

# REKOMENDASI OBAT

Somnium



Final Project Data Science Academy



# ANGGOTA

01

Nadira Eka Rahmaharva

02

Maryesta Apriliani S.

03

Yiesha Reyhani G.

# CONTENT

01

Pendahuluan

02

Data Pre-processing & Cleaning

03

Exploratory Data Analysis (EDA)

04

Feature Engineering

05

Sistem Rekomendasi

06

Evaluasi

07

Penutup

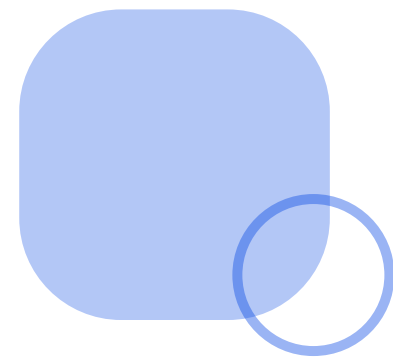


The background of the slide is a close-up, slightly blurred image of numerous light blue, round tablets. Each tablet has a thin, vertical line (score line) running down its center. The tablets are scattered across the frame, creating a textured, medical-themed background.

# PENDAHULUAN

# Latar Belakang

Dalam dunia kesehatan yang terus berkembang, semakin banyak penelitian yang menghasilkan berbagai jenis obat untuk penyembuhan dan pencegahan penyakit. Dengan **banyaknya merk dan jenis obat** yang beredar, seringkali muncul tantangan dalam memilih **obat yang tepat** sesuai dengan kebutuhan pasien. Selain itu, terdapat pula masalah tentang kelengkapan data dari pasien terkait privasinya. Oleh karena itu, diperlukan sistem **rekomendasi obat** yang dapat membantu untuk **mengurangi waktu** proses pencarian obat yang dibutuhkan pasien.



## 01 Inefisiensi Proses Manual dalam Pencarian Obat Pengganti

Dibutuhkan sistem rekomendasi yang dapat mempercepat proses pencarian obat alternatif

## 02 Keterbatasan Ketersediaan Obat

Ketersediaan obat yang seringkali tidak mencukupi, mengakibatkan kebutuhan akan rekomendasi obat dibutuhkan.

## 03 Kompleksitas Pemilihan Obat

Banyaknya merk obat dengan indikasi, fungsi, dan komposisi bahan dasar yang bervariasi, membuat pemilihan obat menjadi kompleks.



# TUJUAN & MANFAAT

Dengan memanfaatkan dataset yang ada, kami bertujuan untuk menggali wawasan serta informasi penting terkait obat-obatan yang beredar, sekaligus mengembangkan sistem rekomendasi yang dapat mendukung peningkatan layanan kesehatan.



## **Mendapat Informasi/Insight Mengenai Obat**

Menganalisis data obat secara mendalam untuk mendapatkan pola, hubungan, dan informasi penting.



## **Sistem Rekomendasi Obat**

Membangun sistem rekomendasi obat yang relevan dan efisien sehingga mempercepat proses pelayanan kesehatan



The background of the image is a dense, overlapping pattern of blue, oval-shaped pills. The pills are semi-transparent, allowing the ones underneath to be visible. They are arranged in a way that creates a sense of depth and texture. The overall color is a vibrant blue.

# **DATA PRE-PROCESSING & CLEANING**

# Dataset

Terdapat dua data yang akan digunakan dalam project ini:

- **Data 250k Medicines Usage, Side Effects and Substitutes** - [sumber](#)

Data ini berisi informasi tentang lebih dari 248.000 obat medis dari semua produsen yang tersedia di seluruh dunia. Data tersebut mencakup rincian seperti nama obat, bahan aktif, kegunaan terapeutik, dosis, efek samping, dan penggantinya. Kumpulan data ini bertujuan untuk menyediakan sumber daya yang berguna bagi peneliti medis, profesional kesehatan, dan produsen obat.

Fitur	Keterangan
name	Nama obat
substitute	Obat alternatif
sideEffect	Efek samping penggunaan obat
use	Kegunaan obat

Fitur	Keterangan
Chemical Class	Pengelompokkan berdasarkan kandungan bahan kimia yang digunakan
Habit Forming	Indikasi kecanduan, Yes or No
Therapeutic Class	Pengelompokkan berdasarkan fungsi obat
Action Class	Pengelompokkan berdasarkan tindakan yang diberikan obat



# Dataset

- **Data A-Z Medicine Dataset of India** - [sumber](#)

Dataset ini berisi semua kemungkinan obat-obatan yang dapat peneliti temukan selama ia melakukan penelitian untuk mengetahui detail obat seperti komposisi obat, jenis obat, ketersediaan pasar, harga, dll.

Fitur	Keterangan
<b>name</b>	Nama obat
<b>price</b>	Harga obat
<b>Is_discontinued</b>	Ketersediaan obat di pasaran per November 2022
<b>manufacturer_name</b>	Perusahaan yang memproduksi obat

Fitur	Keterangan
<b>type</b>	tipe Obat
<b>pack_size_label</b>	Ukuran kemasan dan kuantitas obat
<b>short_composition1</b>	Kandungan pertama obat
<b>short_composition2</b>	Kandungan lainnya

# Langkah - Langkah

1

Data Profiling

2

Deteksi Duplikasi

3

Menghapus Plural Issue pada Kolom pack\_size\_label

4

**Menghapus Redudansi pada Data med:**  
Mempertahankan satu baris untuk setiap nama obat

5

Menghapus Extra Word pada Kolom use0 - use4

6

**Membuat Kolom Baru:**  
type, primary\_comp, entire\_comp, value, entire\_value, main\_subs, other\_subs, sideeffects

7

Merge Dataset

8

Drop Kolom yang Tidak Dibutuhkan

9

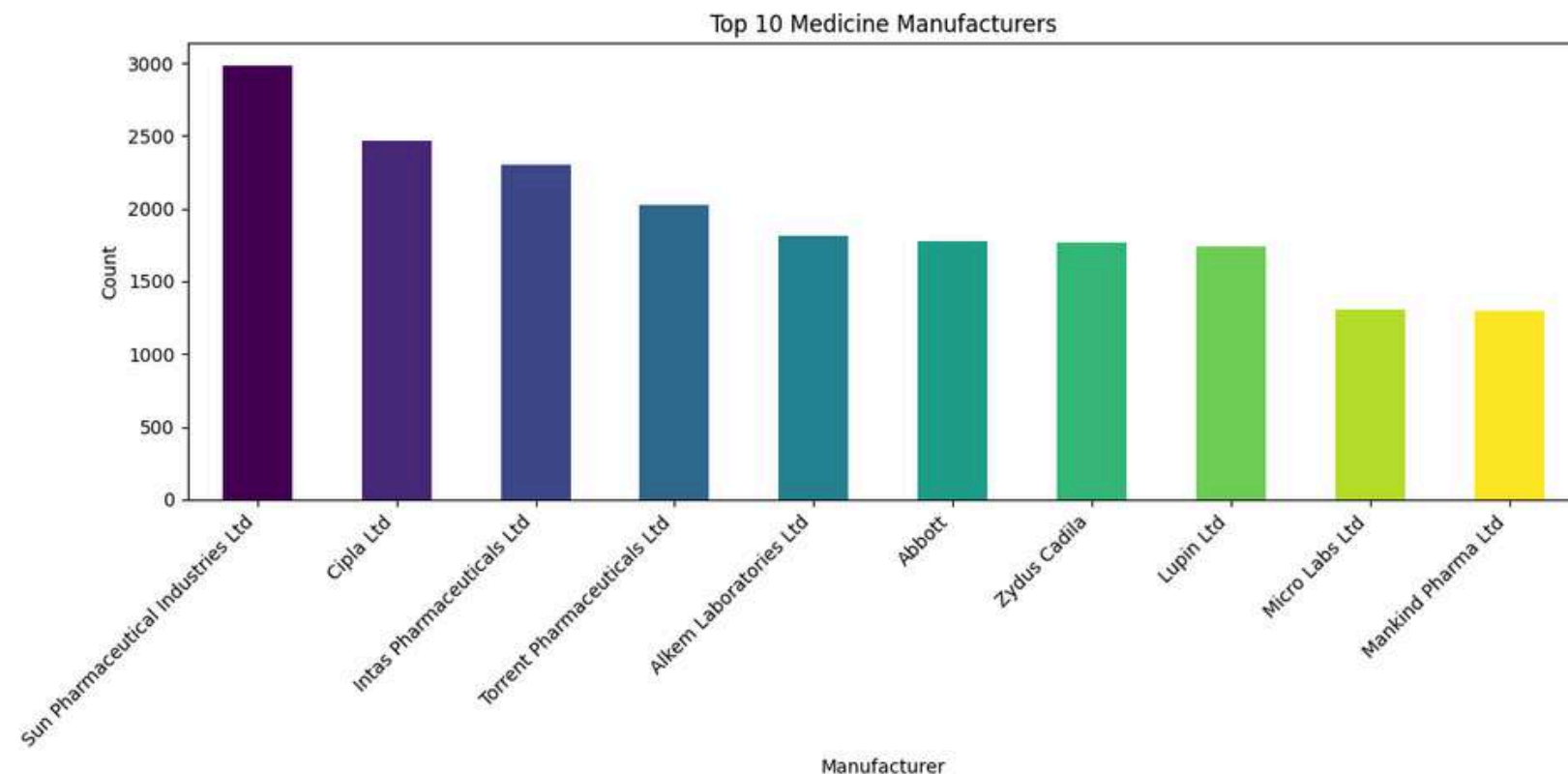
Missing Value

The background of the image is a dense, overlapping field of light blue, circular pills. Each pill has a thin, vertical line or score mark running down its center. The pills are slightly out of focus, creating a sense of depth. The overall color palette is a monochromatic blue, ranging from a very light, almost white blue to a deep, saturated blue.

# **EXPLORATORY DATA ANALYSIS**



# TOP 10 MEDICINE MANUFACTURERS



Gambar di atas menunjukkan 10 perusahaan dengan merk obat paling banyak. Selanjut akan dibahas mengenai 10 perusahaan tersebut dari hasil analisis yang sudah dilakukan.

## **Umumnya obat dalam bentuk tablet**

Semua perusahaan yang berada pada daftar 'Top 10 Medicine Manufacturers' umumnya menjual obatnya dalam bentuk tablet

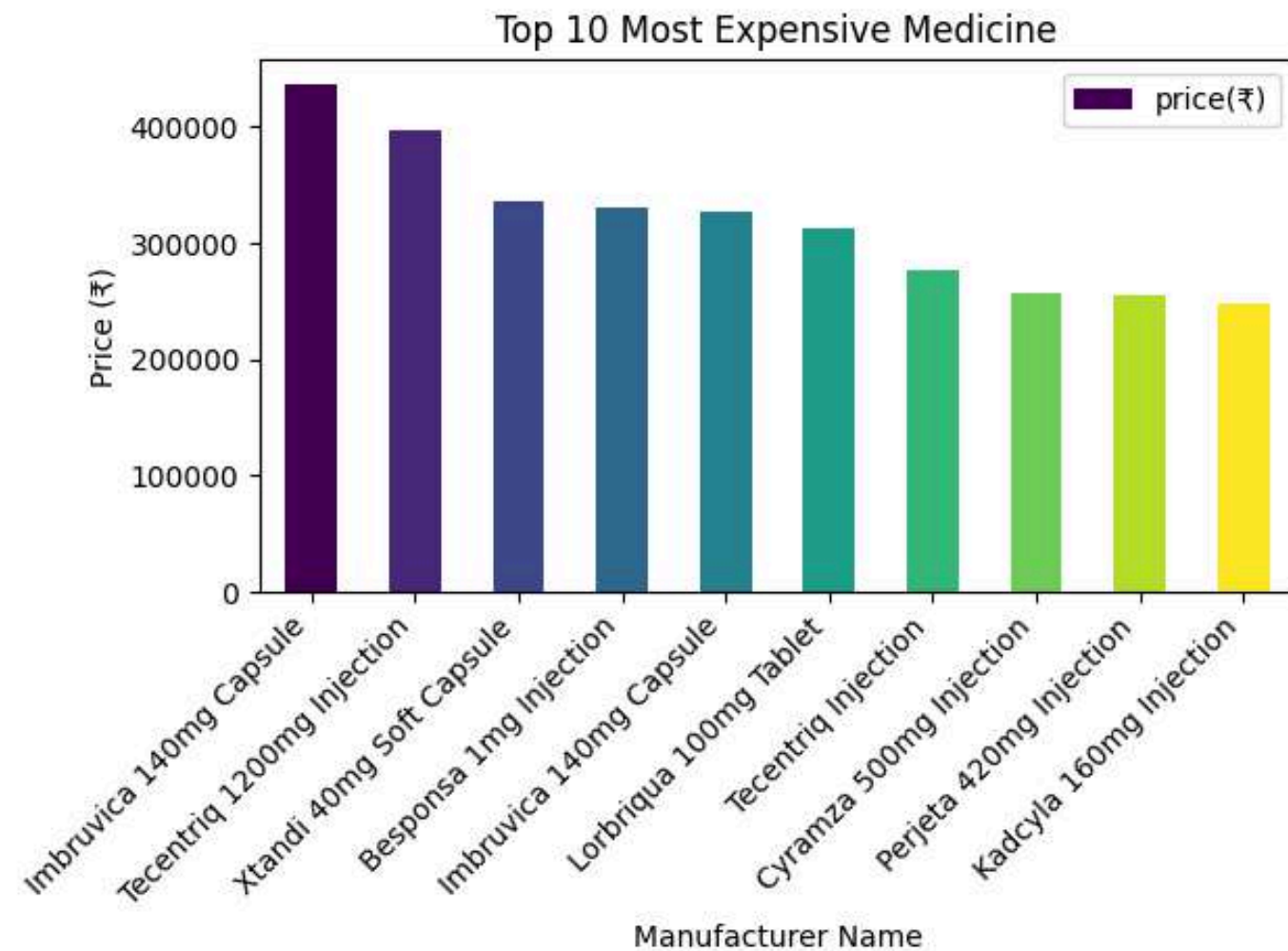
## **Banyak menjual obat untuk mengobati infeksi bakteri**

Semua perusahaan kecuali Intas Pharmaceuticals Ltd dan Torrent Pharmaceuticals Ltd paling banyak menjual obat untuk mengobati infeksi bakteri. Hal ini tentu wajar mengingat banyak penyakit yang disebabkan oleh serangan bakteri

## **Masing-masing perusahaan memproduksi obat yang beragam**

Tidak terdapat kriteria tertentu atau perusahaan yang lebih menonjol dalam satu bidang

# TOP 10 MOST EXPENSIVE MEDICINE



Gambar di atas menunjukkan 10 obat dengan harga paling mahal (dalam Rupee). Selanjut akan dibahas mengenai 10 obat dari hasil analisis yang sudah dilakukan.

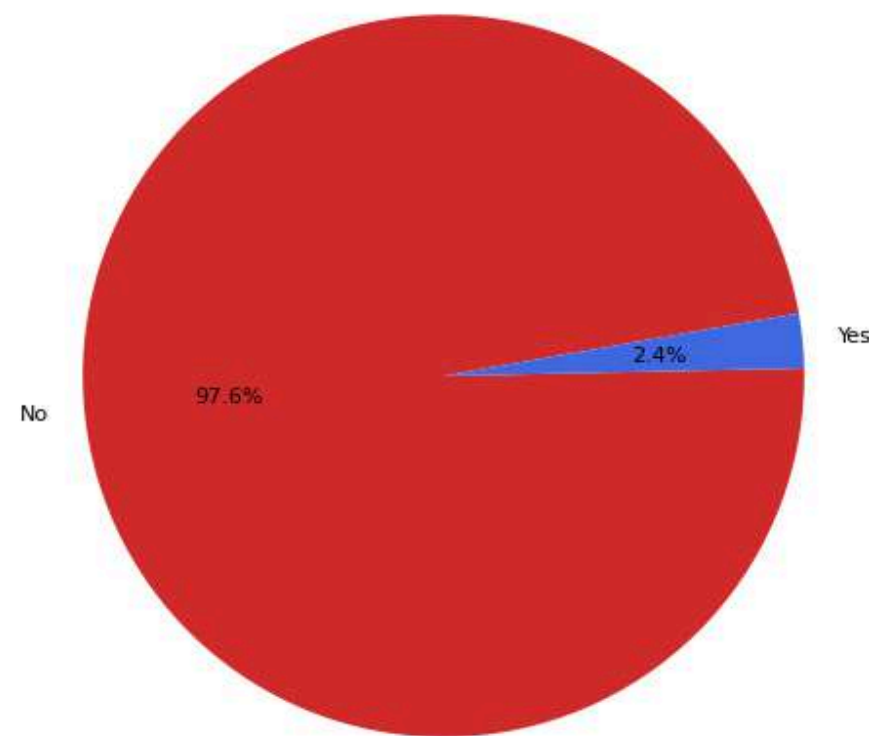
## Obat untuk perawatan kanker

Semua obat yang termasuk ke dalam 'Top 10 Most Expensive Medicine' memiliki kegunaan untuk perawatan kanker.

Kanker termasuk ke dalam penyakit ganas yang sulit disembuhkan. Biaya uji klinis yang mahal serta harga pasar dari obat kanker yang memang di atas rata-rata membuat obat kanker memiliki harga yang tidak murah

# SIFAT KETERGANTUNGAN

Proportion of Habit-Forming Drugs



Gambar di atas menunjukkan proporsi mengenai sifat ketergantungan obat. Dapat dilihat bahwa hanya **2,4% obat pada dataset yang dapat menyebabkan ketergantungan**. Selanjut akan dibahas mengenai obat yang dapat menyebabkan ketergantungan tersebut.

## **Obat untuk suatu keadaan psikologis dan *pain relief***

Obat yang mempunyai kemungkinan ketergantungan adalah obat yang berhubungan dengan keadaan psikologis dan pain relief.

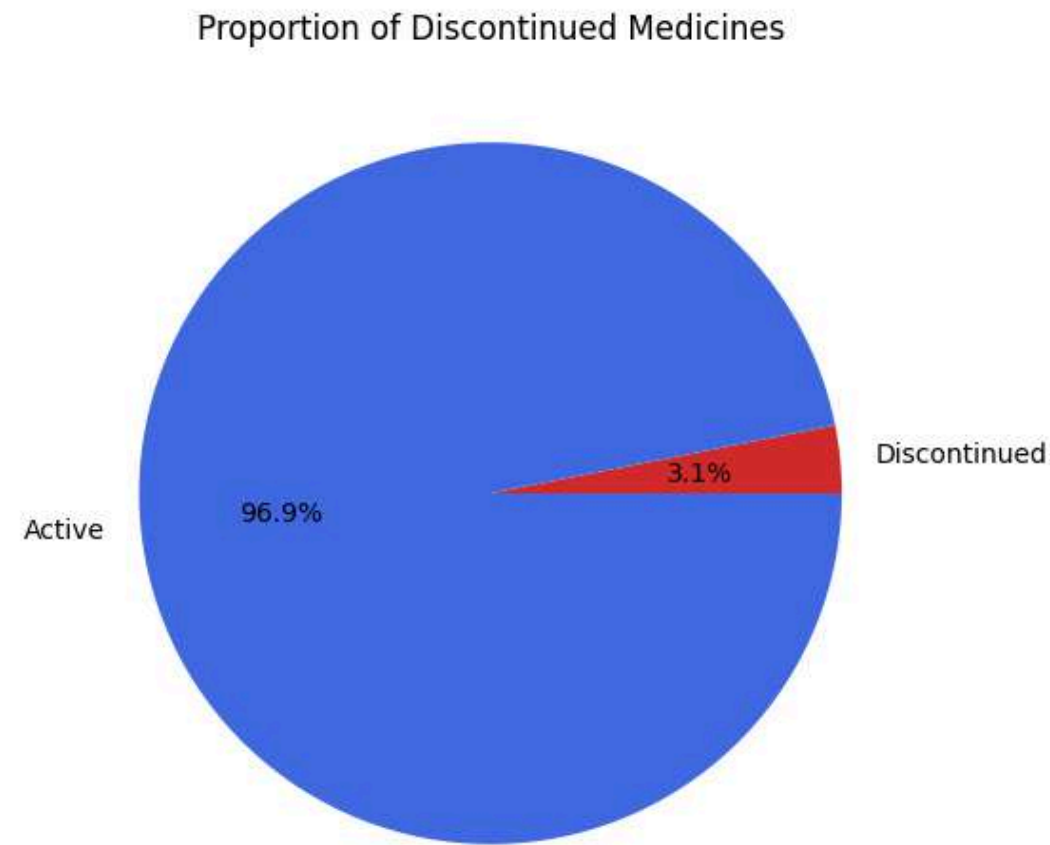
## **Kelas obat yang dapat menyebabkan ketergantungan**

Benzodiazepines, opioids, non-benzodiazepine hypnotics, barbiturate adalah kelas obat yang dapat menyebabkan ketergantungan apabila dikonsumsi dalam jangka panjang atau tidak sesuai dosis. Hal ini disebabkan oleh kandungan zat kimianya.

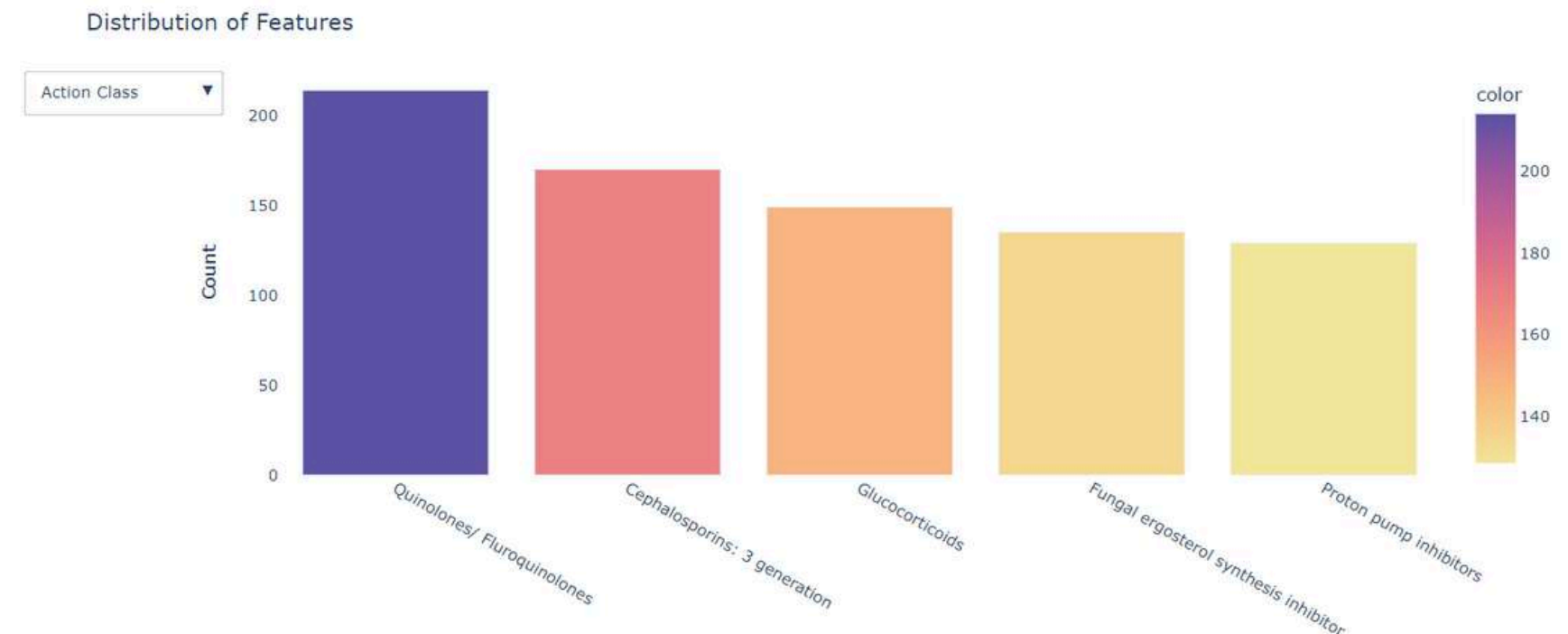


# OBAT DISKONTINU

\*Terhitung sejak November 2022



Gambar di atas menunjukkan proporsi mengenai ketersediaan obat di pasaran **per November 2022**. Dapat dilihat bahwa terdapat **3,1% obat yang tidak tersedia di pasaran per November dari dataset yang digunakan**. Selanjut akan dibahas mengenai obat yang tidak tersedia di pasaran tersebut.



Gambar di atas adalah 5 action class yang paling sering muncul pada obat diskontinu. Urutan pertama ditempati oleh **Fluoroquinolone** dan yang kedua adalah **Cephalosporins: 3 generation**

## **Action class Fluoroquinolone memiliki efek samping yang serius**

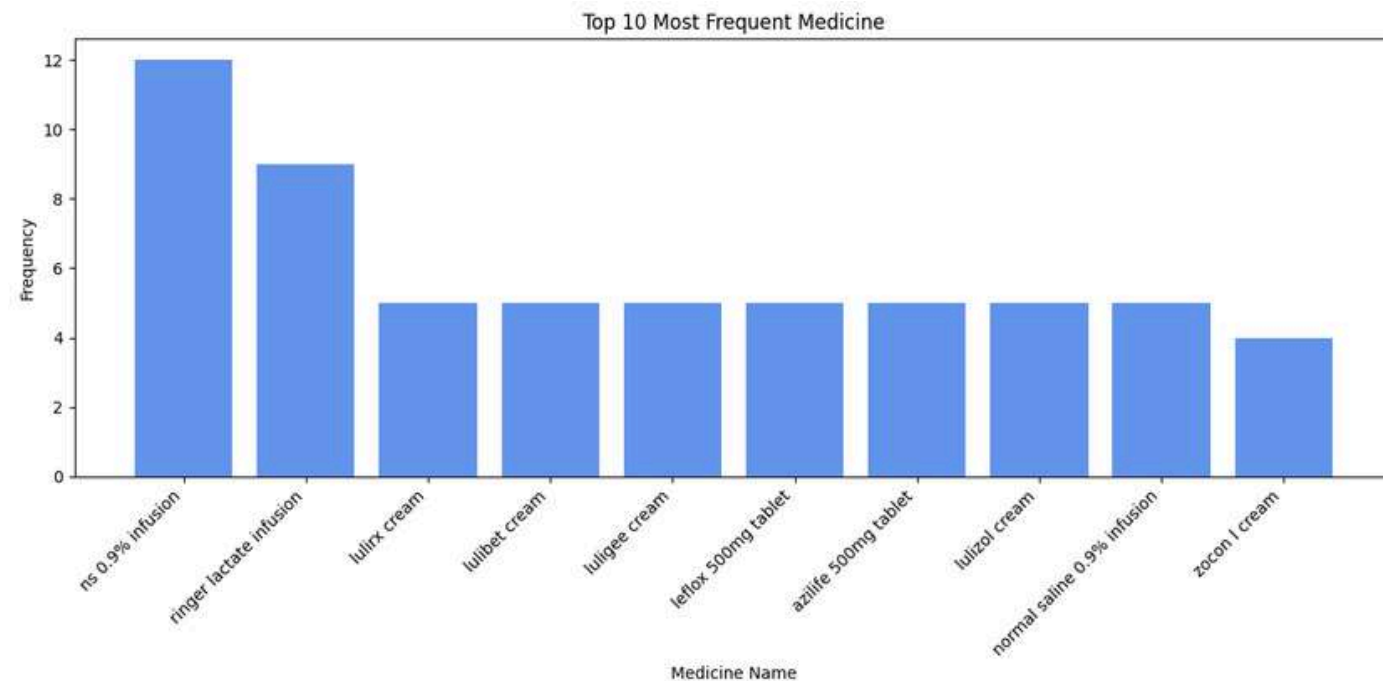
FDA menyarankan bahwa Fluoroquinolone tidak cocok untuk kondisi umum dan hanya boleh dipertimbangkan untuk digunakan jika pengobatan dengan antibiotik lain yang kurang beracun telah gagal.

## **Adanya perubahan dari generasi Cephalosporins yang ingin digunakan**

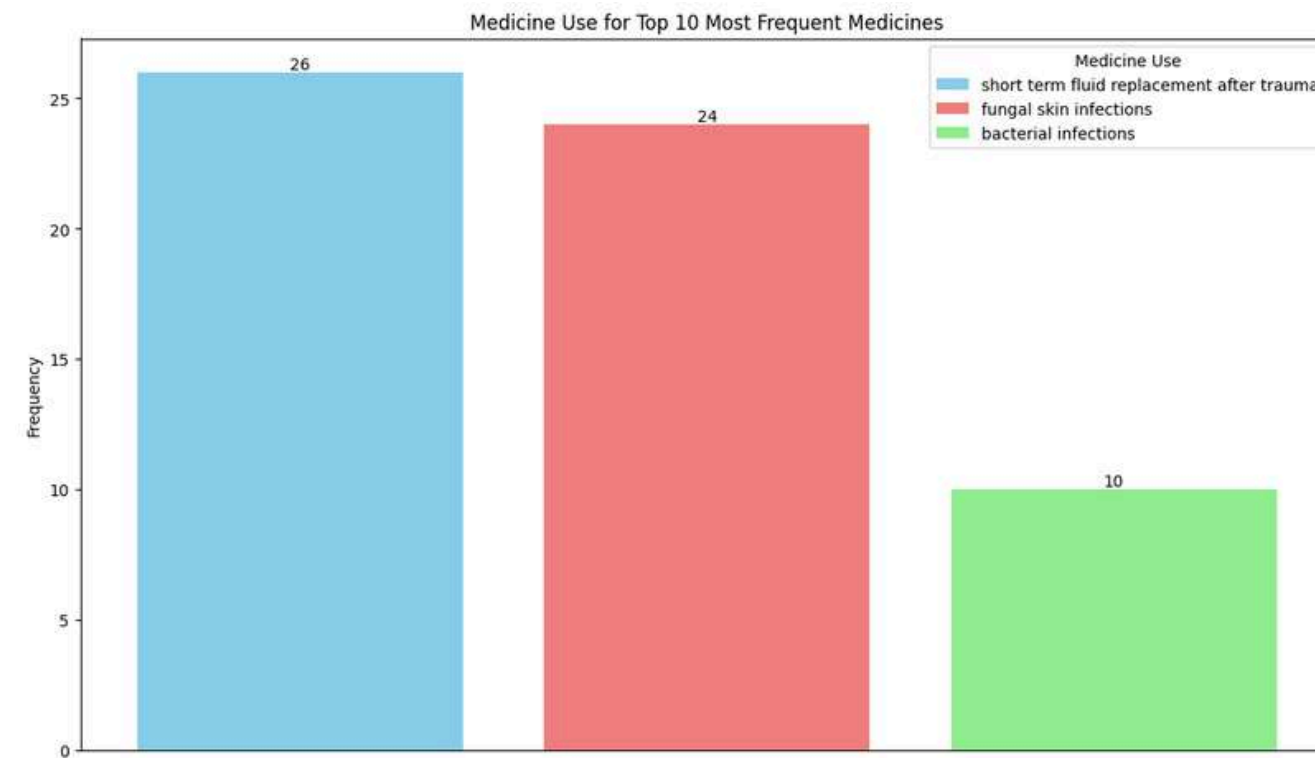
Hingga saat ini, sudah ada 5 generasi cephalosporins. Kemungkinan obat tidak lanjut diproduksi karena adanya perubahan dari generasi Cephalosporins yang ingin digunakan.

# TOP 10 MOST FREQUENT MEDICINE

\*Terhitung sejak November 2022



Gambar di atas menunjukkan 10 obat yang paling sering muncul pada dataset. Selanjut akan dibahas mengenai 10 obat tersebut



Gambar di samping menunjukkan 3 kegunaan yang paling sering muncul dari 10 obat yang sering muncul.

## Obat untuk infeksi bakteri dan infeksi jamur

Obat yang banyak beredar terkait untuk infeksi bakteri dan infeksi jamur pada kulit.

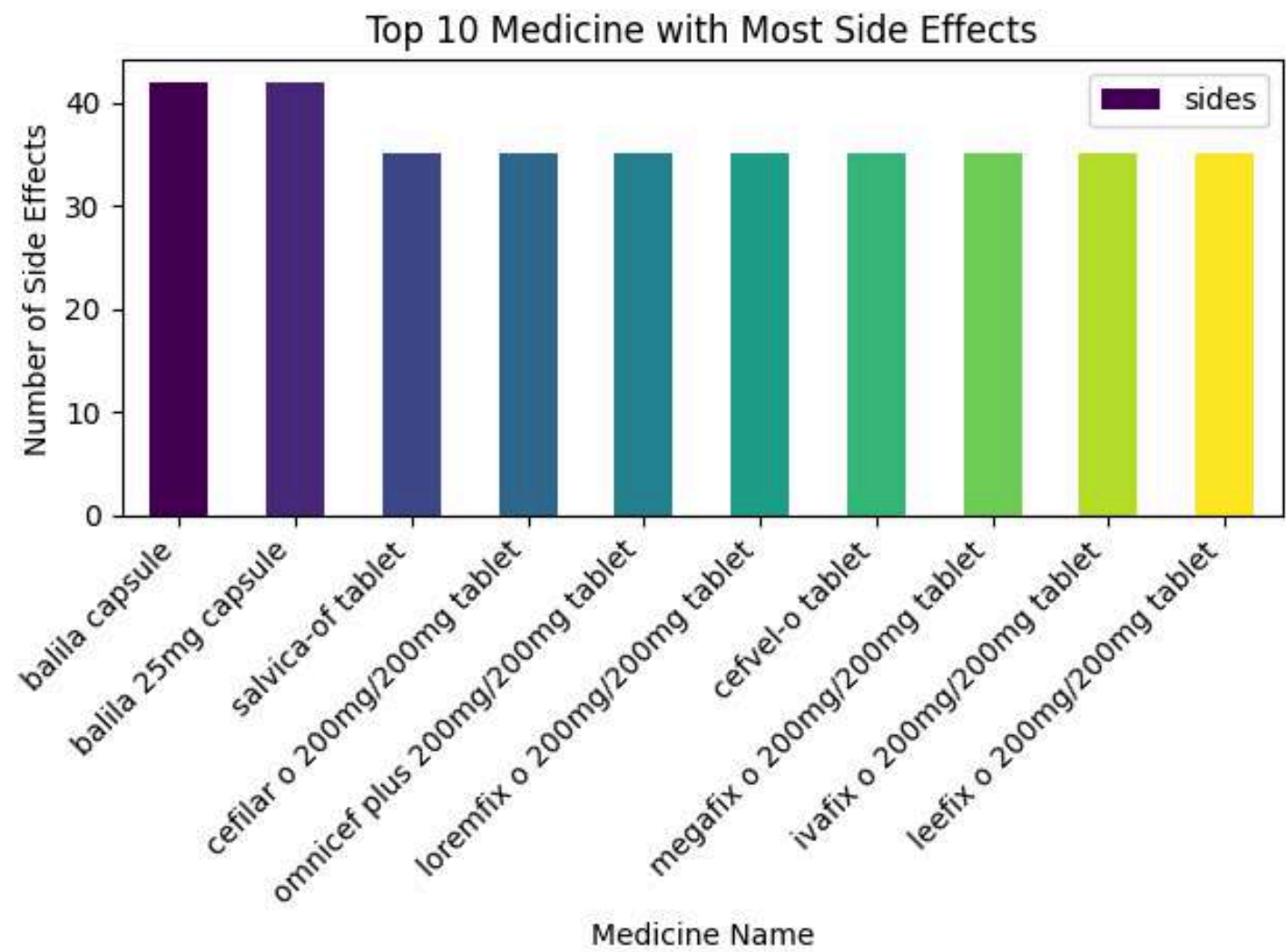
## Obat untuk *short term fluid replacement*

Obat yang sering digunakan untuk kegunaan ini salah satunya adalah NS 0.9% Infusion



# EFEK SAMPING

\*Terhitung sejak November 2022



Gambar di atas menunjukkan **10 obat yang paling banyak memiliki efek samping.**

## Obat untuk infeksi bakteri

8 dari 10 obat yang paling banyak memiliki efek samping terkait dengan obat yang memiliki kegunaan untuk perawatan infeksi bakteri.

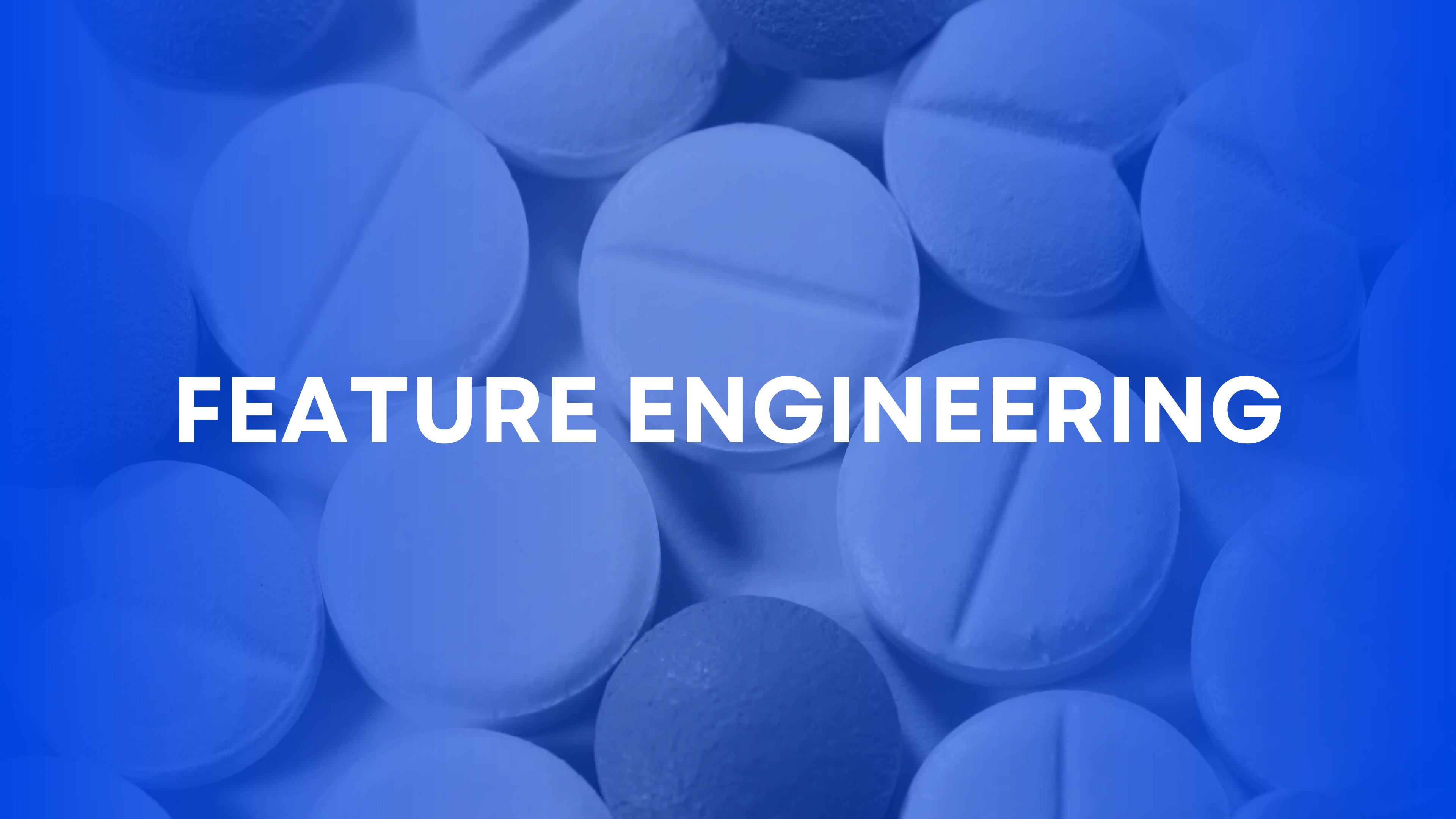
Tabel di samping menunjukkan **efek samping yang paling sering muncul** dari dataset obat.

## Efek samping mual

Sekitar 24,38% obat pada dataset memiliki efek samping berupa mual.

Efek Samping	Frekuensi
Nausea	61930
Diarrhea	21756
Headache	15503
Vomiting	13366
Rash	10083
Dizziness	8704
Hypoglycemia (low blood glucose level)	6629
Sleepiness	6429
Constipation	5520
Application site reactions (burning, irritation, itching and redness)	5496



The background of the image is a dense, close-up shot of numerous light blue, oval-shaped pills. Each pill has a thin, dark horizontal line (a score line) across its center. The pills are scattered across the entire frame, creating a textured, repeating pattern. The lighting is even, highlighting the smooth surface of the pills.

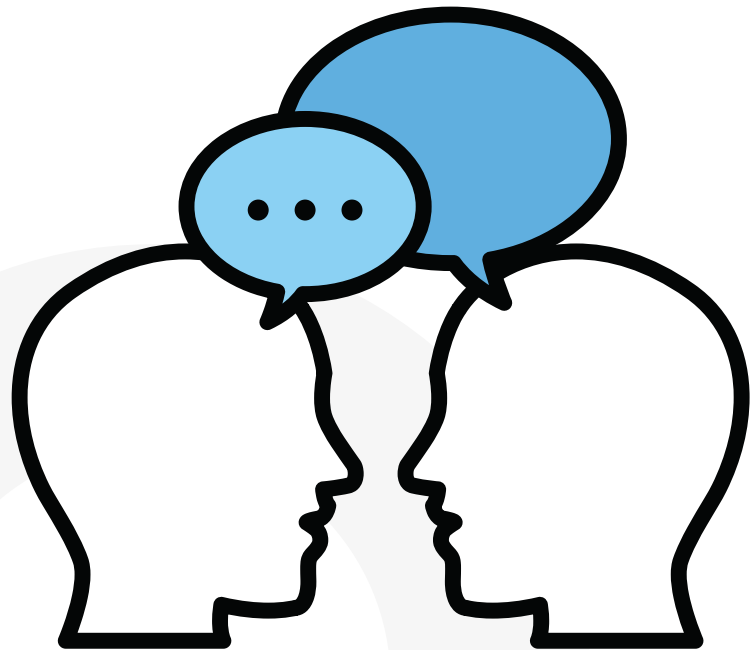
# FEATURE ENGINEERING

# WORD EMBEDDING

Mengonversi **teks** dengan Word2Vec menjadi **vektor numerik** yang mampu menangkap makna dan konteks kata-kata.

Mengaplikasikan *word embedding* pada fitur type, primary\_comp, entire\_comp, value, entire\_value, dan use.

Menambahkan vektor numerik hasil embedding sebagai **fitur baru** pada dataset.





The background of the image consists of numerous white, round pills scattered across a blue gradient. Some pills have a visible score line. The text 'SISTEM REKOMENDASI' is centered in white, bold, uppercase letters.

# **SISTEM REKOMENDASI**



# SISTEM REKOMENDASI

Data yang digunakan pada sistem rekomendasi adalah 10% dari dataset awal secara acak karena adanya keterbatasan perangkat.

Sistem rekomendasi obat dirancang untuk menerima **nama obat** sebagai **input** dan memberikan 10 nama obat lain yang paling mirip berdasarkan skor similaritas.

**Fitur** yang digunakan untuk menghitung kesamaan meliputi jenis obat (seperti sirup atau tablet), komposisi, dosis obat, dan kegunaan obat.

**Penentuan bobot** yang sesuai untuk setiap fitur cukup penting untuk sistem ini. Setiap bobot mencerminkan **signifikansi** tiap-tiap atribut dalam menentukan kesamaan antar obat.

Dengan menggunakan bobot yang telah ditetapkan, sistem menghitung skor **similaritas** antara obat yang diinput dan obat-obat lain dalam dataset. Obat-obat dengan skor similaritas tertinggi direkomendasikan sebagai alternatif yang paling sesuai.

# SISTEM REKOMENDASI



## Cosine Similarity

Mengukur kemiripan antara dua vektor. Vektor yang digunakan adalah vektor numerik hasil embedding.



## Weighted Similarity

Menghitung kemiripan dengan mempertimbangkan bobot dari setiap fitur teks. Fitur yang lebih penting atau berpengaruh diberikan bobot yang lebih besar



## Hyperparameter Tuning

Membuat dictionary "weightage" yang menyimpan bobot untuk masing-masing fitur. Bobot dapat diatur dengan mempertimbangkan fitur mana yang lebih penting.

The background of the image consists of numerous white, round pills scattered across the frame. Each pill has a vertical score line down its center. The entire image is overlaid with a semi-transparent blue filter, giving it a monochromatic appearance. The pills are in various orientations and focus, creating a sense of depth.

**EVALUASI**



# Pengujian Sistem

Total Precision:  
 $420\%/5 = 84\%$

Akan dilakukan pencarian rekomendasi 5 obat dengan kegunaan yang berbeda

- **Laurunam 1000mg Injection** - severe bacterial infection

Pada obat Laurunam 1000mg injection, 10 rekomendasi relevan. Relevan dalam artian 10 rekomendasi memiliki kegunaan yang sama, begitu pula tipe hingga komposisi obatnya. **Precision: 10/10 (100%)**

- **Newtel-h 40 Tablet** - hypertension (high blood pressure)

Pada obat Newtel-h 40 Tablet, 10 rekomendasi relevan. Relevan dalam artian 10 rekomendasi memiliki kegunaan yang sama, begitu pula tipe hingga komposisi obatnya. **Precision: 10/10 (100%)**

- **Prestoheal DS Tablet** - pain relief

Pada obat Prestoheal DS Tablet, 10 rekomendasi relevan. Relevan dalam artian 10 rekomendasi memiliki kegunaan yang sama, begitu pula tipe obat. **Precision: 10/10 (100%)**

- **Dolzy MR 250mg/50mg/325mg Tablet** - pain due to musclespasm

Pada obat Dolzy MR 250mg/50mg/325mg Tablet, 10 rekomendasi relevan. Relevan dalam artian 10 rekomendasi memiliki kegunaan yang sama, begitu pula tipe hingga komposisi obatnya. **Precision: 10/10 (100%)**

- **Bdpoma 1mg Capsule** - multiple myeloma

Pada obat Bdpoma 1mg Capsule terdapat 2 rekomendasi relevan dan 8 rekomendasi yang tidak relevan. Hal ini disebabkan karena data yang dipakai hanya 10% sehingga kemungkinan data obat yang memiliki kegunaan yang sama dengan obat ini tidak terambil. **Precision: 2/10 (20%)**

The background of the image is a close-up, high-angle shot of numerous penicillin tablets. The tablets are circular and light-colored, with a distinct vertical score line on each. They are scattered across the frame, creating a textured, repeating pattern. The entire image is overlaid with a semi-transparent blue filter, which gives it a clinical and professional appearance. The lighting is even, highlighting the texture of the tablets and the sharpness of the score lines.

**PENUTUP**

# Rekomendasi

01

## **Menambahkan variabel interaksi obat**

Informasi tentang potensi interaksi obat dapat sangat penting dalam pengobatan.

02

## **Menambahkan variabel golongan obat**

Mengidentifikasi apakah obat termasuk dalam kategori obat resep atau obat bebas dapat mempengaruhi rekomendasi, terutama dalam konteks aksesibilitas dan regulasi.

03

## **Penggunaan perangkat dengan spesifikasi yang lebih tinggi**

Disarankan untuk menggunakan perangkat dengan spesifikasi yang lebih tinggi guna meningkatkan kinerja dan efisiensi dalam pengolahan data.

04

## **Implementasi sistem rekomendasi obat pada situs web atau aplikasi**

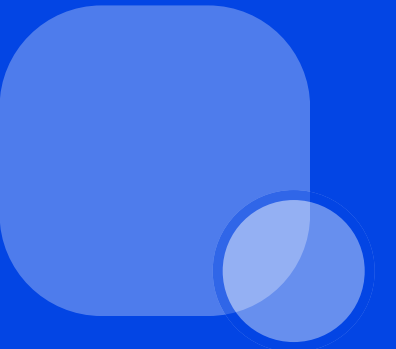
Dengan mengimplementasi sistem rekomendasi obat, penggunaannya dapat digunakan secara luas

05

## **Menggunakan data dengan volume lebih besar dalam sistem rekomendasi**

Untuk meningkatkan relevansi sistem rekomendasi obat, disarankan agar data yang digunakan mencakup volume yang lebih besar.



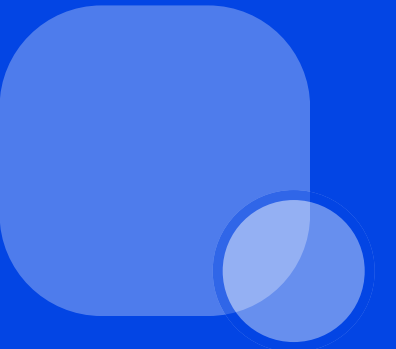


## Rekomendasi untuk Pengambil Kebijakan dan Pemerintah:

- 01 Upaya mengenai mahalnya harga obat perawatan kanker**  
Obat untuk perawatan kanker mendominasi daftar obat dengan harga termahal. Upaya untuk menurunkan harga obat tersebut dapat membantu menjangkau lebih banyak pasien yang membutuhkan.
- 02 Edukasi mengenai efek samping obat**  
Mengingat tingginya prevalensi obat yang dapat menyebabkan efek samping seperti mual, pemerintah dapat menjalankan kampanye edukasi untuk meningkatkan kesadaran masyarakat tentang potensi risiko
- 03 Pengawasan terhadap praktik promosi obat**  
Memperketat regulasi terkait promosi obat untuk memastikan bahwa informasi yang diberikan kepada masyarakat tidak menyesatkan dan berdasarkan data ilmiah yang valid.



# Konklusi



Melalui analisis data obat yang komprehensif, didapatkan informasi-informasi mengenai obat-obatan yang beredar di pasaran. Hasil analisis obat menunjukkan bahwa obat perawatan kanker menjadi yang termahal, sementara beberapa obat yang sudah tidak tersedia di pasaran terindikasi karena memiliki komposisi dengan efek samping serius. Efek samping mual umum terjadi, menekankan pentingnya perhatian terhadap efek samping dalam pemilihan obat. Berdasarkan hasil analisis, disarankan agar pengambil kebijakan mempertimbangkan langkah-langkah seperti meningkatkan edukasi publik mengenai risiko efek samping obat serta memperketat regulasi terkait promosi obat untuk memastikan informasi yang akurat dan terpercaya bagi masyarakat.

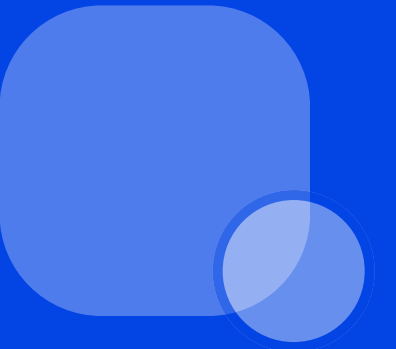
Sistem rekomendasi obat telah berhasil dibangun dengan menggunakan metode *weighted similarity*. Sistem ini menunjukkan performa yang baik dengan nilai presisi sebesar 84% pada pengujian dengan 5 skenario. Untuk memperoleh rekomendasi obat, sistem dipengaruhi oleh jumlah data serta fitur yang digunakan. Semakin banyak data maka semakin kompleks proses yang dijalankan dan membutuhkan waktu yang cukup lama. Tentunya sistem akan bekerja lebih baik apabila data obat yang digunakan lebih lengkap. Dengan peningkatan data dan fitur, sistem ini berpotensi memberikan rekomendasi yang semakin akurat serta bermanfaat bagi pengguna.

# **Link Colab**

<http://ristek.link/SomniumCode>

# **Interactive Dashboard**

<http://ristek.link/SomniumDashboard>



Alkaff, M., Khatimi, H., & Eriadi, A. (2020). Sistem Rekomendasi Buku Menggunakan Weighted Tree Similarity dan Content Based Filtering. MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput, 20(1), 193-202.

Aryanto, A. D., Santoso, J., & Purwanto, D. D. (2021). Sistem rekomendasi obat pengganti menggunakan metode cnn. Jurnal Sistem Cerdas dan Rekayasa (JSCR), E-ISSN: 2656-7504, 3(1), 25-36

Baggio, D., & Ananda-Rajah, M. R. (2021). Fluoroquinolone antibiotics and adverse events. Australian prescriber, 44(5), 161–164. <https://doi.org/10.18773/austprescr.2021.035>

Edinoff, A. N., Nix, C. A., Hollier, J., Sagrera, C. E., Delacroix, B. M., Abubakar, T., Cornett, E. M., Kaye, A. M., & Kaye, A. D. (2021). Benzodiazepines: Uses, Dangers, and Clinical Considerations. Neurology international, 13(4), 594–607. <https://doi.org/10.3390/neurolint13040059>.



# THANK YOU

FOR YOUR ATTENTION

Somnium

