

Capstone Project

Data Classification and Summarization using IBM Granite

Nama: Moch Alwi Maosul

Repo: [GitHub - capstone-covid-granite](#)

Dataset

Dataset

Nama: COVID-19 Indonesia Time Series All (Publik)

Sumber: Kaggle / <https://www.kaggle.com/datasets/hendratno/covid19-indonesia>

Format: CSV:

Ukuran: ~4.000+ baris, multi-provinsi, time-series

Link Dataset:

https://github.com/alwi2323/capstone-covid-granite/blob/main/data/covid_19_indonesia_time_series_all.csv

Project Overview

Tujuan proyek ini adalah untuk menganalisis perkembangan COVID-19 di Indonesia berdasarkan data harian per provinsi.

Proyek ini menggabungkan klasifikasi data manual dengan bantuan AI IBM Granite untuk mendapatkan insight berbasis data.

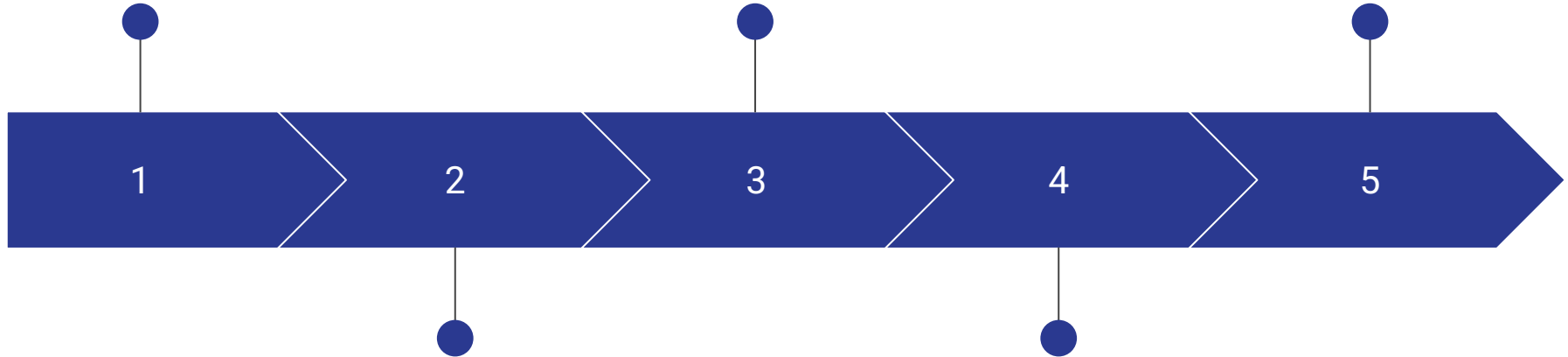
Output yang dihasilkan meliputi visualisasi, klasifikasi tingkat risiko, dan ringkasan insight dengan dukungan AI.

Proses Analisis

Data cleaning dan penggabungan kolom relevan (kasus, kematian, kesembuhan)

Klasifikasi level risiko:
Low, Medium, High

Insight tambahan diperoleh dengan summarization oleh IBM Granite



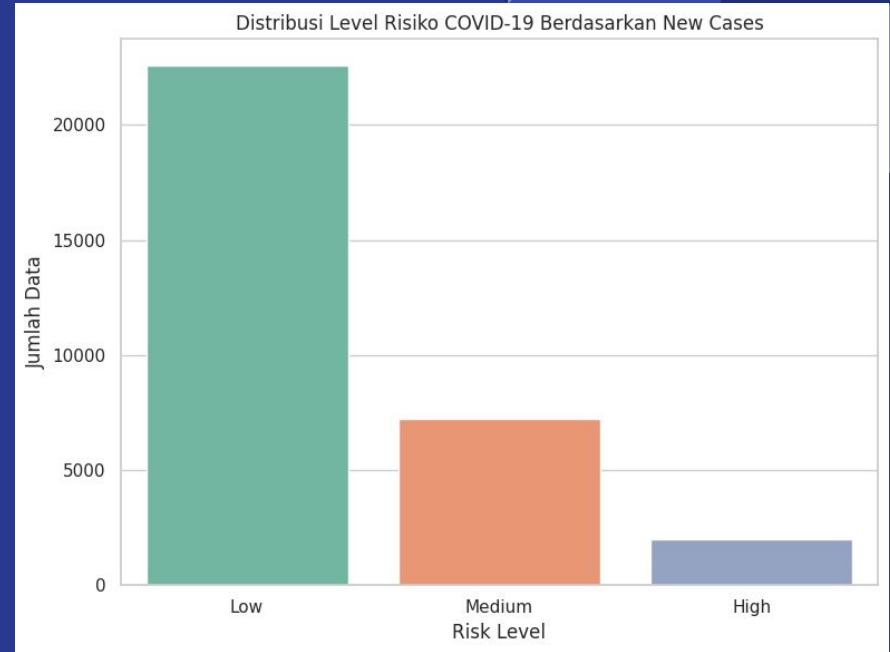
Perhitungan case fatality rate & recovery rate

Visualisasi distribusi risiko menggunakan seaborn

Insight & Visualisasi

Sebagian besar data menunjukkan provinsi dengan level risiko **rendah**, yang mengindikasikan penanganan yang membaik di sebagian besar wilayah.

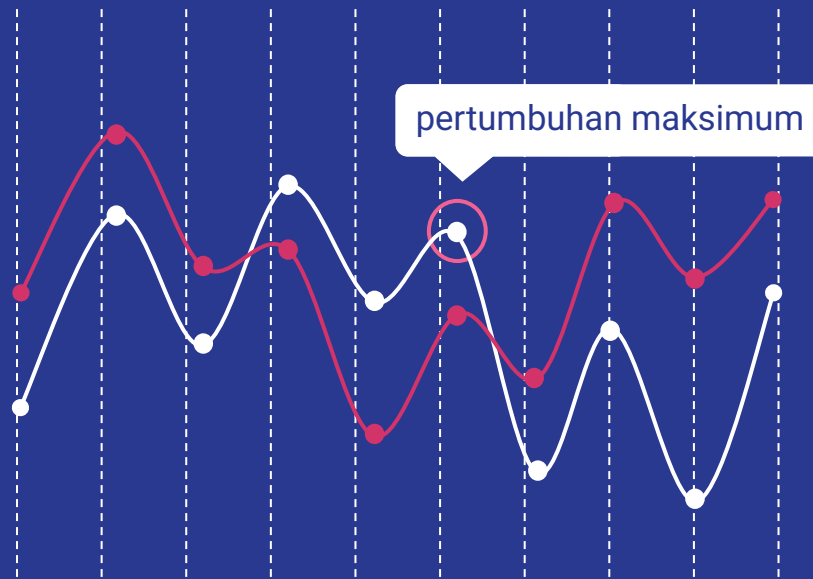
Namun, masih terdapat provinsi dengan level risiko **tinggi**, yang menjadi fokus utama untuk intervensi lebih lanjut.



Insight dari WatsonX Granite (AI Summary)

Dalam bulan Juli 2021, Provinsi DKI Jakarta mencatat perkembangan COVID-19 yang bervariasi. Selama masa tersebut, terdapat 15.480 kasus baru, dengan 540 kematian baru, menunjukkan tingginya jumlah kematian bulanan. Namun, jumlah pasien sembuh baru mencapai 13.920, menunjukkan tingkat kesembuhan yang tinggi sebesar 89.9%. Rata-rata kasus harian di Jakarta tersebut mencapai 499, dan tingkat fatalitas terhadap kasus baru mencapai 3.49%. Perkembangan ini menunjukkan situasi yang kritis, namun juga menandai potensi peningkatan sembuh di provinsi tersebut.

Model: [granite-3-3-8b-instruct](#), WatsonX Prompt Lab



Kesimpulan & Rekomendasi

Kesimpulan

DKI Jakarta memiliki beban kasus tinggi selama Juli 2021 namun juga menunjukkan tingkat kesembuhan yang baik.

Tingkat fatalitas mengindikasikan perlunya perhatian ekstra pada penanganan pasien.

Rekomendasi

- Fokus pada penguatan sistem perawatan pasien kritis
- Edukasi publik untuk mengurangi penyebaran
- Optimalkan kapasitas rumah sakit dan distribusi vaksin

AI Support Explanation

Proyek ini memanfaatkan model AI `granite-3-3-8b-instruct` dari IBM melalui WatsonX Prompt Lab. Model digunakan untuk summarization data provinsi secara natural-language. Hasilnya mendukung analisis manual dan memperkuat narasi data dalam bentuk yang mudah dipahami.

Model: IBM Granite, WatsonX platform

Tugas AI: Summarization berbasis prompt

Bahasa: Natural Language Bahasa Indonesia