**Abstract**

The burgeoning research in the fields of Artificial Intelligence and machine learning has given rise to numerous weather prediction models. But the problem of accurately predicting or forecasting the weather still persists. Numerical weather prediction is taking the existing numerical data on weather conditions and applying machine learning algorithms on it to forecast the weather. This paper is the application of machine learning algorithms, linear regression model from statistics, and two optimization Algorithms, Decision Tree and Linear Regression to predict the weather on the basis of few parameters. Two optimization algorithm have been used to compare the performance of the algorithms.

**Keywords**

—Weather prediction, machine learning, efficiency , Linear Regression, Decision Tree, dataset,, temperature, humidity ,dew-point, error.

**INTRODUCTION**

Weather forecasting means predicting the weather conditions (conditions of atmosphere) of a particular given area or location. More importantly, accurate weather prediction is very important to pursue day-to-day activities. Living and non-living things are dependent on weather predictions. Even after decades of weather forecasting, the weather industry in India is still in its initial stage, facing many obstacles. One of the major obstacles that weather forecasting faces is the arbitrary & ill-suited expectations from the nature. Machine learning is the ability of computer to learn without being explicitly programmed. It allows machines to find hidden patterns and insights. In supervised learning, we build a model based on labeled training data. The model is then used for mapping new examples. So, based on the observed weather patterns from the past, a model can be built and used to predict the weather. Several data mining techniques have been employed in diversified applications such as predicting rainfall, weather, storms and flood. Weather forecasting falls under predictive mining which focuses on the data analysis, formulates the database, and forecasts the features of anonymous data. [1] This research work focuses on solving the weather prediction anomalies and in-efficiency based on Linear regression algorithms . The major contribution of this research work is to formulate an efficient weather prediction model based on the Linear regression algorithms.

**PROPOSED SYSTEM:**

The proposed model will use linear regression, which will predict the high and low temperatures as a linear combination of all the features. Linear regression

does not use weather classification data of each day because this algorithm cannot be used with classification data. Therefore initially in our project only eight parameters

are selected for use which are maximum temperature, minimum temperature, mean humidity, and mean atmospheric.

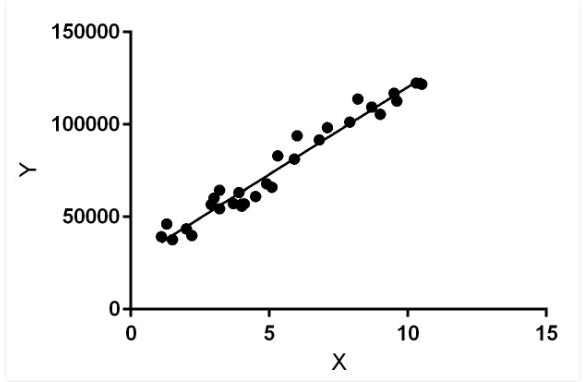
**ALGORITHM:**

**DECISION TREE**

A tree has many analogies in real life, and turns out that it has influenced a wide area of **machine learning**, covering both **classification and regression.** In decision analysis, a decision tree can be used to visually and explicitly represent decisions and decision making. As the name goes, it uses a tree-like model of decisions. Though a commonly used tool in data mining for deriving a strategy to reach a particular goal, its also widely used in machine learning.

**LINEAR REGRESSION**

**Linear Regression** is a machine learning algorithm based on **supervised learning**. It performs a **regression task**. Regression models a target prediction value based on independent variables. It is mostly used for finding out the relationship between variables and forecasting. Different regression models differ based on – the kind of relationship between dependent and independent variables, they are considering and the number of independent variables being used.



Linear regression performs the task to predict a dependent variable value (y) based on a given independent variable (x). So, this regression technique finds out a linear relationship between x (input) and y(output). Hence, the name is Linear Regression.  
In the figure above, X (input) is the work experience and Y (output) is the salary of a person. The regression line is the best fit line for our model.

**SYSTEM REQUIREMENTS**

**SOFTWARE REQUIREMENTS:**

* OS : Windows
* Python IDE : python 2.7.x and above

: Pycharm IDE

* setup tools and pip to be installed for 3.6.x and above

**HARDWARE REQUIREMENTS:**

* RAM : 4GB and Higher
* Processor : Intel i3 and above
* Hard Disk : 500GB: Minimum
* Dt,svm,lr,knn