International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 16s, 2025
https://www.theaspd.com/ijes.php

Stormsense: A Multi-Model Cloudburst Prediction System Using ML And Deep Image Processing

Tanaya Handore[1], Dr. Shubhangi Handore[2], Prof. Mahendra Handore[3]
[1]AISSMS IOIT, Pune, India tanayahandore12@gmail.com
[2]Trinity College of Engineering and Research, Pune, India handore.shubhangi@gmail.com
[3]Trinity College of Engineering and Research, Pune, India handore.mv@gmail.com

Abstract—*Multiple hazards exist due to cloudburst events be- cause of their abrupt surge and destructive nature. The prediction accuracy faces obstacles because of the multiple atmospheric variables that exist in the environment. The proposed system utilizes Random Forest (RF) and Support Vector Machine (SVM) for numerical meteorological data analysis while Convolutional Neural Networks (CNNs) using VGG16 performs satellite image processing for cloudburst prediction purposes. The system en- ables users to obtain automated weather predictions through a web-based application by either providing input data parameters or uploading satellite images. The system gives real-time alerts through the Pushbullet API to notify users when cloudburst predictions are detected. The combination of numerical inputs with satellite imagery analysis in the evaluation proved successful in developing a better and safer prediction system for disaster preparedness.*

KeyWords—*Cloudburst Prediction, Machine Learning, Hybrid Model (RF-SVM), Satellite Image Processing, CNN (VGG16), Disaster Early Warning*

INTRODUCTION

Heavy localized rainfall known as cloudbursts leads to flash floods that produce extensive environment destruction alongside substantial damage to infrastructure. The volatile nonlinear behavior of atmospheric systems makes it extremely difficult to perform accurate forecasts of such events. Standard numerical weather models that operators frequently utilize at present do not reach sufficient levels of accuracy for de- pendable predictions. The deep learning and machine learning techniques have boosted prediction performance by process- ing structured meteorological data with unstructured satellite imagery. This research develops an ML-based prediction system for cloudburst forecasting which merges meteorological data evaluation through Random Forest (RF) and Support Vector Machines (SVM) alongside satellite images analysis with Con- volutional Neural Networks (CNNs) utilizing VGG16. Under the web-based deployment users can activate either manual weather input or automated satellite image processing for prediction services. The system sends real-time alerts through the Pushbullet API platform which allows users to receive notifications about upcoming cloudburst events.

By integrating numerical weather modeling with satellite image processing, the proposed system improves forecasting accuracy and early warning reliability. The study presents an examination of the system which includes design speci- fications along with implementation strategy and evaluation methodology to prove the effectiveness of ML-based hybrid models for cloudburst prediction and disaster readiness.

LITERATURE REVIEW

Faiyaz Ahmad et al., [1] worked on Machine learning- based weather forecasting models which used LSTM and GRU together with Bi-LSTM yet they found problems in feature selection and computational complexity and high error rates. A combined RF-SVM model has been developed as an answer to current issues by integrating Random Forest to identify significant weather elements while Support Vector Machine increases prediction precision. The implementation method addresses both non-linearities in data effectively along with cost optimization while it provides better predictions for future short-term results.

T. Akilan et al., [2] studied how IoT technology pairs with CNNs to build a weather prediction system that monitors agricultural fields more effectively. The use of deep learning models with CNNs enhances forecasting through spatial and temporal pattern recognition however the systems demand high processing power which leads to difficulties in feature se- lection and error propagation. The proposed solution combines the RF intelligently selected features with SVM-classification accuracy improvement to create a hybrid RF-SVM model.

Supriya S et al., [3] introduces an investigation of hybrid weather forecasting through network integration of LSTM neural networks with Random Forest models to enhance prediction accuracy. Standard models encounter issues with selecting features as well as both complexity in computation and complications with propagation errors. The combination of LSTM neural networks with Random Forest proves beneficial since both systems effectively address different aspects of the forecasting process. The proposed improvement for perfor- mance creation involves an RF-SVM hybrid model.

Vigneshwaran. The paper by Vigneshwaran. V et al., [4] analyzes IoT-based rainfall prediction through combined usage of ANN and Intensified LSTM models. Deep learning provides better accuracy but several difficulties persist due to weather variable selection and high processing needs and prediction mistakes.

The research of Mar Mar Soe [5] predicts rainfall through SoftMax logistic regression by analyzing past weather data from eleven cities in Myanmar. Numerical and statistical prediction models encounter limitations when dealing with in- tricate patterns along with choosing essential features because they require powerful computing systems. Machine learning employs Random Forest as well as Decision Trees and SVM to increase results accuracy but deep learning models need significant processing resources to function properly. The sug- gested solution to handle these barriers involves implementing a combined RF-SVM framework.

Amol Ashok Patil et al., [6] introduces a rainfall prediction system that merges Global Forecast System (GFS) models with machine learning algorithms in order to enhance forecasting accuracy. The accuracy of numerical weather prediction mod- els remains limited because they face different types of errors in combination with computational complexity alongside bi- ases. To enhance the process of rain classification and intensity prediction relies on support vector machines coupled with artificial neural networks and XGBoost as machine learning models. Multiple hurdles exist during the selection of features together with model optimization and accuracy of predictions. The study by D.Vasudeva Rayudu et al., [7] evaluates accurate rainfall prediction between artificial neural networks (ANN) and deep learning neural networks (DNN). Weather forecasting methods encounter three main obstacles while being proposed including non-linear features and difficulty in computational processing and finding suitable attributes. Research has been conducted to improve accuracy levels by applying ANN, DNN and hybrid models from machine learning techniques. Deep learning models need large compu- tational resources at the same time they can develop specific patterns that lead to overfitting issues. The addressed issues in these challenges can be managed through a combination of RF-SVM models.

The random forest algorithm gets studied by Shivam Mishra et al. [8] to supervise weather-dependent operations throughout agriculture as well as industry. Current weather forecasting technologies experience three main operational limitations by being unreliable regarding data integrity and needing super- computer power along with delayed response times. Artificial neural networks (ANN) and support vector machines (SVM) and decision trees together with other machine learning models help predict systems but need high processing capacity from deep learning models.

Overview of the Model

RF-SVM works as a data processing platform that special- izes in solving meteorological problems through its feature set management capabilities for extensive data volumes along with sophisticated connection patterns. RF offers an ensemble-tree approach for feature selection by discarding weak information attributes to enhance the model performance and make it more resistant to noise effects. SVM accepts an optimized feature subset from the initial selection because kernel optimization

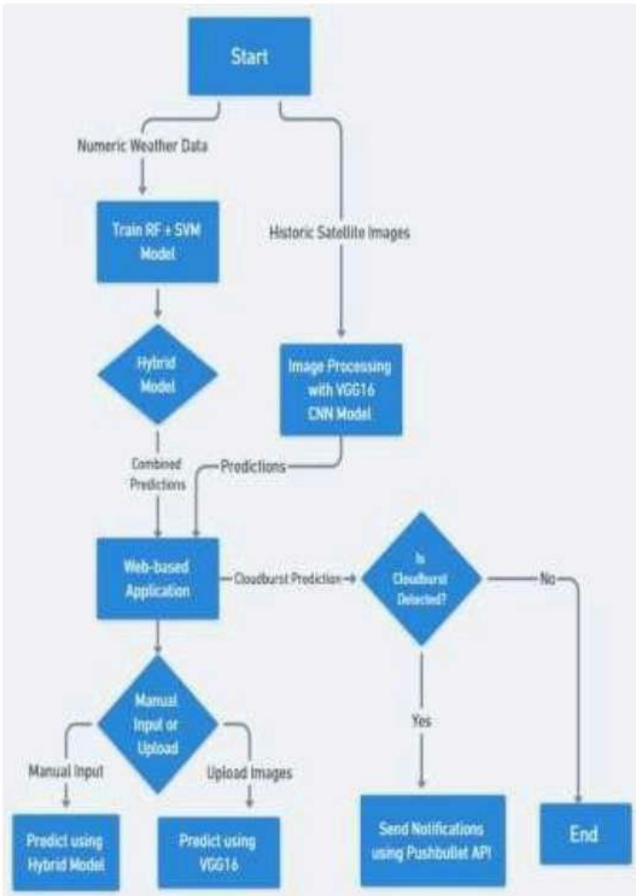

Fig. 1. Flowchart of the proposed system.

of its decision boundary enables it to perform robust nonlin- ear classification. The predictive accuracy of SVM increases through maximizing class separation and simultaneously it prevents model overfitting. RF-SVM unifies into one system that manages bias-variance equilibrium for high-quality and high-efficiency cloudburst forecasting.

A satellite imagery analysis system operates with the VGG16-based CNN architecture to analyze numerical meteo- rological data automatically. The VGG16 deep convolutional model pre-trained on large-scale image datasets, modifies its settings in order to recognize critical meteorological signatures such as storm formation and cloud shape changes during detection of cloudburst events. Regional patterns together with textural irregularities find their detection through the hierarchi- cal feature extraction capabilities of the CNN framework and its convolutional and max-pooling layers when used on remote sensing data. The system can detect precise cloudburst events in specific domains by utilizing transfer learning and model fine-tuning techniques.

The hybrid system utilizes RF to examine satellites images for non-linear correlations before SVM optimizes bound- aries while extracting spatial-temporal information with CNN. Through its combination approach the system delivers superior forecast accuracy and provides both operational stability and real-time quote alerts on a large scale.

SYSTEM ARCHITECTURE

A multi-paradigm computational framework in our system brings together diverse meteorological data sources using ensemble machine learning operations which achieve peak

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 16s, 2025
https://www.theaspd.com/ijes.php

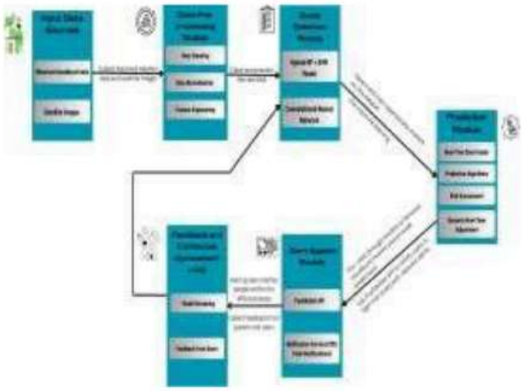

Fig. 2. System Architecture Overview.

accuracy and warning time effectiveness. This system uses an architecture designed with modular, independently functioning components that enable seamless data flow and supports flex- ible data transfer alongside parallel processing, and real-time decision-making. Each architectural unit performs specialized tasks, ranging from data acquisition and feature extraction to model analysis and the final dissemination of localized alerts.

A. Input Data Sources

Historical Meteorological Dataset: Numerical historical records containing 96,453 records with 12 meteorolog- ical attributes, including precipitation intensity together with data of temperature whole humidity, atmospheric pressure, wind speed and wind bearing allow the system to find patterns and anomalies that affect cloudburst occurrences.

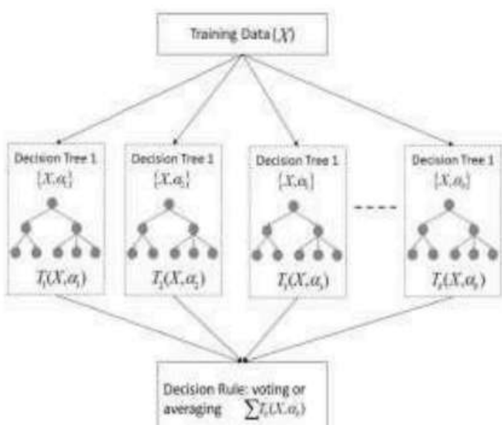

Fig. 3. Working of RF.

C. Model Training and Selection

The system uses an ensemble learning architecture which merges Random Forest (RF) + Support Vector Machine (SVM) and Convolutional Neural Network (CNN - VGG16) for combining their analytical powers.

1) Random Forest (RF) + Support Vector Machine (SVM) Hybrid Model: RF is an ensemble machine learning algorithm developed by adding additional layers of decision trees through random sampling of training data and subsets of features while splitting nodes. The decision trees $T_i(x)$ receive their data from randomly chosen subsets of the sample set and random feature subsets. The final prediction is made by aggregating the outputs from all trees:

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 16s, 2025
https://www.theaspd.com/ijes.php

1 Σ n

- Historical Satellite Imagery Dataset: The model was fine- tuned on the Roboflow Cloud Burst dataset which uses
- $\hat{y}=$

n

i=1

$T_i(x)$ (1)

high-resolution meteorological satellite images that have been properly selected to monitor and analyze both cloud evolution and severe weather patterns.

B. Data Preprocessing Module

This application retrieves current meteorological observa- tions from Firebase while obtaining weather archive data from CSV dataset. The system maintains an uninterrupted data supply which helps perform continuous model updates. The Data Preprocessing Module tidies up raw datasets by standardizing their format and by transforming data to enhance the efficiency of utilized models. Statistical and interpolation techniques help detect problematic data points followed by valuation insertion that replaces missing or faulty invaluable contents. Multicomponent meteorological data undergoes nor- malization procedures that create equivalent ranges to ensure data consistency from various sources. The process of feature engineering includes extracting predictors from the domain which consist of convective cloud patterns together with temperature gradients and precipitation intensity distributions and wind shear dynamics since these factors strongly affect cloudburst development.

where $T_i(x)$ represents the prediction from the ith decision tree, and n is the total number of trees in the forest.

It has emerged as a quite efficient and robust algorithm for our system that can deal with feature selection, even with higher numbers of features. It was additionally, especially proficient while managing missing data, large data without preprocessing, and rescaling. Unlike other black-box algo- rithms, the RF is trained by a bootstrap aggregating (bagging) algorithm; it improves model stability, accuracy of individual trees, reduces the variance, and overfitting, thus enhancing the accuracy of prediction.

Following feature selection by RF, the SVM model pro- cesses the refined feature set to perform non-linear classi- fication. SVM establishes the best fitting hyperplane which basically maximizes the margin
2 to categorize different classes throughout the feature space. SVM implements radial basis functions among other kernel functions to process non- chevron separable data types which ar el typical in cloudburst prediction. SVM functions to achieve maximum separation between different groupings of data as:

$$y_i(w \cdot x_i + b) \geq 1, \qquad \forall i \qquad (2)$$

where w represents the weight vector, x_i the input features, and y_i the class labels.

2) *Convolutional Neural Network (CNN - VGG16):* The satellite imagery contains multi-spectral cloud structures that undergo processing through the VGG16 implementation of the CNN model. The convolution operation F applies to input images I through the following mathematical formula:

$$F(I) = I * K \qquad (3)$$

The CNN framework starts with convolutional layers that lead to pooling layers which together extract increasing levels of image features from the input image. Each convolutional layer activates a ReLU function before pooling layers reduce feature map resolution:

$$A = \mathrm{ReLU}(W \cdot X + b) \quad (4)$$

Where A is the output of the convolutional layer, W is the filter weight, X is the input feature map, and b is the bias.

Transfer learning enables the VGG16 model to use pre- trained weights which become adjusted based on the new dataset for cloudburst detection. During this process the model updates its weights through backpropagation and gradient descent algorithm with L as the loss function to reduce wrong prediction errors. Weight adjustments become possible through the gradient calculations of the loss function against model parameters to optimize performance on the target dataset.

$$\theta_{\text{new}} = \theta_{\text{old}} - \eta\nabla_\theta L(\theta) \quad (5)$$

where θ represents the model parameters (weights), η is the learning rate, and $\nabla_\theta L(\theta)$ is the gradient of the loss function.

D. Alert Dissemination and Risk Communication Module

Through the Alert System, organizations achieve time- sensitive risk communication that distributes information over various channels. The system distributes alerts using the Push- Bullet API which enables notifications through push messages. Warnings become active only when the defined thresholds for risk factors meet their criteria, thus decreasing false alarm activations.

E. Feedback and Continuous Learning Module

The system employs a feature that implements several methods to boost system adaptability and predictive accuracy which includes:

- **User Feedback Loop:** Aggregates accuracy assessments from meteorologists and disaster response teams to refine model predictions. The system updates its knowledge base periodically by using new labeled data for sustained forecasting accuracy over time. The cloudburst prediction system continually updates itself through its self-learning ability so it can adapt properly to changing climatic patterns.

TABLE I COMPARATIVE ANALYSIS

Models	**Dataset**	**Results**
This paper (Hybrid RF + SVM)	96,453 records, 12 meteoro- logical attributes	99% Accuracy
This paper (VGG16, CNN)	320 Satellite Image Dataset	83.33% Accuracy
Ref.[3] (LSTM + RF)	Historical weather data	MAE: 0.0126
Ref.[5] (SoftMax Regression)	Weather data from 11 cities	83% Accuracy
Ref.[1] (LSTM, GRU, Bi-LSTM)	Time-series dataset	RMSE: 0.1232(LSTM)

KEY DIFFERENTIATORS

- **Hybrid Learning Architecture:** Hybrid RF + SVM ensemble utilizes RF's low variance in addition to SVM's low bias, which addresses the common model weakness of selecting between fitting and underfitting.

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 16s, 2025
https://www.theaspd.com/ijes.php

- **Superior Generalization Capability:** The integration of randomized feature selection (RF) with optimal margin hyperplanes (SVM) results in better classification accu- racy because it enables effective complex non-linear re- lationship and high-dimensional decision boundary mod- eling.
- **Multi-Source Data Integration:** The system combines numerical weather data with satellite images through a unified framework which strengthens predictive accuracy by creating strengthened interpretative and robust extrac- tion features.
- **Dynamic Decision Boundary Adaptation:** The pro- posed hybrid model dynamically modifies its decision boundaries for complex data sets, thus it helps decrease classification errors and enhances cloudburst prediction accuracy.
- **Scalable Real-Time Implementation:** The built-in web- based solution provides immediate system access for real-time evaluation of meteorological conditions, which supports early disaster protection through proactive strate- gies.

OBSERVATIONS AND RESULTS

The Hybrid RF+SVM model, trained on historical meteoro- logical data, achieved a high accuracy of 99%, demonstrating its ability to generalize effectively. Meanwhile, the VGG16- based CNN model, designed for satellite image classification, demonstrated an 83.33% accuracy.

The proposed Hybrid RF+SVM Model reaches a 99% clas- sification accuracy, thus outperforming other current method- ologies. The method exceeds LSTM-based forecasting [1] ca- pabilities by using SVM for boundary optimization while Ran- dom Forest contributes to reducing variance, thus improving robust classification performances. The ensemble framework of our approach outperforms SoftMax Regression [2] because it manages complex meteorological patterns effectively and

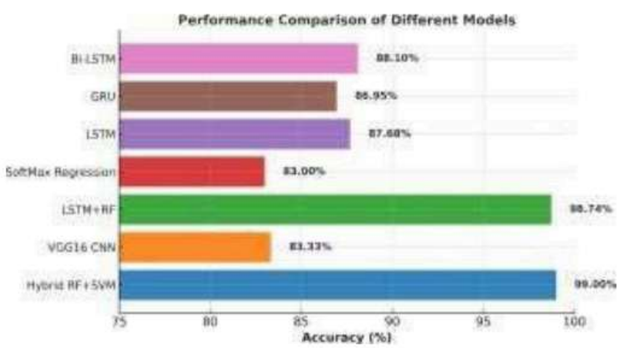

Fig. 4. Comparison of Classification Models.

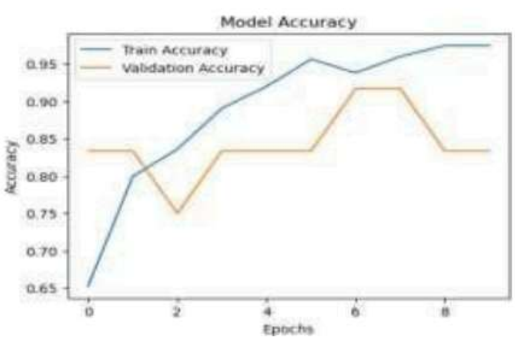

Fig. 5. Training and Validation Accuracy.

strikes the best balance between bias and variance, leading to more reliable predictions. Activities based on deep learning models including LSTM, GRU, and Bi-LSTM offer excellent series prediction [3] yet

they do not fulfill essential real- time satellite image processing needed for cloudburst warning and emergency systems. The main asset of our system stems from its ability to combine numerical meteorological data with image data from different sources, leading to joint- validated predictions. Our CNN-based VGG16 model shows excellence in feature extraction from satellite imagery through meteorological attributes because it reaches an accuracy level of 83.33%. This exceeds traditional time-series forecasting models.
As shown in Figure 5, the validation accuracy stabilizes around 83.33%, confirming the model's robustness in real- world cloudburst detection scenarios.

CONCLUSION

The study presents an advanced Cloudburst Prediction Sys- tem which employs Random Forest (RF) and Support Vector Machine (SVM) jointly with deep learning-based image clas- sification (VGG16) to produce enhanced predictions through their shared benefits and to detect cloudburst events based on satellite images. The web-based application deployment pro- vides real-time access to meteorological agencies for making decisions based on the analysis.

FUTURE SCOPE

Future research needs to integrate dimensional reduction methods, feature engineering principles and model inter- pretability approaches in order to enhance class identification capabilities and minimize computational load. Through the combination of advanced augmentation techniques together with hyperparameter optimization, researchers can optimize the performance achieved by CNN-based classification mod- els. Predicting cloudburst events becomes more effective by leveraging the time-dependent weather patterns through LSTM networks or Transformer-based models. Variations in proba- bility scores together with occasional misclassifications occur due to inconsistencies in image orientation that has been introduced. The future development should add automated rotation correction together with edge cropping and adaptive resizing for enhancing feature representation uniformity, which results in improved detection accuracy of cloudburst events.

REFERENCES

[1] F. Ahmad et al., "Weather Forecasting Using Deep Learning Algo- rithms," *Proc. 2023 Int. Conf. Recent Advances in Electrical, Electronics & Digital Healthcare Technologies (REEDCON)*, 2023, pp. 498–502, doi: 10.1109/REEDCON57544.2023.10150439.

[2] T. Akilan et al., "Enhanced IoT-Based Weather Forecasting and Field Monitoring System Utilizing Multiple CNN Classification Models," *Proc. 2023 5th Int. Conf. Adv. Comput., Commun. Control Netw. (ICAC3N)*, 2023, pp. 961–966, doi: 10.1109/ICAC3N60023.2023.10541855.

[3] S. Supriya et al., "Hybrid Weather Forecasting: Integrating LSTM Neural Networks and Random Forest Models for Enhanced Ac- curacy," *Proc. 2024 Int. Conf. Recent Adv. Electrical, Electron- ics, Ubiquitous Commun. Comput. Intell. (RAEEUCCI)*, 2024, doi: 10.1109/RAEEUCCI61380.2024.10547997.

[4] V. Vigneshwaran et al., "IoT Based Weather Forecasting for Rain- fall Prediction Using Intensified Long Short-Term Memory Artificial Neural Network," *Proc. 2024 Int. Conf. Signal Process., Comput., Electron., Power Telecommun. (IConSCEPT)*, 2024, doi: 10.1109/ICON- SCEPT61884.2024.10627908.

[5] M. M. Soe, "Rainfall Prediction Using Regression Model," *Proc. 2023 IEEE Conf. Comput. Appl. (ICCA)*, Yangon, Myanmar, Feb. 27-28, 2023,
pp. 113–117, doi: 10.1109/ICCA51723.2023.10182116.

[6] A. A. Patil et al., "A Hybrid Machine Learning - Numerical Weather Prediction Approach for Rainfall Prediction," *Proc. 2023 IEEE In- dia Geosci. Remote Sens. Symp. (InGARSS)*, 2023, doi: 10.1109/In- GARSS59135.2023.10490397.

[7] D. V. Rayudu et al., "Accurate Weather Forecasting for Rainfall Predic- tion Using Artificial Neural Network Compared with Deep Learning Neural Network," *Proc. 2023 Int. Conf. Artif. Intell. Knowl. Dis- cov. Concurrent Eng. (ICECONF)*, 2023, pp. 1–6, doi: 10.1109/ICE- CONF57129.2023.10084252.

[8] S. Mishra et al., "Controlling Weather Dependent Tasks Using Random Forest Algorithm," *Proc. 2020 Third Int. Conf. Adv. Electron., Comput. Commun. (ICAECC)*, 2020, doi: 10.1109/ICAECC50550.2020.9339508.