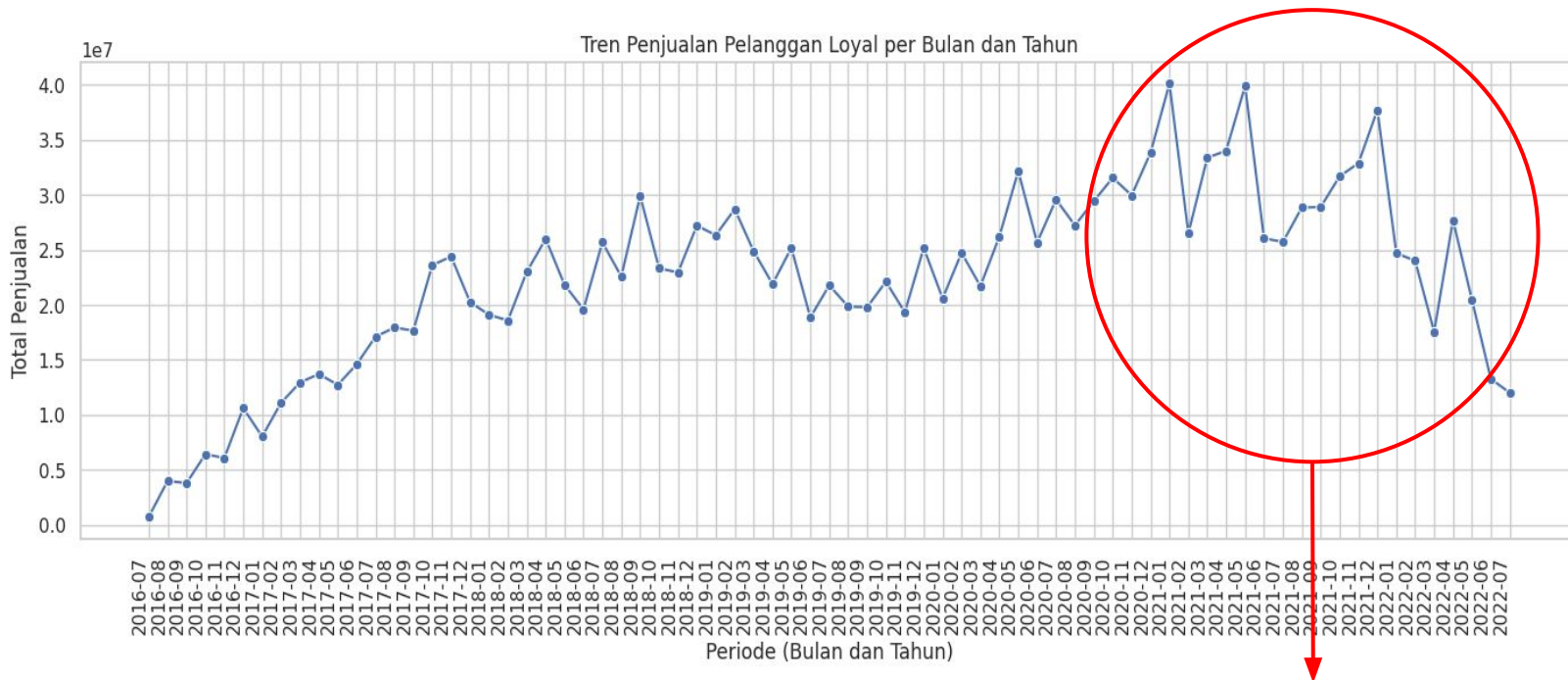
EDA



Penjualan **cenderung meningkat setiap tahunnya**, terutama **pada bulan 3 dan 7** (diidentifikasi ada event menyambut bulan ramadhan dan masa libur panjang)

Namun, **pada tahun 2022 terjadi penurunan penjualan** yang sangat ekstrim, **bahkan pada bulan yang seharusnya merupakan puncak peningkatan penjualan.**

Penurunan pembelian juga terjadi pada pelanggan yang tetap.



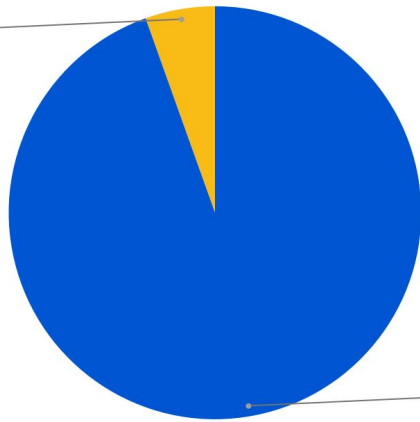
Perusahaan **gagal mempertahankan** retensi pelanggan.

EDA

Persentase **pelanggan yang sudah melakukan transaksi** dan **belum melakukan transaksi**.

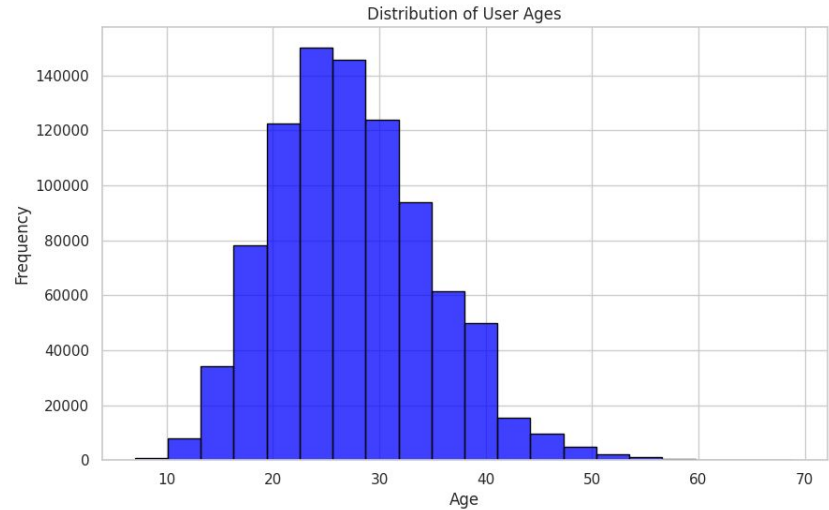
Persentase Transaksi Pelanggan

Belum Bertransaksi
5,5%



Sudah Bertransaksi
94,5%

Distribusi **usia pelanggan** di e-commerce.



EDA

Churn, merupakan kondisi dimana **pelanggan memutuskan berhenti berlangganan dan berpindah ke perusahaan lain** untuk berbelanja

Dihitung dari 3 bulan terakhir pelanggan tidak melakukan transaksi sama sekali.

Kegagalan memahami dan mempertahankan preferensi pelanggan menjadi faktor utama dalam masalah ini

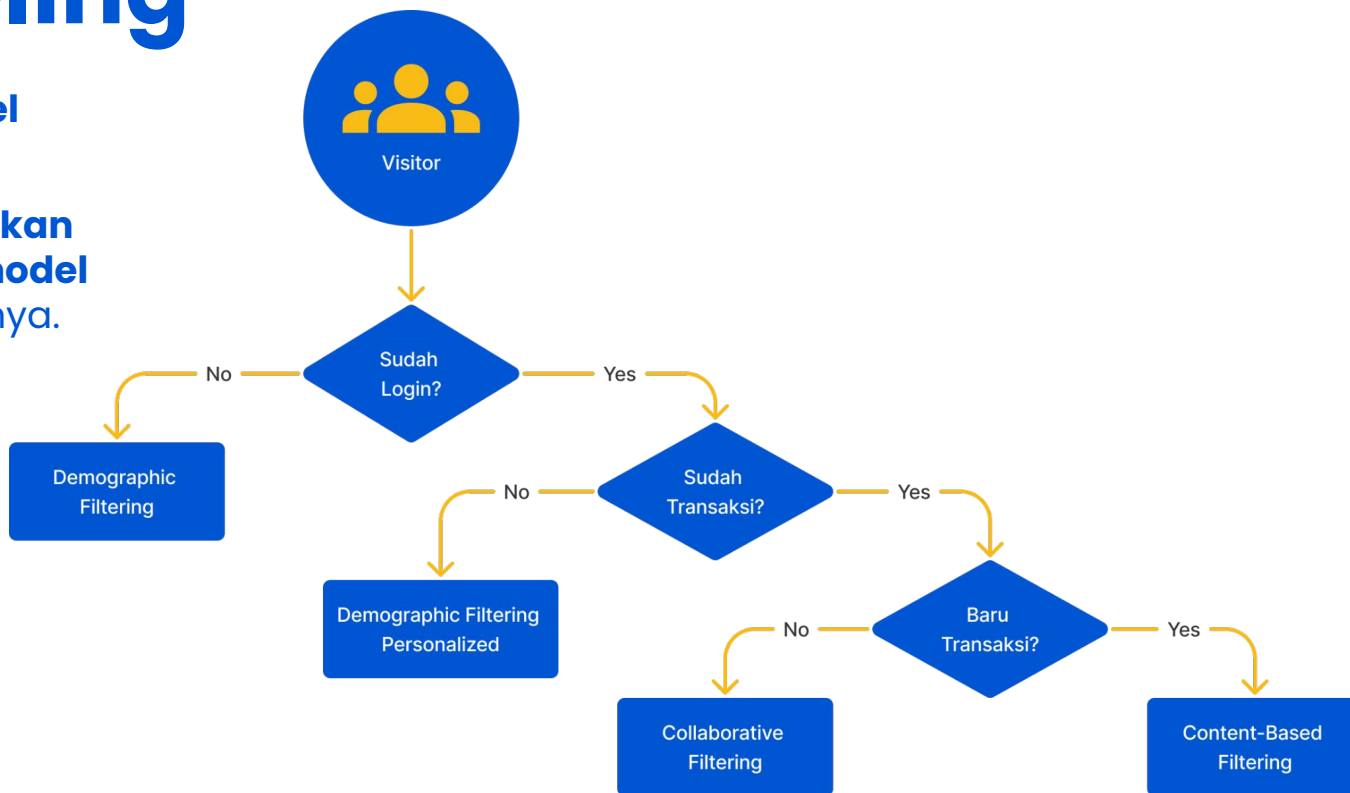
Faktanya **Perusahaan banyak kehilangan pelanggannya.**

Period_Quarter	2016	2017	2018	2019	2020	2021	2022
2016Q2	1	0	0	0	0	0	0
2016Q3	632	0	0	0	0	0	0
2016Q4	727	0	0	0	0	0	0
2017Q1	0	838	0	0	0	0	0
2017Q2	0	791	0	0	0	0	0
2017Q3	0	1316	0	0	0	0	0
2017Q4	0	1227	0	0	0	0	0
2018Q1	0	0	1319	0	0	0	0
2018Q2	0	0	1353	0	0	0	0
2018Q3	0	0	2026	0	0	0	0
2018Q4	0	0	1843	0	0	0	0
2019Q1	0	0	0	1930	0	0	0
2019Q2	0	0	0	1995	0	0	0
2019Q3	0	0	0	2853	0	0	0
2019Q4	0	0	0	2631	0	0	0
2020Q1	0	0	0	0	2720	0	0
2020Q2	0	0	0	0	2746	0	0
2020Q3	0	0	0	0	3588	0	0
2020Q4	0	0	0	0	3244	0	0
2021Q1	0	0	0	0	0	3413	0
2021Q2	0	0	0	0	0	3488	0
2021Q3	0	0	0	0	0	3728	0
2021Q4	0	0	0	0	0	3575	0
2022Q1	0	0	0	0	0	0	1807
2022Q2	0	0	0	0	0	0	0
2022Q3	0	0	0	0	0	0	0
Total	1360	4172	6541	9409	12298	14204	1807

Modeling

Schema model

Schema ini
menggambarkan
bagaimana model
bekerja nantinya.



Modeling | Demographic Filtering

Metode ini merekomendasikan **produk berdasarkan popularitas (terjual paling banyak)** kepada pelanggan. Kami mengaplikasikan metode ini ketika **belum memiliki informasi mendalam tentang pengunjung dan preferensinya**.

Groupby product_id -> sum quantity -> sorting descending.

	product_id	Total Quantity Sold	masterCategory	subCategory	articleType	baseColour	season	usage	productDisplayName
0	39731	87	Apparel	Bottomwear	Skirts	Grey	Summer	Formal	Arrow Woman Grey Skirt
1	13992	86	Footwear	Shoes	Sports Shoes	Grey	Fall	Sports	Puma Women Body Train Grey Sports Shoes
2	10445	85	Apparel	Bottomwear	Jeans	Blue	Fall	Casual	Flying Machine Men Midrise Blue Jeans
3	42953	83	Accessories	Belts	Belts	Black	Summer	Casual	French Connection Men Black Leather Belt
4	2329	82	Apparel	Topwear	Tshirts	Grey	Fall	Casual	Reebok Men Force India Grey T-shirt
5	8828	80	Apparel	Topwear	Shirts	White	Fall	Casual	Spykar Men Ranger Yellow N Blue Stripes White ...

Modeling | Demographic Filtering Personalized

Metode ini memberikan **rekomendasi personalisasi berdasarkan popularitas produk** kepada **pelanggan yang sudah login tetapi belum melakukan transaksi**. Rekomendasi dapat di filter berdasarkan gender, kategori, dan produk terpopuler pada bulan tertentu.

Groupby product_id -> sum quantity -> **filter by inputan** -> sorting descending.

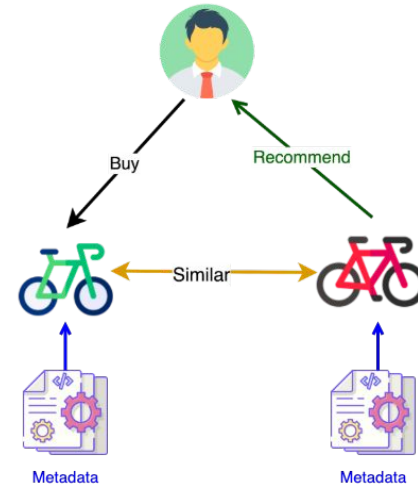
	product_id	masterCategory	subCategory	articleType	baseColour	season	usage	productDisplayName	total_sold
0	18261	Apparel	Topwear	Kurtas	Blue	Fall	Ethnic	Aneri Women Anshu Blue Kurta	42
1	40667	Apparel	Topwear	Tops	Red	Summer	Casual	Lee Women Red Top	26
2	51501	Apparel	Bottomwear	Jeans	Blue	Summer	Casual	Denizen Women Blue Jeans	22
3	40373	Apparel	Topwear	Tops	White	Summer	Casual	Tokyo Talkies Women White Top	20
6	6472	Apparel	Topwear	Tshirts	White	Summer	Casual	ADIDAS Wmen's Free Soul White T-shirt	20
4	34957	Apparel	Saree	Sarees	Beige	Summer	Ethnic	Fabindia Beige Jaipur Print Cotton Sari	19
12	24153	Apparel	Topwear	Kurtas	Black	Fall	Ethnic	Aurelia Women Black Printed Kurta	19
5	47066	Apparel	Bottomwear	Patiala	Teal	Fall	Ethnic	Shree Women Teal Printed Patiala	18

Modeling | Content-Based Filtering

Content-Based Filtering merupakan metode **rekomendasi produk yang dibuat berdasarkan data historis pelanggan.**

Produk direkomendasikan **berdasarkan kesamaan dengan barang yang telah dibeli** sebelumnya oleh pelanggan.

Diharapkan **pelanggan tertarik untuk melakukan pembelian lagi**, mungkin karena adanya opsi harga yang berbeda atau layanan tambahan yang menarik perhatian mereka.



Modeling | Content-Based Filtering

Selected feature → combined feature → inputan id item to check → calculation of similarity between items.

Metode yang digunakan untuk memfilter similar content adalah **Cosine Distance**.

8 Kategori produk yang digunakan untuk mengidentifikasi kesamaan produk: masterCategory, subCategory, genderProduct, articleType, baseColour, season, usage, ProductDisplayName.

product_id			combined_features	quantity
39901	53893	Accessories Bags Women Handbags Green Summer C...		48
39715	53654	Accessories Bags Women Handbags Green Summer C...		45
22066	28444	Accessories Bags Women Handbags Green Summer C...		42
39724	53666	Accessories Bags Women Handbags Green Summer C...		28
39758	53720	Accessories Bags Women Handbags Green Summer C...		25
39731	53673	Accessories Bags Women Handbags Green Summer C...		21
40070	54182	Accessories Bags Women Handbags Green Summer C...		18
23007	29737	Accessories Bags Women Handbags Green Summer C...		36
23005	29735	Accessories Bags Women Handbags Green Summer C...		35
23003	29733	Accessories Bags Women Handbags Green Summer C...		26

Modeling | Collaborative Filtering

Konsep rekomendasi ini **berfokus pada kesamaan preferensi antara pelanggan.**

Jika Customer A memiliki preferensi serupa dengan Customer B, produk yang dibeli oleh A akan direkomendasikan juga kepada B.










Tujuan utamanya adalah **memanfaatkan pembelian satu pelanggan untuk menarik pelanggan lainnya.**

Diagram illustrating user-user similarity and item-item similarity in collaborative filtering.

The diagram shows a grid of similarity scores between users and items.

Item-Item Similarity: Indicated by a horizontal double-headed arrow above the grid.

User-User Similarity: Indicated by a vertical double-headed arrow to the left of the grid.

					
	5		1	1	3
		2		4	
	5		1		3
	3	1	4	4	

Modeling | Collaborative Filtering

Input customer id -> Products not purchased -> scoring (predict customer id with product id for product not bought).

Filter Similar User and Content with SVD technique

Kami menggunakan **Alternating Least Squares (ALS)** untuk **meminimalkan kesalahan kuadrat antara prediksi dan nilai real**. Terdapat juga model **Singular Value Decomposition (SVD)** dalam **mengidentifikasi pola pembelian pelanggan**.

	product_id	score
7252	15792	5.000000
6581	51982	5.000000
20625	4826	5.000000
42651	32751	5.000000
31403	47155	5.000000
...
22041	1813	4.697274
33783	30944	4.689405
6697	39858	4.688437
33050	37941	4.686190
25382	22736	4.679422

100 rows × 2 columns

Model Evaluation

Model Evaluation Collaborative Filtering (Model yang lain tidak dilakukan evaluasi karena didasarkan pada proses filtering dan sorting saja)

Dalam mengevaluasi kita menggunakan nilai **Root Mean Square Error (RMSE)** dengan hasil sebagai berikut:

RMSE	1.648
Std.	0.0126

-> model **bekerja dengan cukup baik**

Rata-rata **waktu pelatihan 13.42 detik**, dan rata-rata **waktu pengujian 1.66 detik**.

Menunjukkan model mampu bekerja menggunakan waktu yang cukup cepat untuk data yang besar. -> dan ini **penting dan merupakan keuntungan karena responsivitas sistem rekomendasi penting**.

Insight

- Implementasi sistem rekomendasi ini dapat membantu meningkatkan penjualan, memperbaiki pengalaman berbelanja, dan menghemat biaya 5-10% untuk pemasaran yang bisa tidak dilakukan karena berhasil mempertahankan pelanggan lama.
- Strategi ini merupakan **metode yang efektif**, dengan kondisi perusahaan baru saja mengalami **kerugian yang signifikan**.
- Dalam rekomendasi, **tak ada jawaban benar atau salah; preferensi pengguna adalah kunci**. Model berupaya **memudahkan** dengan rekomendasi dari data historis, **meningkatkan kenyamanan berbelanja**, dan membantu perusahaan **meningkatkan penjualan**.

Recommendation

- Untuk meningkatkan retensi pelanggan yang lebih, dapat menggunakan **model model rekomendasi produk yang terkini dari aktivitas pengguna** (seperti add to chart, search keyword, dan lainnya).
- Direkomendasikan pula untuk **menambahkan fitur rating product** (yang benar-benar menunjukkan kualitas produk dari penilaian customer).
- Jika hasil implementasi tidak sesuai target, metode lain seperti **Deep Learning, Context-Aware Recommendation, Association Rule Mining, Factorization Machines, Word Embeddings**, atau lainnya dapat diuji.
- Perlu dilakukan **monitoring dan pembaruan secara berkala** untuk **menjaga dan meningkatkan retensi pelanggan, mengikuti perubahan perilaku pelanggan dan tren produk**.

Thank you~

Team Thalassa Adieu



MENTOR
Muhammad
Anwar Sanusi



FASILITATOR
Jasmine Azalia



Adliyan Rizky
Pratama



Rofik



Gustian



Muslim
Musthofa Febri



Siti Salwa
Nurhadiva