**Real-Time Weather Estimation and Forecasting Using Hybrid Machine Learning Approaches**

**ABSTRACT**

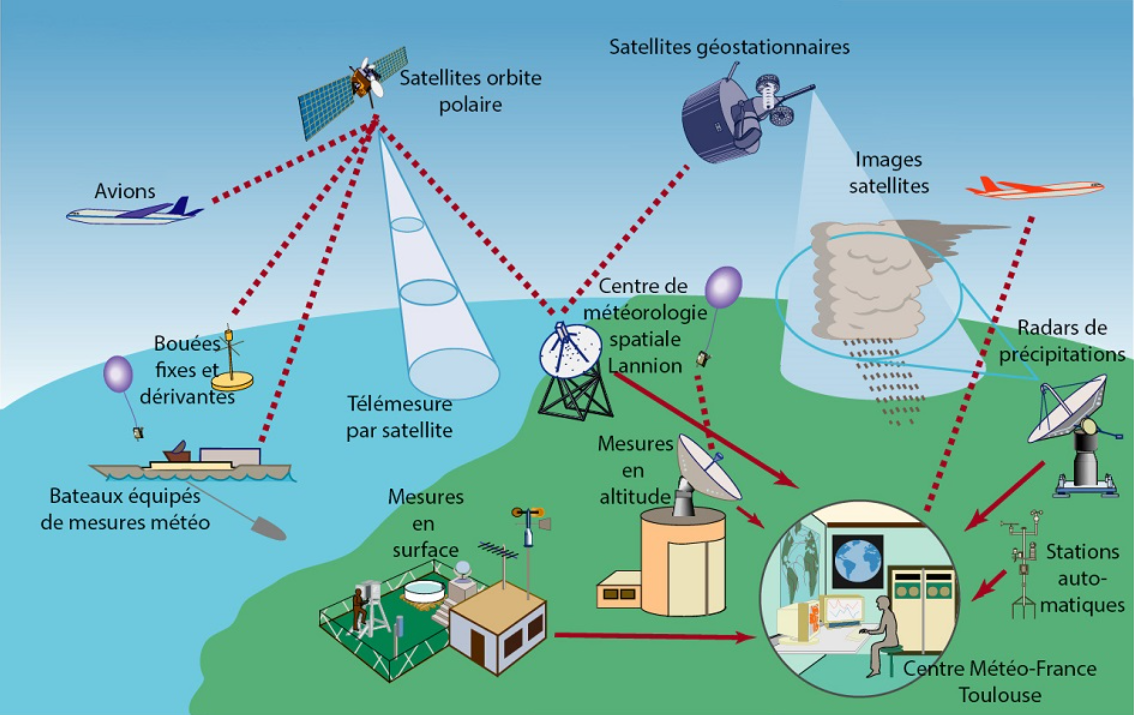
In real-time forecasting, estimation is the most crucial measurement for atmospheric observations by time and area. Present-day weather conditions observations are most important for many things, like agriculture, and different types of calamities that are unexpected in normal life. Machine learning techniques are used for weather forecasting. Among machine learning regression techniques, a key role is played in this problem. It monitors continuously from time to time. Temperature predictions also include weather forecasting by using regression techniques. Multiple regression techniques were used in our research. We measure the MSE rate for every technique. Multiple variety of metrics were tested. We consider short MSE to indicate for best forecasting outcome.

**Keywords:** Weather forecasting, Regression Techniques, Prediction, MSE

1. **INTRODUCTION**

Weather forecasting to forecasting the state of the atmosphere by location and time. But in the past, this procedure was measured manually based on pressure and weather conditions. New computer technologies were introduced to measure weather conditions based on multiple factors [1]. The linear relationship was established between the input to output attributes. It is different from nonlinear predictions of weather attribute relations. Generally, weather predictions are based on past trends, these past data are connected to present situations of the atmosphere. New technologies to predict weather situations analysis are most valuable for farmers, industry, and the general community for security [2]. From 2018 onwards, machine learning-based weather forecasting started. Different types of machine learning algorithms are used for weather forecasting. They predict the different types of results they predict for weather conditions in research, this research plays a vital role in different research domains [3].

Ensemble learning-based weather prediction provides primary knowledge for predicting better quality [4]. These technologies take less time and fewer possessions in a single effort. Emerging technologies provide better predictions for weather forecasting the world [5]. This procedure of training, testing, and authentication of data monitors endlessly predicts the precise outcome of weather situations. Accurate predictions are the most significant for the real-life [6].



**Figure 1:** operative weather conditions [5]

Figure 1 demonstrates the arrangement for the weather prediction investigation of information from the external atmosphere. Dissimilar arenas associated with weather forecasting effectively perform their work at appropriate times without any loss.

1. **PROPOSED SYSTEMS AND ARCHITECTURE STYLE**

Our proposed approach mainly focuses on regression techniques of machine learning for continuous monitoring of weather conditions [7]. It's based on numerical concepts for predictions. Multiple algorithms used to predict weather conditions refer to class issues [8]. This paper provides weather forecasting predictions and temperature predictions every time using machine learning techniques (Regression Techniques). Here MSE measures for every method. The succeeding figure 2 demonstrates the planned manner. Figure 2 shows the architecture of the proposed machine learning model.

Cataloguing (Classification)

Atmosphere

Weather Radars

Model Development

Weather States

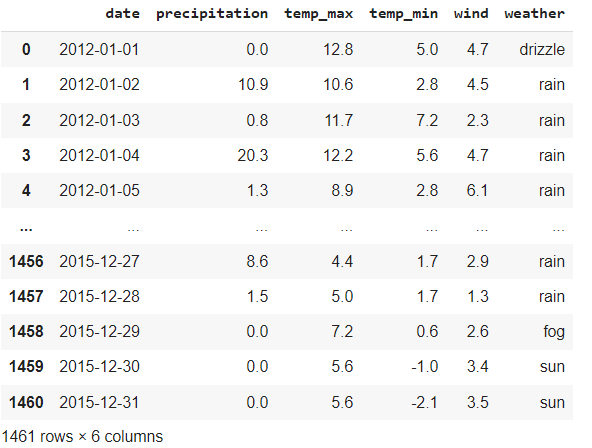
**Figure 2:** Architecture of projected system [9]

1. **STATISTICAL RESULTS AND INVESTIGATION**

## 3.1 Dataset Report

We focus on 1461 records in our dataset with six attributes. This is a sample dataset, but the actual dataset size is very high. The dataset was collected from the Kaggle dataset. Multiple weather forecasting datasets are available on Kaggle. Because different types of weather are available throughout the world.

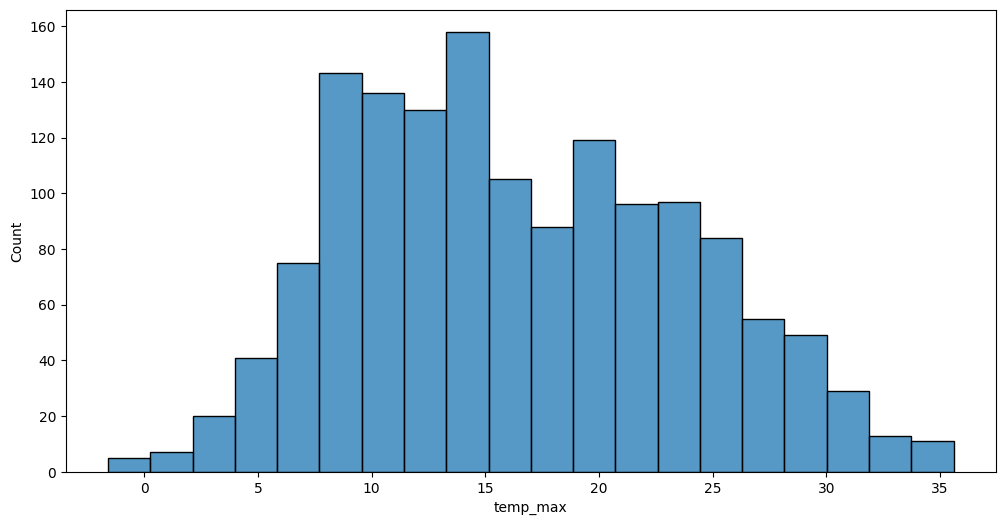
**Table 1:** Sample Dataset [10]



## 3.2 Data Preprocessing

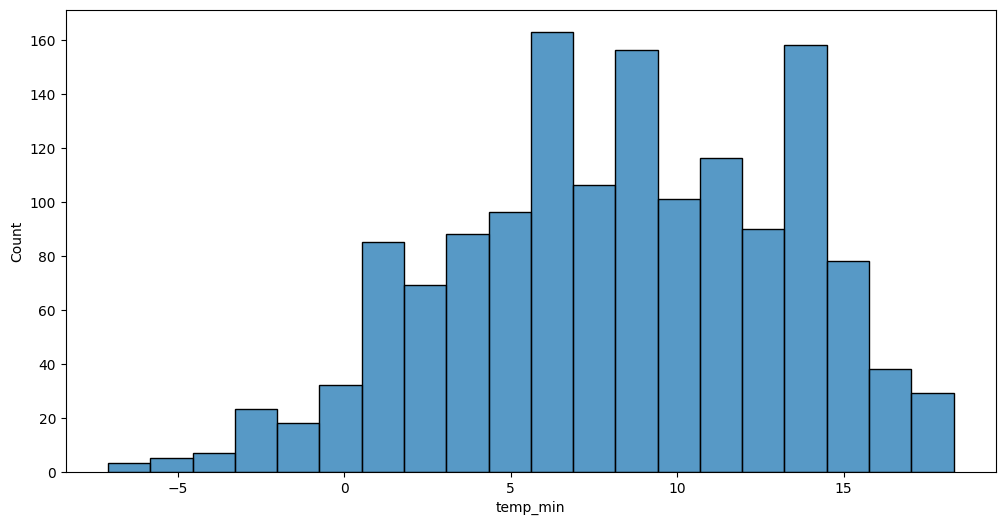
Data preprocessing to predict the pure data from raw data. This approach is crucial for predicting accurate results [11]. If not pre-processed, targeted results are not predicted. Global temperatures are generally warmer near the equator and cooler towards the poles, with significant variations influenced by factors like latitude, altitude, proximity to water, and prevailing winds.

**3.3 Data Picturing**



**Figure 3:** Data Picturing for temperature

## weather, the state of the atmosphere at a particular place during a short period of time. It involves such atmospheric phenomena as temperature, humidity, precipitation (type and amount), air pressure, wind, and cloud cover. The succeeding figures 3 and 4 demonstrate the information about temperature at the least and extreme. [12]



**Figure 4:** Data picturing the lowest infection rate of temperature

with 2023 being the warmest year on record, exceeding previous records like 2016 and 2020. The resulting Figure 5 epitomizes the extreme temperatures in every month from 2019 to 2023 [13].

### 

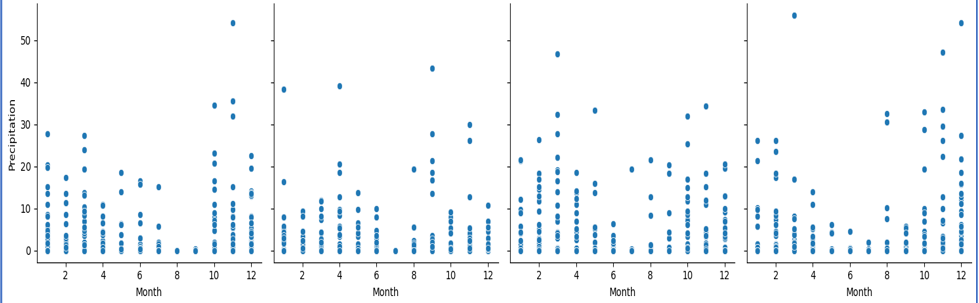
**Figure 5:** Data picturing each month in every year over temperature

Every month from June to December was record warm for the respective month. The resulting Figure 6 denotes the lowest temperature in every month in every year from 2019 to 2023.

## 

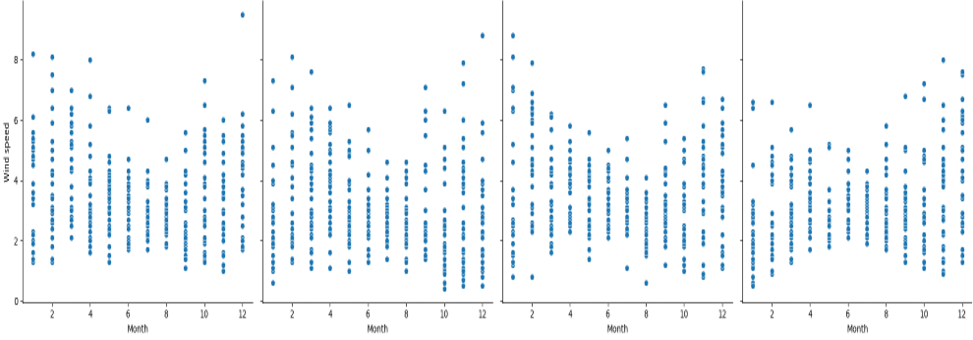
**Figure 6:** Data picturing min temperature

Every month of 2023 ranked among the warmest for that month, especially in the second half of the year. Figure 7 expresses the overall rain in a piece month in a separate year.

****

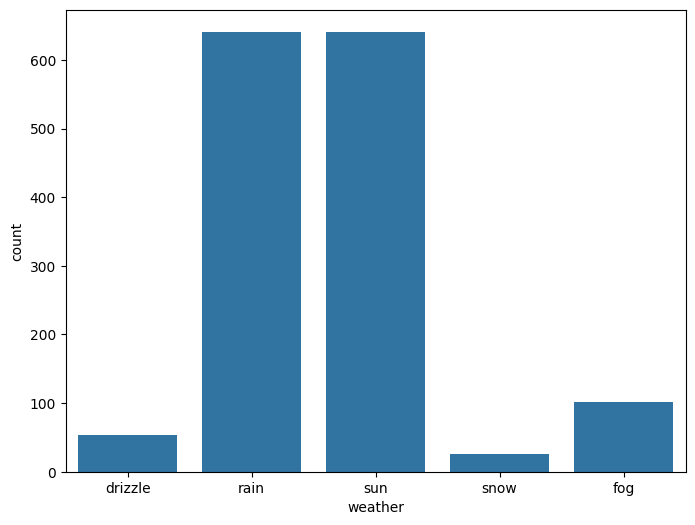
## Figure 7: Sleet in each month of year

## 



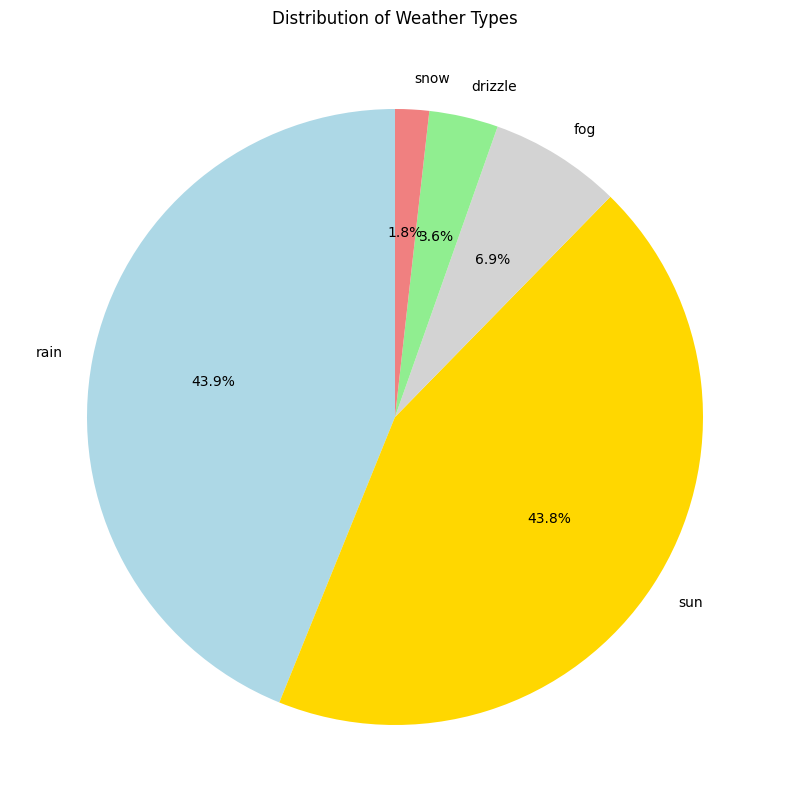
## Figure 8: Representation of (wind speed)

Figure 8 displays thedata imagining of dissimilar weather environments like fog, snow, rain, and sun. Figure 9illustrates thedata for the delivery of weather types in percentage.



**Figure 9:** Data picturing of dissimilar weather situations

Dissimilar weather situations can include a wide range of conditions, from sunny and clear to rainy, stormy, or snowy, and even extreme events like heat waves, cold snaps, and droughts. Deserts are expanding, while heat waves and wildfires are becoming more common. Amplified warming in the Arctic has contributed to thawing permafrost, retreat of glaciers and sea ice decline.

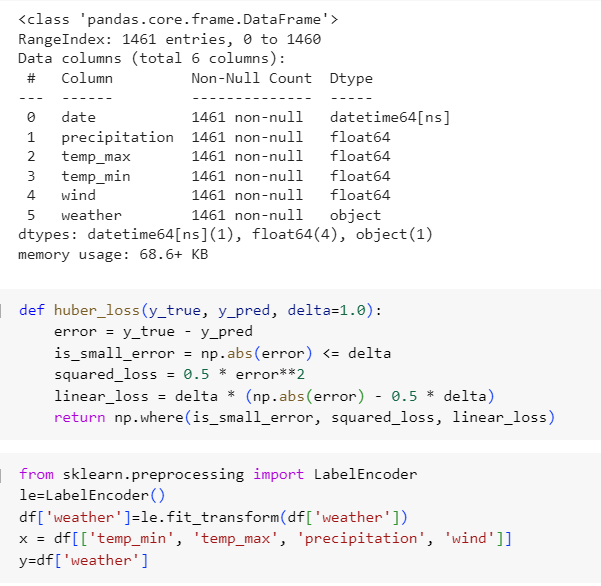


**Figure 10:** Data picturing of dissimilar weather situations

# 3.4 Regression Analysis of the Dataset

In machine learning, Regression analysis is used to study the relationships between dependent and independentvariables. Regression analysis is a statistical method used to model the relationship between a dependent variable and one or more independent variables, allowing for predictions and understanding of the strength and direction of these relationships.

# Fig. 11. Regression Analysis of the Dataset

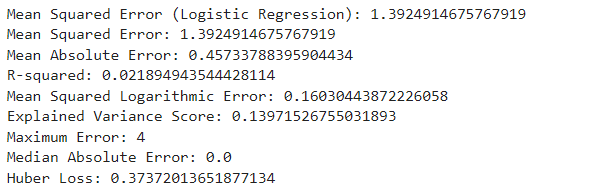


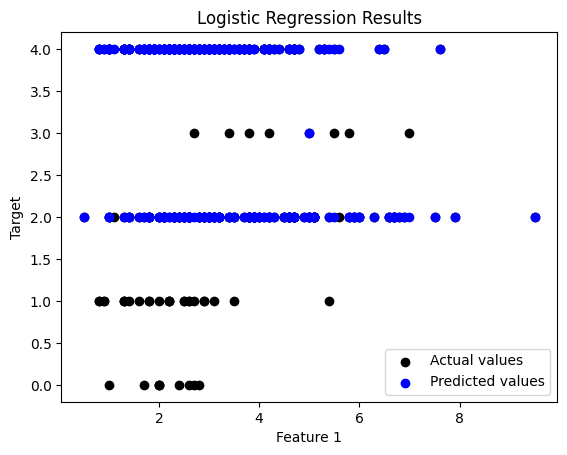
## 3.5 Ordinary Least Squares Regression

## The following method provides 1.07 as MSE, 0.83 as MAE, and 0.24 as MSLE. Max error is 3.28, MAE is 0.67, and Huber loss is 0.44.

## 3.6 Logistic Regression

MSE 1.40, MAE 0.46, MSLE 0.15, Max Error is 3.9, MAE 0, and Huber Loss 0.37.

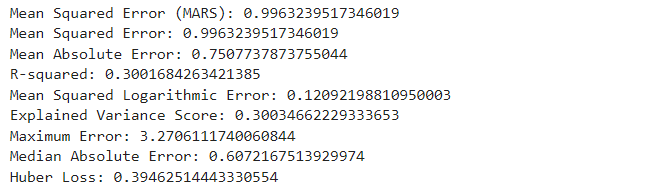




**Figure 12:** Actual vs Target Values over Logistic Regression

* 1. **Multivariate Adaptive Regression Splines (MARS)**

MSE 1.99, MAE 0.75, MSLE 0.11, Max Error 3.26, MAE 0.59 and Huber loss 0.4.



## 

**Figure 13:** MARS outcomes for actual vs target values

## 3.8 Locally Estimated Scatterplot Smoothing (LOESS)

## The MSE 1.25, MAE 0.98, MSLE 0.12, R squared 0.12, Variance 0.12, Maximum error 3.01, MAE 0.94, and Huber loss 0.54.

## 

## 

**Figure 14:** LOESS outcome for actual vs target values

LOESS is a non-parametric method for smoothing data by fitting simple regression models to localized subsets of data, resulting in a smooth curve that adapts to the underlying structure when comparing actual vs. target values. Among all models, Logistic Regression gets the lowest error rate.

**CONCLUSION**

Machine learning techniques are used for weather forecasting. Among machine learning regression techniques play a key role in this problem, it monitors continuously from time to time. Temperature predictions also include weather forecasting by using regression techniques. Multiple regression techniques were used in our research. We measure the MSE rate for every technique. Multiple variety of metrics were tested in this research. Logistic regression gets 0.45 for the MAE value. This is less compared to other models.

**Disclaimer (Artificial intelligence)**

Option 1:

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc) and text-to-image generators have been used during writing or editing of manuscripts.

Option 2:

Author(s) hereby declare that generative AI technologies such as Large Language Models, etc have been used during writing or editing of manuscripts. This explanation will include the name, version, model, and source of the generative AI technology and as well as all input prompts provided to the generative AI technology

Details of the AI usage are given below:

1.

2.

3.

**REFERENCES**

[1] Imran Maqsood, Muhammad Riaz Khan, and Ajith Abraham, “An ensemble of neural networks for weather forecasting”, Neural Computing & Application (2004) 13: 112–122.

[2] Amanpreet Kaur, JK Sharma, and Sunil Agrawal, “Artificial neural networks in forecasting maximum and minimum relative humidity”. International Journal of Computer Science and Network Security, 11(5):197-199, May 2011.

[3] https://rp5.ru/aRussianwebsitewithweatherdata of airport weather stations from all around the world

[4] https://colab.research.google.com/ Google’s free cloud service for AI developers for educational or research purposes.

[5] <https://www.encyclopedie-environnement.org/en/air-en/introduction-weather-forecasting/>

[6] Culclasure, Andrew, "Using Neural Networks to Provide Local Weather Forecasts" (2013). Electronic Theses and Dissertations 32. https://digitalcommons.georgiasouthern.edu/etd/32

[7] Trebing, K., Staǹczyk, T. and Mehr kanoon, S., 2021. Statuette: Precipitation nowcasting using a small attention-unet architecture. Pattern Recognition Letters, 145, pp.178-186.

[8] Mr. Sujan Reddy, Ms. Renu Sri, andSubhani Shaik,” Sentimental Analysis using Logistic Regression”, International Journal of Engineering Research and Applications (IJERA), Vol 11, Series-2, July 2021.

[9] Ms. Mamatha, Srinivasa Datta, and Subhani Shaik,” Fake Profile Identification using Machine Learning Algorithms”, International Journal of Engineering Research and Applications (IJERA), Vol 11, Series-2, July 2021.

[10] R. Vijaya Kumar Reddy, Subhani Shaik,B. Srinivasa Rao, “Machine learning based outlier detection for medical data” Indonesian Journal of Electrical Engineering and Computer Science, Vol. 24, No. 1, October 2021.

[11] [M. Jagan Chowhaan](https://journalajrcos.com/index.php/AJRCOS/article/view/339),  [D. Nitish](https://journalajrcos.com/index.php/AJRCOS/article/view/339),  [G. Akash](https://journalajrcos.com/index.php/AJRCOS/article/view/339), [Nelli Sreevidya](https://journalajrcos.com/index.php/AJRCOS/article/view/339), Dr. Subhani Shaik, “Machine Learning Approach for House Price Prediction” Asian Journal of Research in Computer Science, Volume 16, Issue 2, Page 54-61, June 2023.

[12] Neeraja, Anupam, Sriram, Subhani Shaik, and V. Kakulapati,” Fraud Detection of AD Clicks Using Machine Learning Techniques”, Journal of Scientific Research and Reports, Volume 29, Issue 7, Page 84-89, June 2023.

[13] P. Pranathi, V. Revathi, P. Varshitha, Subhani Shaik and Sunil Bhutada,” Logistic Regression Based Cyber Harassment Identification”, Journal of Advances in Mathematics and Computer Science, Volume 38, Issue 8, Page 76-85, June 2023.