VISVESVARAYA TECHNOLOGICAL UNIVERSITY

Jnana Sangama, Machhe, Belagavi, Karnataka 590018

A
Project Synopsis
on

IMPACT OF TEMPERATURE VARIATIONS OVER THE BAY OF BENGAL ON THE CLIMATE OF EASTERN COAST OF INDIA.

Submitted in partial fulfillment of the requirement
for the award of the degree of

Bachelor of Engineering
in

Information Science & Engineering

by
Bhavana (1BG14IS008)
H Sudhanva (1BG14IS017)
Sai Navaneeth V (1BG14IS041)
Satish Kumar M S (1BG14IS045)

Under the Guidance of

Dr. Saritha Chakrasali
Professor
Dept. of ISE
B.N.M.I.T

Vidyaya Amrutham Ashnuthe

B.N.M. Institute of Technology

12th Main, 27th Cross, Banashankari II Stage, Bangalore 560 070.
Department of Information Science and Engineering
2017-2018

Vidyaya Amrutham Ashnuthe

CERTIFICATE

Certified that the synopsis of project work entitled **"Impact of temperature variations over the Bay of Bengal on the climate of Eastern coast of India"** is carried out by Ms. **Bhavana (1BG14IS008)**, Mr. **H Sudhanva (1BG14IS017),** Mr. **Sai Navaneeth V (1BG14IS041)**, **Mr. Satish Kumar M S (1BG14IS045)**, the bonafide students of **B.N.M Institute of Technology** in partial fulfillment for the award of **Bachelor of Engineering** in **Information Science & Engineering** of the **Visvesvaraya Technological University**, Belagavi during the year 2017-2018. It is certified that all corrections / suggestions indicated for Internal Assessment have been incorporated in the report deposited in the departmental library. The project synopsis has been approved as it satisfies the academic requirements in respect of Project work prescribed for the said Degree.

Dr. Saritha Chakrasali
Professor, Dept. of ISE
BNMIT

CONTENTS

Chapter 1

ABSTRACT OF THE PROJECT

The increasing 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. Weather forecasting has traditionally been done by physical models of the atmosphere, which is unstable to perturbations, and thus is inaccurate for large periods of time. Weather is a continuous, data-intensive, multidimensional, dynamic and chaotic process and these properties make weather prediction a big challenge. Machine Learning techniques are more robust to perturbations, it would be more ideal to explore their applications in the field of weather forecasting and to potentially generate more accurate forecasts of weather for large periods of time. This project attempts to predict the changes in the temperature of cities in the coastal region using Machine Learning algorithms, by analyzing the statistical climate data of the cities in eastern coast of India such as Chennai, Visakhapatnam etc., along with the corresponding sea water temperature of Bay of Bengal. This work aim at studying the relationship and establishing a pattern between the climatic changes in the land and its associated sea water temperature using Data Mining techniques and Machine Learning algorithms.

Chapter 2

INTRODUCTION

Artificial intelligence (AI) traditionally refers to an artificial creation of human-like intelligence that can learn, reason, plan, perceive, or process natural language. Artificial intelligence is further defined as "narrow AI" or "general AI". Narrow AI, which we interact with today, is designed to perform specific tasks within a domain (e.g. language translation). General AI is hypothetical and not domain specific, but can learn and perform tasks anywhere. This project focuses on advances in narrow AI, particularly on the development of new algorithms and models referred to as *machine learning*. Computer algorithms organize enormous amounts of data into information and services, based on certain instructions and rules. It's an important concept to understand, because in machine learning, learning algorithms – not computer programmers – create the rules. Instead of programming the computer every step of the way, this approach gives the computer instructions that allow it to learn from data without new step-by-step instructions by the programmer.

Data analytics refers to qualitative and quantitative techniques and processes used to enhance productivity and business gain. Data is extracted and categorized to identify and analyze behavioral data and patterns, and techniques vary according to organizational requirements

Weather simply refers to the condition of air on earth at a given place and time. The application of science and technology are to predict the state of the atmosphere in future time for a given location which is important due to its effectiveness in human life. Today, weather forecasts are made by collecting quantitative data about the current state of the atmosphere and using scientific understanding of atmospheric processes to project how the atmosphere will evolve. The chaotic nature of the atmosphere implies the need of massive computational power required to solve the equations that describe the atmospheric conditions.

Various papers related to weather forecasting using Machine Learning algorithms and other data mining concepts were studied. The relevant papers have been discussed below.

2.1. LITERATURE SURVEY

The related papers for this project are discussed below.

- Mark Holmstrom, Dylan Liu, Christopher Vo, in their paper "Machine Learning Applied to Weather Forecasting" [1] explore the applications of machine learning techniques to weather forecasting and potentially generate more accurate weather forecasts for large period of time. This paper predicts the maximum temperature and the minimum temperature for seven days, given weather data for the past two days of a particular region using linear regression model. The linear regression model implemented in this paper is the base algorithm used in this project.
- Siddharth S. Bhatkhande, Roopa G. Hubballi in their paper "Weather Prediction Based on Decision Tree Algorithm Using Data Mining Techniques" [2] investigate the use of data mining techniques in forecasting attributes like maximum temperature, minimum temperature, which is carried out using Decision Tree algorithms and meteorological data collected between 2012 and 2015 from the different cities. On available datasets, they apply the Decision Tree Algorithm for deleting the inappropriate data. The concepts of data mining implemented here will be used as reference in this project.

Further, the other related papers are discussed below:

- C. Johansson, M. Bergkvist, O. De Somer, D. Geysen, N. Lavesson and D. Vanhoudt in their paper, "Operational demand forecasting in district heating systems using ensembles of online machine learning algorithms" [7], state that Heat demand forecasting is in one form or another an integrated part of most optimization solutions for district heating and cooling (DHC). This paper presents the current status and results from extensive work in the development, implementation and operational service of online machine learning algorithms for demand forecasting.
- Aastha Sharma, Setu Chaturvedi and Bhupesh Gour in their paper, "A Semi-Supervised Technique for Weather Condition Prediction using DBSCAN and KNN" [5], propose a semi-supervised weather prediction technique to

validate the predictions done for certain atmospheric parameters taken for four years on a day wise basis in a certain city.

- Aditya Grover, Ashish Kapoor and Eric Horvitz in their paper, "A Deep Hybrid Model for Weather Forecasting" [3], state that weather forecasting is a canonical predictive challenge that has depended primarily on model-based methods. They study specifically the power of making predictions via a hybrid approach that combines discriminatively trained predictive models with a deep neural network that models the joint statistics of a set of weather-related variables.

- John K. Williams and D. A. Ahijevych, C. J. Kessinger, T. R. Saxen, M. Steiner and S. Dettling in their paper, "A machine learning approach to finding weather regimes and skillful predictor combinations for short-term storm forecasting" [4], it is shown that the random forest machine learning approach provides a tool for identifying a set of skillful predictors for thunderstorm initiation as well as providing a performance benchmark.

- Mrs. C. Beulah Christalin Latha, Dr. (Mrs.) Sujni Paul, Dr.E.Kirubakaran and Mr. Sathyanarayanan in their paper, "A Service Oriented Architecture for Weather Forecasting Using Data Mining" [6], treat weather as a continuous, data-intensive, multidimensional,dynamic process, that makes weather forecasting a formidable challenge.Their paper proposes a novel method to develop a service oriented architecture for a weather information system and forecast weather using data mining techniques.

- Daniel Bejarano and Adriano Quiroga in their project, "Wind Prediction: Physical model improvement through support vector regression" [8], concentrates on wind speed prediction through the combination of support vector regression and the weather research and forecast model was explored.

- Emilcy Hern´andez, Victor Sanchez-Anguix, Vicente Julian, Javier Palanca, and N´estor Duque in their paper, "Rainfall prediction: A Deep Learning approach" [9], they introduce an architecture based on Deep Learning for the prediction of the accumulated daily precipitation for the next day. More specifically, it includes an autoencoder for reducing and capturing non-linear relationships between attributes, and a multilayer perceptron for the prediction task.

- Kiran Kumar. R and Usha Rani. R in their paper, "Weather Prediction through Machine Learning" [10], design an effective rainfall prediction agent model using support vector machine and multiple linear regressions.

2.2 LIMITATIONS OF EXISTING SYSTEM

The following limitations were observed after the literature survey:
- No correlation between climatic condition of the land and sea surface temperature is being established.
- The linear regression model is inherently a high variance model and thus, would be unstable to outliers.
- The functional regression model is high bias and traditionally requires a large dataset, the weather data of only past two days is insufficient to capture Any trends in the weather.

2.3 MOTIVATION

The motivation behind this project is to be able to recognize and establish a pattern between sea and land climate conditions upon considering a large data set for a specific location. Thus, being able to forecast weather more efficiently while improving the process of doing so.

2.4 PROBLEM STATEMENT

The project aims to achieve the following objectives:
- To analyze related weather data using data mining techniques.
- To obtain correlation between sea surface temperature to corresponding coastal temperature.
- Implementing supervised machine learning algorithm on the available dataset.
- To forecast the weather conditions of the respective city.

2.5 PROPOSED SYSTEM

Data collection allows us to gather records of measurements that are already conducted by data mining techniques. The data mining stage is divided into three phases. At each phase the algorithms are used to analyze the available meteorological datasets. The testing method adopted for this research is percentage split, that train on a percentage of the dataset, cross validate on it and test on the remaining percentage. Thereafter interesting patterns representing knowledge will be identified. With the identified patterns, we apply the supervised machine learning techniques such as linear regression and functional regression to build a Machine Learning model as shown in Fig 1. This Machine Learning model will be able to predict climatic changes which affect the temperature of Chennai city in coastal region due to corresponding change in Bay of Bengal temperature.
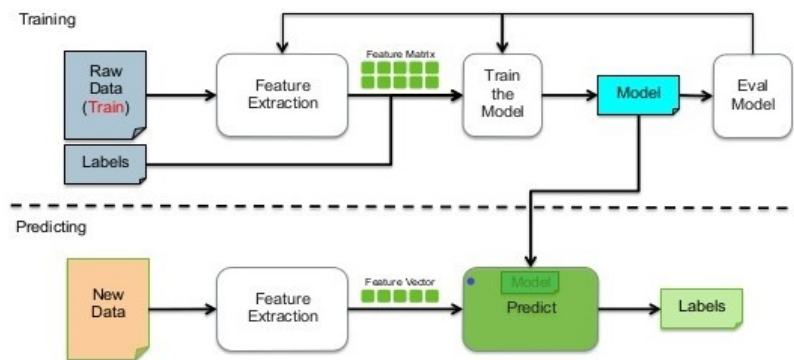

Fig 1: Supervised Learning Workflow retrieved from [13]

The dataset that will be used for this work is obtained from the following sources:
- Land temperature [11]
- Sea surface temperature [12]

2.6 SYSTEM REQUIREMENTS

Hardware requirements:

- **System:**   Intel core 7th Gen i7 Processor.
- **Hard Disk:** 150Gb Solid State Drive
- **RAM:**   8Gb recommended
- **Monitor:**   15 VGA Colour
- **GPU:**   Nvidia GT 630M 1Gb VRAM

Software requirements:

- **O/S:**   Windows 7 and above++
- **Language:**   python, R and MATLAB
- **Additional Packages:**   python scikit, python anaconda, tensorflow

REFERENCES

[1]   Mark Holmstrom, Dylan Liu, Christopher Vo. Machine Learning Applied to Weather Forecasting. Stanford University. 2016.

[2]   Siddharth S. Bhatkhande, Roopa G. Hubballi. Weather Prediction Based on Decision Tree Algorithm Using Data Mining Techniques. *IJARCCE* (2016).

[3]   Aditya Grover, Ashish Kapoor and Eric Horvitz. A Deep Hybrid Model for Weather Forecasting.

[4]   John K. Williams and D. A. Ahijevych, C. J. Kessinger, T. R. Saxen, M. Steiner and S. Dettling. A machine learning approach to finding weather regimes and skillful predictor combinations for short-term storm forecasting. National Center for Atmospheric Research, Boulder, Colorado.

[5]   Aastha Sharma, Setu Chaturvedi and Bhupesh Gour. A Semi- Supervised Technique for Weather Condition Prediction using DBSCAN and KNN.International Journal of Computer Applications (0975 – 8887) Volume 95– No. 10, June 2014.

[6]   Mrs. C. Beulah Christalin Latha, Dr. (Mrs.) Sujni Paul, Dr.E.Kirubakaran and Mr. Sathianarayanan. A Service Oriented Architecture for Weather Forecasting Using Data Mining.Int. J. of Advanced Networking and Applications. Volume: 02, Issue:02, Pages:608-613. 2010.

[7]   C. Johansson, M. Bergkvist, O. De Somer, D. Geysen , N. Lavesson and D. Vanhoudt. Operational demand forecasting in district heating systems using ensembles of online machine learning algorithms.The 15th International Symposium on District Heating and Cooling September 4-7, 2016, Seoul, Republic of Korea (South Korea).

[8]   Daniel Bejarano and Adriano Quiroga. Wind Prediction: Physical model improvement through support vector regression.Stanford University. December 2013.

[9]   Emilcy Hern´andez Victor Sanchez-Anguix Vicente Julian Javier Palanca and N´estor Duque.Rainfall prediction: A Deep Learning approach.

[10]   Kiran Kumar. R 1, Usha Rani. R.Weather Prediction through Machine Learning.Dept of Computer science, Krishna University, Machilipatnam. AP, India.

[11]   Wunderground.com, 'Weather forecast & Reports – Long Range and Local'. [Online]. Available: http://www.wunderground.com.

[12]   Seatemperature.info, 'Sea water temperature. Water surface temperatures on the coast of the oceans and seas'. [Online]. Available: https://seatemperature.info

[13]   En.profit.me, 'Blog about Linux, Python, Vim and healthy lifestyle'. [Online]. Available: http://en.proft.me/2015/12/24/types-machine-learning-algorithms