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 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.

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.

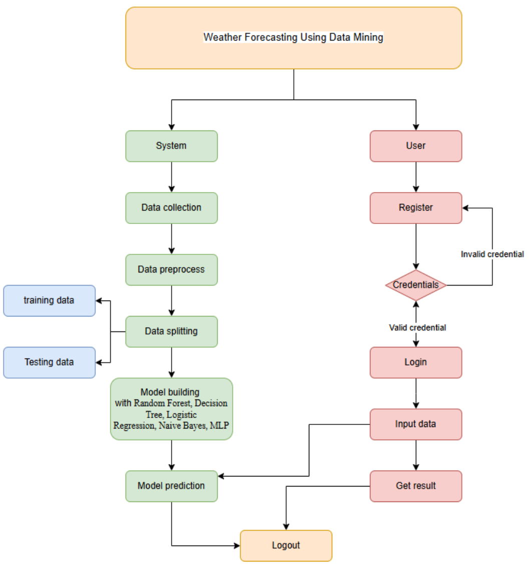
e. Random Forest

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.

[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.