International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 12 | December - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

Crop Yield Prediction using GELM (Gaussian Extreme Learning Machine)

TNVSD. Murali Krishna
2111CS020289
2111CS020289@mallareddyuniversity.ac.in

D. Nagamani
2111CS020290
2111CS020290@mallareddyuniversity.ac.in

M. Shashank
2111CS020291
2111CS020291@mallareddyuniversity.ac.in

M. Nandeeswar
2111CS020292
2111CS020292@mallareddyuniversity.ac.in

S. Nandini
2111CS020293
2111CS020293@mallareddyuniversity.ac.in

G. Nandini
2111CS020294
2111CS020294@mallareddyuniversity.ac.in

Prof. Dr. Komati Satish
Department of CSE(AIML)
MALLA REDDY
UNIVERSITYHYDERABAD

Abstract:* *Accurate prediction of crop yield is crucial for optimizing agricultural practices and ensuring food security. This project presents a novel approach, titled "Soil Factors-Driven Crop Yield Prediction with Optimized GELM (Gaussian Extreme Learning Machine)", aimed at improving crop yield prediction by leveraging soil factors. In this study, we propose an optimized version of the Gaussian Extreme Learning Machine (GELM) algorithm to effectively model the complex relationship between soil factors and crop yield.

I. INTRODUCTION

Modern agriculture is becoming increasingly reliant on data-driven approaches to enhance crop management practices and optimize resource allocation. In this context, machine learning techniques have shown great promise in predicting crop outcomes based on diverse environmental and historical factors. The Gaussian Extreme Learning Machine (GELM) proposed in this research represents a cutting-edge advancement in this domain, offering a fusion of the computational efficiency of ELM with the non-linear modeling capabilities afforded by Gaussian activation functions.

The Extreme Learning Machine (ELM) framework, known for its rapid learning speed and simplicity, forms the basis of our model. By introducing Gaussian activation functions in the hidden layer, we aim to elevate the model's ability to discern complex patterns inherent in agricultural datasets.

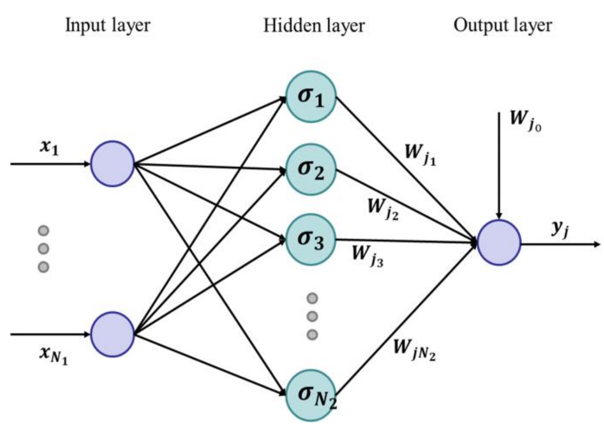

Fig.1: Extreme Learning Machine

This choice is motivated by the desire to better capture the intricate relationships between environmental conditions, soil quality, weather patterns, and historical crop data that influence crop outcomes.

The optimization of the GELM model is a crucial aspect of this research, emphasizing the meticulous calibration of hyperparameters and the selection of relevant features. Through these optimization techniques, we seek to not only enhance the predictive accuracy of the model but also to ensure its applicability in real-world agricultural scenarios.

The primary objectives of this research encompass the development, optimization, and validation of the proposed GELM model for crop prediction. We anticipate that the outcomes of this study will contribute significantly to the advancement of precision agriculture, empowering farmers, researchers, and policymakers with robust predictive tools to make informed decisions about crop management and resource allocation.

II. EXISTING SYSTEM

Traditional Machine Learning Approaches: Support Vector Machines (SVMs): SVMs are a type of supervised learning model that works well for classification tasks. They aim to find a hyperplane that best separates different classes in the feature space.

Decision Trees: Decision trees are tree-like structures that recursively split the dataset based on feature values, leading to a hierarchy of decisions. They are intuitive and easy to interpret but can be sensitive to noise.

Demerits:
Limited Non-Linearity Handling: Traditional machine learning approaches like Support Vector Machines (SVMs) and Decision Trees may struggle to capture intricate non-linear relationships present in complex agricultural datasets.

Sensitivity to Feature Engineering: These methods often require meticulous feature engineering, making them sensitive to the quality and relevance of input features.

Neural Network-Based Approaches: Multilayer Perceptrons (MLPs): MLPs are a type of feedforward neural network where information flows from the input layer through hidden layers to the output layer. They are versatile but may require careful tuning of hyperparameters.

Recurrent Neural Networks (RNNs): RNNs are designed to capture sequential dependencies in data. They are suitable for time-series data but can suffer from the vanishing gradient problem.

Demerits:
Computational Complexity: Neural network architectures, particularly deep learning models like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), can be computationally intensive, requiring significant resources for training and inference.

Data Intensiveness: Deep learning models may demand large amounts of labeled data, which can be a constraint in agricultural scenarios where data collection might be challenging.

Black Box Nature: The interpretability of neural networks can be limited, making it difficult to understand the underlying factors influencing predictions**.**

III. PROPOSED SYSTEM

Gaussian Extreme Learning Machine (GELM): ELM Architecture with Gaussian Activation Functions: The proposed system is built upon the Extreme Learning Machine (ELM) framework, known for its simplicity and rapid learning. The innovation lies in the incorporation of Gaussian activation functions within the hidden layer of the ELM architecture.

Enhanced Non-Linearity Handling: The introduction of Gaussian activation functions enhances the model's capacity to capture complex and non-linear relationships within agricultural datasets, providing a more nuanced representation of the underlying patterns.

Optimization Techniques: Fine-Tuning of Hyperparameters: The proposed system employs optimization techniques to fine-tune hyperparameters, ensuring that the model is configured for optimal performance on diverse agricultural datasets.

Strategic Feature Selection: The inclusion of feature selection methods enhances the model's efficiency by focusing on the most relevant input factors, potentially mitigating issues related to data dimensionality.

Advantages of the Proposed System: Computational Efficiency: Leveraging the inherent advantages of ELM, the proposed system maintains the fast-learning speed characteristic of this framework, contributing to computational efficiency.

Non-Linearity Capture: The utilization of Gaussian activation functions enhances the model's ability to capture intricate non-linear relationships, addressing a common limitation of traditional machine learning approaches.

Applications in Precision Agriculture: Informed Decision-Making: The proposed system has the potential to empower farmers, agricultural researchers, and policymakers with accurate predictions, facilitating informed decision-making in areas such as crop management, resource allocation, and planning.

Precision Agriculture: By providing precise insights into crop outcomes based on various environmental factors, the proposed system aligns with the principles of precision agriculture, where targeted interventions can optimize resource usage.

Future Research Avenues:

Scalability and Generalization: Future research could focus on assessing the scalability of the proposed GELM model to handle larger datasets and evaluating its generalization across diverse agricultural environments.

Integration with Emerging Technologies: Exploring the integration of the GELM model with emerging technologies such as remote sensing, IoT, and advanced data collection methods could further enhance its applicability.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 12 | December - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

IV. ARCHITECTURE

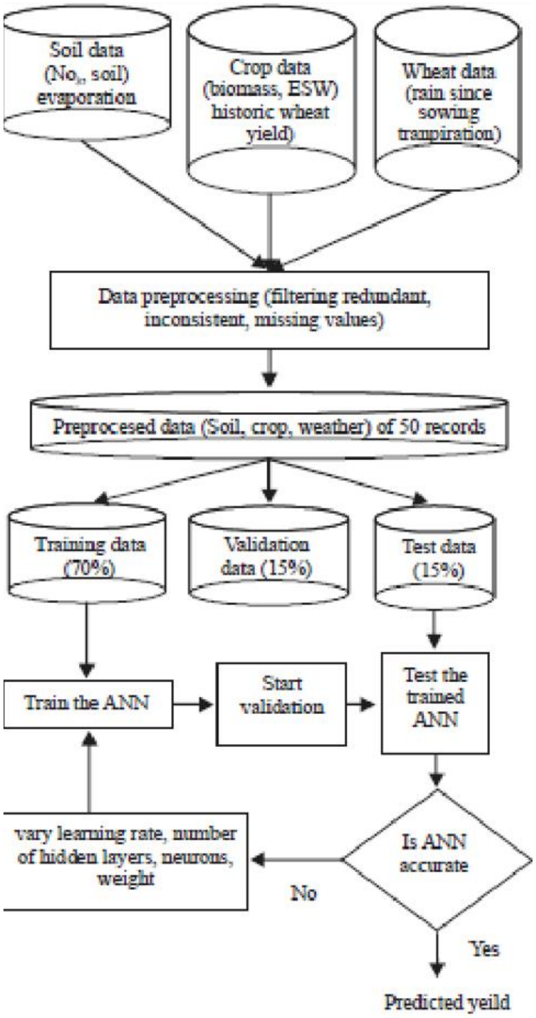

Fig.2. Architecture

Data Collection and Databases:

Data Sources: Gather agricultural data from diverse sources, including environmental sensors, historical crop records, soil quality databases, and meteorological data. Database Integration: Establish a comprehensive agricultural database that incorporates relevant features such as temperature, humidity, soil composition, and past crop yields.

Data Preprocessing:

Cleaning and Imputation: Handle missing values, outliers, and inconsistencies in the dataset through data cleaning and imputation techniques.

Preprocessed Data Splitting:

Training Data (70%): Allocate 70% of the preprocessed data for training the model. This set will be used to teach the GELM the underlying patterns and relationships in the agricultural data.

Validation Data (15%): Dedicate 15% for validation to fine-tune hyperparameters and monitor the model's performance during training.

Test Data (15%): Reserve the remaining 15% for testing the trained GELM's generalization on unseen data, evaluating its predictive capabilities.

Testing the Trained GELM:

Test Data Evaluation: Assess the GELM's performance on the reserved test dataset to evaluate its generalization and predictive accuracy on unseen agricultural data.

Performance Metrics: Employ appropriate performance metrics such as Mean Absolute Error (MAE) or Root Mean Squared Error (RMSE) to quantify the accuracy of the GELM predictions.

Crop Yield Prediction:

Deployment: Once the GELM model is trained and validated, deploy it for real-time or future predictions.

Input Data: Provide relevant input features, such as current environmental conditions and soil attributes, to the trained GELM for predicting crop yield.

Output Prediction: Obtain the predicted crop yield as the output of the GELM, offering valuable insights for decision-making in agriculture.

V. FLOW DIAGRAM

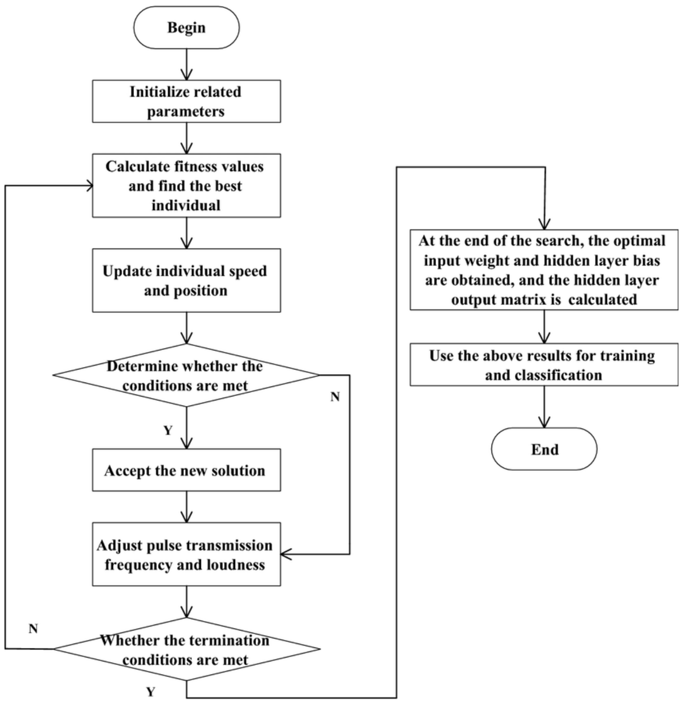

Fig.3.Flow Diagram

V. CONCLUSION

In conclusion, the proposed project, centered around the development of an Optimized Gaussian Extreme Learning Machine (GELM) for crop prediction, represents a significant stride toward enhancing the precision and efficiency of agricultural decision-making. The integration of Gaussian activation functions within the Extreme Learning Machine (ELM) architecture addresses the inherent limitations of traditional machine learning approaches, allowing the model to capture intricate non-linear relationships within complex agricultural datasets.

The architectural framework outlined for the project encompasses a robust pipeline, starting from the aggregation of diverse agricultural data sources to the final prediction of crop yield. The meticulous process of data preprocessing ensures the quality and relevance of the input data, paving the way for effective training, validation, and testing phases. The strategic split of datasets into training, validation, and test subsets facilitates comprehensive model evaluation, promoting the reliability and generalization of the GELM.

As agriculture continues to grapple with the challenges of environmental variability and the need for sustainable practices, the Optimized GELM emerges as a promising tool. Its potential impact on crop prediction can empower farmers, agricultural researchers, and policymakers with timely and accurate information, fostering a data-driven approach to address the complexities of modern farming.

With a continuous improvement cycle and adaptability to changing conditions, the proposed system stands as a beacon for the integration of advanced machine learning techniques in agriculture, ushering in a new era of efficiency and sustainability.

REFERENCES

[1]Python Documentation:

https://docs.python.org/3/library/tkinter [2]MySQL

Documentation: https://dev.mysql.com/doc/

[3] VS Code Documentation:

https://code.visualstudio.com/docs

[4] "Python GUI Programming with Tkinter" by Alan D. Moore

[5] "Python Projects for Beginners" by Connor P. Milliken

[6] "Python and Tkinter Programming" by John E. Grayson

[7] "Tkinter GUI Application Development Blueprints" by Bhaskar Chaudha