AIR QUALITY ANALYSIS AND PREDICTION

IN TAMILNADU

Problem Statements:-

* Tamil Nadu faces significant challenges related to

air quality ,with various regions

experiencing high levels of pollution due to

industrial activities , vehicular emissions and

natural factors.

* Poor air quality poses severe health and

environmental risks,making it imperative to

develop innovative solutions.

Innovative Solutions:-

* AirQuality Montoring Solutions
* Machine Learning Models
* Air Quality Index (AQI) Alerts
* Community-based Air Quality Monitoring
* Public Awareness Campaigns
* Early Warning Systems

Dataset Link:-

Columns Details:-

* Stn Code
* State
* City/Down/village
* Location of Monitoring Station
* Agency
* Type of Location
* SO2
* NO2
* RSPM/PM10

SPM

**Explanation about column details:**

* **Std Code:STD stands for Subscriber Trunk Dialling.To find STD code of any city from drop down select box.**
* **State:In this dataset we use the state TamilNadu.In That state We find the air quality analysis and prediction.**
* **City/Town/Village:In the dataset we use the city, town,villages like chennai,coimbatore,madurai,salem,thoothukudi.**
* **Location of Montoring Station: The location of montoring station in the dataset is Kathivakkam,Govt.High Secondary School,Thiruvottiyur,Madras Medical,etc...**
* **Agency:The agency we use in the datasets is TamilNadu,Thiruvottiyur Municipal Office, Chennai etc..**
* **Type of Location:In this datasets we use the location likes Industries Area,National Environmental Engineering Research Institute,Tamilnadu State Pollution Control Board etc..**

Libraries Used :

In this we use a libraries like Numpy,Pandas,Matplotlib,Seaborn and import the packages like warnings and fliters.

* Pandas:Pandas is a python package that is used for data analysis and manipulation.Is a open source libaries that is built over numpy. It can be installed as “pip install pandas”
* Seaborn:Is a python data visualization library based on matplotlib.It can be installed as “pip install seaborn”.
* Scikit-Learn (sklearn): Provides tools for machine learning tasks like regression (for prediction), classification, and clustering. You can use regression models to predict air quality parameters.

Train and test :

* Model selection:Choose a suitable machine learning or deep learning model for air quality prediction. Common choices include linear regression, decision trees, random forests, support vector machines, or neural networks.
* Model trainig:Train the selected model on the training dataset using appropriate algorithms and hyperparameter tuning.Monitor the model's performance on the validation set to prevent overfitting.
* Model Evaluation:Evaluate the model's performance on the test set using appropriate evaluation metrics like Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), or R-squared (R2) score.Use time series-specific evaluation techniques if applicable, such as time-series cross-validation or rolling origin validation.
* Model Deployment:Once you have a satisfactory model, deploy it to make real-time air quality predictions. This could involve creating a web-based dashboard or integrating it into existing monitoring systems.

Rest of Explanation:

* Launch public awareness campaigns to educate residents about the health risks associated with poor air quality. Promote sustainable practices such as reducing vehicle emissions, planting more trees, and reducing waste burning.

Accuracy:

* Improving the accuracy by splitting the data on heavy varations
* From sklearn import the metrics ad mean square error.
* Mean squared error (MSE): MSE is a measure of the average squared difference between the predicted and actual values.