**Rainfall Prediction using Multiple Linear Regressions Model**

Meteorological scientists always try to find means to understand the atmosphere o f the Earth, and to develop accurate weather prediction models. Several methods have been used in weather prediction. Recently, machine learning methods are assumed to be accurate techniques and have been widely used as an alternative to classical methods for weather prediction. The rainfall rate is one of the essential phenomena in the weather system, which has a direct influence on the agriculture and biological sectors. This paper aims to develop a multiple linear regression model in order to predict the rate of precipitation (PRCP), i.e., rainfall rate, for Khartoum state. It is based on some weather parameters, such as temperature, wind speed, and dew point. The data used in this research has been provided from the website o f the National Climatic Data Center. A Python code using the Pytorch library has been written to develop the model, which applies Artificial Neural Networks. The efficiency of the model has been measured by comparing the average value of the mean square error of the training data with the test data. The obtained results show that the average of the mean square error has been improved by 85% during test time, when the same amount o f data is used during the training and test phases. However, it drops to 59% when the amount o f data at the test phase exceed the amount of training phase data .

**EXISTING SYSTEM:**

In the Existing system used back propagation neural network for rainfall prediction. This model used by Xianggen Gan and he was tested using the dataset from 1970 to 2000 which has 16 meteorological parameters. During network training the target error is set as 0.01 and learning rate is set as 0.01. This model implemented on mat lab neural network. ,Several methods are used to generate weather forecasting, each of which differ in its accuracy and efficiency. There are three important steps that must precede the process of weather forecasting, which are to collect atmospheric data as much as possible, to understand the data and its inter-relation to determine the behavior of the atmosphere, and to use it in numerical models to predict the future state of the atmosphere.Genetic Programming (GP) and MCRP were compared on 21 different datasets of cities across Europe. Daily rainfall data for 10 years were taken as training data and one year rainfall data were taken as testing data.

**DISADVANTAGES OF EXISTING SYSTEM:**

* The disadvantage of MCRP is that it predicts accurate only for annual rainfall when compared with monthly rainfall prediction.
* The assumptions which are made by the multiple linear regression are: linear relationship between the both the descriptive and independent variables, the highly correlated variables are independent variables, yi is calculated randomly.
* Weather is extremely difficult to forecast correctly.
* It is expensive to monitor-so many variables from so many sources.
* The computers needed to perform the millions of calculations necessary are expensive.
* **Algorithm**: Markerbased motion Capture (MoCap) Framework.

**PROPOSED SYSTEM:**

Applying machine learning techniques in weather forecasting can compensate complex meteorological physics model. With the availability of metrological data set, the two authors were encouraged to select supervised learning method, which is multiple linear regression, instead of unsupervised learning or reinforcement learning [1]. There are different regression types used in machine learning, such as linear regression, logistic, polynomial regression. The simpler and most frequent method is linear regression, which is used for prediction [4].The aim of this paper is to develop a multiple linear regression model to predict the rainfall rate in Khartoum state, which depends on many variables. The remainder of this paper is organized as follows. Section II provides a brief survey about related work, The process of data cleaning was done manually using an Excel program. It was done in four steps, which understanding the data set and the correlation between variables, deleting unwanted factors from the data set, dealing with missing data and outliers, and treating data to facilitate handling

**ADVANTAGES OF PROPOSED SYSTEM:**

* Linear regression is one type of the supervised learning techniques to predict a numeric value (dependent variable) form a set of features (predictors). Likewise, it is about finding a function that maps inputs . It forms a prediction by computing a weighted sum of the input features.
* **Algorithm:** Linear Regression, Machine Learning. Artificial Neural Networks.

**SYSTEM REQUIREMENTS:**

**HARDWARE REQUIREMENTS:**

* System : Intel Core i3.
* Hard Disk : 1TB.
* Monitor : 15’’ LED
* Input Devices : Keyboard, Mouse
* Ram : 8GB.

**SOFTWARE REQUIREMENTS:**

* Operating system : Windows 10.
* Coding Language : Python
* Tool : PyCharm, Visual Studio Code
* Database : SQLite

**REFERENCE:**

**Date of Conference:**26 February 2021 - 01 March 2021

**Date Added to IEEE *Xplore*:**17 May 2021

**ISBN Information:**

**Electronic ISBN:**978-1-7281-9111-9

**Print on Demand(PoD) ISBN:**978-1-7281-9112-6

**INSPEC Accession Number:**20613422

**DOI:**[10.1109/ICCCEEE49695.2021.9429650](https://doi.org/10.1109/ICCCEEE49695.2021.9429650)

**Publisher:**IEEE

**Conference Location:**Khartoum, Sudan