EARLY WARNING FOR PRE AND POST FLOOD RISK MANAGEMENT

2021-124

Project Proposal Report

Vinobaji Selvaratnam

Project Proposal Report

BSc (Hons) in Information Technology Specializing in

Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

February 2021

EARLY WARNING FOR PRE AND POST FLOOD RISK MANAGEMENT

2021-124

Project Proposal Report

BSc (Hons) in Information Technology Specializing in

Information Technology

Department of Information Technology

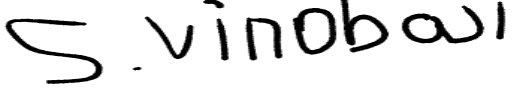
Sri Lanka Institute of Information Technology

Sri Lanka

February 2021

DECLARATION

I declare that this our own work and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of our knowledge and belief it does not have any material previously published or written by another person except where the acknowledgement is made in the text.

Name	Registration Number	Signature
Vinobaji. S	IT17181648	

The above candidates are carrying out research for the undergraduate Dissertation under my supervision.

Signature of the supervisor: …… Date……

ABSTRACT

Predicting rainfall is one of the most difficult aspects of weather forecasting. Accurate and timely rainfall forecasting can be extremely useful in preparing for ongoing construction projects, transportation activities, agricultural jobs, flight operations, and flood situations, and other aspects. The Sri Lankan meteorological department provides the necessary historical data for the prediction model.

This research-based on data sets that one normally used for rainfall prediction and how to improve the accuracy by suitable small data set change. When using each dataset, this system is intended to verify the different accuracy of rainfall by using a machine learning algorithm. The primary goal of this model's design is to find out best the accuracy dataset, based on data sets used for the training model. This system predicts rainfall and rainfall volume (mm) by using linear regression. The rainfall prediction data will be used in the main flood prediction model.

Keywords – **Machine Learning Algorithm, Rainfall Prediction, Linear Regression, Flood Prediction Model.**

Table content

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviations	Description
ML	Machine Learning
LR	Linear Regression
R	Rainfall Prediction
FPM	Flood Prediction Model

1 INTRODUCTION

1.1 Background and Literature survey

Rainfall is one of the natural events which one happens in the water cycle. Rainfall prediction is one the most common prediction in metrology. Four main factors are most commonly used to predict rainfall, Such as

1. Temperature
2. Relative humidity
3. Pressure
4. Wind speed

There are many methods used to predict rainfall. Using machine learning (ML) is also one of the methods to predict rainfall. There are a lot of algorithms that can be used to predict rainfall in machine learning.

Even there is a lot of methods to predict rainfall, but not enough accurate methods to predict rainfall volume. There are a lot of exciting researches there to done to increase the accuracy of rainfall prediction by changing machine learning models and algorithms and some researches using a lot of factors to predict rainfall. In machine learning better to use minimum factors and get the best prediction. Here needed to avoid many factors because it will be complicated and costly when using many factors in the machine learning model.

So better to use the main four rainfall factors as usual and find other two or three replacement factors that can able to cover other small factors which are causing rainfall, when looking for replacement factors, a better solution as seven bases unites related factors. So, we can consider here location and time (day, month) as replacement factors.

The research paper “Rainfall Prediction for Udaipur, Rajasthan using Machine Learning Models Based on Temperature, Vapour Pressure, and Relative Humidity,” [1] by Jitendra Shreemali and others provides a thorough examination of machine learning-based approaches to rainfall forecasting. This paper provides information on current rainfall prediction techniques, state-of-the-art techniques in various rainfall prediction applications, and a collection of frequently used weather datasets (temperature, vapor pressure, and relative humidity) in previous studies. This paper also highlights information on the studies that were carried out to evaluate and compare the performance of the weather datasets used in rainfall prediction. Because this is a survey paper, they look at the accuracy of various machine learning

methods for rainfall prediction and identify some future work that needs to be done with those methods and datasets (geographic locations and datasets longer than a decade suggested for more accuracy). Some of the current techniques' difficulties include the algorithm's running time being too long due to the rise in datasets. This paper presents a survey of different rainfall prediction methods using machine learning and deep learning. Also, the limitations of these methods are outlined in this paper.

The paper "Weather Forecasts Based on Rainfall Prediction Using Machine Learning Methodologies," [2] by *Adalya J.* Analyzes the different accuracy of each machine learning model from existing research papers. They test some models such as Logistic regression (LR), Decision Trees (DCT), Random Forest (RF), Support Vector Machine (SVM). For this rainfall forecast, the Random Forest model is both competent and reasonable. The data used as feedback for prediction and classification has a significant impact on the percentage of accuracy and prediction. All models have advantages and disadvantages, and the most difficult part is deciding which model is the best. After reviewing all of the supervised learning models listed above, they found that the Random Forest (RF) classification model has a high level of accuracy and acceptance for their weather dataset. Implementing a hybrid prediction model, in which several machine learning models are assembled to work, may improve the accuracy of the model's forecast. They intend to use the hybrid prediction model in our future work to achieve greater and greater accuracy.

The paper "Rainfall Prediction using Machine Learning and Neural Network," [3] *by* K. Dutta and P. Gouthaman. This project aims to use machine learning and neural networks to predict rainfall. The project conducts a comparative analysis of machine learning and neural network methods and then depicts the most effective method for rainfall prediction. They use a machine learning method for forecasting rainfall is discussed. It uses two kinds of errors, RE and RMSE, to determine the precision of the machine learning strategy. There are two methods for forecasting rainfall. The first is a machine-learning strategy that uses LASSO regression. The neural network technique is the second. Two methods are being compared to find an efficient way to predict rainfall. The LASSO regression technique is used in the first one., and the second is an approach based on artificial neural networks. The future improvement of this project will be a method of reducing the proportion of errors that are present. Also, one of the most significant improvements would be a reduction in the ratio of train data to test data.

The paper “Rainfall Prediction Using Machine Learning Techniques and an Analysis of the Outcomes of These Techniques,” [4] by S. B and J. K.S. This research has been working to increase the accuracy of rainfall prediction by optimizing and incorporating Machine Learning approaches. In this case, supervised learning and unsupervised learning are the two basic types of Machine Learning methods that are used to construct predictive models, supervised learning algorithms are used. The classification algorithms ANN, Logistic Regression, Nave Bayes, and Random Forest being tested and compared. When compared to other classification algorithms, Random Forest delivers the best rainfall forecast results with an accuracy of 87.76 percent and also has the highest values in Recall and F-Measure. They proposed that the model's prediction accuracy could be improved by creating a hybrid prediction model that combines multiple machine learning algorithms.

1.2 Research Gap

When we compared some exiting research to the proposed project

Rainfall prediction using Machine learning(researches)	**Data used**	**Models**
"Rainfall Prediction for Udaipur, Rajasthan using Machine Learning Models Based on Temperature, Vapour Pressure, and Relative Humidity," [4]	Temperature, Vapour pressure, Relative Humidity	Regression, Generalized regression, Linear –AS, LSVM, Random tree, Linear, XGBoost Linear, Tree-AS, Random Forest, Neural Network
S. B and J. K.S, "Rainfall Prediction Using Machine Learning Techniques and an Analysis of the Outcomes of These Techniques," [2]	Temperature, humidity, Sea Level Pressure, Windy	Artificial neural network, Logistic Regression, Naïve Bayes, Random forest
K. Dutta and P. Gouthaman, "Rainfall Prediction using Machine Learning and Neural Network," [1]	Temperature, humidity, Wind speed, Sunshine Duration	Neural network, LASSO Regression
"Weather Forecasts Based on Rainfall Prediction Using Machine Learning Methodologies," *Adalya J.*, [3]	Min Temperature, Max Temperature, Rainfall, Sunshine, Evaporation & Pressure, Humidity, Cloud, Windspeed in the different period	Decision Tree, SVM, LR, RF
Proposed Research	Temperature, Humidity, Pressure, Wind speed, Location, Time (day, month)	Linear Regression

Table 1: Research Gap

1.3 Research Problem

Each natural event causing by many factors. The rainfall is also caused by many factors, but when we come to consider rainfall prediction using any type of method include machine learning, we have some limitations - cost and time. So, we can not use lot factors to predict rainfall because if we added many factors to prediction, we need to gather big data and it will become complicated and costly. But on other hand, we need to ensure the accuracy of the rainfall prediction. So better we select time replacement factors which one can cover wide range small factors. Time (day, month) and location are natural factors that cover a wide range of small factors. If we with time and location we can cover a lot of small factors such as monsoon, sunshine duration, sun direction, deforestation, height from sea level, etc.

This proposed research is based on only check secondary weather factor contribution in causing rainfall by using location and time (day, month). End of the research we can conclude small factor contribution in rainfall.

2 OBJECTIVES

2.1 Main Objective

To find out effective data set for predict rainfall based on machine learning and identify which dataset contributes to predicting flooding.

2.2 Specific Objectives

1. Analysis of weather historical data (temperature, relative humidity, pressure, wind speed) and predict rainfall.

2. Analysis of weather historical data (temperature, relative humidity, pressure, wind speed) based on Location and time (day, month) predict rainfall.

3. Checking accuracy different between each model.

3 METHODOLOGY

This research mainly considers two main data sets for predict rainfall. Such as

Data set 1

1. Temperature
2. Relative humidity
3. Sea level pressure
4. Wind speed

Data set 2

1. Dataset 1
2. Time (Day, Month)
3. Location(place)

Data set one is the main factors that are causing rainfall and that is already using in exiting research. Data set two data set one and two replacement factors time and locationally. This weather historical data for the past 10 years will be collected from the metrology department of Sri Lanka.

This the overview of the system diagram

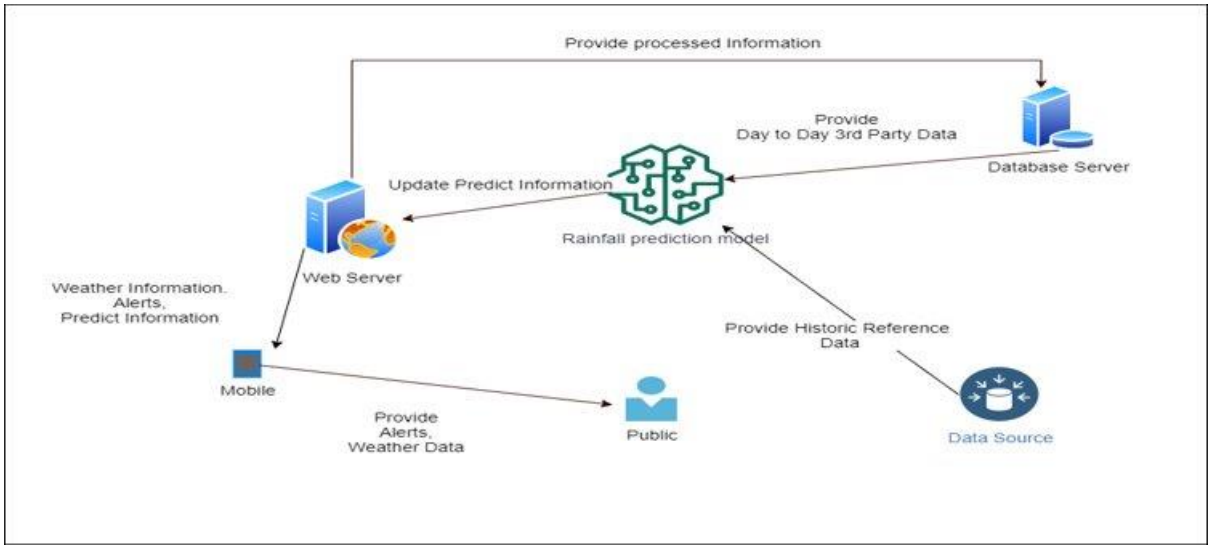


Figure 1 : System Overview Diagram

In this system, the machine learning method Linear regressing will be used to predict rainfall and rainfall amount (mm) by using both data sets and checking accuracy differences. After the observation of the prediction accuracy can make assumptions and test it with needed evaluation in the training model again and again. We can make some conclusions based on the accuracy difference of each data set. At end of the research find out the best dataset for the prediction model and predict the rainfall and rainfall amount.

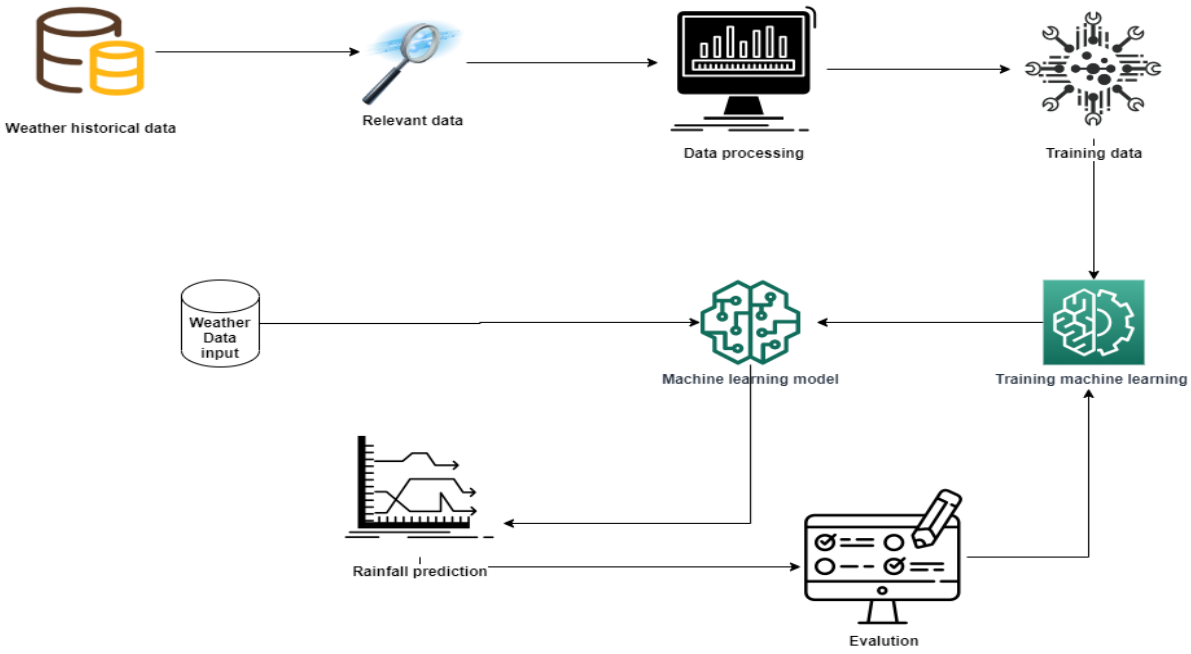


Figure 2 : Process Steps

3.1 Technologies

- Minimum system requirement for development – Windows XP, Linux
- Language – Java, python
- IDE – Eclipse 4.2.1, PyCharm
- Simulator - Android Virtual Devices (AVDs) that run on the Android Emulator.
- Framework - Flask

4 Requirement

4.1 Functional Requirement.

- Accuracy of rainfall prediction
- Provide rainfall prediction through the mobile application.

4.2 Non-Functional Requirement.

- Usability- This focuses on the appearance of the user interface and how people interact with it to achieve required goals effectively and efficiently.
- Performance – Good performance for maintaining a good user experience
- User satisfaction

4.3 Gantt chart

Task	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov
1. Project Initiation												
- Group Discussion												
- Topic Selection												
- Project Assessment Form												
2. Feasibility Study												
- Charter Form												
- Resource Evolution												
- Project Proposal Document												
- Proposal Presentation												
3. Project Planning												
- System Planning												
- Assembling Necessary data												
- Selecting tech & tools												
- Research Paper												
- Evaluate Design												
4. Implementation												
- Data Extraction												
- Classification and recognition												
- Experimental Analysis												
5. Testing												
- Component Testing												
- Integration Testing												
- System Testing												

6. Finalization												
- Final Presentation												
- Final VIVA session												
- Final Report												

Table 2:Gantt chart

4.4 Work Breakdown Structure

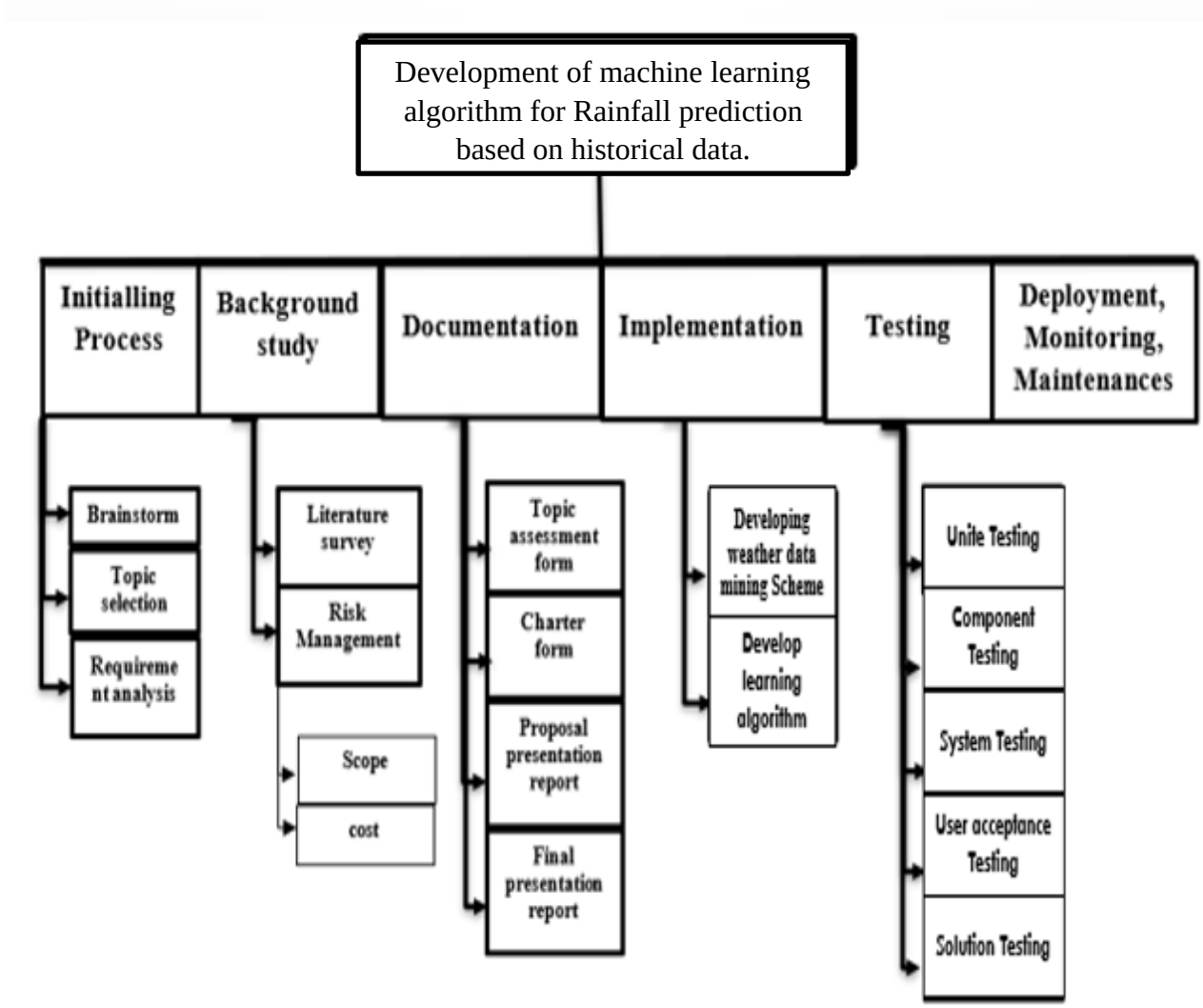


Figure 3 : Work Breakdown Structure

4.5 Budget

Component	Amount (LKR)
Historical data of weather (last 10 years from 3 different district)	35550.00
Database	3900.00
Other	1500.00
Total	40,950.00

Table 3: Budget

REFERENCES

[1] "Rainfall Prediction for Udaipur, Rajasthan using Machine Learning Models Based on Temperature, Vapour Pressure and Relative Humidity," *Int. J. Recent Technol. Eng.*, vol. 8, no. 6S, pp. 133–137, 2020, doi: 10.35940/ijrte.f1024.0386s20.

[2] "Weather Forecasts Based on Rainfall Prediction Using Machine Learning Methodologies," *Adalya J.*, vol. 9, no. 6, 2020, doi: 10.37896/aj9.6/009.

[3] K. Dutta and P. Gouthaman, "Rainfall Prediction using Machine Learning and Neural Network," *Int. J. Recent Technol. Eng.*, vol. 9, no. 1, pp. 1954–1961, 2020, doi: 10.35940/ijrte.a2747.059120.

[4] S. B and J. K.S, "Rainfall Prediction Using Machine Learning Techniques and an Analysis of the Outcomes of These Techniques," *Int. J. Eng. Appl. Sci. Technol.*, vol. 04, no. 09, pp. 365–371, 2020, doi: 10.33564/ijeast.2020.v04i09.047.

[5] S. Zainudin, D. S. Jasim, and A. A. Bakar, "Comparative analysis of data mining techniques for malaysian rainfall prediction," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 6, no. 6, pp. 1148–1153, 2016, doi: 10.18517/ijaseit.6.6.1487.

[6] D. Tien Bui, T. C. Ho, B. Pradhan, B. T. Pham, V. H. Nhu, and I. Revhaug, "GIS-based modeling of rainfall-induced landslides using data mining-based functional trees classifier with AdaBoost, Bagging, and MultiBoost ensemble frameworks," *Environ. Earth Sci.*, vol. 75, no. 14, p. 910, 2016, doi: 10.1007/s12665-016-5919-4.

[7] M. S. Sreekanth, R. Rajesh, and J. Satheeshkumar, "Extreme learning machine for the classification of rainfall and thunderstorm," *Journal of Applied Sciences*, vol. 15, no. 1. pp. 153–156, 2015, doi: 10.3923/jas.2015.153.156.

[8] J. Refonaa, M. Lakshmi, A. C. S. Kiran, and A. R. Teja, "Rainfall prediction using Apriori Algorithm," *Int. J. Recent Technol. Eng.*, vol. 8, no. 2 Special Issue 3, pp. 593–596, 2019, doi: 10.35940/ijrte.B1109.0782S319.

[9] J. Refonaa, M. Lakshmi, R. Abbas, and M. Raziullha, "Rainfall prediction using regression model," *Int. J. Recent Technol. Eng.*, vol. 8, no. 2 Special issue 3, pp. 543–546, 2019, doi: 10.35940/ijrte.B1098.0782S319.

[10] S. Poornima and M. Pushpalatha, "Prediction of rainfall using intensified LSTM based recurrent Neural Network with Weighted Linear Units," *Atmosphere (Basel).*, vol. 10, no. 11, 2019, doi: 10.3390/atmos10110668.

[11] M. Kühnlein, T. Appelhans, B. Thies, and T. Nauss, "Improving the accuracy of rainfall rates from optical satellite sensors with machine learning - A random forests-based approach applied to MSG SEVIRI," *Remote Sens. Environ.*, vol. 141, pp. 129–143, 2014, doi: 10.1016/j.rse.2013.10.026.

[12] Indrabayu, S. Aditama, A. A. Prayogi, S. Pallu, A. Achmad, and I. S. Areni, "Spatial-temporal approach for predicting rainfall in tropical country," *ICIC Express Lett.*, vol. 13, no. 2, pp. 113–118, 2019, doi: 10.24507/icicel.13.02.113.

[13] S. Cramer, M. Kampouridis, A. A. Freitas, and A. K. Alexandridis, "Kent Academic Repository Full text document (pdf) Enquiries Citation for published version Link to record in KAR An extensive evaluation of seven machine learning methods for rainfall prediction in weather derivatives," *Expert Syst. Appl.*, vol. 85, pp. 1–27, 2017.

[14] I. Cholissodin and S. Sutrisno, "Prediction of Rainfall using Simplified Deep Learning based Extreme Learning Machines," *J. Inf. Technol. Comput. Sci.*, vol. 3, no. 2, p. 120, 2018, doi: 10.25126/jitecs.20183258.

[15] S. Aftab, M. Ahmad, N. Hameed, M. S. Bashir, I. Ali, and Z. Nawaz, "Rainfall prediction in Lahore City using data mining techniques," *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 4, pp. 254–260, 2018, doi: 10.14569/IJACSA.2018.090439.