



PAVAI
ENGINEERING COLLEGE

Namakkal, Tamil Nadu – 637018, India

Data-Driven Rainfall Prediction Using Ensemble Learning on Long-Term IMD Dataset (2001 – 2024)

Dr. Thirumala Lakshmi K^{*1}, Deepak D², Aravindan V²

¹ Assistant Professor, Department of Electronics and Communication Engineering,
Paavai Engineering College, Namakkal, Tamil Nadu – 637018, India

² UG Students, Department of Electronics and Communication Engineering,
Paavai Engineering College, Namakkal, Tamil Nadu – 637018, India



1. INTRODUCTION



- Accurate rainfall prediction is essential for agriculture, water-resource management, flood preparedness and disaster management.
- Rainfall patterns are influenced by several meteorological variables and show nonlinear and temporal variations.
- This study investigates machine learning and ensemble learning techniques for rainfall prediction using long-term IMD data for Tamil Nadu.



2. OBJECTIVES



- To analyse long-term rainfall patterns over Tamil Nadu.
- To develop machine learning models for rainfall prediction.
- To compare the performance of different ML algorithms.
- To develop an ensemble learning model combining multiple ML techniques.
- To improve rainfall prediction accuracy and robustness.
- To evaluate model performance using precision, recall and F-score.



3. DATASET

Indian Meteorological Department (IMD)

Study Region	Tamil Nadu
Data Source	IMD
Study Period	2001 – 2024
Training Period	2001 – 2020
Validation / Testing	2021 – 2024
Data Type	Historical rainfall & meteorological data

Meteorological Features



Rainfall



Temperature



Humidity

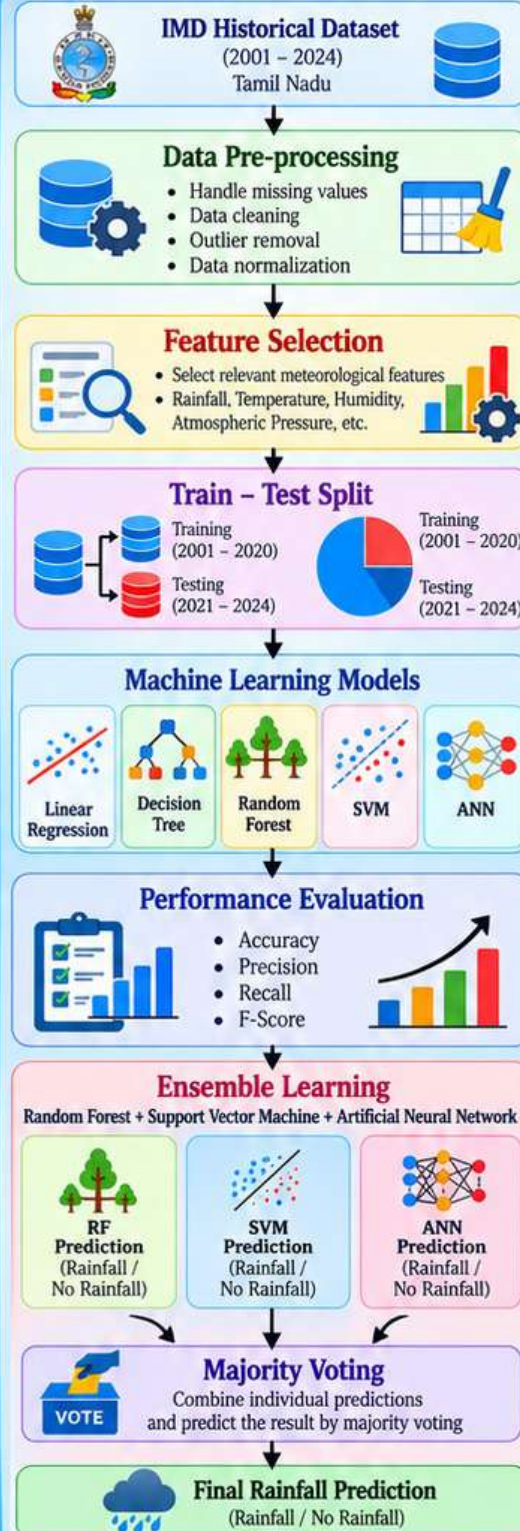


Atmospheric Pressure

These parameters are used as predictors for rainfall forecasting.



4. PROPOSED METHODOLOGY



5. RESULTS

Performance Comparison of ML Models

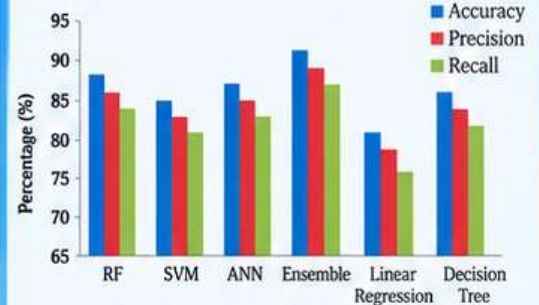


Fig. 1. Accuracy, precision, and recall of machine learning models for rainfall prediction.

Model	Accuracy (%)	Precision (%)	Recall (%)
Linear Regression	80	78	75
Decision Tree	85	83	81
Random Forest	87	85	83
SVM	84	82	80
ANN	86	84	82
Ensemble	90	88	86



Key Finding

The proposed ensemble model achieved approximately **90% accuracy** in rainfall prediction and outperforms individual models.



6. KEY OBSERVATIONS

Individual Models

- Each ML algorithm captures different rainfall patterns.
- Performance varies due to nonlinear relationships.
- Individual models may produce model-specific prediction errors.

Ensemble Model

- Combines predictions from RF, SVM and ANN.
- Captures diverse data patterns.
- Reduces individual model limitations.
- Provides improved prediction robustness.
- Achieves ~90% accuracy.



7. CONCLUSION

- This study demonstrates the potential of machine learning and ensemble learning for rainfall prediction using long-term IMD data.
- The RF-SVM-ANN ensemble model provides improved prediction performance compared with individual models.
- Ensemble learning effectively captures nonlinear rainfall patterns and reduces prediction errors.
- The proposed framework can support agricultural planning, water-resource management, flood preparedness and climate-sensitive decision-making.



8. FUTURE SCOPE

- Incorporation of additional meteorological parameters.
- Integration of satellite and remote-sensing data.
- Development of deep learning models such as LSTM and GRU.
- District-level and location-specific rainfall prediction.
- Real-time rainfall forecasting.
- Development of an early-warning decision-support system.



9. KEYWORDS

Rainfall Prediction, IMD, Machine Learning, Ensemble Learning, Random Forest, Support Vector Machine (SVM), Artificial Neural Network (ANN), Weather Forecasting