Future City Digital Twin – Project Report

A Capstone Project Report

Abstract

This project, titled 'Future City Digital Twin', is focused on developing a predictive and analytical framework for urban air quality monitoring. The system leverages machine learning (Random Forest) and deep learning (LSTM) models to forecast pollution levels. Additionally, the project demonstrates a real-time deployment using Streamlit and Ngrok, allowing interactive visualization and accessibility through the web. The purpose is to help city planners, environmentalists, and policymakers make informed decisions.

Introduction

With the rise of smart cities, the concept of a digital twin has gained significant importance. A digital twin is a virtual replica of physical systems that allows simulation, prediction, and monitoring. In this project, a digital twin of a city environment has been designed to analyze and predict air quality trends. The project integrates past datasets with real-time future predictions, giving an edge in disaster preparedness and environmental planning.

Dataset Description

The dataset used in this project consists of both historical (past) and forecasted (future) data for air quality in urban environments. Key parameters include PM2.5 levels, temperature, humidity, and wind speed. The past dataset was used for training the ML/DL models, while the future dataset was utilized for testing and validation.

Methodology

1. Data Preprocessing: Cleaning missing values, normalization, and feature engineering.
2. Machine Learning Model: Random Forest Classifier was used for risk prediction.
3. Deep Learning Model: LSTM model was applied to forecast PM2.5 concentration levels.
4. Visualization: Streamlit was used to create an interactive app for real-time analysis.
5. Deployment: Ngrok was used to host the application and share it externally.

Results and Analysis

The results indicate that the models were able to capture trends in pollution levels effectively. The graph below represents the historical vs. predicted air quality levels in the city.

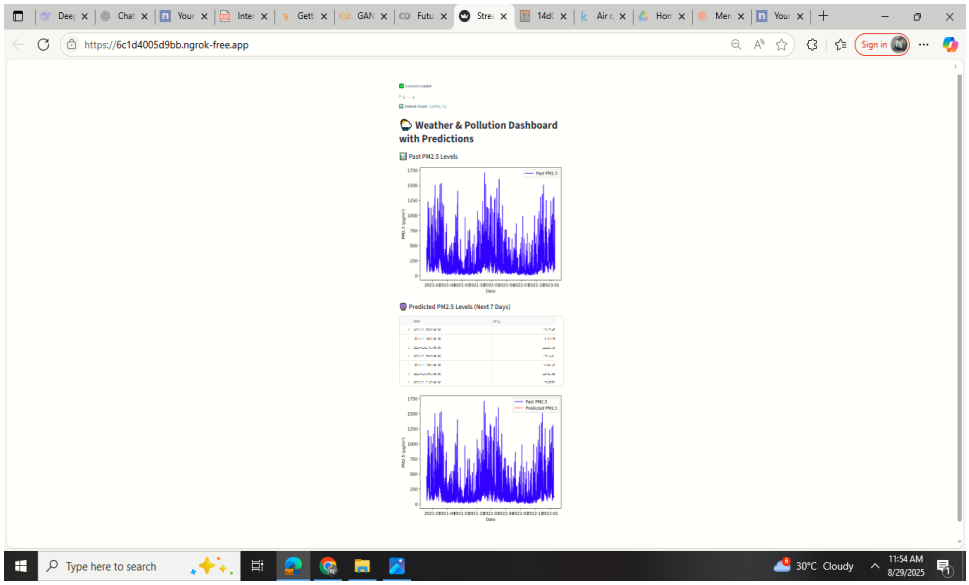

Figure 1: Visualization of past and future dataset trends (PM2.5 Levels).

Discussion

The project demonstrates how combining machine learning and deep learning approaches can enhance urban analytics. Random Forest provided robust classification for multi-risk assessment, while LSTM captured sequential dependencies for forecasting. The integration with Streamlit makes the solution accessible for stakeholders.

Future Work

1. Integrating more environmental factors such as traffic and industrial emissions.
2. Scaling the system to handle multiple cities.
3. Deploying on cloud infrastructure (AWS, GCP, Azure) for reliability.
4. Adding reinforcement learning for adaptive policy simulation.

Conclusion

The Future City Digital Twin project successfully demonstrates the potential of AI-driven urban monitoring. By predicting air quality, the project provides insights for planning and mitigation strategies. It serves as a foundation for building large-scale smart city applications.

References

[1] Nishit K. Sinha, Introduction to Data Interpretation.
[2] Research papers on Digital Twin and Smart Cities.
[3] Streamlit and Ngrok official documentation.
[4] Python libraries: Pandas, Numpy, Matplotlib, TensorFlow, Scikit-learn.