**PROJECT NAME** – **VISIBILITY PREDICTION FOR AIR TRAFFIC CONTROL**

**PREPARED BY** – **YATRIK SHAH**



**DEPARTMENT OF ARTIFICIAL INTELLIGRNCE AND MACHINE LEARNING**

**SCHOOL OF EMERGING SCIENCE AND TECHNOLOGY**

**GUJARAT UNIVERSITY**

1. **INDEX**

* **Abstract**
* **Introduction**
* **Data Description**
* **Workflow Steps**
* **Methodology**
* **Real Life Applications**
* **Expected Outcome**
* **Deployment Link**
* **References**

1. **ABSTRACT**

In today’s world, the problem of human safety is increasing day by day due to absorption of light to reduce the clarity and colours of what you see. This affects the visibility in the environment, and it goes down. Many factors can cause visibility in the environment like haze, air pollution, climate change etc. Considering this, I have decided to make the machine to predict the level of visibility in the environment taking various features like temperature, wind speed, humidity etc. It becomes very important in cities like Delhi, where the air pollution remains at peak level, or the cities where there is high amount of haze which can cause the air traffic problem. My model will predict the level of visibility level in the environment so we can avoid these problems before it arises.

1. **INTRODUCTION**

Visibility is a measure of the distance at which an object or light can be clearly perceived. We predict the visibility of an object from a particular distance in kilometers on the basis of: Dry Bulb Temp, Wet Bulb Temp, Relative humidity and many more. Visibility is affected by certain environmental factors and may vary according to the direction and angle of view and the height of the observer.

1. **REAL LIFE APPLICATIONS**

* Key value in deciding if the flight would take off or land or not.
* Flight's Visibility is an important factor in all phases of flight, but especially when an aircraft is maneuvering on or close to ground, at that time weather has to be clear in order to avoid accidents.
* In addition, it has very important value in knowing distance from the land while docking a ship at a dockyard.
* Car Racing is not allowed if there isn't proper visibility.
* Useful in day-to-day life as well while driving vehicles.

1. **Data description**
2. VISIBILITY - Distance from which an object can be seen.
3. DRYBULBTEMPF-Dry bulb temperature (degrees Fahrenheit). Most commonly reported standard temperature.
4. WETBULBTEMPF-Wet bulb temperature (degrees Fahrenheit).
5. DewPointTempF-Dew point temperature (degrees Fahrenheit).
6. Relative Humidity-Relative humidity (percent).
7. Wind-Speed-Wind speed (miles per hour).
8. WindDirection-Wind direction from true north using compass directions.
9. StationPressure-Atmospheric pressure (inches of Mercury; or ‘in Hg’).
10. SeaLevelPressure- Sea level pressure (in Hg).
11. PreciP Total-precipitation in the past hour (in inches).
12. **WORKFLOW STEPS**

* Raw Data Collection
* Data Pre-processing
* Eda and Data Visualization
* Feature Engineering
* Logging
* Maintain Training Data in Database
* Cross Validation and Hyper Parameter Tuning
* Model Selection
* Prediction Database Maintaining
* Creating api
* Integrating with WEB FRAMEWORK - flask
* Model Deployment

1. **Tools used in the Project**

Logo

Description automatically generated**A picture containing text, clipart

Description automatically generatedA picture containing text, clipart

Description automatically generated**

**A picture containing text, clipart

Description automatically generated** **Text

Description automatically generated with medium confidence** A picture containing graphical user interface

Description automatically generated

**Icon

Description automatically generated A picture containing text

Description automatically generated**

1. **METHODOLOGY**

**8.1 DATA PRE-PROCESSING AND EDA**

* **Observation** – As a first step, the data was observed and all the features were taken into consideration.
* **Null Values** – In the dataset, it was checked whether there is the null values or not and if it is, how many records it contains.
* **Trends and Patterns** – Visualized some important graphs and charts to observe the considerable trends and some patterns to know the relationship between the features.

**8.2 FEATURE ENGINEERING**

* **Scaling** – Standardization was used as a method of scaling to bring down all the features to a common scale without distorting the differences in the range of the values.
* **Multi-collinearity** – Removed multi-collinearity by creating heatmap of correlation metrics.
* **Handling null values** – Null values was critically observed and it was removed from all the features in the dataset.

**8.3 LOGGING**

* Logging is very important in end-to-end projects.
* Maintained logs for saving and get informed each and every information, warnings, errors in any program.
* Logging is very useful specially in case of program crashing.
* It is an essential part of troubleshooting application and infrastructure performance.

**8.4 MAINTAINED TRAINING DATA IN DATABASE**

* After data cleaning, exploratory data analysis and feature engineering, dumping the final data for model training in the database.
* This data can directly be taken for model training purpose.
* MySQL database is maintained.

**8.5 CROSS VALIDATION AND HYPER PARAMETER TUNING**

* Cross validation and Hyperparameter tuning is one of the most important steps in Machine Learning, since the performance of the machine learning model is affected upon the hyperparameters those are selected for model training.
* Performed Cross validation and Hyperparameter tuning on all the models and selected the best models those performs best on the validation sets according to the given parameter grid.
* Found and selected the best hyperparameters of all the models for model training on Full data.

**8.6 CREATING API**

* A real-world Data Science based does not end just after successfully building a machine learning model that performs well on the test dataset.
* In real-world industrial end-to-end Artificial Intelligence or Data Science based projects creating Front-End API is also having equal importance as back-end tasks like- Model Training or Model Evaluation
* The Front-End is created with the help of HTML (Hypertext Markup Language) and CSS (Cascading Style Sheets), which are very powerful tools.
* Flask is used as WEB – API in python that integrates back-end and front-end.
* Flask depends on the Jinja template engine and the Werkzeug WSGI toolkit.

**8.7 MODEL DEPLOYMENT**

* Generally Model Deployment is considered as the last step in any Machine Learning project.
* After successfully creating front-end API with HTML and CSS and using Flask as web Framework only Model Deployment is left.
* Here HEROKU platform is chosen as Model Deployment.
* Heroku is a container-based cloud Platform as a Service (PaaS). Developers use Heroku to deploy, manage, and scale modern apps.