Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 4, April : 2023

EFFECTIVE FINE-GRAINED WEATHER FORECASTING MODEL USING MACHINE LEARNING

Dr.B.ASHA LATHA, Associate Professor, Department of CSE, S.R.K INSTITUTE OF TECHNOLOGY, Andhra Pradesh., India.

B . JYOTHI, K.SAI SAMPATH, P. SANDEEP, K. MANOJ, Student, Department of CSE, S.R.K INSTITUTE OF TECHNOLOGY, Andhra Pradesh., India.

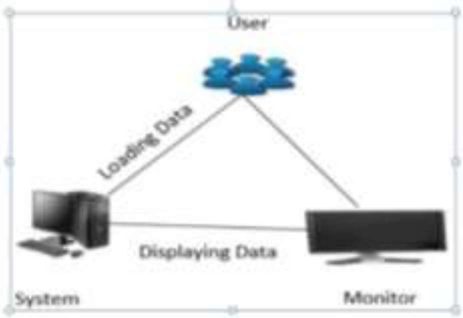
ABSTRACT: We propose this system which uses Machine Learning Algorithms like SVR, Random Forests, Decision Tree, Ridge, Linear Regression and ARIMA for Weather Prediction. Weather forecasting is the application of science and technology to predict the conditions of the atmosphere for a given location and time. It is well-known that numerical weather prediction (NWP) models require considerable computer power to solve complex mathematical equations to obtain a forecast based on current weather conditions. In this project, we propose a novel light weight data-driven weather forecasting model by exploring temporal modelling approaches of More specifically Standard Linear Regression (SR), Ridge Regression (RR), Support Vector Regression (SVR), and Random Forest Regressor (RF), Decision Tree Regressor (DT) are implemented as the classical machine learning approaches, and Autoregressive Integrated Moving Average (ARIMA) is implemented as the statistical forecasting approaches.

INTRODUCTION

Weather forecasting and Prediction is the process of predicting the state of the atmosphere based on the temperature values and specific time and locations. Numerical weather prediction (NWP) utilizes computer algorithms to provide a forecast based on current weather conditions.

Machine learning models and time-series forecasting are used for predicting and seeing the data according to the time and atmosphere parameters and it will build the models on its superior performance.

Weather forecasts are made by collecting as much data as possible about the current state of the atmosphere (particularly the temperature, humidity, and wind) and using understanding of atmospheric processes (through meteorology) to determine how the atmosphere evolves in the future.



PROPOSED SYSTEM

In existing system, Observational data collected by Doppler radar, Radiosonde, weather satellites, buoys and other instruments are fed into computerized NWS numerical forecast models. The models use equations, along with past weather data, to provide forecast guidance to our meteorologists.

DISADVANTAGES:

- The primary problem with Numerical Weather Prediction (NWP) models is it takes a long time toproduce its results.
- Forecasts are Never Completely Accurate.
- Forecasts are never 100% and it is almost impossible to predict the future with certainty.
- Problems concern availability, timeliness, and quality of observational data.

We propose this system which uses Machine Learning Algorithms like SVR, Random Forests, Decision Tree, Ridge, Linear Regression and ARIMA for Weather Prediction. These techniques can determine the possibility of weather condition. Due to limitation in to get live data we done forecasting on the old data.

ADVANTAGES:

- Reduce weather-related losses and enhance societal benefits.
- Protection of life and property.
- Public health and safety.
- Avoid forest fires.

ALGORITHMS

RANDOM FOREST ALGORITHM

Random Forest is a popular machine learningalgorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML.

Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset.

The greater number of trees in the forest leads to higher accuracy and prevents the problem of overfitting.

DECISION TREE ALGORITHM

Decision Tree is a Supervised learning technique that can be used for both classification and Regression problems, but mostly it is preferred for solving Classification problems

It is a tree-structured classifier, where internal nodes represent the features of a dataset, branches represent the decision rules and each leaf node represents the outcome.The decisions or the test are performed on the basis of features of the given dataset.

LITERATURE SURVEY

[1] **Basak D, Pal S, Patranabis D (2007) "Support vector regression." Neural Inf Process Lett Rev 11(10):203–224**

Instead of minimizing the observed training error, Support Vector Regression (SVR) attempts to minimize the generalization error bound so as to achieve generalized performance. The idea of SVR is based on the computation of a linear regression function in a high dimensional feature space where the input data are mapped via a nonlinear function. SVR has been applied in various fields – time series and financial (noisy and risky) prediction, approximation of complex engineering analyses, convex quadratic programming and choices of loss functions, etc. In this paper, an attempt has been made to review the existing theory, methods, recent developments, and scopes of SVR.

[2] **Kavitha S, Varuna S, Ramya R (2016) "A comparative analysis on linear regression and support vector regression." In: 2016 online international conference on green engineering and technologies (IC-GET), November 2016, pp 1–5.**

In business, consumer's interest, behavior, product profits are the insights required to predict the future of business with the current data or historical data. These insights can be generated with the statistical techniques for the purpose of forecasting. The statistical techniques can be evaluated for the predictive model based on the requirements of the data. The prediction and forecasting are done widely with time series data. Most of the applications such as weather forecasting, finance and stock market combine historical data with the current streaming data for better accuracy. However the time series data is analyzed with regression models.

[3]. Sanchez-Fernandez M, de-Prado-Cumplido M, Arenas-Garcia J, Perez-Cruz F (2004) "SVM multi regression for nonlinear channel estimation in multiple-input multiple-output systems." IEEE Trans Signal Process 52(8):2298–2307.

This paper addresses the problem of multiple-input multiple-output (MIMO) frequency nonselective channel estimation. We develop a new method for multiple variable regression estimation based on Support Vector Machines (SVMs): astate-of-the-art technique within the machine learning community for regression estimation. We show how this new method, which we call M-SVR, can be efficiently applied. The proposed regression method is evaluated in a MIMO system under a channel estimation scenario, showing its benefits in comparison to previous proposals when nonlinearities are present in either the transmitter or the receiver sides of the MIMO system.

[4]. Sharaf A, Roy SR (2018) "Comparative analysis of temperature prediction using regression methods and back propagation neural network." In: 2018 2nd international conference on trends in electronics and informatics (ICOEI), May 2018, pp 739–742.

Weather forecasting is very important for our day-to-day life as the prediction of climatic parameters like temperature, humidity, rainfall, etc. is important for agriculture, forestry, commercial companies etc. It is known that the climatic conditions are irregular in nature and may be often unpredictable. In previous years most of the work was done on finding a linear equation to predict climatic parameters. It is now known that climatic conditions are nonlinear in nature and thus Artificial Neural Network is now considered to be an efficient nonlinear method for weather prediction. In this paper a comparative analysis has been done between regression methods and nonlinear method like artificial neural network to analyze the difference between the performance of linear and nonlinear models for prediction of temperature.

IMPLEMENTATION

1. System:

Store Dataset:

The System stores the dataset given by the user.

Model Training:

The system takes the data from the user and fed that data to the selected model.

Model Predictions:

The system takes the data given by the user and predict the output based on the givendata.

Graph:

For the visualization purpose and forecasting using ARIMA we have used the graphical representation.

2. User:

Load Dataset:

The user can load the dataset he/she want to work on.

View Dataset:

The User can view the preprocessed train test split dataset.

Selection model :

User can apply the model to the dataset for accuracy/performance.

View Result:

User can evaluate the model performance and predict the desired output.

SAMPLE RESULTS

CONCLUSION

In this project, Weather forecasting/ prediction model we will implement it by next semester by using Support Vector Regressor, Random Forest, Simple Linear

Regression, Ridge Regression and Decision Tree with the MAE, MSE and R2 Score. ARIMA is used to

Forecast the data of past 10 years and through graph we are analysing the result, LSTM isto predict next 30 days data using past data and a flask-based web application is developed to predict the weather. We will implement it by next semester.

REFERENCES

[1]M. Wiston, M. Km

Weather forecasting: from the early weather wizards to modern-day weather predictions

[2]J. Han, M. Kamber, J. Pei

Data mining: concepts and techniques

J. Han, M. Kamber, J. Pei (Eds.), Data mining: concepts and techniques, Morgan Kaufmann, San Francisco, CA, itd (2012),

[3]D. Siuta, G. West, H. Modzelewski, R. Schigas, A.R. Stull

Viability of cloud computing for real-time numerical weather prediction

Weather
Forecast, 31 (6) (2016), 10.1175/WAF-D-16-0075.1

View PDF

[4]D.K. Krishnappa, D. Irwin, E. Lyons, M. Zi nk

CloudCast: cloud com-puting for short-term mobile weather forecasts

[5]J.T. de
Castro, G.M. Salistre, Y.C. Byun, B.D. Gerard o

Flash flood prediction model based on multiple regression analysis for decision support system

[6]S. Kothapalli, S.G. Totad

A real-time weather forecasting and analysis

IEEE international conference on power, control, signals and instrumentation engineering (ICPCSI), vol. 2017 (2017), pp. 1567-