

RAINFALL PREDICTION USING MACHINE LEARNING

¹B.Gopi, ²Mr. K Subhash Chandra

¹Student, Dept. of Master of Computer Applications, Amrita Sai Institute of Science and Technology, Paritala, Andhra Pradesh, 521180, India.

²Asst.Prof, Dept. of Computer Science & Engineering, Amrita Sai Institute of Science and Technology, Paritala, Andhra Pradesh, 521180, India.

ABSTRACT

The ability to forecast rainfall is very important for farmers, as India's economy relies on successful crops and rain. If rainfall is accurately predicted, planning how to use water and grow crops becomes much easier. The research aims to establish ways to use mining techniques and science to predict the chances of rainfall. In predicting rainfall, researchers use data mining algorithms which include Random Forest, K-Nearest Neighbours, Logistic Regression, Support Vector Machines (SVM) and Decision Trees. The research investigates the algorithms to choose the one most accurate for forecasting rainfall.

Keywords: *Rainfall Prediction, Data Mining, Crop Productivity, Water Resources, Random Forest, K-Nearest Neighbors, Logistic Regression.*

1. INTRODUCTION

Rapid urbanization and the increasing demand for parking spaces has made it vital to develop smarter parking solutions that can handle scarcity of parking resources effectively. An Intelligent Parking Management System (IPMS) plans to solve this problem by exploring a smart method for parking space management and usage, therefore reducing traffic congestion and providing better user experience in parking facilities. Therefore, this paper makes an effort to present a smart system for parking management which is automatically connected to a wide range of parking spaces, taking information from the surroundings in order to optimize traffic flow and control parking space through machine learning, which not only allows to know in real, the parking availability, but also to be able to estimate when you will have a greater or lesser probability of finding available an empty parking.

By analyzing traffic flow and parking demand trends, the system uses data like vehicle entry and exit times, sensor readings and parking usage history. Predictions of future parking availability can be made utilizing machine learning models, and forecast of the popular places to park can also be made. It is this predictive quality that allows parking spaces to be best distributed, cutting down on time spent looking for places to park, environmental impact and ensuring a more fluid urban mobility.

The methodology employed in this system is to use algorithms which are In statistics, the methods include Random Forest (a group of decision trees working together to create a final prediction), Decision Tree (a decision-support tool using a tree-like model for several consequences and outcomes) and Support Vector Machine (SVM) (a machine learning model that separates groups of data by making them line up on either side of a boundary).These algorithms cooperate and hence, enhance the efficiency and performance of the parking management system.

1.1 Overview

Knowing the weather can give important details needed for running farms, managing water supplies and preventing disasters. The project centres on creating algorithms that use data from past weather records to predict if it will rain. The predictions are made using supervised learning models including Random Forest, SVM and Logistic Regression.

Through its use on the web, this project allows farmers and planners to get the rainfall information they need in real time.

1.2 Project Objectives

My focus is on creating a model that can predict rainfall and can be updated over time to replace the usual ways of doing so. The objective will be to achieve accurate predictions by examining the results of several supervised algorithms.

Necessity:

This model of forecasting rainfall is necessary for dealing with problems involving crops and understanding the weather in agricultural areas. It makes it easy for users with its friendly interface and is reliable in several circumstances which helps improve efficiency. This application is created to help users save time, as it delivers data quickly and accurately.

Software Development Method:

Software is being developed for the project using the Waterfall Model. Because this method is systematic, you complete every phase of planning and development before moving on to the next. It also applies test cases and case software strategies to support the project's reliability and overall performance.

1.3 Project Scope

This project focuses on looking into the impact of various supervised binary classification methods on predicting rainfall. The model is assessed by using precision, sensitivity and the F1-score. Some steps involved in the project are:

- **Data Acquisition:** Finding the proper data about the weather.
- This step includes organising and preparing the dataset to be looked at.
- By using visualisation, it is possible to better examine and explore the data.
- Using several machine learning algorithms on the data to make predictions.
- I then generate a pickle file for the model that performed best.
- Using Flask, you can make a website where users can access and view the final results.

2. LITERATURE SURVEY

Many studies have adopted ML and statistics to boost the accuracy of predicting rainfall. Scientists discovered that Decision Trees, Support Vector Machines (SVM) and Random Forests can be successfully used for forecasting weather. Besides, integrating various sets of data and applying ensemble models has proven to increase how accurate the predictions are. Besides, using a combination of hybrid approaches and deep learning has proven effective in processing nonlinear atmospheric data. This allows machine learning to solve issues related to climate change more accurately and efficiently.

For many years, meteorologists have found it vital to forecast rainfall, as this area has important results in farming, water planning and getting prepared for emergencies. Initially, using methods such as linear regression and time series, experts tried to estimate the weather, though these tools failed to account for the erratic shifts occurring in temperatures. Because of this, recent work has used Random Forest, SVM, Decision Trees and ANNs, among other techniques, to help machines make more accurate predictions. It is important to note that researchers Rani et al. (2019) proved that machine learning performed better than usual approaches using data sets available on IMD and Kaggle. Deep learning is being investigated for its impact on prediction, but they need a lot of computing resources. Researchers generally agree that good pre-processing, feature selection and model tuning help ensure a model is successful.

3. SYSTEM ANALYSIS

3.1 Existing System

The main role of agriculture in India's economy is that it depends on farmers waiting for the monsoon to plant crops. To produce crops effectively, you must ensure the soil is fertile, use suitable fertilisers and ensure the climate is favourable. Since sudden changes in weather can greatly affect farmers, weather forecasting is very important for them. It is a difficult task to accurately predict the weather. This paper aims to predict the weather using techniques such as classification, clustering, decision trees and neural networks. For proper predictions, experts in meteorology use data on rainfall, wind speed, temperature and humidity.

Disadvantages:

Long-term planning is not helpful with the current system since it cannot give accurate forecasts.

3.2 Proposed System

The system is meant to combine different datasets from different places to form a complete one. Once the data is collected, machine learning models analyse it and make reliable predictions.

DETAILS

The proposal details the plan, necessary steps and resources that will be used to complete the project. It describes the procedures needed for the project, the software needed and the possible advantages and threats. All necessary actions are included in this project, starting from the development phase until it is tested and any dangers or dependencies are noted.

Mission:

On the web, applications using machine learning are gaining popularity as people look for these tools. As rainfall prediction helps farmers, they can prepare water devices ahead, consider crop output and prevent water waste. Using this method, experts can forecast rainfall quickly and accurately.

Goal:

The objective is to design a machine learning model that can forecast rainfall.

Getting the data in the right format.

The 145,460 records in the dataset come from Kaggle and are classified into two categories.

- ⇒ Rainy day
- ⇒ Sunny day

4. FEASIBILITY STUDY

4.1 Economical Feasibility

Since the proposed system relies on open-source software like Python, Flask, NumPy and scikit-learn, it is highly cost-effective. Since you can work with open-source tools instead, the start-up price is also significantly lower. No complicated hardware is needed; the system can easily be handled on a mid-range PC that supports recent processing standards. Since the system is made up of smaller parts, changes and updates typically won't require replacing everything. Both Heroku and AWS's free tiers can be used to help minimise costs in deploying on the cloud. This system is sustainable and accessible to many because it saves money for the government and provides durable solutions for agriculture, government planning and the education sector.

4.2 Operational Feasibility

The system is easy to use for practical purposes. Even those with no prior training can access GIS because it is easy to use. Basic weather information can be entered by users, who receive clear predictions of rain shortly afterward to help with farming and watering plants. The system works equally well under any circumstances, be it on local computers or in cloud settings. Because the backend is reliable and well-designed, it runs smoothly and requires very little upkeep. This allows users to accept these products and use them easily in many areas and communities each day.

4.3 Technical Feasibility

Its technical aspects are solid, because its technologies are reputable, well-known and easy to adjust. The Python platform is very flexible and it offers machine learning functions through scikit-learn. Because the system requires only very little, Flask supports easy scaling and gives an easy way to set up web hosting. The Emacs system can be used on Windows, Linux and Mac operating systems. Additionally, it can be connected to external APIs or sensors in the future, giving more chances to improve the system. With various resources and help from others, it is not too challenging to use these technologies which in turn prevents the system from facing issues and allows it to remain stable and work well even in limited circumstances.

5. STRUCTURE OF THE PROJECT

The system for predicting rainfall is arranged so that it remains easy to read, manage and upgrade. There are several stages in the project and every stage has a function and purpose.

In the beginning, data is collected and prepared by dealing with missing information, changing category data into numbers and adjusting the data to a standard scale. They confirm that the data is of good quality and can be further studied for better results.

The following stage is to carry out exploratory data analysis and visualise data. Tools such as Matplotlib and Seaborn are used to cheque the data’s distribution and note any connexions between various features. At this stage, valuable information is gained that directs the process for building the next model.

Once the data insights are gathered, the project proceeds to model development, during which many machine learning algorithms are used. After defining the model, Logistic Regression, K-Nearest Neighbours, Decision Tree, Random Forest and Support Vector Machine (SVM) algorithms are applied and their results are evaluated. Some of the measurements used to assess how the model works are accuracy, precision, recall and F1-score. The most effective model is selected to be put into use.

After serialising the model with Pickle, it is placed into a Flask-based web application. The user interface lets you provide data on the weather and get predictions for rainfall in binary format. HTML collects data with the forms and then Python processes the data through the chosen machine learning algorithm.

Since each module does its own job, the entire system can be easily checked for errors, tested and modified. The preparedness of this platform means new data or features can be conveniently added in future versions.

6. SYSTEM REQUIREMENTS

6.1 Hardware Requirements

Component	Specification
Processor	Intel
RAM	16 GB
Hard Disk	512 GB

TABLE1: HARDWARE REQUIREMENTS

6.2 Software Requirements

Component	Specification
Operating System	Windows
Framework	Flask
Technology	Machine Learning with Python
Web Browsers	Chrome, Microsoft Edge
Code Editors	Visual Studio Code, Google Colab, Anaconda, Jupyter Notebook

TABLE2: SOFTWARE REQUIREMENTS

6.3 UML Diagram

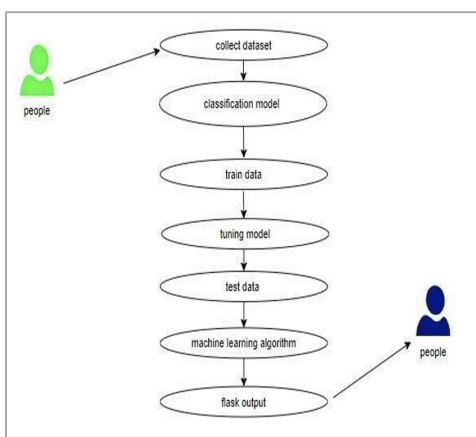


Fig1: Uml diagram

6.4 Class Diagram

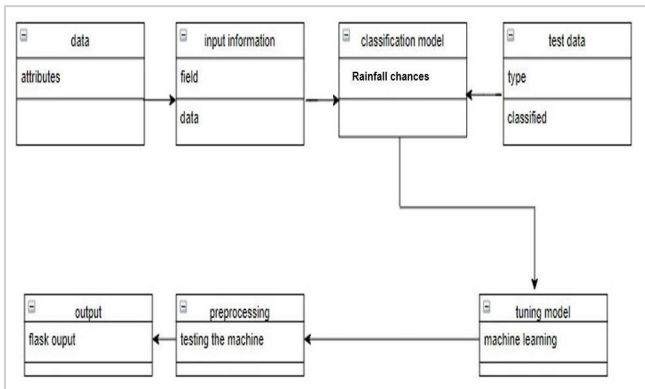


Fig2: Class diagram

6.5 Entity Relationship Diagram (ERD)

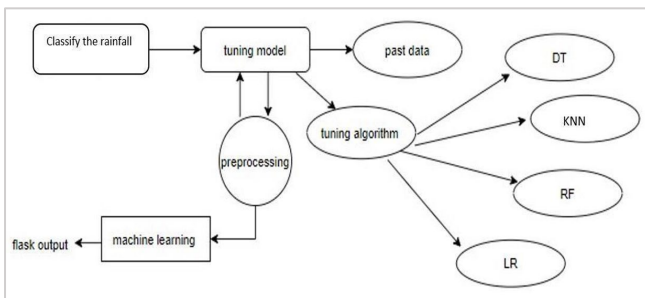


Fig3: ERD

6. TESTING

6.1 Accuracy

Algorithm	Accuracy
Logistic Regression	0.83
RandomForest	0.86
KNeighborsClassifier	0.84
DecisionTreeClassifier	0.78

Table: Algorithms Accuracy

7. CONCLUSION

The project demonstrates how machine learning is used in predicting rainfall. Using records of the past and different supervised techniques, the system can accurately predict weather conditions. Comparing Logistic Regression, K-Nearest

Neighbours, Random Forest and SVM algorithms led to the identification of the strongest algorithm by analysing its accuracy, precision and recall.

The final application, developed with Flask, allows users to use the system with ease. For this reason, the system is beneficial for both farmers, researchers and environmentalists. You can add new data as you like, as the system is organized to easily adapt to new weather events.

With its predictions, the project benefits agriculture, emergency teams and organizations managing water resources. Because it uses open-source software, only basic hardware and can run anywhere, adoption of this solution is likely to increase. Thanks to this project, further progress can be made by connecting to live weather APIs, mobile apps or monitoring the environment with IoT devices.

ACKNOWLEDGEMENT

I am thankful to the Management of Amrita Sai Institute of Science and Technology for giving me an opportunity to work with his project.

I would like to thank **Dr. M. Sasidhar**, Principal, Amrita Sai institute of science and technology, for his constant encouragement and support during the progress of this work.

I am deeply grateful to **Dr. P. Chiranjeevi**, Professor and Head of the Department, for his valuable guidance and consistent support during the course of the project.

A special note of thanks to my internal guide, **Mr. K. Subhash Chandra (M.Tech)**, for his exceptional guidance, constant motivation, and continuous encouragement, which played a crucial role in the successful completion of this project.

B.GOPI

REFERENCES

- [1] P. Singh, "Indian summer monsoon rainfall (ISMR) forecasting using time series data: A fuzzy entropy-neuro based expert system," *Geoscience Frontiers*, vol. 9, no. 4, pp. 1243-1257, 2018.
- [2] S. Cramer, M. Kampouridis, A. Freitas, and A. Alexandridis, "An extensive evaluation of seven machine learning methods for rainfall prediction in weather derivatives," *Expert Systems with Applications*, vol. 85, pp. 169-181, 2017.
- [3] S. Pour, S. Shahid, and E. Chung, "A hybrid model for statistical downscaling of daily rainfall," *Procedia Engineering*, vol. 154, pp. 1424-1430, 2016.
- [4] N. Manjunath, B. R. Muralidhar, S. Kumar, K. Vamshi, and P. Savitha, "Rainfall prediction using machine learning and deep learning techniques," *IRJET*, vol. 8, no. 8, 2021. [Online]. Available: <https://www.irjet.net/archives/v8/i8/irjet-v8i850.pdf>. [Accessed: Jan. 20, 2022].
- [5] T. Patil and Dr. K. Shah, "Weather forecasting analysis using linear and logistic regression algorithm," *IRJET*, vol. 8, no. 6, 2021. [Online]. Available: <https://www.irjet.net/archives/v8/i6/irjet-v8i6454.pdf>. [Accessed: Jan. 20, 2022].
- [6] N. D. Prabha and P. Radha, "Prediction of weather and rainfall forecasting using classification techniques," *IRJET*, vol. 6, no. 2, 2019. [Online]. Available: <https://www.irjet.net/archives/v6/i2/irjet-v6i2154.pdf>. [Accessed: Jan. 20, 2022].
- [7] D. Waghmare, "Machine learning technique for rainfall prediction," *International Journal for Research in Applied Science and Engineering Technology*, vol. 9, no. VI, pp. 594-600, 2021.

[8] V. Yashasathreya, V. Vaishalibv, S. Sagark, and S. Srinidhihr, "Flood prediction and rainfall analysis using machine learning," *IRJET*, vol. 8, no. 7, 2021. [Online]. Available: <https://www.irjet.net/archives/v8/i7/irjet-v8i7432.pdf>. [Accessed: Jan. 20, 2022].