

Machine Learning Assignment – GDGoC Selection

Air Quality Data Analysis Project

Name : Revania Azzahra

Role Applied : Machine Learning

Date : 24 – 12 – 2007

A. Links

- GitHub Repository

<https://github.com/RevaNia12/MachineLearning-Air-Quality-Assignment>

- Dashboard Streamlit

<https://machinelearning-air-quality-assignment-8hulmtv9yqnyhzhkqvj7s2t.streamlit.app/>

B. Project Overview

Project ini merupakan analisis data menggunakan Air Quality Dataset secara end-to-end. Proses analisis ini meliputi data prepare, data cleaning, exploratory data analysis (EDA), visualisasi data, dan penyampaian insight melalui dashboard Streamlit.

C. Tools

Alat yang digunakan dalam project ini:

- Python
- Pandas
- Matplotlib
- Seaborn
- Streamlit

D. Hasil Analisis

Berdasarkan hasil analisis data dengan EDA pada dataset Air Quality pada tahun 2004-2005 ditunjukkan bahwa suhu udara memiliki hubungan yang signifikan dengan beragam variabel di lingkungan, khususnya yaitu kelembapan relatif (RH) dan juga konsentrasi polutan seperti CO₂, NO₂ dan lain-lain. RH ditunjukkan memiliki korelasi negatif terhadap suhu yang berarti semakin suhu meningkat maka semakin turun pula kelembapan relatifnya. Di variabel yang lain, beberapa polutan memiliki pola perubahan terhadap waktu yang dapat divisualisasikan dengan bentuk grafik.