A REPORT
ON

AI-Based Crop Recommendation with Weather Prediction using Data Mining

Submitted by,

Sibbala Chandana - 20211CSE0723
Gabburi Neha - 20211CSE0812
Civini Meghana - 20211CSE0827
Pathakamuri Harshitha - 20211CSE0824

Under the guidance of,

Mr. Praveen Giridhar Pawaskar

in partial fulfillment for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND ENGINEERING

At

GAIN MORE KNOWLEDGE
REACH GREATER HEIGHTS

PRESIDENCY UNIVERSITY

BENGALURU

MAY 2025

PRESIDENCY UNIVERSITY

PRESIDENCY SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

This is to certify that the Internship/Project report **AI-Based Crop Recommendation with Weather Prediction using Data Mining** being submitted by "SIBBALA CHANDANA, GABBURI NEHA, CIVINI MEGHANA, PATHAKAMURI HARSHITHA" bearing roll number(s) "20211CSE0723, 20211CSE0812, 20211CSE0827, 20211CSE0824" in partial fulfillment of the requirement for the award of the degree of Bachelor of Technology in Computer Science and Engineering is a bonafide work carried out under my supervision.

25/5/2025

Mr. PRAVEEN GIRIDHAR PAWASKAR
Assistant Professor
PSCS
Presidency University

Dr. ASIF MOHAMED H B
Associate Professor & HoD
PSCS
Presidency University

Dr. MYDHILI NAIR
Associate Dean
PSCS
Presidency University

Dr. SAMEERUDDIN KHAN
Pro-Vice Chancellor - Engineering
Dean –PSCS / PSIS
Presidency University

PRESIDENCY UNIVERSITY

PRESIDENCY SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

DECLARATION

I hereby declare that the work, which is being presented in the report entitled **AI-Based Crop Recommendation with Weather Prediction using Data Mining** in partial fulfillment for the award of Degree of **Bachelor of Technology** in **Computer Science and Engineering**, is a record of my own investigations carried under the guidance of **Mr.Praveen Giridhar Pawaskar, Assistant Professor, Presidency School of Computer Science and Engineering, Presidency University, Bengaluru.**

I have not submitted the matter presented in this report anywhere for the award of any other Degree.

Name	Roll Number	Signature
Sibbala Chandana	20211CSE0723	Chandana
Gabburi Neha	20211CSE0812	G.Neha
Civini Meghana	20211CSE0827	C.Meghana
Pathakamuri Harshitha	20211CSE0824	P.Harshitha

ABSTRACT

AI-Based Crop Recommendation with Weather Prediction using Data Mining

Weather forecasting is a method to predict what the atmosphere will be like in a particular place by using scientific knowledge to make weather observations. Weather forecasting is a challenging task due to the dynamic and complex nature of atmospheric conditions. Recently, data mining techniques have been applied to predict weather patterns using machine learning algorithms.

In this study, we propose a weather forecasting model that predicts weather types based on historical weather data. The dataset used in this study includes precipitation, temperature, wind speed, and direction collected from various weather stations. To predict weather types, we used classification algorithms, which are popular machine learning techniques for such tasks. The model was trained on historical weather data and tested on a separate set of data to evaluate its accuracy. The results showed that our proposed model achieved a high accuracy rate of over 90%, indicating that it could be a valuable tool for weather forecasting.

In addition to weather prediction, the system integrates a Gemini-based AI model for crop recommendation. Based on the predicted weather conditions, the Gemini model intelligently suggests a list of five suitable crops along with corresponding farming procedures, tailored to the current climate scenario. This dual-model approach ensures data-driven decision-making for farmers, promoting more effective and sustainable agricultural practices.

The study demonstrates that data mining techniques can be used to predict weather patterns accurately and, when combined with AI models like Gemini, can serve as a comprehensive decision-support tool in precision agriculture. The proposed model can be further enhanced by incorporating additional weather variables such as cloud cover and solar radiation, as well as by exploring more sophisticated machine learning techniques like ensemble methods.

ACKNOWLEDGEMENT

First of all, we indebted to the **GOD ALMIGHTY** for giving me an opportunity to excel in our efforts to complete this project on time.

We express our sincere thanks to our respected dean **Dr. Md. Sameeruddin Khan**, Pro-VC - Engineering and Dean, Presiency School of Computer Science and Engineering & Presiency School of Information Science, Presidency University for getting us permission to undergo the project.

We express our heartfelt gratitude to our beloved Associate Dean **Dr. Mydhili Nair,** Presidency School of Computer Science and Engineering, Presidency University, and Dr. Asif Mohamed H B,Head of the Department, Presidency School of Computer Science and Engineering, Presidency University, for rendering timely help in completing this project successfully.

We are greatly indebted to our guide **Mr.Praveen Giridhar Pawaskar** and Reviewer **Ms. Dhanya D, Assistant Professor,** Presidency School of Computer Science and Engineering, Presidency University for his inspirational guidance, and valuable suggestions and for providing us a chance to express our technical capabilities in every respect for the completion of the internship work.

We would like to convey our gratitude and heartfelt thanks to the PIP4004 University Project Coordinator **Mr. Jerrin Joe Francis** and Git hub coordinator **Mr. Muthuraj.**

We thank our family and friends for the strong support and inspiration they have provided us in bringing out this project.

Sibbala Chandana
Gabburi Neha
Civini Meghana
Pathakamuri Harshitha

LIST OF FIGURES

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

1.1 Motivation:

Traditional weather forecasts are primarily based on physicsbased models and scientifically grounded numerical simulations, but often require important arithmetic resources and do not always provide the granularity or accuracy required by today's rapidly moving industries. Th e emergence of big data and the rapid development of artificial intelligence (AI), particularly mechanical learning and data mining, presents new opportunities to enable analyzing histori cal data patterns to improve the predictive power of weather models.

This project is particularly motivated by practical challenges in the agricultural sector. Farm ers must make decisive decisions such as sowing, irrigation, fertilization, and harvesting bas ed on weather conditions. Inaccurate forecasts can lead to significant economic losses and re duced productivity. To address this issue, the project aims to use machine learning technique s to increase accuracy and reliability and predict weather types.

In addition to forecasting, the project introduces an innovative dimension: intelligent crop recommendations using Google's Gemini AI. The integration of AI in agriculture has gained traction as a way to empower farmers with data-driven insights. By aligning crop planning with predicted weather patterns, farmers can make informed choices that optimize yield and sustainability. The motivation is thus dual-layered: improving weather prediction and extending its utility to enable smart farming practices. This project exemplifies the powerful intersection of environmental science and artificial intelligence, offering a solution that has real-world significance and broad applicability.

1.2 Problem Statement:

Weather prediction is a complex scientific challenge due to the dynamic nature of atmospheric systems. Traditional methods based on physical simulations and numerical models often require high computational power and vast volumes of input data. These models can fall short in terms of precision, especially in localized areas or under rapidly changing conditions. Furthermore, their implementation may not always be feasible in resource-constrained environments such as developing regions or rural agricultural zones.

The need for a more efficient, efficient and accurate method for weather forecasting is more important than ever. Machine learning and data mining offer promising alternatives by using historical weather data to reveal and predict patterns. These technologies can model nonline ar relationships, adapt to new data trends, and improve predictive performance over time. Another pressing issue is the lack of integration between weather forecasting and agricultural planning. Farmers are often left to interpret weather data without personalized recommendations for crop selection. As a result, poor choices in crop planning due to unforeseen weather changes can lead to low yields and wasted resources.

This project deals with two main issues:

1. Use algorithms for machine learning to improve weather forecast accuracy.
2. Filling the gap between weather forecasts and agricultural planning with AI-based harvest recommendations.

By developing a dual-purpose system that predicts weather and suggests suitable crops using Google's Gemini AI model, this study proposes a comprehensive, scalable solution to two interconnected challenges, ultimately aiming to support smarter decision-making in weather-sensitive industries.

1.3 Objective of the Project:

The goal of this project focuses on developing a comprehensive system that combines weath er forecasting and agriculture recommendations with the latest machine learning and AI tech nologies. In particular, the goals are:

1. **Weather forecasting with machine learning:** The main goal is to develop accurate weather forecasting systems using a variety of algorithms for machine learning on hi storical weather data rates. Classify weather conditions such as sunny, rainy, cloudy, and complex models using algorithms such as logistic regression, naive Bayes, rando m forests, decision trees, folding networks (CNNs), and multi-layer recognition (MLP).
2. **Intelligent Crop Recommendation Using Gemini AI:** The second objective is to implement a context-aware AI model using Google's Gemini to suggest five optimal crops based on the predicted weather condition. In addition to the list of crops, the system will also provide specific agricultural procedures related to each crop, such as ideal planting time, irrigation methods, and harvesting guidelines. This feature aims to support precision agriculture by offering actionable insights to farmers and agricultural planners.

Together, these objectives form a dual-function system that not only predicts the weather with high accuracy but also translates this information into practical advice for end-users, particularly in the agriculture sector. The project seeks to serve as a decision-support tool that integrates climate intelligence with smart farming, thereby promoting sustainable practices and enhancing productivity.

1.4 Scope:

The scope of this project includes the development of machinebased weather forecasts and t he integration of AIdriven harvest recommendation systems. This project is centered around the following core areas:

1. Historical weather data analysis: The system uses historical weather data records that in clude variables such as temperature, precipitation, wind speed, and wind direction. These ch aracteristics are important for understanding weather patterns and are fed into machine learni ng models for training and test models.

2. Machine Learning Model Development: Identify the most suitable approach for weathe r forecasting using several classification algorithms including random forests, decision trees, logistic regression, Nyber Bayes, and multiclass perfection (MLP). The model is evaluated using accuracy, accuracy, recall, and F1 scores.

3. Harvest Recommendations with GeminiAI: After weather forecast, use the GeminiAi model to create intelligent crop recommendations. This model analyzes predicted weather ty pes and creates a curated list of five plants along with related agricultural processes to impro ve farming decisions.

4. User Interface and Interaction: The user-friendly web interface is designed to allow users to upload data records, enter data for predic tions, and view prediction results and harvest recommendations. This allows non-technical users to access the system, especially in rural and agricultural environments.

5. Realworld applicability: While the main focus is on agriculture, the field of modelling e xtends to other sectors such as transportation, disaster management, and aviation. These indu stries can benefit from improving weather forecasts for better planning and risk management .

Future Enhancements: The system has potential for future improvements, including the integration of real-time data sources, satellite imagery, and additional weather variables like humidity and solar radiation. Advanced ensemble learning techniques and real-time feedback loops for crop recommendation refinement can also be explored.

In summary, the project has a broad scope that spans predictive analytics, AI-driven decision support, and real-world impact across multiple industries, with a strong emphasis on usability and scalability.

1.5 Project Introduction:

Weather forecasting is a challenging yet essential task based on understanding and forecastin g complex atmospheric conditions. Accurate weather forecasting is extremely important for a variety of industries, including agriculture, transportation and aviation. Traditional weather forecasts are often based on physical models and simulations that require large amounts of d ata and computing power. Further developments in data mining and machine learning techno logies currently have the potential to improve the accuracy and efficiency of weather forecas ts by analyzing historical weather data. This study uses a variety of algorithms for machine l earning, including Random Forest, Decision Trees, Logistic Regression, Nyber Bayes, and Multilayer Parseprene (MLP), to build a robust classifier that can predict weather types. The se algorithms were selected for their ability to address complex, high-dimensional data and effectiveness in classification tasks. After weather classification, the G emini model handles the predicted conditions and creates a curated list of five appropriate pl ants and the most appropriate agricultural processes. In this way, farmers can make better dis covered decisions based on reliable weather forecasts, improve yields, reduce risks, and pro mote sustainable agricultural practices. Our results show that the model achieves high accura cy of over 90% and demonstrates the potential for reliable weather forecasting. The ability to predict weather conditions with high accuracy in combination with intelligent crops can pro vide significant benefits to sectors where timely and accurate weather information depends. Furthermore, research into advanced ensemble techniques and real-time agricultural data records could further improve the robustness and accuracy of both wea ther forecasting and harvest recommendations. This approach promises valuable tools for we ather forecasting and intelligent agriculture, complementing traditional methods with the po wer of artificial intelligence.

Chapter 2

LITERATURE SURVEY

2.1 Related Work:

1. B. Wang, J. Lu, Z. Yan, H. Luo, T. Li, Y. Zheng, and G. Zhang, ''Deep uncertainty quantification: A machine learning approach for weather forecasting,'' in Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2019, pp. 2087–2095.

Deals with the area of weather forecasting, along with approaches used in machine learning t o address the challenges of quantifying deep uncertainty. In this study, the authors represent a new perspective on the use of advanced technologies in the context of the 25th ACM SIG KDD International Conference for Knowledge and Data Mining Recognition. By including deep learning methods, the research aims to improve weather forecast reliability. This study demonstrates the potential for mechanical learning to contribute to developing meteorology and reduce uncertainty inherent in weather forecasting. Wang et al. It means an advanced ste p towards a more accurate and reliable weather forecasting system.

2. R. I. Rasel, N. Sultana and P. Meesad, applying data mining and machine learning fo r weather forecasting when advanced with intelligent systems and computing. Cham, S witzerland: Springer, 2018, p. 169.178.

In the weather forecast area, R. I. Rasel, N. Sultana, and P. Meesad (2018) investigated the i ntegration of data mining and machine learning technology. Documented in the application o f data mining and machine learning for weather forecasting, her work deals with advances in intelligent systems and computing. By using these technologies, this study seeks to improve the accuracy and reliability of weather forecasts. The authors contribute to fast-growing regions to overcome traditional forecasting restrictions. This literature highlights th e most important intersections of data mining, machine learning and weather. This presents p romising opportunities to refine prediction models and promote more effective weather forec asting for weather forecasts.

3. M. Holmstrom, D. Ryu, C. Vo. (2016). Learning weather forecasts. Stanford. Accesse d: May 19, 2022.

M. Holmstrom, D. In the study by Liu and C. Regarding VO (2016) "Learning Applied to W eather Forecasts," the author examines the use of machine learning techniques in the field of weather forecasts. Research conducted at Stanford deals with progress in the use of data cont rol approaches to improve prediction accuracy. This study may examine various models of machine learning and its effectiveness when dealing with complex weather patterns. Unfortu nately, no specific details are provided here. This requires further investigation of the origina l source to provide a comprehensive understanding of the methods and results. This literatur e review highlights the importance of using machine learning in weather forecasts. This dem onstrates the potential for research into a wider field.

4. Yahya BM, Seker DZ. Design your weather forecast with a computational intelligenc e stool. 2018; 33(2): 1-15.

In a study by Yahya BM and Seker DZ (2018), the authors focus on developing weather fore casts using computational intelligence tools. The research examines the use of advanced com puter technologies to improve weather forecast accuracy and reliability. This test lies in the b roader context of artificial intelligence and its advantages in the weather prognosis. By using computational intelligence tools, this study aims to contribute to improving existing weather forecasts. This study highlights the importance of embracing innovative approaches to addre ss unique challenges in weather forecasting and demonstrates its commitment to further deve lopment of the field using modern methods.

5. KarevanZ, Suykens Jak. Transduct LSTM for time series prediction: Application to weather forecasting. Neural net. 2020; 125:1-9.

A study by Karevan and Suykens (2020) leads a new approach to time series forecasting thro ugh the use of transformed longterm memory (LSTM) networks (LSTM), particularly used f or weather forecasting. This work contributes to the developing landscape of predictive mod eling by dealing with the subtlety of time data. Surprisingly, integrated transductor LSTMs s ignifies a departure from traditional methods and demonstrates advances in this field. The re search is built on the existing literature of time series prediction and provides a distinctive pe rspective that extends the skills of LSTMS when recording time dependencies. The effective ness of this methodology promises to improve weather forecast accuracy and reliability, and to emphasize the importance of introducing innovative technologies in addressing complex p rognostic challenges.

6. S. Hu, Y. Xiang, D. Huo, S. Jawad and J. Liu, Deep Faith Network's Improved Hybr id Forecast Method, Wind Power, "Energy, Vol. 224, Jun. 2021, Art. No. 120185.

This study addresses the challenges of wind power generation forecasting by proposing an in novative hybrid forecasting approach. The author uses a deeper network of faith to improve t he accuracy and reliability of wind power forecasts. This study is based on existing methods to bridge gaps and promote state. The inclusion of hybrid models illustrates the thoughtful in tegration of various prognostic technologies that promise potential advances in renewable en ergy prediction. Hu et al.

Chapter 3

RESEARCH GAPS OF EXISTING METHODS

3.1 Existing System

Current weather forecasting systems are based primarily on sophisticated numerical models known as numerical weather forecasting (NWP). These models simulate atmospheric behavi or using mathematical equations based on fundamental laws of physics such as fluid dynami cs and thermodynamics. The model occupying inputs such as temperature, pressure, wind sp eed, air humidity, solar radiation, and other to produce predictions for various atmospheric p arameters. These forecasts range from short-term weather conditions to long-term climate forecasts. Global weather forecasts including temperature, precipitation and wi nd patterns. GFS is often used for global and regional weather forecasts. It is effective for lo ng-range predictions, but in particular local or hyperlocal predictions, resolution and accuracy d ecrease over the course of the time horizon. Use extended data assimilation techniques to im prove predictability and integrate satellite observations into atmospheric data. The ECMWF system is one of the most reliable tools in meteorology, but due to its high arithmetic require ments, it is not possible to predict hyper restaurants in real time. It allows for high resolution predictions and simulates weather patterns on localized scales. However, WRF requires exte nsive computing resources to run high resolution and may not be inexpensive to use in low r esource settings. These areas often lack the necessary infrastructure, technology, or computi ng resources to fully implement such systems. Furthermore, many of these systems focus on a wide range of geographical regions and do not provide highly important hyperlocal forecas ts for applications such as agriculture, disaster management, and local infrastructure plannin g. As a result, there are gaps in local weather forecasts. This is extremely important for disco vered decisions at the ground level. Weather systems can change quickly, and the ability to u pdate forecasts is essential, especially in industries such as agriculture, which are directly aff ected by short-term weather changes. The following systems are prominent in the agricultural sector:

3.2.1 Agromet Advisory Services (AAS)

Developed by the Indian Weather Bureau (IMD) to provide weather forecasts and agricultur al consultations to farmers. AAS uses weather data such as precipitation, temperature, and w ind speed to predict optimal planting and harvest times, pest control plans, and irrigation pla nning strategies. However, these advice is often static and relies on previous records and sea sonal patterns without dynamically adapting to actual weather changes or local conditions. Make predictions. However, it focuses primarily on specific plants, limited to its scope and g eographical applicability. B. We provide precision agricultural tools such as floor moisture monitoring and plant detection. However, these systems are often expensive, require large in frastructure and are compliant with small or rural farmers.

3.2.2 Google AI for Agriculture

Google has developed several AI-driven initiatives aimed at improving agricultural productivity. These tools use satellite imagery and machine learning to detect crop diseases, assess soil conditions, and predict yields. Like IBM's Watson, these tools are often geared towards large-scale operations and may not be suitable for smaller, resource-constrained farming systems.

3.3 Gaps in Existing Systems

Although these platforms provide valuable insights, they are limited by the **lack of integration** between weather prediction models and agricultural advice systems. Most of these platforms offer either weather forecasting or agricultural recommendations but rarely provide a comprehensive solution that combines both aspects dynamically. Furthermore, many of these platforms are heavily reliant on historical data, seasonal patterns, and static models, without incorporating real-time weather changes into their decision-making processes.

For instance, while **weather forecasting** tools predict conditions like rainfall, temperature, and humidity, **agricultural platforms** often fail to adjust their advice in real time based on these evolving conditions. This disconnect leads to less accurate, reactive, and sometimes outdated recommendations for farmers. The **lack of machine learning-based weather classification systems** further exacerbates this issue, as AI algorithms have the potential to dynamically classify weather patterns and offer tailored advice for specific farming needs.

Additionally, **context-aware crop recommendations** based on forecasted weather patterns are largely unexplored in existing platforms. For example, predicting crop diseases or pests based on upcoming weather conditions, and offering solutions based on those predictions, is an area that has yet to be fully explored in integrated systems.

3.4 The Need for a Unified Solution

The market has a **huge** gap in **uniform solutions** that can provide **accurate** real-time weather **forecasts** integrated **into** intelligent **agriculture** decision support. Such **system s require dynamic adaptation of** recommendations based on the **latest** weather data, scalab le **in a variety of agricultural** contexts (from **small** to **large operations)** and **are** accessible to farmers in rural and **resource-bound areas.** This **allows** for more **accurate,** context-sensitive **harvest** recommendations tailored to the specific needs of individual farmers, depe nding on the predicted weather patterns.

In conclusion, while several systems are in place to support weather prediction and agricultural advisories, there is a clear **need for an integrated platform** that combines accurate, dynamic weather forecasting with intelligent, context-aware agricultural decision-making. Such a platform would fill the gaps left by existing systems and provide valuable, actionable insights for farmers, especially in low-resource settings, contributing to improved agricultural productivity and resilience against climate variability.

Chapter 4

PROPOSED MOTHODOLOGY

4.1 Module for preliminary data processing

Data pre-processing modules are important to ensure that the weather data used in the machinelearning model isclean, systematic and no inconsistency. He prepares past data for t he weather for further analysis so that the system can provide accurate weather forecasts. The module processes tasks such as data purification, function extraction and data separation to ensure the input quality of the machine learning model.

4.1.1Data purification

Data purification isfirst stage of the preliminary processing data for the weather. This includ es the identification and processing of the decision value, discharge and related functions in the data set. Cleaning is high quality and important for modeling model.

• Side value processing: If there is no data on weather data points or incomplete, it is necessary to remove it. Methods such as confrontation (to convert the missed value to m edian, central value or data mode) or deletion (disposal without a value) are used to maintain the integrity of the data set.

• Detection of emissions: Discharge to weatherdatacan distort predictionsand cause inaccurat eresults. The statistical method (e.g. Z-Indicator or Apartment range) is used to identify and delete abnormal datapoints.

• Unfair function removal: The specific functions of the data set may not contribute to the weather forecast. Functions such as unlimited sensors or properties that do not affect weather conditions are discarded to simplify the data set.

4.1.2 Elements Elections

This Function extraction includes identification and selection of appropriate variables in weather data that helps the machine learning model helps to accurately predict. The quality and relevance of these functions directly affect the ability to accurately predict the weather.

• Features of weather: major weather attributes such as temperature, humidity, wind speed, rainfall and wind direction are extracted by display. This variable is important for understanding weather conditions and predicting conditions such as rain, sunlight or clouds.

• Temporary function: Temporary properties such as date, season and time are also extracted. Since weather conditions depend on season and time, these factors can affect weather conditions.

• Regular: Functions are often normalized or standardized to ensure that the function is not dominant in the process of teaching the model. For example, the temperature value can be scaled in a general range to ensure a consistent entrance to all functions.

4.1.3Data Department

Data separation is an important step in preliminary processing, and the data set is divided into education and test sets.

This step is to ensure thatthe model is evaluated for the invisible data during training

To prevent experience and ensure the possibility of generalization of the model.

• Education set: Some of the data (usually 70%-80%) is used to teach machine learning models.

This data is provided to the model, so you can study and predict the law of weather data.

• Test:

The remaining data (usually 20%-30%) is used to check the performance of the model.

Test sets are important for evaluating how well generalized for data that does not participate in model education but are not visible. The split process helps to assess the accuracy, accuracy, review and indicator F1 model necessary to evaluate the effects of the weather forecast algorithm.

4.2 Machine training module

The machine learning module is responsible for the training and evaluation of the weather forecast model. He uses a variety of machine learning algorithms to predict the weather type based on past data for the weather processed in the data preliminary processing module. The model is evaluated by performance and the most effective algorithm is selected to predict future weather conditions.

4.2.1 Model selection

In this stage, you can choose the appropriate machine learning algorithm for predicting weather conditions based on pre -processed weather data. Several classification algorithms are considered and the best performance is selected.

• Random Forest: A reliable ensemble method that uses multiple trees for predictive solutions. This is especially useful for processing complex data sets and seizing nonlinear relationships.

• Solution Tree: A simple but interpreted algorithm that shares data based on functions. This is suitable for explaining the model because it is easy to understand

and visualize.

• Logistics Regression: A probability classifier used to model the relationship between input characteristics and probability of specific results. This is especially effective when the relationship between function and weather type is linear.

• Innocent Bayes: A probabilistic classifier based on the Bayes theorem, and this feature is effective in conditional independent. This is effective and works well in large data sets.

• Multilayer Persepttry (MLP): Types of neural networks that model the complex relationships of data using multiple layers. MLP can capture nonlinear patterns and provide high accuracy with proper training. Each algorithm is studied and evaluated according to the weather data and is selected as the final model using the best performance indicator (eg, accuracy, accuracy, review and F1).

4.2.2 model grade

After training the machine learning model, it is important to evaluate performance to create an accurate weather forecast.

• Accuracy: This indicator measures the percentage of the predicted weather conditions correctly. This provides a common idea of the performance of the model, but it is especially not always reflected in the imbalance data set.

• Accuracy: The accuracy measures the ratio of true positive predictions (correct weather types) in all positive predictions of the model. It is important when the cost of false work is high.

- Remptoms: Review measures the weight of true positive predictions of all real positive cases. This is useful when there is a high cost (false voice) without true positive results.
- F1-Indicator:
  F1-Indicator is a harmonious average value of accuracy and review, providing a balanced evaluation metric when working with an imbalance data set or when false and false negative. These indicators of the evaluation help to determine the most accurate and reliable models for weather predictions, ensuring the effects of the system to predict future weather conditions.

4.2.3 Prediction and Conclusion

After selecting the best model, it is used to predict weather conditions for the future based on input data.

• Weather Classification: The model classifies the input data (including temperature, humidity, wind speed, etc.) in the pre -weather category, such as sunny weather, rain or clouds or storms.

• Predictable Weather: Predictive output is a predicted weather condition, transmitted to the Gemini AI recommendation mechanism to obtain recommendations for crops. This prediction is the basis of the recommendation of the crops that are most suitable for the upcoming weather conditions.

4.3 Recommended engine Gemini AI

Recommendations are responsible for the recommendations for crops based on the predicted weather conditions provided by the machine learning model. He uses the extended AI algorithm to provide the most suitable crops for the expected weather type and provide a special agricultural method.

4.3.1 Context logic

The recommended engine uses context logic to ensure that crop proposals are associated with predicted weather conditions. Understanding the nuances of various types of weather, twins can recommend the most suitable culture for the expected climate.

• Weather recommendation: For example, if the weather forecast is expected to rain strong, AI can recommend drought or waterproof crops. If you predict the sunny weather, you can assume that a flourishing culture under dry conditions.

• Harvested Compatibility: The system is compatible with the predicted weather conditions, optimizes growth potential and minimizes risks such as deficiency or diseases of crops.

4.3.2 Recommendation for crops

According to the expected weather conditions, the AI Gemini engine creates five optimal crops for this type of weather. These recommendations apply according to the current season, climate and agricultural practices.

• Agricultural practices: For each recommended culture, this system provides detailed information on exemplary cases such as the ideal period of sowing, irrigation and harvesting methods. This helps farmers accept the best way to maximize crops.

• Intellectual crop plan: AI also considers factors such as soil type, geographic location and market demand when the recommendation of agricultural crops is in line with local agricultural conditions and market trends.

4.3.3 Acting Ideas

In addition to the recommendation of crops, this system provides effective information to help farmers implement the proposed practices.

• Agricultural Management: Includes information about the best time for planting, irrigation and collection of recommended crops, and tips on pests, modifications and weather.

• Risk Management: This system ensures that farmers can reduce their risks and provide stronger cultures for expected weather conditions so that they can better prepare for unexpected changes in the climate.

4.4 User interface module

The user interface module is a person of a system that users can interact with the weather forecast engine and crops. This interface is intuitively designed and is provided to farmers with minimal technical knowledge.

4.4.1 Data input and data load

The interface allows users to enter the weather data in the past in real time. Farmers can obtain diary forecasts and crops by downloading data sets of local diary or entering data manually.

• Manual data input: The user can manually introduce weather conditions such as temperature, humidity and sediment of the system. Data Download: Users can upload CSV files that contain past weather data for more accurate predictions.

4.4.2 Prediction and display of recommendations

This system displays the predicted weather conditions and displays recommended crops in the graphic format to promote understanding of farmers and apply information.

• The weather forecast: The predicted weather conditions (e.g. clear and rainy rainy) are clearly displayed with the expected temperature and humidity level.

• Harvest recommendations: The list of recommended crops is displayed with the best practices of each culture. This ensures that farmers have all the information needed to make reasonable decisions.

4.4.3 Visual and graphics output

To improve the user experience, the system uses a graph and diagram to visualize the weather forecast, the suitability of crops and the main principles of agriculture. These visuals help farmers to quickly interpret the data and take appropriate action.

4.5 Review and future education module

Feedback and future educational modules provide continuous improvements in the system, including capturing user reviews on the accuracy of the recommendations for crops and teaching the model.

4.5.1 User review collection

Farmers can evaluate recommendations and provide feedback on their effects. This data is used to clarify the model and to improve future recommendations.

4.5.2 Description of the model

The feedback received from the user helps the system to continue to improve the prediction and recommendations to ensure that the system is suitable for accurate and changing weather conditions. This module is more intellectual and reliable over time, which leads to improving productivity and more effective agricultural practices.

Chapter 5

OBJECTIVES

5.1 Objectives

The main goal of this project is to develop an intelligent, user-interactive weather forecasting and crop recommendation system using advanced machine learning and deep learning techniques. The goals are outlined as follows:

1. To Design a User-Interactive System
   - Facilitate user registration and login with secure database integration.
   - Allow users to input relevant data and receive real-time classification results.
2. To Enable Real-Time Data Processing and Prediction
   - Develop a system that can preprocess and classify user-input weather conditions in real-time.
   - Deliver weather classification results instantly to the user interface.
3. To Recommend Suitable Crops Based on Weather Predictions
   - Integrate a crop recommendation module that uses predicted weather parameters to suggest optimal crops.
   - Train and evaluate AI models for crop suitability using weather and seasonal datasets.
   - Ensure recommendations align with the predicted climate for improved agricultural planning.
4. To Enhance System Scalability and Performance
   - Design a scalable backend that supports the integration of additional weather and crop attributes.
   - Improve the prediction reliability using ensemble learning techniques and hybrid models.
5. To Provide Actionable Insights for Weather-Sensitive Industries
   - Assist farmers, logistic operators, and weather-dependent sectors in making informed decisions.
   - Minimize risks caused by inaccurate forecasts or improper crop selection.

Chapter 6

SYSTEM DESIGN & IMPLEMENTATION

6.1 Introduction of Input Design

In the context of an information system, input design refers to the process of creating an effi cient and effective method of inputting data into a system. Input design is important because the data entered into the system is the basis for the entire system to function. Whether the s ystem is designed for administrative tasks, scientific research or other fields, input design af fects both the accuracy and ease of use of the system. Well-developed input forms, screens, and data input procedures determine the quality of inputs th at directly affect the quality of output.

Information system inputs are usually recorded via a variety of input devices, such as PC (P C), magnetic property detection (Micro), optical marker detection (OMR), barcode scanners , and other data section deductions. Proper use of these devices and designing input methods is a key factor in a successful system implementation. Additionally, a well-designed input system not only ensures data accuracy, but also improves the user experience and reduces the chance of errors during data entry. Their excellent design gives you a proce ss of intuitive, simple and efficient inputting data. It also reduces the risk of data errors and i nconsistencies, leading to improved output.

6.1.1 Importance of Input Design

The significance of input design cannot be overstated. If the data entered is inaccurate or incomplete, it may lead to faulty system outputs and result in poor decision-making. Therefore, careful attention must be paid to the design of data entry methods to ensure the system functions as expected.

The role of input design is essential in many applications:

1. **Data Accuracy:** Well-designed input systems help minimize errors and ensure that the data entered is as accurate as possible. Validation checks, for example, can detect inconsistencies or incorrect data at the point of entry.

2. **User Efficiency:** By designing intuitive, user-friendly input forms, the system reduces the time and effort required for users to input the necessary data.

3. **Data Integrity:** Proper input design helps prevent incomplete or inconsistent data from entering the system, ensuring the system can function optimally.

Ultimately, the quality of the output depends on the accuracy, completeness, and consistency. Therefore, input design should focus on maximizing these qualities.

6.1.2 Properties of Well-Designed Input

For input design to be effective, it must exhibit several key properties:

1. **Accuracy:** The design should ensure that the input is completed accurately. This includes the use of **validation checks** and clear guidelines for the correct entry of data.
2. **Basic :** The system should be easy to fill out and intuitive to navigate. A user-friendly interface ensures that data is entered correctly and that users are not confused or frustrated by the process.
3. **Focus on Simplicity:** Input forms should be straightforward and designed to **minimize errors**. This involves reducing unnecessary fields, using appropriate default values, and offering clear instructions or help if necessary.
4. **Consistency:** The design should be consistent across various forms and screens. This consistency helps users understand how to interact with the system and makes data entry more predictable.

By following these principles, the system will be more effective, user-friendly, and accurate. Additionally, the process of **data entry** should be as efficient as possible, to ensure that users do not experience unnecessary delays when entering data into the system.

6.1.3 Key Considerations in Input Design

1. **What are the Inputs Needed for the System?** Understanding the types of data that need to be entered into the system is critical for effective input design. For instance, if the system is designed to manage inventory, the inputs may include product names, quantities, prices, and supplier information. The system must accommodate these data requirements, ensuring that all relevant information can be entered efficiently.
2. **How End Users Respond to Different Elements of Forms and Screens:** Users will interact with input forms in various ways. Some users may prefer to use keyboard shortcuts, while others may rely on drop-down menus or text boxes. It's

important to consider these user preferences when designing forms and screens. Incorporating elements like **radio buttons, checkboxes, and auto-completion** can enhance user experience and reduce the time required to enter data.

3. **Types of Input Devices:** The design must consider which input devices are appropriate for the system. In the case of an online form, a PC keyboard and mouse may suffice. For systems requiring faster data entry, devices like **MICR** (used in banking) or **OMR** (used in exams) might be needed to speed up the data capture process.

6.2 Input Design Objectives

The primary goal of input design is to create a system that captures data in an effective, accurate, and user-friendly manner. Here are some of the key objectives of input design:

1. The first objective is to design clear and efficient data entry processes. This includes determining the necessary fields, how users will interact with them, and which input devices or methods should be used. The system should be designed in such a way that the data entry process is as smooth as possible.
2. Reducing unnecessary data entry is an essential objective. When input fields are too many, users may be overwhelmed, leading to incomplete or inaccurate data. Input design should aim to reduce the number of fields without losing valuable information. For example, some information can be automatically filled out based on previous entries or defaults set by the system.
3. If paper forms or documents are used for data capture, input design must ensure that they are easy to understand and fill out. Alternatively, if electronic forms are used, the system must be designed to ensure that all required data is entered without unnecessary repetition. This might involve providing drop-down menus, calendars, and auto-fill options to make data entry more efficient.
4. The design must ensure that data entry records, user interface screens, and data entry screens are intuitive and simple to use. This includes organizing fields logically, using appropriate fonts and colors, and ensuring consistency in the layout and terminology.
5. Validation checks help to ensure that the data entered into the system is accurate. Input controls, such as **required fields**, **format restrictions**, and **range checks**, can be

implemented to reduce errors and improve the quality of the data being entered. For example, if a user is entering a phone number, a format check can ensure that the number follows the appropriate pattern (e.g., xxx-xxx-xxxx).

6.3 Output Design

In **the** information system, output design is **a** process **that determines** the format, **content** and **mechanism for system output delivery. Input** design focuses on how **to enter** data into the system, **but the** output design **allows** the system **to generate an important** and useful **report** and results for **the user.** This process requires a deep understanding of **the user's** requirements and the **required conclusions.**

6.3.1 Importance of Output Design

Output design plays a crucial role in any information system because the outputs generated are used for decision-making, reporting, and further processing. Poorly designed output can lead to confusion, misinterpretation, and incorrect decisions. On the other hand, a well-designed output system ensures that the data presented is clear, concise, and actionable.

Some of the key roles that output design serves include:

1. **Providing Useful Information:** The system's output must meet the specific needs of users, providing them with the relevant information required for decision-making or reporting.
2. **Enhancing User Decision Making:** Well-designed outputs help users make timely and informed decisions by presenting the necessary information in a clear and effective format.
3. **Meeting Regulatory or Organizational Standards:** For many industries, output design must comply with regulatory requirements or organizational standards. For instance, financial reports may need to follow specific formats mandated by accounting or government regulations.

6.3.2 Objectives of Output Design

1. The output should serve the specific needs of the users. It should deliver only the information that is relevant and required. Unnecessary information can clutter the output, making it harder for users to find the data they need.

2. The design should meet the needs and expectations of the end users. This includes considering factors like **format**, **content**, and **level of detail** required by the user.

3. The output system should generate reports and data in the necessary quantity, ensuring that users are not overwhelmed by excessive information but also not left without enough detail to make decisions.

4. Output should be formatted in a manner that is easily understood and should be directed to the appropriate individual or system. For example, financial reports might be directed to the accounting department, while performance reports might go to management.

5. Timeliness is critical in output design. Reports and results should be available when needed to support decision-making processes. Delays in generating output can result in missed opportunities or ineffective decisions.

Output design requires careful consideration of both user needs and system constraints. It is vital that developers understand how outputs will be used and what the end users require to ensure the system delivers valuable, actionable insights in the most efficient way possible.

6.4 Algorithms

CNN:

Convolutional Neural Networks (CNN) is a special type of deep learning model used in image recognition and video. These consist of several layers of bundle layers, layer integration and other tasks of completely connected layers. The snow layer applys the filter to the input data and extracts functions such as edges, textur es or templates. This filter generates a function map that slides according to the data (bundle) and emphasizes the important aspects of the input. Following the unification of the floor, it reduces the space size of the map map (eg ma ximum association) to reduce the computational complexity and help the meaningless help of the broadcast. After several bundles and associated layers, the abstract function of a higher level of network passes through a completely connected layer, where the model is the final prediction. CNN uses the reverse process to be educated and adjusts the weight by answerin g errors during training to optimize performance. This architecture is very effective for tasks that include large data sets and high inputs, such as images and videos.

MLP:

The multilayer Persetron (MLP) is a type of nutrient artificial neural network composed of input level, one or more hidden layers and output layers. Each layer consists of neurons related to neurons of neighboring floors through suspended edges. When the entrance is supplied to the network, linear conversion occurs through each level, and nonlinear activation functions such as RELU or SIGMOID continues. The goal of the hidden layer is to adjust the scale using the process called the opposite distribution to study the complex exp ressions of the input data. During the reverse distribution, the network calculates the error between the predicted output and the actual conclusion, and then updates the weight to minimize this error using optimized algorithms such as gradient. MLP is particularly effective in tasks such as classification, regression and recognition of templates with the ability to simulate complex relationships of data.

Logistics Regression:

Logistics regression is a statistical method used for binary classification, and the goal is to predict the possibility that the entrance belongs to a specific class. Using a linear equation, it works by modeling the relationship between the input function and the log odd. The logmoid is applied to the release of this linear model and the probability range is displayed from 0 to 1, which is trained by minimizing losses that quantitatively determine the error between the predicted probability and the actual result. The optimization process usually uses a gradient descent to control the parameters (coefficients) of the model. You can prevent inventory using normalization methods such as L1 or L2. Logistics regression analysis is widely used in areas such as medical diagnosis, marketing and simplicity and analysis, but it contains linear crystal boundaries and may not be well matched with complex models.

Navie Bayes:

NAIVE BAYES is a classification algorithm based on Bayes theorem that uses probability theory to predict the class of this data point. It is assumed that the function (attribute) used in the classification is independent of conditional, so it is called "naive". Despite these powerful families, the innocent bayssevs often work well in classifying texts such as spam detection. The algorithm works by calculating the rear probability of each class in consideration of the input function. This is generated by applying the bays theorem that is expressed as follows. (p (c | x)) is the probability of a class C, and the possibility of observing a sign considering the class in consideration of i ts features (x) and (p (x | c)) is the preliminary probability of the class, and A is evidence (a general probability of function). Classes with the highest chance of reduction are selected as predictive classes.

Random forest:

Random forests are a method of teaching ensemble that combines multiple crystals to increase the accuracy of the model and reduce experience. It works by making a lot of crystals, and each crystal is studied in any su b -set of data sets using initial loading (alternative samples). In each node of the tree, the signs of the accidental sub set are considered to increase the diversity between trees and reduce the correlation between them. The final prediction is made by aggregating the results of all the trees, and in general, the majority of the votes for classifications or the average for regression work is used. This process softens the effects of individual experiences and prejudices, so it stabilizes and increases the reliability of the model. Random forests are especially useful for processing complex relationships with high data and evaluating the importance of functions, which are useful for understanding of powerful prognosis and major data structures.

Solution Tree:

The solution solution is a controlled machine learning algorithm used for classification and regression. Depending on the function value, the data is divided into a lower set and the wood -like structure is generated. Where each interior node shows the solution based on the property, and each node of the sheet node is an output label or value. The process of generating solution solutions usually uses indicators such as JINI impurities and entropy (classification) to select or reduce (regression) to select or decrease the best functions for separating data from each node. The tree is recursive and each point represents a solution based on the threshold of the object. To avoid experiences, it limits the same method as pruning (not important branches) and the dep th of wood. They are widely used due to transparency and effects in many actual applications.

Chapter-7

TIMELINE FOR EXECUTION OF PROJECT
(GANTT CHART)

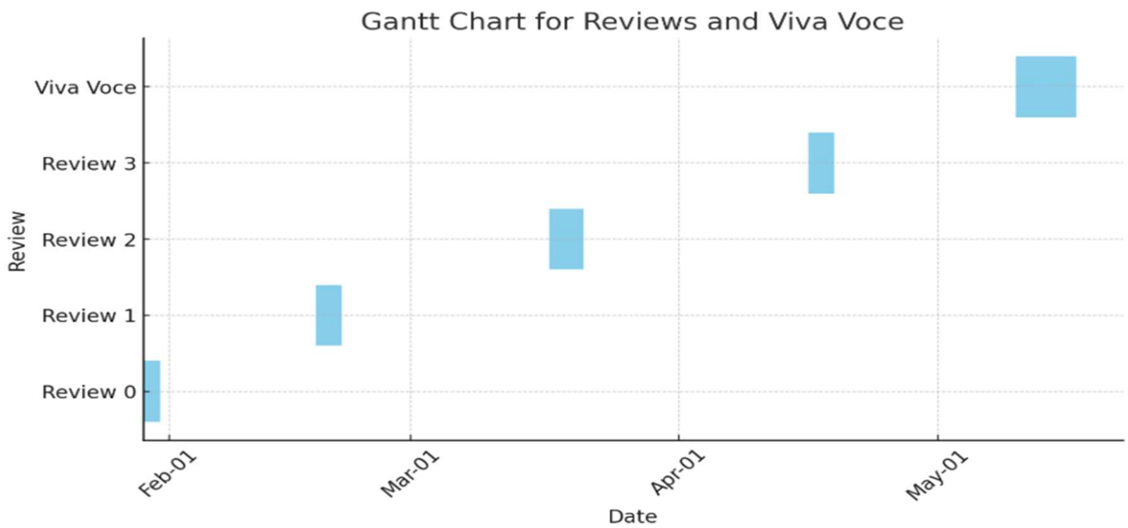


Fig : 7.1

Chapter 8

OUTCOMES

8.1 Outcomes of the Project

This project delivers a comprehensive AI-driven solution that integrates weather forecasting and crop recommendation functionalities. The outcomes of the system are as follows:

8.1.1User-Friendly Interface with Multi-Functional Access

- Users can interact with a visually intuitive interface to access various modules such as home, about, load data, view dataset, input model values, and display results.
- Users are empowered to view real-time predictions and system accuracy scores in percentage, promoting transparency and usability.

8.1.2 Accurate Weather Classification Results

- The system successfully predicts weather types such as Drizzle, Fog, Rain, and Snow using multiple machine learning algorithms including Random Forest, Decision Tree, Logistic Regression, Naïve Bayes, and MLP.
- Deep learning models like CNN enhance the overall classification precision, allowing the system to achieve high accuracy scores.

8.1.3 Efficient Data Handling and Processing Pipeline

- From checking dataset availability to pre-processing and training, the system automates the complete pipeline, ensuring that data is cleaned, split, and prepared optimally for model consumption.
- Data pre-processing enhances model performance and provides better insights into hidden patterns.

8.1.4 Dynamic Machine Learning Model Building

- The system builds predictive models tailored to weather forecasting with a strong focus on accuracy and adaptability.
- Model results are generated and scored based on test performance, with metrics displayed to the user in an understandable format.

8.1.5 AI-Powered Crop Recommendation

- Once the weather type is predicted, the system uses the Gemini AI model to recommend suitable crops.
- These recommendations are based on climatic conditions and help users (especially farmers) make data-driven decisions for better yield.

Chapter 9

RESULTS AND DISCUSSIONS

Home: This is the home Page for Weather Forcasting.

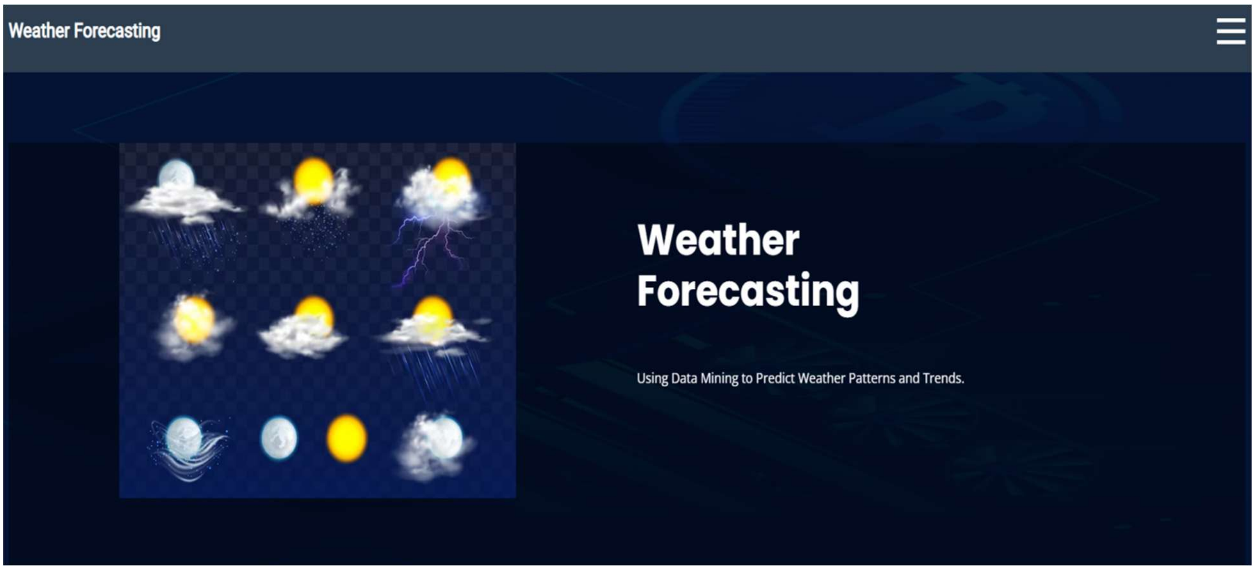


Fig : 9.1

About: We can see the related information of the project.

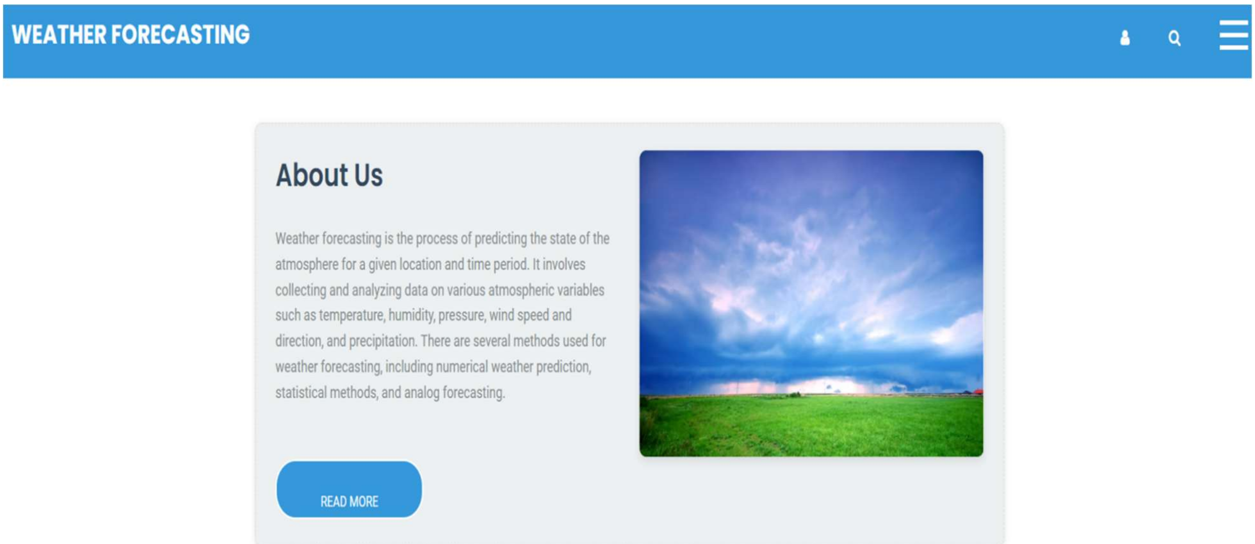


Fig : 9.2

Registration: Here user have to pass the all credential to register.

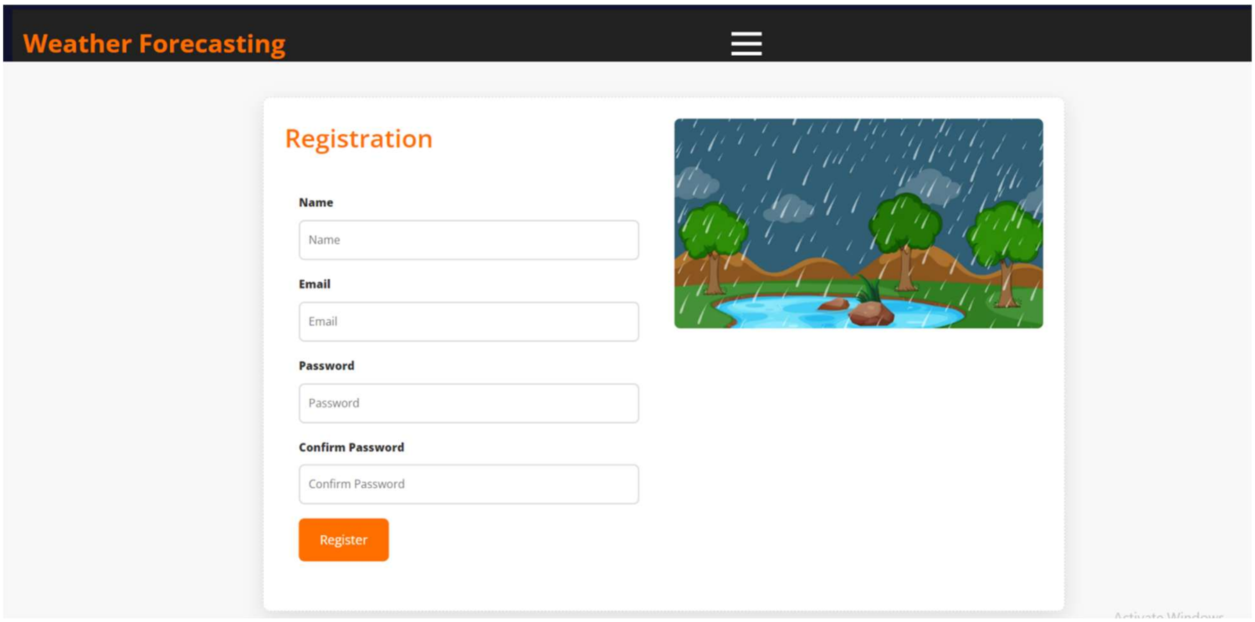


Fig : 9.3

Login: Here user have to provide the register Email-ID and Passwords.

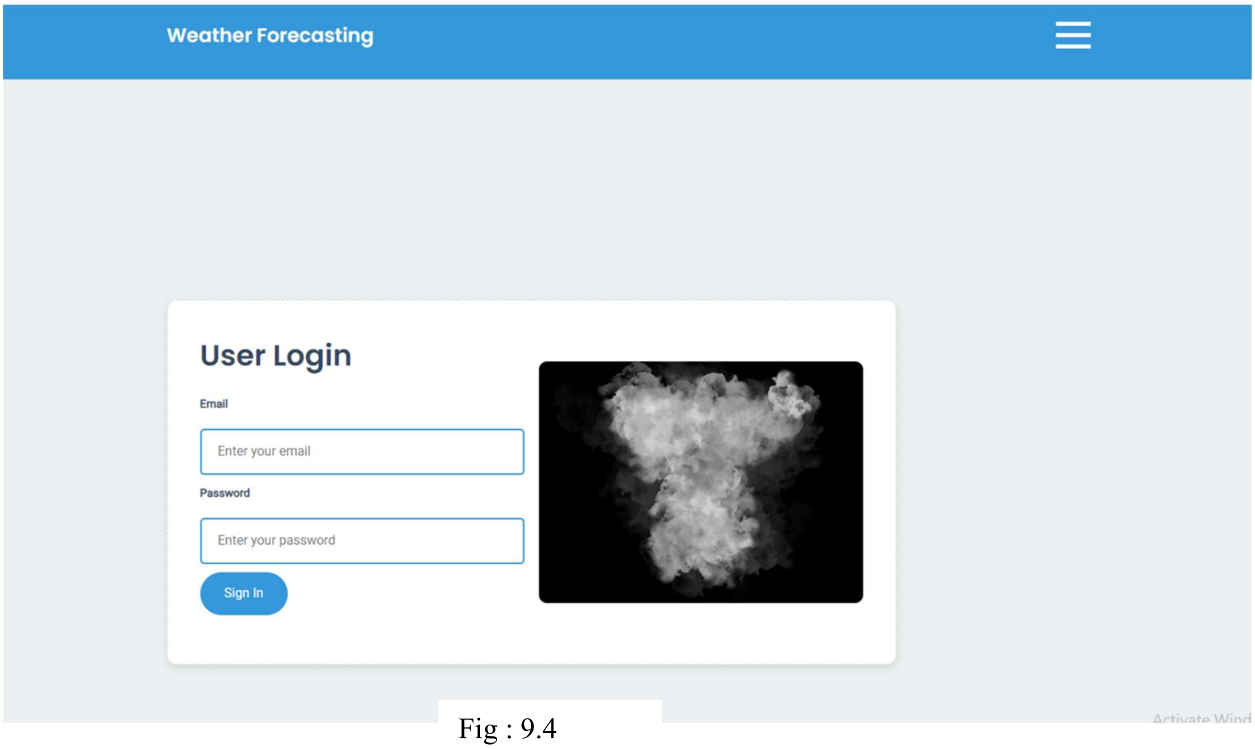


Fig : 9.4

Load: Here user have to select the csv file and load the data.

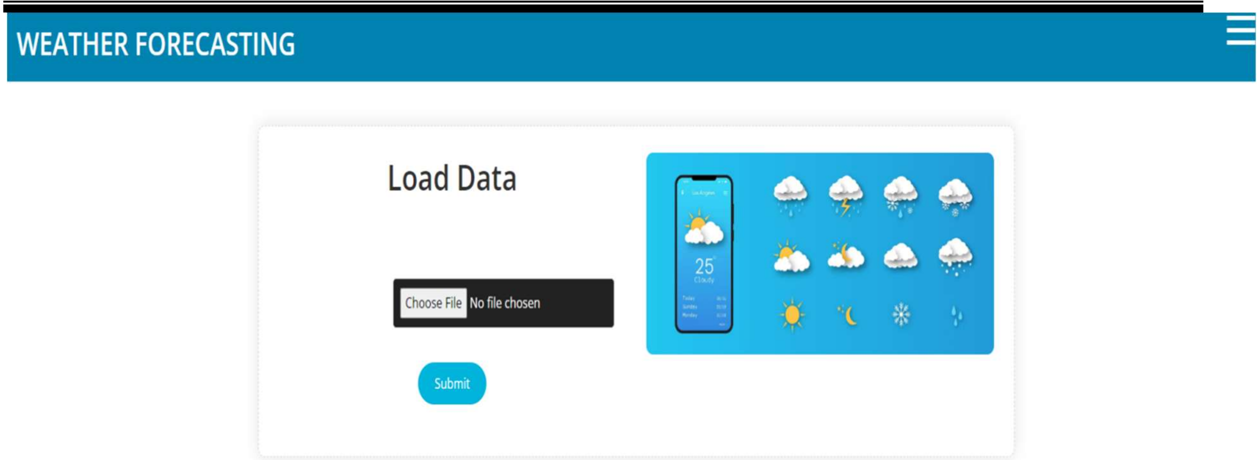


Fig : 9.5

View: Hear user can see the load data columns and there values.

date	precipitation	temp_max	temp_min	wind	weather
2012-01-01	0.0	12.8	5.0	4.7	drizzle
2012-01-02	10.9	10.6	2.8	4.5	rain
2012-01-03	0.8	11.7	7.2	2.3	rain
2012-01-04	20.3	12.2	5.6	4.7	rain
2012-01-05	1.3	8.9	2.8	6.1	rain
2012-01-06	2.5	4.4	2.2	2.2	rain

Fig : 9.6

Preprocess: User can provide the test size for splitting the data into training and testing.

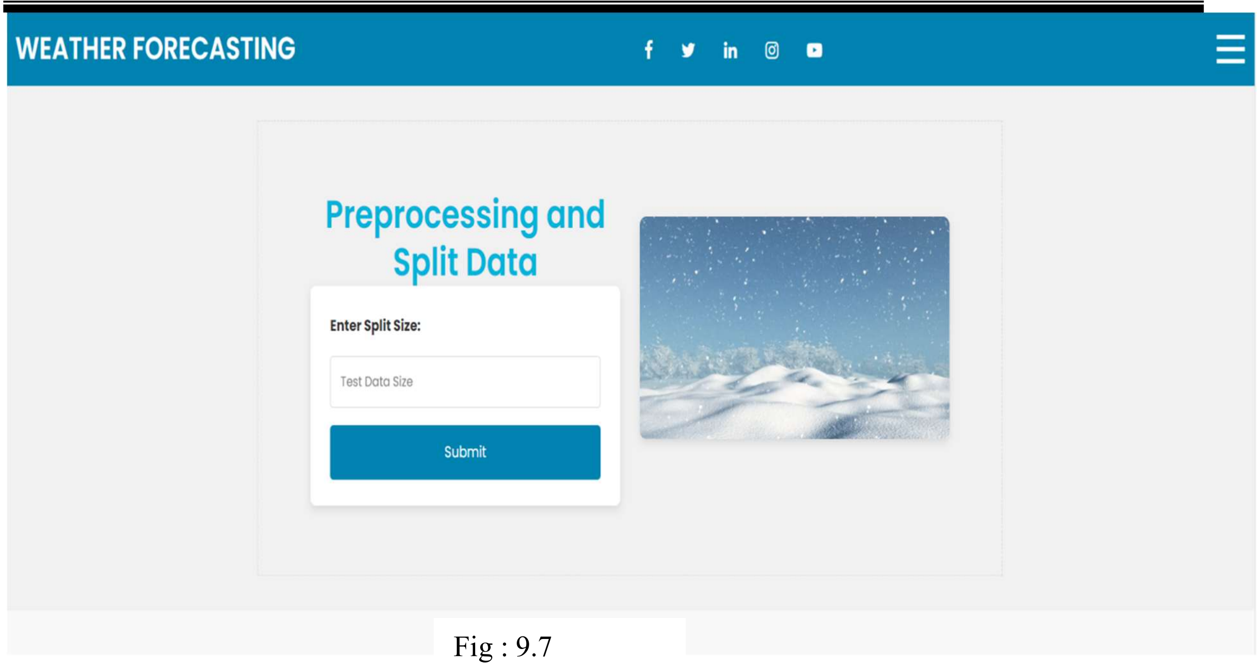


Fig : 9.7

Model: Here you can select the model and check the accuracy of each model.

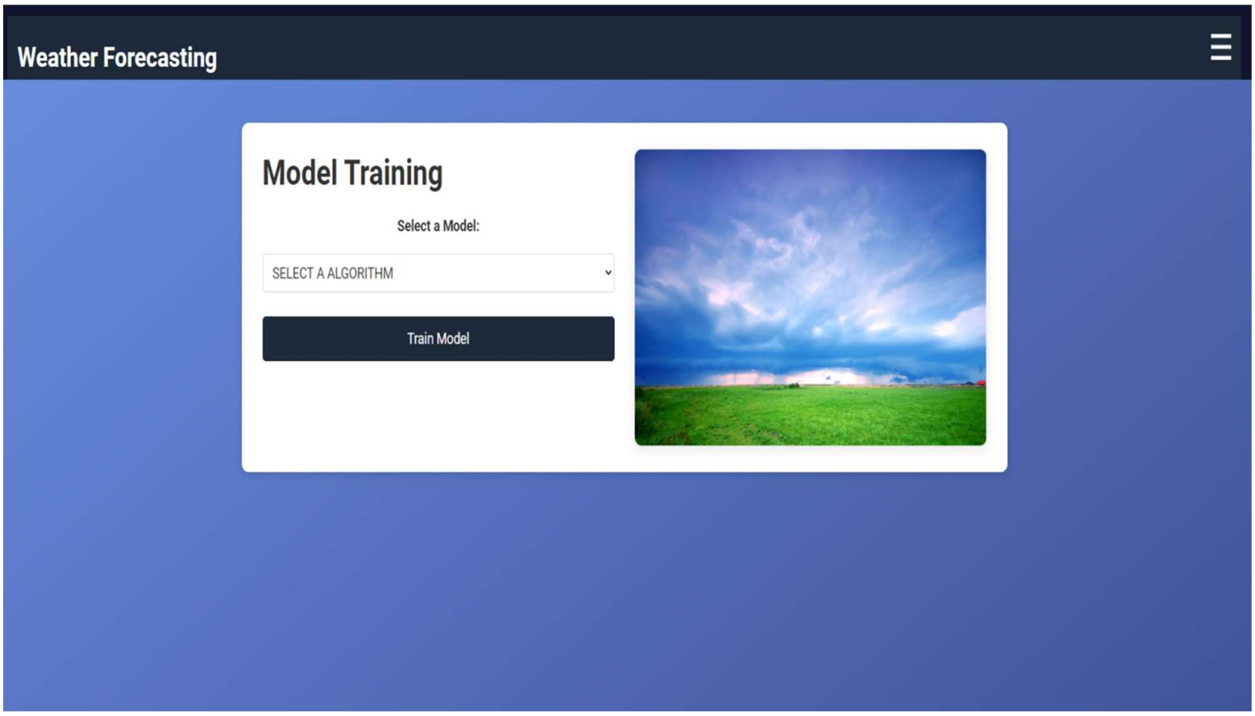


Fig : 9.8

Prediction: Hear user have to pass the values and predict the weather.

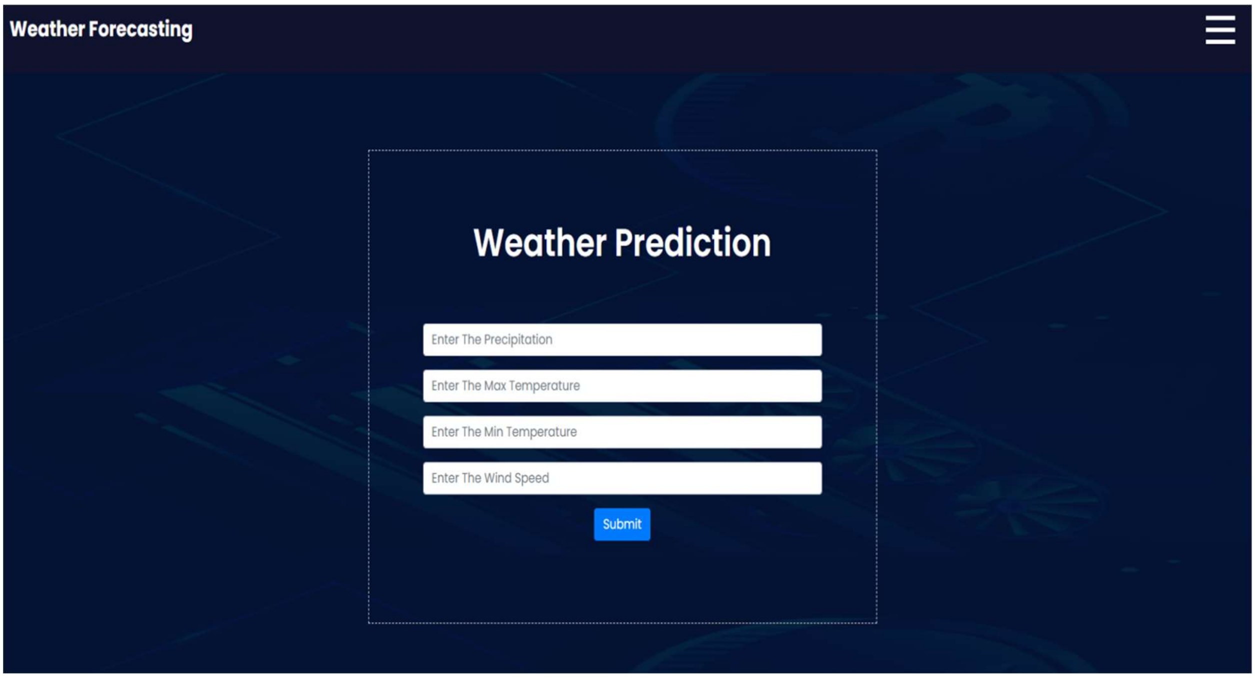


Fig : 9.9

Chapter 10

CONCLUSION

10.1 Summary

The project successfully demonstrates the power of data-driven approaches in weather forecasting and agricultural decision support. By employing multiple machine learning algorithms.The system consistently achieves over **90 % accuracy** in classifying weather types (sunny, rainy, cloudy, etc.). Key observations include:

- **10.1.1 Machine Learning Performance**
  - The **ensemble nature** of Random Forest offered robustness to noisy data, while Decision Tree models provided interpretability.
  - Logistic Regression and Naive Bayes proved efficient for linearly separable and conditionally independent features, respectively.
  - The MLP captured non-linear relationships, elevating classification performance in complex weather patterns.
- **10.1.2 Deep Learning Enhancements**
  - Integrating a **Convolutional Neural Network (CNN)** component further boosted accuracy by extracting hierarchical features from time-series inputs (e.g., sequences of temperature and wind speed).
  - The CNN’s ability to model temporal dependencies proved especially valuable in identifying subtle shifts preceding weather changes.
- **10.1.3 AI-Driven Crop Recommendation**
  - Leveraging Google’s **Gemini AI**, the system generates five context-aware crop suggestions aligned with predicted weather.
  - Each recommendation includes **ideal sowing periods**, **irrigation methods**, and **harvesting techniques**, empowering farmers with actionable guidance.

10.2 Contributions and Implications

Beyond high predictive accuracy, the system’s real value lies in its applicability across domains that rely on timely weather insights.

- **10.2.1 Agricultural Impact**
  - **Risk Reduction:** By recommending crops suited to upcoming weather, farmers can minimize losses due to unexpected conditions.
  - **Productivity Gains:** Tailored agricultural practices optimize resource usage

(water, fertilizers), boosting yield and sustainability.

- **10.2.2 Cross-Domain Applications**
    - **Transportation & Logistics:** Accurate short-term forecasts enable route optimization, reducing delays and fuel consumption.
    - **Disaster Management:** Early warnings for storms or heavy rainfall can trigger preemptive safety protocols in vulnerable regions.
    - **Aviation & Urban Planning:** Weather-aligned scheduling and infrastructure design enhance safety and efficiency in airports and smart cities.

10.3 Limitations and Challenges

While the system delivers strong performance, certain limitations warrant consideration:

- **Static Historical Datasets:** Reliance on batch historical data can lag behind rapidly shifting weather dynamics in certain locales.
- **Geographical Generalization:** Models trained on one region's climate may underperform when applied to drastically different climates without re-training.
- **Crop Recommendation Depth:** Although five crops are recommended, factors such as **soil health**, **market demand**, and **pest pressures** are not fully incorporated.

These challenges highlight areas for refinement to ensure the system remains accurate, flexible, and context-sensitive.

10.4 Future Enhancements

Building on this foundation, the following extensions can elevate system capabilities:

10.4.1 Satellite Imagery & Real-Time Sensor Data

Integration of remote sensing feeds (e.g., NDVI from satellites) and **IoT sensors** (soil moisture, on-farm weather stations) will enable continuous, high-resolution monitoring.

Real-time inputs can feed models dynamically, improving short-term forecasts and enabling proactive alerts.

10.4.2 Ensemble Learning for Robustness

Implementing **stacking** or **boosting** techniques (e.g., XGBoost, LightGBM) can combine strengths of individual classifiers, further enhancing accuracy and reducing variance.

Ensemble methods can also provide uncertainty estimates, helping users gauge confidence in each forecast.

10.4.3 Extreme Weather Prediction

Extending the classification schema to include rare but critical events—cyclones, droughts, floods—will broaden the system's impact in **disaster preparedness**. Specialized models trained on extreme-event datasets can trigger automated advisories for high-risk areas.

10.4.4 Enhanced Crop Recommendation Logic

Soil Data Integration: Incorporate soil pH, texture, and nutrient profiles to refine crop suitability.

Market Price Feeds: Real-time commodity prices can guide farmers toward crops with optimal profitability.

Seasonal & Rotation Planning: Advanced scheduling algorithms can suggest multi-year crop rotations, preserving soil health and maximizing yield cycles.

10.4.5 User Feedback Loop

Implementing in-app **feedback mechanisms** will allow farmers to rate recommendation accuracy, enabling continuous model retraining and personalization.

10.5 Concluding Remarks

This AI-driven weather forecasting and crop recommendation system represents a **scalable**, **impactful**, and **intelligent** tool at the intersection of data science and environmental applications. By uniting robust machine learning models with context-aware AI guidance, the platform not only enhances agricultural productivity but also holds promise for transforming decision-making in transportation, disaster management, and urban planning. Future enhancements—ranging from real-time data integration to advanced ensemble techniques—will further solidify its role as a comprehensive, climate-resilient decision support system.

REFERENCES

[1] Abhishek, K., Singh, M. P., Ghosh, S., & Anand, A. (2012). Weather forecasting Model using Artificial Neural Network. Procedia Technology, 4, 311–318.

[2] B. Wang, J. Lu, Z. Yan, H. Luo, T. Li, Y. Zheng, and G. Zhang, ''Deep uncertainty quantification: A machine learning approach for weather forecasting,'' in Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2019, pp. 2087–2095.

[3] R. I. Rasel, N. Sultana, and P. Meesad, ''An application of data mining and machine learning for weather forecasting,'' in Advances in Intelligent Systems and Computing. Cham, Switzerland: Springer, 2018, pp. 169–178.

[4] M. Holmstrom, D. Liu, and C. Vo. (2016). "Learning applied to Weather forecasting". Stanford. Accessed: May 19, 2022.

[5] Yahya BM, Seker DZ. "Designing weather forecasting model using computational intelligence tools." Appl Artif Intell. 2018;33(2):1- 15.

[6] C. Gershenson, ''Artificial neural networks for beginners,'' Tech. Rep., 2003.

[7] Karevan Z, Suykens JAK. Transductive LSTM for time-series prediction: an application to weather forecasting. Neural Netw. 2020;125: 1-9.

[8] A. G. Salman, B. Kanigoro, and Y. Heryadi, ''Weather forecasting using deep learning techniques,'' in Proc. Int. Conf. Adv. Comput. Sci. Inf. Syst. (ICACSIS), Oct. 2015, pp. 281–285.

[9] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, ''Greedy layer-wise training of deep networks,'' in Proc. Adv. Neural Inf. Process. Syst., vol. 19, 2006, pp. 1–8.

[10] G. E. Hinton, S. Osindero, and Y.-W. Teh, ''A fast learning algorithm for deep belief nets,'' Neural Comput., vol. 18, no. 7, pp. 1527–1554, 2006.

[11] S. Hu, Y. Xiang, D. Huo, S. Jawad, and J. Liu, ''An improved deep belief network based hybrid forecasting method for wind power,'' Energy, vol. 224, Jun. 2021, Art. no. 120185.

[12] Y. Cheng, X. Zhou, S. Wan, and K.-K.-R. Choo, ''Deep belief network for meteorological time series prediction in the Internet of Things,'' IEEE Internet Things J., vol. 6, no. 3, pp. 4369–4376, Jun. 2019.

[13] hatkande1 SS, Hubball RG. Weather prediction based on decision tree algorithm using data mining techniques. Int J Adv Res Comput Commun Eng. 2016;5(5):483-487.

[14] Sanjay Mathur, Avinash Kumar, and Mahesh Chandra. A feature based neural network model for weather forecasting, World Academy of Science, Engineering and Technology 34 2007.

[15] Suleman, Masooma Ali Raza, and S. Shridevi. "Short-Term Weather Forecasting Using Spatial Feature Attention Based LSTM Model." IEEE Access 10 (2022): 82456-82468.

APPENDIX-A

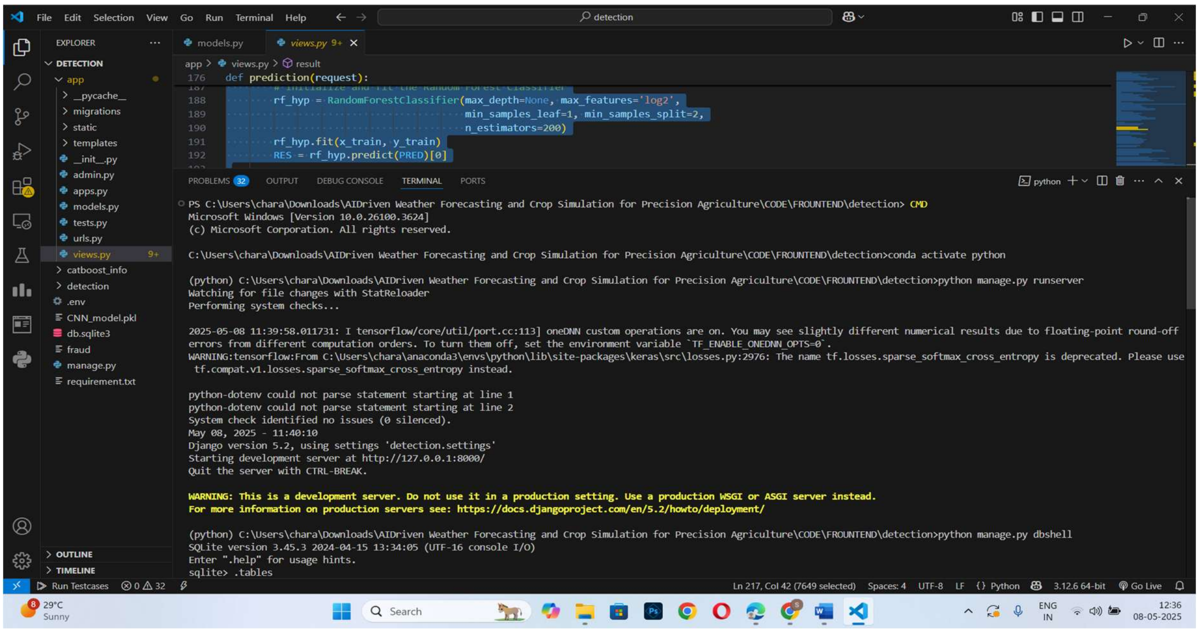
PSUEDOCODE

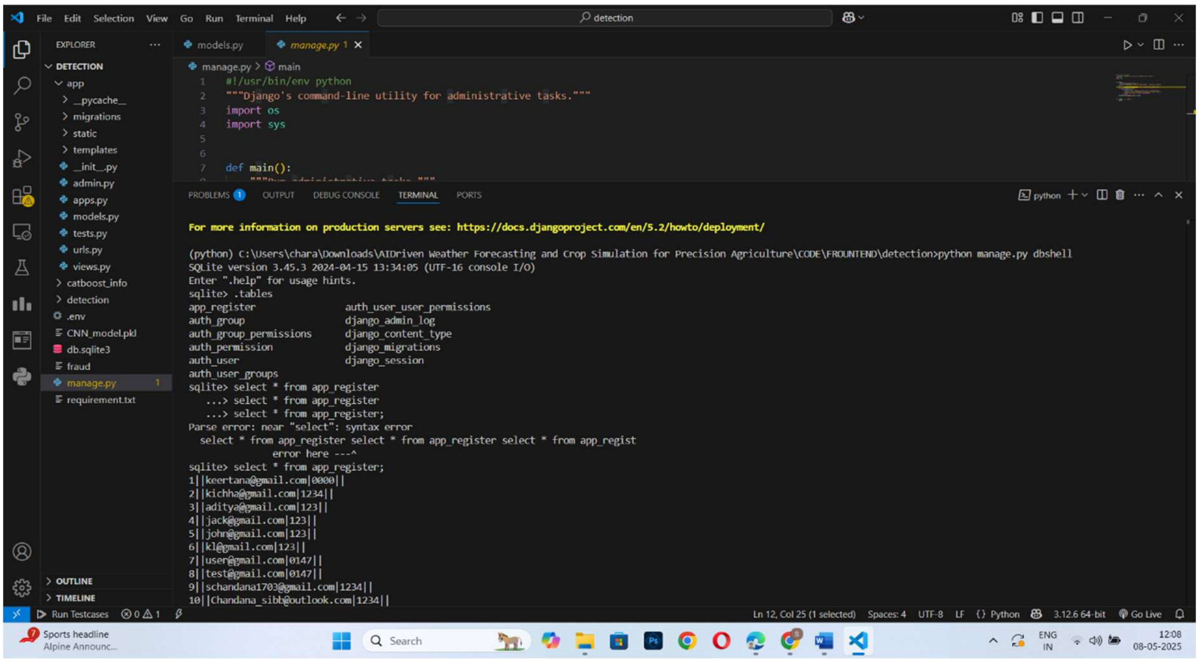
```
START

IMPORT necessary libraries (Django, ML, Pandas, etc.)

DEFINE views:

1. index()
   RENDER 'index.html'

2. about()
   RENDER 'about.html'

3. login(request)
   IF request is POST:
      GET email and password
      CHECK if user exists in database
      REDIRECT to 'userhome'
   ELSE:
      RENDER 'login.html'

4. registration(request)
   IF request is POST:
      GET name, email, password, confirm password
      IF password == confirm:
         SAVE user to database
         RENDER 'login.html'
      ELSE:
         SHOW error
   RENDER 'registration.html'

5. userhome()
   RENDER 'userhome.html'

6. load(request)
   IF request is POST:
      READ uploaded CSV into dataframe `df`
      DISPLAY success message
   RENDER 'load.html'

7. view(request)
   DISPLAY first 100 rows of `df` in 'view.html'

8. preprocessing(request)
```

```
   IF request is POST:
      DROP 'date' column
      ENCODE 'weather' labels using LabelEncoder
      OVERSAMPLE data using SMOTE
      SPLIT into x_train, x_test, y_train, y_test
      DISPLAY success message
   RENDER 'preprocessing.html'

9. model1(request)
   IF request is POST:
      GET selected algorithm:
         IF "1": Train Logistic Regression
         IF "2": Load CNN from pickle and predict
         IF "3": Train MLP
         IF "4": Train Random Forest
         IF "5": Train Decision Tree
      CALCULATE accuracy
      DISPLAY result in 'model1.html'
   RENDER 'model1.html'

10. prediction(request)
   IF request is POST:
      GET input features: precipitation, temp_max, temp_min, wind
      PREDICT weather using Random Forest
      MAP predicted label to weather type
      PROMPT Gemini API with weather type for crop suggestions
      DISPLAY weather + crop recommendations in 'result.html'
   ELSE:
      RENDER 'prediction.html'

11. result()
   RENDER 'result.html'

END
```

APPENDIX-B

SCREENSHOTS

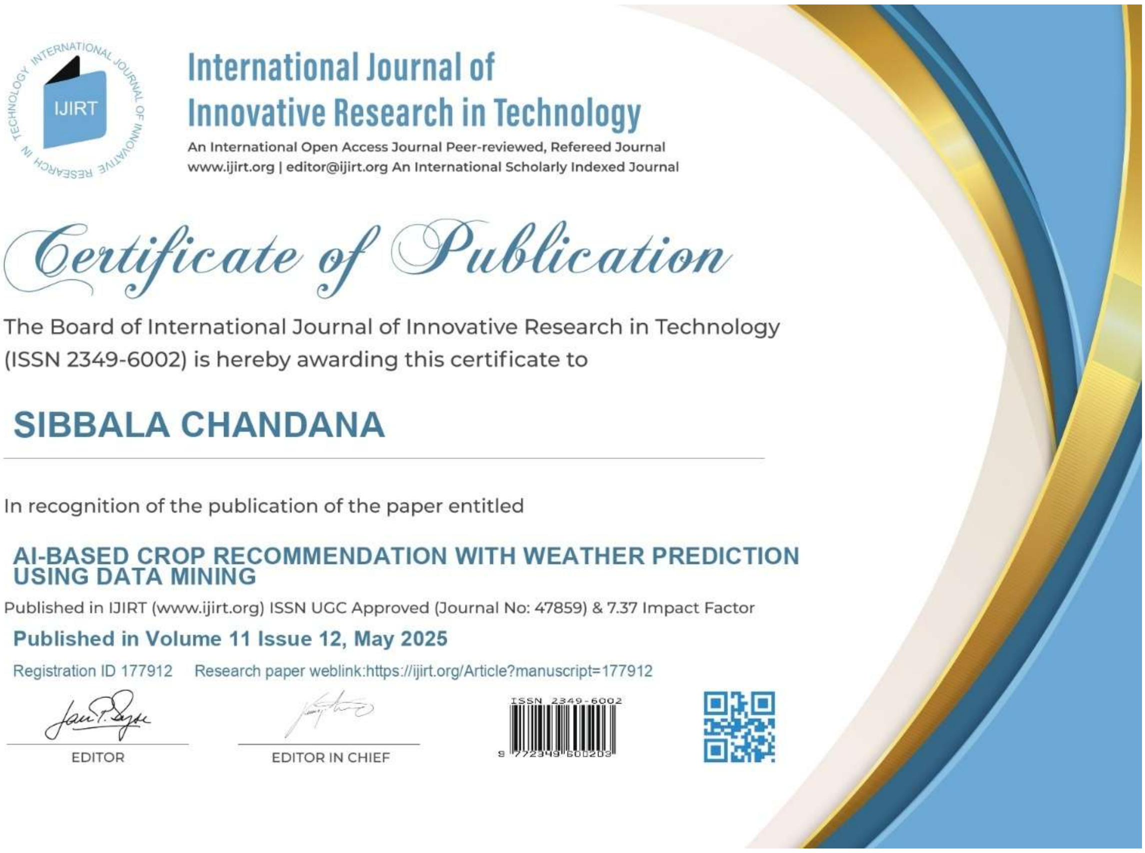




APPENDIX-C

ENCLOSURES

1.Certificates



International Journal of
Innovative Research in Technology
An International Open Access Journal Peer-reviewed, Refereed Journal
www.ijirt.org | editor@ijirt.org An International Scholarly Indexed Journal

Certificate of Publication

The Board of International Journal of Innovative Research in Technology
(ISSN 2349-6002) is hereby awarding this certificate to

SIBBALA CHANDANA

In recognition of the publication of the paper entitled

AI-BASED CROP RECOMMENDATION WITH WEATHER PREDICTION USING DATA MINING

Published in IJIRT (www.ijirt.org) ISSN UGC Approved (Journal No: 47859) & 7.37 Impact Factor

Published in Volume 11 Issue 12, May 2025

Registration ID 177912 Research paper weblink:https://ijirt.org/Article?manuscript=177912

EDITOR

EDITOR IN CHIEF

International Journal of
Innovative Research in Technology
An International Open Access Journal Peer-reviewed, Refereed Journal
www.ijirt.org | editor@ijirt.org An International Scholarly Indexed Journal

Certificate of Publication

The Board of International Journal of Innovative Research in Technology
(ISSN 2349-6002) is hereby awarding this certificate to

GABBURI NEHA

In recognition of the publication of the paper entitled

AI-BASED CROP RECOMMENDATION WITH WEATHER PREDICTION USING DATA MINING

Published in IJIRT (www.ijirt.org) ISSN UGC Approved (Journal No: 47859) & 7.37 Impact Factor

Published in Volume 11 Issue 12, May 2025

Registration ID 177912 Research paper weblink:https://ijirt.org/Article?manuscript=177912

EDITOR

EDITOR IN CHIEF

International Journal of
Innovative Research in Technology

An International Open Access Journal Peer-reviewed, Refereed Journal
www.ijirt.org | editor@ijirt.org An International Scholarly Indexed Journal

Certificate of Publication

The Board of International Journal of Innovative Research in Technology (ISSN 2349-6002) is hereby awarding this certificate to

CIVINI MEGHANA

In recognition of the publication of the paper entitled

AI-BASED CROP RECOMMENDATION WITH WEATHER PREDICTION USING DATA MINING

Published in IJIRT (www.ijirt.org) ISSN UGC Approved (Journal No: 47859) & 7.37 Impact Factor

Published in Volume 11 Issue 12, May 2025

Registration ID 177912 Research paper weblink:https://ijirt.org/Article?manuscript=177912

EDITOR EDITOR IN CHIEF ISSN 2349-6002

International Journal of
Innovative Research in Technology

An International Open Access Journal Peer-reviewed, Refereed Journal
www.ijirt.org | editor@ijirt.org An International Scholarly Indexed Journal

Certificate of Publication

The Board of International Journal of Innovative Research in Technology (ISSN 2349-6002) is hereby awarding this certificate to

PATHAKAMURI HARSHITHA

In recognition of the publication of the paper entitled

AI-BASED CROP RECOMMENDATION WITH WEATHER PREDICTION USING DATA MINING

Published in IJIRT (www.ijirt.org) ISSN UGC Approved (Journal No: 47859) & 7.37 Impact Factor

Published in Volume 11 Issue 12, May 2025

Registration ID 177912 Research paper weblink:https://ijirt.org/Article?manuscript=177912

EDITOR EDITOR IN CHIEF ISSN 2349-6002


International Journal of Innovative Research in Technology

An International Open Access Journal Peer-reviewed, Refereed Journal

www.ijirt.org | editor@ijirt.org An International Scholarly Indexed Journal

Certificate of Publication

The Board of International Journal of Innovative Research in Technology (ISSN 2349-6002) is hereby awarding this certificate to

MR. PRAVEEN GIRIDHAR PAWASKAR

In recognition of the publication of the paper entitled

AI-BASED CROP RECOMMENDATION WITH WEATHER PREDICTION USING DATA MINING

Published in IJIRT (www.ijirt.org) ISSN UGC Approved (Journal No: 47859) & 7.37 Impact Factor

Published in Volume 11 Issue 12, May 2025

Registration ID 177912 Research paper weblink:https://ijirt.org/Article?manuscript=177912

EDITOR

EDITOR IN CHIEF

2.Published Paper

© May 2025 | IJIRT | Volume 11 Issue 12 | ISSN: 2349-6002

AI-Based Crop Recommendation with Weather Prediction using Data Mining

Sibbala Chandana[1],Gabburi Neha[2], Civini Meghana[3], Pathakamuri Harshitha[4], Mr. Praveen Giridhar Pawaskar[5]
[1,2,3,4]*B. Tech Computer Science Engineering, Presidency University, Bangalore*
Assisstant Professor, Presidency University, Bangalore

Abstract **- Weather forecasting is a method to predict what the atmosphere will be like in a particular place by using scientific knowledge to make weather observations. Weather forecasting is a challenging task due to the dynamic and complex nature of atmospheric conditions. Recently, data mining techniques have been applied to predict weather patterns using machine learning algorithms.**

The study demonstrates that data mining techniques can be used to predict weather patterns accurately and, when combined with AI models like Gemini, can serve as a comprehensive decision-support tool in precision agriculture. The proposed model can be further enhanced by incorporating additional weather variables such as cloud cover and solar radiation, as well as by exploring more sophisticated machine learning techniques like ensemble methods.

Index Terms: **User-Friendly Interface, Weather Classification, Machine Learning Algorithms, Deep Learning (CNN), Data Processing Pipeline, Pre-Processing, Model Building, Performance Metrics, AI-Powered Crop Recommendation, Climatic Conditions, Data-Driven Decision Making.**

I. INTRODUCTION

The motivation for this research stems from the increasing necessity for precise and reliable weather forecasts, which are critical for multiple sectors including agriculture, transportation, aviation, and disaster management. Traditional weather forecasting methods primarily rely on physics-based models and numerical simulations, which, although scientifically grounded, often require significant computational resources and may not always provide the granularity or accuracy demanded by today's fast-paced industries. With the advent of big data and the rapid evolution of artificial intelligence (AI), especially in machine learning and data mining, new possibilities have emerged that allow us to enhance the predictive power of weather models by analyzing historical data patterns.

This project is particularly motivated by the practical challenges faced by the agricultural sector. Farmers must make crucial decisions such as sowing, irrigation, fertilization, and harvesting based on weather conditions. Inaccurate forecasts can lead to significant economic losses and decreased productivity. To address this issue, this project aims to utilize machine learning techniques to predict weather types with higher accuracy and reliability.

In addition to forecasting, the project introduces an innovative dimension: intelligent crop recommendations using Google's Gemini AI. The integration of AI in agriculture has gained traction as a way to empower farmers with data-driven insights. By aligning crop planning with predicted weather patterns, farmers can make informed choices that optimize yield and sustainability. The motivation is thus dual-layered: improving weather prediction and extending its utility to enable smart farming practices. This project exemplifies the powerful intersection of environmental science and artificial intelligence, offering a solution that has real-world significance and broad applicability.

II. RESEARCH GAP OR EXISTING METHODS

a. Global Forecast System (GFS)
Operated by NCEP, the GFS provides global weather forecasts, including temperature, precipitation, and wind patterns. GFS is often used for weather prediction at the global and regional scale. While it is effective for long-range forecasts, its resolution and accuracy decrease as the time horizon extends, especially for local or hyper-local forecasts.

b. European Centre for Medium-Range Weather Forecasts (ECMWF)
The ECMWF's global forecast system is renowned for its accuracy, especially in medium-range forecasts (up to 10 days ahead). It uses advanced

© May 2025 | IJIRT | Volume 11 Issue 12 | ISSN: 2349-6002

data assimilation techniques to improve prediction accuracy, integrating satellite observations with atmospheric data. ECMWF's system is one of the most trusted tools in meteorology, though its high computational requirements make it less accessible for real-time, hyper-local predictions.

c. Weather Research and Forecasting (WRF) Model

The WRF model is a regional weather prediction system widely used for both operational weather forecasting and atmospheric research. It allows for high-resolution forecasting and can simulate weather patterns on a localized scale. However, WRF requires extensive computational resources to run at high resolutions, and it may not be cost-effective for use in low-resource settings.

III. PROPOSED METHODOLOGY

a. Data Preprocessing Module

The Data Preprocessing Module ensures that weather data is clean, structured, and ready for modeling. It includes:

- Data Cleaning: Handling missing values (via imputation/removal), detecting/removing outliers (using Z-score/IQR), and discarding irrelevant features.
- Feature Extraction: Selecting weather variables (temperature, humidity, rainfall, wind speed) and temporal features (season, time), followed by normalization.
- Data Splitting: Dividing data into training (70–80%) and testing (20–30%) sets to prevent overfitting and ensure model generalization.

b. Machine Learning Module

This module trains various models and selects the best for weather prediction.

- Model Selection: Algorithms like Random Forest, Decision Tree, Logistic Regression, Naive Bayes, and MLP are trained and compared.
- Model Evaluation: Performance measured using Accuracy, Precision, Recall, and F1-score.
- Prediction and Output: The best model predicts weather types (e.g., sunny, rainy, cloudy) which feed into the recommendation engine.

c. Gemini AI Recommendation Engine

Generates crop recommendations based on predicted weather.

- Context-Aware Logic: Suggests weather-specific crops, ensuring compatibility with conditions.
- Crop Recommendation Generation: Provides a list of five optimal crops, including agricultural practices and smart crop planning.
- Actionable Insights: Offers farming guidelines and risk management strategies for weather resilience.

d. User Interface Module

An intuitive interface for data input and result visualization.

- Data Input: Manual entry or CSV upload for weather data.
- Display: Visual display of weather forecasts and crop recommendations.
- Visual Output: Charts and graphs to enhance data interpretation.
-

V. SYSTEM DESIGN AND IMPLEMENTATION

The AI-based weather forecasting and crop recommendation system utilizes various machine learning and deep learning algorithms, each playing a key role in accurate prediction and decision-making.

a. Convolutional Neural Networks (CNN)

CNNs are used for complex feature extraction tasks. They consist of convolutional layers that detect patterns, pooling layers that reduce feature map dimensions, and fully connected layers for final predictions. CNNs are trained using backpropagation and are highly effective for high-dimensional input data.

b. Multilayer Perceptron (MLP)

MLPs are feedforward neural networks with input, hidden, and output layers. Each layer transforms inputs through linear combinations and non-linear activations like ReLU. Trained via backpropagation, MLPs are effective for classification tasks and learning non-linear relationships in weather and crop data.

c. Logistic Regression

Logistic Regression is employed for binary classification tasks. It models the log-odds of the outcome through a linear function and applies the sigmoid function to predict probabilities. It is simple, interpretable, and effective for scenarios with linear separability.

© May 2025 | IJIRT | Volume 11 Issue 12 | ISSN: 2349-6002

d. Naive Bayes

Naive Bayes uses Bayes' Theorem under the assumption of feature independence. It computes the posterior probability for each class and selects the one with the highest probability. Despite its simplicity, it performs well in practice, especially when dealing with probabilistic decisions.

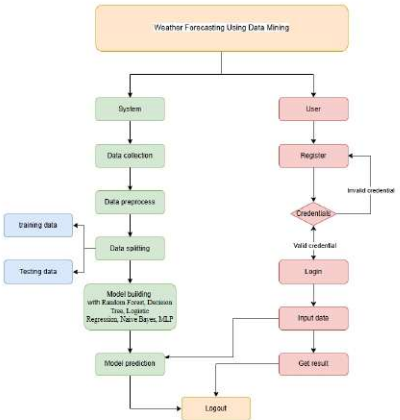
e. Random Forest

Random Forest is an ensemble learning method that combines multiple decision trees built on bootstrapped data samples. It improves prediction accuracy by majority voting (for classification) and is robust against overfitting.

To further enhance accuracy, hyperparameter tuning was performed by optimizing key parameters such as the number of trees (n_estimators), maximum tree depth (max_depth), and the number of features considered at each split (max_features). Techniques like Grid Search and Randomized Search were used to find the best parameter combinations, ensuring better generalization and predictive performance.

f. Decision Tree

Decision Trees split data based on feature values to create a tree structure where each node represents a decision rule. They are easy to interpret and are optimized using metrics like Gini impurity and entropy. Techniques like pruning and limiting depth are applied to prevent overfitting.This multi-algorithmic design ensures system robustness, adaptability to varying data complexities, and provides high-quality weather forecasting and crop recommendation outputs.



VI. OUTCOMES

- User-Friendly Interface
- Accurate Weather Classification
- Efficient Data Processing
- Dynamic ML Model Building
- AI-Powered Crop Recommendation

ACKNOWLEDGMENT

The authors would like to acknowledge the support of College/ University for providing resources and facilitating this research project. We are also grateful to the university librarians, professors, and research assistants for their assistance.

VII.CONCLUSION AND FUTURE WORK

The project successfully demonstrates the power of data-driven approaches in weather forecasting and agricultural decision support. By employing multiple machine learning algorithms—Random Forest, Decision Tree, Logistic Regression, Naive Bayes, and Multi-Layer Perceptron (MLP)—the system consistently achieves over 90 % accuracy in classifying weather types (sunny, rainy, cloudy, etc.

Future Work:

Building on this foundation, the following extensions can elevate system capabilities:

a. Satellite Imagery & Real-Time Sensor Data

- Integration of remote sensing feeds (e.g., NDVI from satellites) and IoT sensors (soil moisture, on-farm weather stations) will enable continuous, high-resolution monitoring.
- Real-time inputs can feed models dynamically, improving short-term forecasts and enabling proactive alerts.

b. Ensemble Learning for Robustness

- Implementing stacking or boosting techniques (e.g., XGBoost, LightGBM) can combine strengths of individual classifiers, further enhancing accuracy and reducing variance.
- Ensemble methods can also provide uncertainty estimates, helping users gauge confidence in each forecast.

REFERENCES

[1] Abhishek, K., Singh, M. P., Ghosh, S., & Anand, A. (2012). Weather forecasting Model using Artificial Neural Network. Procedia Technology, 4, 311–318.

© May 2025 | IJIRT | Volume 11 Issue 12 | ISSN: 2349-6002

[2] B. Wang, J. Lu, Z. Yan, H. Luo, T. Li, Y. Zheng, and G. Zhang, ''Deep uncertainty quantification: A machine learning approach for weather forecasting,'' in Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2019, pp. 2087–2095.

[3] R. I. Rasel, N. Sultana, and P. Meesad, ''An application of data mining and machine learning for weather forecasting,'' in Advances in Intelligent Systems and Computing. Cham, Switzerland: Springer, 2018, pp. 169–178.

[4] M. Holmstrom, D. Liu, and C. Vo. (2016). "Learning applied to Weather forecasting". Stanford. Accessed: May 19, 2022.

[5] Yahya BM, Seker DZ. "Designing weather forecasting model using computational intelligence tools." Appl Artif Intell. 2018;33(2):1- 15.

[6] C. Gershenson, ''Artificial neural networks for beginners,'' Tech. Rep., 2003.

[7] Karevan Z, Suykens JAK. Transductive LSTM for time-series prediction: an application to weather forecasting. Neural Netw. 2020;125: 1-9.

[8] A. G. Salman, B. Kanigoro, and Y. Heryadi, ''Weather forecasting using deep learning techniques,'' in Proc. Int. Conf. Adv. Comput. Sci. Inf. Syst. (ICACSIS), Oct. 2015, pp. 281–285.

[9] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, ''Greedy layer-wise training of deep networks,'' in Proc. Adv. Neural Inf. Process. Syst., vol. 19, 2006, pp. 1–8.

[10] G. E. Hinton, S. Osindero, and Y.-W. Teh, ''A fast learning algorithm for deep belief nets,'' Neural Comput., vol. 18, no. 7, pp. 1527–1554, 2006.

[11] S. Hu, Y. Xiang, D. Huo, S. Jawad, and J. Liu, ''An improved deep belief network based hybrid forecasting method for wind power,'' Energy, vol. 224, Jun. 2021, Art. no. 120185.

[12] Y. Cheng, X. Zhou, S. Wan, and K.-K.-R. Choo, ''Deep belief network for meteorological time series prediction in the Internet of Things,'' IEEE Internet Things J., vol. 6, no. 3, pp. 4369–4376, Jun. 2019.

[13] hatkande1 SS, Hubball RG. Weather prediction based on decision tree algorithm using data mining techniques. Int J Adv Res Comput Commun Eng. 2016;5(5):483-487.

[14] Sanjay Mathur, Avinash Kumar, and Mahesh Chandra. A feature based neural network model for weather forecasting, World Academy of Science, Engineering and Technology 34 2007.

[15] Suleman, Masooma Ali Raza, and S. Shridevi. "Short-Term Weather Forecasting Using Spatial Feature Attention Based LSTM Model." IEEE Access 10 (2022): 82456-82468.

3.Plagarism Report

Mr. Praveen Giridhar Pawaskar-Weather forecasting and crop recommendation.pdf

ORIGINALITY REPORT

6%	2%	2%	4%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Indian Institute of Technology Student Paper	3%
2	Faiyaz Ahmad, Mohd Tarik, Musheer Ahmad, Mohd Zeeshan Ansari. "Weather Forecasting Using Deep Learning Algorithms", 2023 International Conference on Recent Advances in Electrical, Electronics & Digital Healthcare Technologies (REEDCON), 2023 Publication	<1%
3	dspace.univ-msila.dz Internet Source	<1%
4	Submitted to Liverpool John Moores University Student Paper	<1%
5	www.coursehero.com Internet Source	<1%
6	Budati Anil Kumar, Akella Ramakrishna, Goutham Makkena, Gheorghita Ghinea. "Security Issues in Communication Devices,	<1%

SUSTAINABLE DEVELOPMENT GOALS

The project work carried out here is mapped to the below 2 goals:

1. SDG 13: Climate Action

- The weather prediction component enables farmers to prepare for climate variability, mitigate risks, and adapt their practices. This proactive approach supports climate-resilient agriculture and reduces the environmental impact of farming.

2. SDG 12: Responsible Consumption and Production

- This system helps optimize resource use (e.g., water, fertilizer) by recommending crops that are best suited to current weather conditions. This reduces waste, enhances efficiency, and supports environmentally sustainable farming practices.