AUIQ Technical Engineering Science

Manuscript 1025

Classification of Daily Weather Conditions using Decision Tree-Based Machine Learning Models: A Case Study of Kabul, Afghanistan

Ahmad Bilal Ahmadullah

Ahmad Shah Irshad

Basir Ahmad Khaled

AUIQ TECHNICAL ENGINEERING SCIENCE 2025;2:62–82

Scan the QR to view the full-text article on the journal website

ORIGINAL STUDY

Classification of Daily Weather Conditions using Decision Tree-Based Machine Learning Models: A Case Study of Kabul, Afghanistan

Ahmad Bilal Ahmadullah [a,*], Ahmad Shah Irshad [a], Basir Ahmad Khaled [b]

[a] Department of Energy Engineering, Engineering Faculty, Kandahar University, Kandahar 3801, Afghanistan
[b] Department of Hydraulics and Hydraulic Structures, Faculty of Environmental and Water Resources Engineering, Kabul Polytechnic University, Kabul 1010, Afghanistan

ABSTRACT

As global climate variability intensifies, the need for accurate and reliable weather forecasting becomes increasingly important. This study aimed to classify daily weather conditions in Kabul, Afghanistan, by comparing two decision-tree-based machine learning (ML) models that includes Decision Tree Classifier (DTC) and Extra Trees Classifier (ETC). A complete year dataset consisting of 366 daily meteorological observations collected from a central weather station in the region for 2024 was used. Results revealed that the DTC model consistently outperformed the ETC model, obtained an overall accuracy of 99% in both the training and testing phases, compared to the ETC model's accuracy of 96% (training) and 89% (testing). Specifically, the DTC model showed almost perfect weighted-average precision (0.99), recall (0.99), and F1-scores (0.99) for both phases training and testing respectively, whereas ETC demonstrated lower metrics in testing phase with weighted-average precision of 0.90, recall of 0.89, and F1-score of 0.88. Furthermore, sensitivity analysis demonstrated that precipitation probability is 40%, cloud cover is 31%, snow is 18%, temperature fleeks like max is 8%, and solar radiation is 3% as the most impactful variables in weather classification. Scientifically, this study contributes to enhancing the effectiveness of localized weather prediction, providing critical support for urban planning, agriculture, and disaster management decisions in regions with similar climatic conditions.

Keywords: Climate analytics, Predictive meteorology, Environmental decision-making, Advanced weather forecasting, Regional climate modeling

1. Introduction

Weather condition classification is a fundamental aspect of meteorological science [1], leveraging the predictive capabilities of various analytical models to enhance forecasting accuracy [2]. As global climate variability increases, the need for precise and reliable weather prediction becomes more critical [3, 4]. Traditional meteorological methods have evolved with the integration of ML techniques [5, 6], which offer significant improvements in the processing and interpretation of large datasets typical in weather analysis [2, 7]. These advancements have enabled more nuanced classifications of weather patterns, essential for timely decision-making in areas ranging from aviation to agriculture and disaster management [8, 9]. Further, the integration of ML techniques in meteorological applications has not only optimized the predictive accuracy of weather models but also broadened the scope of their applicability across different environmental conditions [10, 11]. The ability of ML to handle complex, nonlinear relationships between meteorological factors is transforming weather forecasting into a more dynamic and responsive tool [12, 13], paving the way for innovations in how we understand and react to atmospheric phenomena.

Received 9 February 2025; revised 6 April 2025; accepted 10 April 2025.
Available online 30 April 2025

* Corresponding author.
E-mail addresses: bilal.kafg@gmail.com (A. B. Ahmadullah), irshad786.kdru@gmail.com (A. S. Irshad), basir.a.khaled@gmail.com (B. A. Khaled).

https://doi.org/10.70645/3078-3437.1025

Numerous studies have explored the application of ML techniques in weather classification, each contributing unique insights and methodologies to address the challenges of meteorological forecasting. For instance, a study conducted in Jiangyin, Jiangsu, China, employed autoencoders (AEs) for classifying dry and wet periods using commercial microwave link (CML) data and rain sensors [14]. The researchers achieved high true positive rates (TPR) and true negative rates (TNR), effectively leveraging signal attenuation data from CMLs [14]. However, the study's dependency on accurate classification of dry periods and the risk of overfitting on unbalanced datasets presented challenges, particularly for light rain events [14]. Similarly, a research based in Chengdu, China, explored a multi-model fusion approach using ResNet50, ResNet101, and DenseNet121 for classifying weather images into nine types, including rain, dew, snow, and fog [15]. While the study achieved a notable classification accuracy of 81.25%, the limited dataset size impacted the robustness and scalability of the models, highlighting the importance of larger datasets for generalization [15]. Further, in another notable study, conducted in Hefei, China, authors used convolutional neural networks (CNNs) integrated with Mask R-CNN for edge extraction, significantly improving the accuracy of weather classification across four types encompassing sunny, foggy, rainy, and snowy [16]. This approach achieved impressive accuracy rates, such as 97.94% for sunny and 98.22% for snowy conditions, but its performance under nuanced or mixed weather conditions remained unexplored [16]. Additionally, a study leveraging Himawari-8 infrared data in Langfang, China, employed XGBoost to classify cloud types with obtaining an overall accuracy of 86.22% [17]. Although the model demonstrated robust performance across temporal and seasonal variations, its reliance on Himawari-8 data limited its applicability to other satellite systems or regions with different data availability [17]. A more detailed review of literature is summarized in Table 1, signifying the diverse applications and advancements in ML for weather and environmental classification.

While significant progress has been made in applying ML techniques for weather classification, previous studies primarily employed generalized models without explicitly addressing localized daily weather prediction in areas characterized by complex climates. Specifically, regions such as Kabul, Afghanistan, with distinct climatic variability and topographical complexity, remain underrepresented in prior literature, leaving a notable gap regarding practical and locally-adapted ML applications. To bridge this gap, this study explicitly compared the DTC and ETC, leveraged daily meteorological data specific to Kabul to get improved accuracy, interpretability, reliability, and reproducibility in localized weather predictions. Consequently, this research significantly contributed to the scientific understanding by introducing a comparative decision tree-based modeling approach clearly adopted for Kabul's unique climatic and environmental conditions. Beyond local applications in agriculture, urban planning, and disaster management, this study offers a replicable methodological framework to guide similar ML implementations in other regions experiencing limited data availability, high climatic variability, or complex environmental conditions.

This study aims to address key gaps in the field of weather classification by leveraging ML techniques to tackle region-specific challenges, with the objectives of: (i) classifying weather conditions in Kabul, Afghanistan, using decision tree-based ML models trained on a comprehensive dataset; (ii) evaluating the accuracy and performance of these models in predicting local weather patterns and identifying meteorological complexities specific to the region; and (iii) contributing to the local and global in terms of scientific and practical in understanding of how tailored ML approaches can be applied effectively in underrepresented regions with unique climatic and environmental characteristics.

2. Methods and materials

2.1. Models development

The development of the ML models in this study involved a structured and systematic process comprising several critical phases, as illustrated in Fig. 1. Initially, the methodology started with defining the study region (Kabul Province, Afghanistan), considering its unique meteorological significance. Subsequently, daily meteorological data from a centrally located weather station in Kabul were collected and analyzed comprehensively for quality and consistency, ensuring reliability for model training. After the preliminary data analysis, the dataset was randomized to remove any biases and then divided into two subsets of 70% for model training and 30% kept for model testing. During the training phase, two decision-tree-based ML models including DTC and ETC were developed and trained. Each model underwent systematic hyperparameter tuning, leveraging a grid search approach with cross-validation to identify optimal parameter configurations that maximized predictive accuracy. Upon completing model training, each model's performance was assessed using the testing dataset. Performance metrics included accuracy, precision, recall, F1-score, ROC-AUC, and

Table 1. Summary of literature review on ML applications in weather and environmental classification.

Reference	Study Region	Data Type	Inputs	Outputs	ML Models Used	Key Findings	Limitations
[14]	Jiangyin, Jiangsu, China	CML data and rain sensor data including rain gauge and OTT disdrometer data.	Signal attenuation data from CMLs	Classification of periods as either 'dry' or 'wet'	AEs	Autoencoders effectively classified dry and wet periods using only dry period data, with slightly better performance than the rolling standard deviation method. TPR and TNR for period classification.	Dependency on accurate classification of dry periods due to model training on dry period data only. Potential issues with model overfitting due to unbalanced datasets and challenges in handling light rain events effectively.
[18]	Piscataway, NJ, USA; Chongqing, P.R.C	Weather images	Weather images	Weather condition classification (sunny, cloudy)	Deep CNNs	Achieved 82.2% normalized classification accuracy, significantly outperforming the state of the art with a 54.8% relative improvement. Demonstrated that CNNs effectively capture non-linear mappings between input images and weather categories.	The study is limited to binary classification (sunny, cloudy) and may not address other complex weather conditions. Also, the model's reliance on large sets of training data might limit its applicability in environments where such data are not readily available.
[19]	University of Essex, UK; University of Birmingham, UK	Weather images	Weather images	Classification of weather conditions into rain, snow, and fog	CNNs	Created a new open-source dataset called RFS (Rain Fog Snow) and proposed a novel data augmentation technique using superpixel delimiting masks, achieving improved classification results with several CNN architectures.	The study focused on general-purpose images and may not effectively address specific challenges like varying lighting and angles which are critical in practical applications like driving assistance systems.
[15]	Chengdu, China	Weather images	Weather images of 9 types including rain, dew, snow, frost, fog, ice, hail	Classification of multiple weather types	ResNet50, ResNet101, DenseNet121	The study proposed a multi-model fusion approach using ResNet and DenseNet for weather classification, achieving a classification accuracy of up to 81.25% on a diverse dataset.	Limited by the size of the dataset, which affects the robustness and scalability of the model. Deep models require large datasets to generalize effectively across different weather phenomena.

(Continued)

Table 1. Continued

Reference	Study Region	Data Type	Inputs	Outputs	ML Models Used	Key Findings	Limitations
[17]	Langfang, China	Himawari-8 infrared data	Latitude/longitude, 10 infrared channels, 5 brightness temperature differences	Cloud type classification	XGBoost	The model achieved an overall accuracy of 86.22% and demonstrated strong performance across all-day temporal, daytime/ nighttime, and seasonal scenarios.	Limited to Himawari-8 data, may not generalize to other satellite systems or regions without similar data availability.
[20]	Australia	Weather observations from the Australian Government's Bureau of Meteorology	Weather observations including 'RainToday'	Prediction of 'RainTomorrow'	Logistic Regression, SVM, Random Forest	Random Forest outperformed SVM and Logistic Regression in terms of ROC-AUC scores, demonstrating superior performance in classifying whether it will rain tomorrow. ROC-AUC for Random Forest was 0.98, compared to SVM (0.89) and Logistic Regression (0.88).	The study is limited to three ML models and does not explore the potential of more complex or ensemble ML techniques that could further enhance prediction accuracy.
[16]	Hefei, People's Republic of China	Weather images	Visible-light weather images (sunny, foggy, rainy, snowy)	Classification of weather conditions into sunny, foggy, rainy, snowy	CNNs with Mask R-CNN for edge extraction	The proposed CNN model incorporating Mask R-CNN for edge extraction improved classification accuracy by 5.06% over traditional methods, achieving high accuracy rates for sunny (97.94%), rainy (95.90%), foggy (88.86%), and snowy (98.22%) days.	Focuses primarily on clear distinctions among major weather types but may struggle with nuanced or mixed weather conditions; performance in variable lighting or under less ideal imaging conditions not fully explored.
[21]	Dataset sourced from Kaggle; no specific region mentioned	Weather data (19 variables including temperature, humidity, events)	Temperature (max, min), precipitation, and event data (hot, cold, rain)	Classification of weather into "Hot," "Cold," and "Rain"	Decision Trees (CART), RF, K-Nearest Neighbors (K-NN) with Neural Networks	RF outperformed other models, achieving 80% accuracy, solving the overfitting problem effectively. Decision Tree and K-NN achieved 55% and 61% accuracy, respectively.	The dataset is limited to specific events (hot, cold, rain) and does not cover more complex weather phenomena. Limited model comparison excludes advanced techniques like SVM or DL.

(Continued)

Table 1. Continued

Reference	Study Region	Data Type	Inputs	Outputs	ML Models Used	Key Findings	Limitations
[22]	Guilan Province, North Iran	Accident data (2017–2019) including vehicle and pedestrian accidents	Environmental (rainy weather, lighting), driver (age, gender), vehicle, and road condition variables	Accident severity prediction (fatal, injury, property damage)	Artificial Neural Networks (ANNs), Logit Model, Factor Analysis	ANN models outperformed statistical methods (Logit model) in accuracy, achieving up to 92.4% accuracy for pedestrian accidents and 89.6% for vehicle accidents. Key factors influencing accidents included rainy weather, driver age (30–50), lighting conditions, and road surface quality.	Limited geographic scope (single highway) and reliance on existing accident data, which may omit unreported incidents. ANN models require significant data pre-processing and computing resources.
[23]	Taiwan (Kaohsiung, Keelung, and Yangming-shan National Park)	Surveillance videos under various weather conditions (e.g., glare, haze, rain)	Video frames; pixel contrast and Hue-Saturation-Value (HSV) parameters	Vehicle detection and classification (sedans, motorcycles, trucks, buses, bicycles, etc.)	YOLOv3 with a visibility complementation module	Improved vehicle detection accuracy in adverse weather conditions without retraining. Detection recall rates increased (e.g., from 85.22% to 89.82% under glare, 66.92% to 87.27% under rain). Achieved real-time performance (30 fps) across test scenarios.	Limited computational efficiency for high-resolution images or under highly complex raindrop conditions. Requires further refinement to handle noise in haze scenarios effectively.

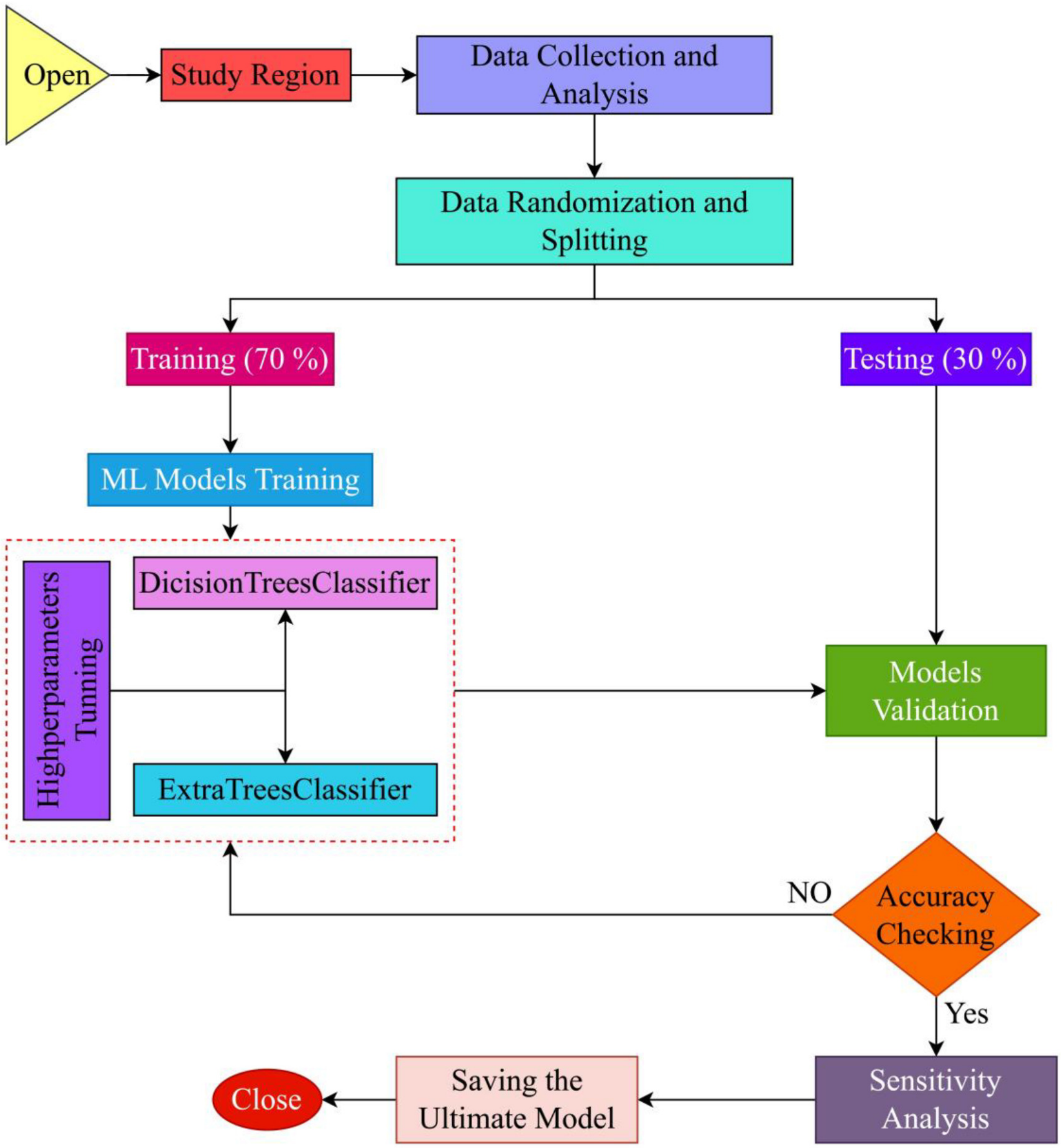


Fig. 1. Models' development process flowchart.

confusion matrices. A validation step was implemented, examining model predictions against actual observed weather conditions to ensure robustness and generalizability. Following validation, model accuracy was critically assessed to determine suitability. Hyperparameters for both DTC and ETC models were optimized through iterative manual tuning. Different parameter settings were systematically tested, and model accuracy was evaluated separately for the training and testing datasets. The final hyperparameters chosen were those providing the highest overall accuracy on the independent testing dataset. When the models demonstrated sufficient accuracy, the best-performing model (the "ultimate model") was selected and saved for practical applications. Finally, sensitivity analysis was performed on the ultimate model to quantify the importance and contribution of each meteorological parameter to prediction accuracy. This analysis identified the critical predictors influenced weather condition classification, facilitating practical decision-making for meteorological applications in the study area.

2.2. Study region

The selected study region is Kabul province, the capital of Afghanistan [10], spans an area of 6563 km^2 located in the eastern part of the country. Selected for its meteorological significance and rapid urban growth [24], Kabul presents unique challenges and opportunities for weather-related studies [25]. Situated in an arid to semi-arid climatic zone, the city experiences diverse weather patterns that significantly influence both its ecological balance and urban development [10]. Kabul receives an average annual rainfall of 330 mm, emphasizing the necessity to study precipitation trends crucial for urban planning and agricultural activities [26]. The importance of predicting weather conditions such as clear days, cloudy, rain, snow, and partly cloudy skies in Kabul is underscored by its expanding population and limited natural water resources [27]. Accurate weather predictions are vital for managing water supply, which is a critical issue in the region due to the rapid depletion of groundwater levels and increasing demands from a growing population [28]. These forecasts are essential for effective agricultural planning, disaster management, and mitigating the impacts of climate variability on the city's infrastructure [29]. Further, Kabul's geographic coordinates, bound between latitudes 34.697 and 34.515 and longitudes 69.857 and 69.845, place it in a strategic location for meteorological studies. The diverse topography of the region, including plains, hills, and mountains, further complicates weather patterns, making the study of meteorology here both challenging and essential [30]. Fig. 2 illustrates the location of Afghanistan on a world map, focusing into Kabul province on Afghanistan's map and figuring out the detailed map of Kabul province, underscoring the strategic importance of this study area for weather classification and prediction in a region facing significant environmental challenges.

2.3. Data collection and analysis

The data used in this study was carefully collected over the period of January 1, 2024, to December 31, 2024, covering a complete year with 366 records to account for the leap year. The weather data was obtained from Visual Crossing's historical weather API for Kabul, Afghanistan [31]. The data was obtained from a central station at latitude 34.5331 and longitude 69.1022, located almost in the center of Kabul province. This central location is pivotal as it serves as a representative sampling point for the overall regional weather conditions, reducing the variability that might arise from using multiple stations. This is due to the change that is generally possible in weather conditions from station to station. Further, the statistical characteristics of the input data are comprehensively detailed in Table 2, indicating variability and trends across several meteorological parameters. For instance, temperatures range widely from severe lows to moderate highs, illustrating Kabul's climatic extremes. The precipitation data indicate sporadic rainfall, which is critical for understanding seasonal water availability in this arid region.

Moreover, a frequency analysis of key meteorological parameters took place, where visually summarizing the distribution of data points. For example, the wind speed histogram reveals a predominance of moderate wind conditions, crucial for modeling weather patterns that affect air quality and temperature regulation. The distribution of solar radiation and cloud cover also provides insights into the solar exposure and potential for solar energy harvesting in Kabul, which is essential for energy resource planning. Overall, the frequency distributions of all input parameters are visualized in Fig. 3. The classification of weather conditions into categories such as clear days, cloudy, and rain, detailed in Table 3, reflects the predominant weather patterns. This categorization is crucial for the predictive modeling process, helping to enhance the accuracy of weather forecasts, essential for agricultural planning and disaster management. This methodical approach to data collection and analysis ensures a reliable framework for understanding the dynamic meteorological conditions in Kabul, providing a robust scientific basis for the conclusions drawn in this study.

2.3.1. Data quality assessment

To ensure the reliability of the collected data, a quality assessment was performed using conventional statistical methods, including completeness analysis, consistency checks, and outlier detection. Data completeness was calculated using the following formula [32]:

$$\text{Completeness }(\%) = \frac{N_{\text{valid}}}{N_{\text{total}}} \times 100 \tag{1}$$

Where N_{valid} is the number of valid observations and N_{total} is the total expected number of observations. The dataset demonstrated 100% completeness, as all 366 daily observations were recorded without missing values.

The consistency of the dataset was checked by performing correlation analysis among key meteorological variables (temperature, humidity, and solar radiation) to detect any anomalous deviations. Pearson's correlation coefficient (r) was calculated as [33]:

$$r = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sqrt{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 \sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2}} \tag{2}$$

where X_i and Y_i are paired meteorological variables, and $\bar{X}$, $\bar{Y}$ denote their means, respectively. The results confirmed high internal consistency among the variables (e.g., temperature and feels-like temperature

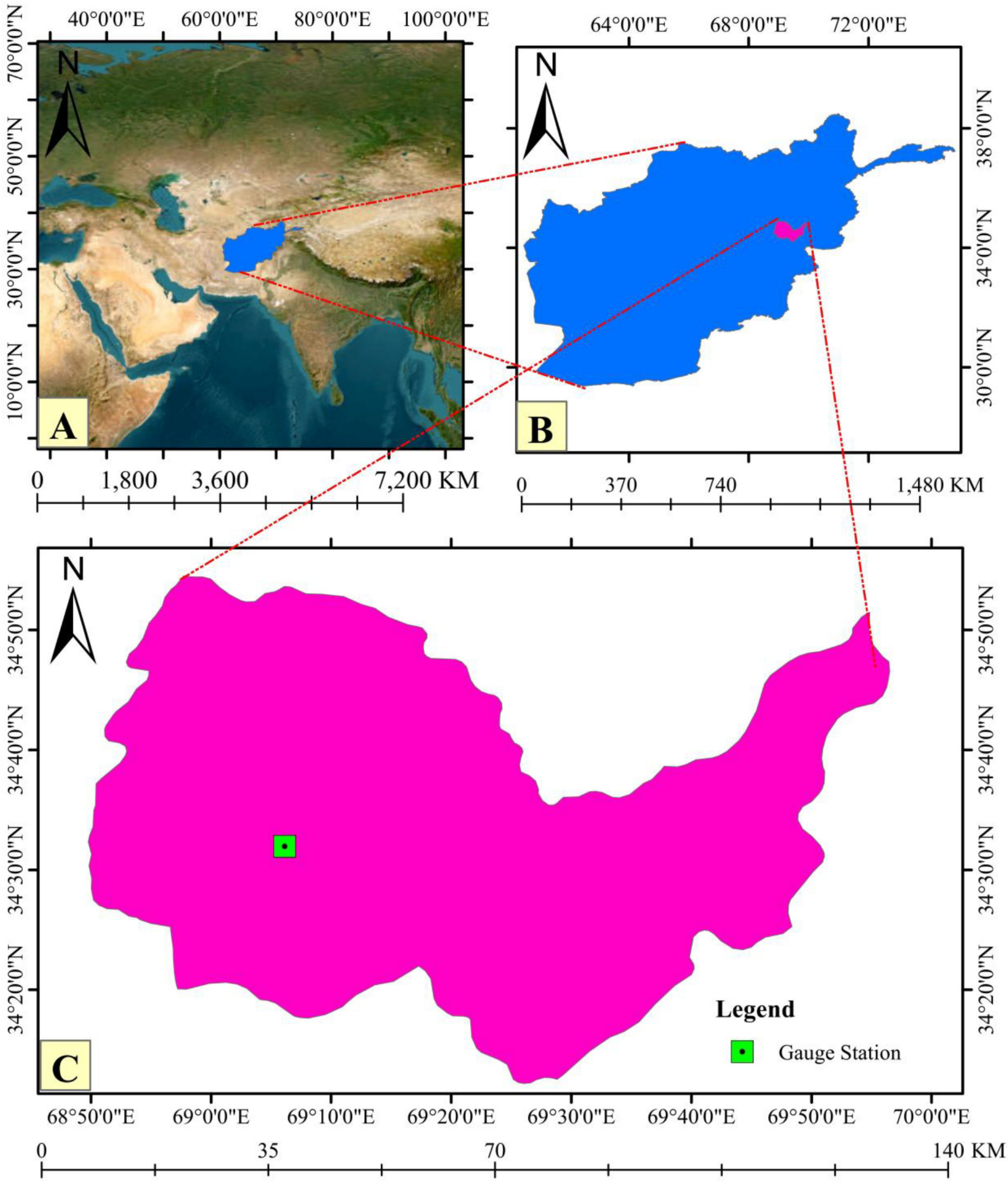


Fig. 2. Study Region: (A) Afghanistan's location on world map, (B) Kabul province location on Afghanistan's map, (C) Kabul province map.

showed $r = 0.998$, temperature and solar radiation $r = 0.663$), demonstrated data reliability (Table 4).

2.4. Applied machine learning (ML) models

In this study, DTC and ETC models were selected for their power in handling non-linear data and their capacity to model complex decision boundaries effectively and accurately, making them particularly suitable for meteorological data which often contains complex and non-linear relationships between variables. These models are preferred over others because of their interpretability and efficiency in processing large datasets, which are important for obtaining accurate and timely weather predictions.

Table 2. Statistical characteristics of the input data.

Factors	count	mean	std	min	25%	50%	75%	max	skew	kurt
tempmax (°C)	366	15.89	10.10	−5.70	7.00	17.00	25.18	32.40	−0.17	−1.25
tempmin (°C)	366	2.83	9.04	−21.30	−4.48	4.55	9.90	17.20	−0.53	−0.45
temp (°C)	366	9.29	9.70	−13.60	0.40	10.70	17.70	24.30	−0.31	−1.01
feelslikemax (°C)	366	15.53	9.94	−5.70	6.73	17.00	25.18	30.10	−0.25	−1.28
feelslikemin (°C)	366	1.36	10.45	−27.00	−6.50	3.70	9.85	17.20	−0.61	−0.33
feelslike (°C)	366	8.60	10.37	−17.80	−0.15	10.35	17.68	23.70	−0.42	−0.85
dew (°C)	366	−0.85	8.09	−23.70	−6.08	0.50	5.50	16.50	−0.51	−0.54
humidity (%)	366	54.11	15.96	17.00	43.05	53.00	63.38	98.60	0.44	−0.02
precip (mm)	366	1.59	4.60	0.00	0.00	0.00	0.80	44.90	5.03	32.68
precipprob (%)	366	39.62	48.98	0.00	0.00	0.00	100.00	100.00	0.43	−1.83
precipcover (%)	366	13.63	25.63	0.00	0.00	0.00	16.67	100.00	2.22	4.14
snow (mm)	366	0.29	1.63	0.00	0.00	0.00	0.00	18.20	7.82	67.95
snowdepth (mm)	366	6.75	16.22	0.00	0.00	0.00	0.00	74.30	2.72	6.77
windgust (km/h)	366	26.50	7.74	12.20	20.50	24.50	31.30	48.60	0.72	−0.21
windspeed (km/h)	366	8.88	2.61	4.00	6.80	8.60	10.10	23.00	1.10	2.07
winddir (degrees)	366	192.53	134.57	0.30	38.93	274.05	303.63	359.40	−0.30	−1.75
sealevelpressure (hPa)	366	1015.13	6.47	998.70	1010.05	1016.45	1020.10	1027.90	−0.38	−0.78
cloudcover (%)	366	27.38	29.43	0.00	3.13	16.95	42.05	100.00	1.09	0.07
visibility (km)	366	22.68	4.28	0.00	24.10	24.10	24.10	24.10	−3.68	13.60
solarradiation (W/m^2)	366	235.42	79.31	36.60	163.50	238.30	303.75	364.60	−0.16	−1.11
solarenergy (MJ/m^2)	366	20.33	6.86	3.10	14.03	20.70	26.30	31.50	−0.16	−1.11
uvindex	366	7.94	1.91	2.00	6.00	8.00	10.00	10.00	−0.67	−0.28
moonphase	366	0.49	0.29	0.00	0.25	0.50	0.75	0.98	−0.03	−1.22

Table 3. Desired classes for prediction with their respective data points.

Used class in paper	Actual class	No of data points
Class A	Clear day	166
Class B	Cloudy day	2
Class C	Partly cloudy day	53
Class D	Rainy-day	114
Class E	Snowy day	31

Table 4. Pearson's correlation among selected meteorological variables.

Variable Pair	Correlation Coefficient (r)
Temp. max vs. Feels-like Temp. max	0.998
Temp. vs. Feels-like Temp.	0.998
Temp. min vs. Feels-like Temp. min	0.997
Temp. min vs. Dew point	0.908
Temp. vs. Dew point	0.858
Dew point vs. Humidity	0.137
Precipitation vs. Humidity	0.602
Precipitation Probability vs. Cloud Cover	0.560
Solar Radiation vs. Temperature	0.636
Wind speed vs. Wind gust	0.728
Sea Level Pressure vs. Temperature	−0.874
Visibility vs. Precipitation	−0.724
Visibility vs. Cloud cover	−0.607
UV Index vs. Solar Radiation	0.963

The subsequent sections provided more details and the way of each model performance, exploring their workability and effectiveness in classifying various weather conditions. The detailed hyperparameters, configurations, and structural specifics of both DTC and ETC models employed in this study are comprehensively summarized in Table 5. This information provides clarity and facilitates reproducibility of the modeling framework.

2.4.1. Decision trees classifier (DTC)

The Decision Trees Classifier is renowned for its robustness and simplicity, making it highly effective for categorizing complex and heterogeneous datasets like meteorological data [34]. This model constructs a tree-like structure where each node represents a decision rule that splits the data based on the most significant features [35]. It operates using the formula for Information Gain, defined as [36, 37]:

$$\text{Gain} = \text{Entropy (parent)} - \left[\frac{N_{\text{left}}}{N} \times \text{Entropy(left child)} + \frac{N_{\text{right}}}{N} \times \text{Entropy(right child)}\right] \quad (3)$$

Where, 'parent' represent the total dataset at the current node, entropy measures the impurity of the dataset, N_{left} and N_{right} are the number of samples in the left and right child nodes, respectively, and N is the total number of samples at the current node [38]. This approach optimizes the decision-making process by maximizing information gain and reducing uncertainty at each split [39]. The effectiveness and efficiency of this classifier in handling complex datasets make it ideal for weather prediction, which involves intricate and variable data inputs [40].

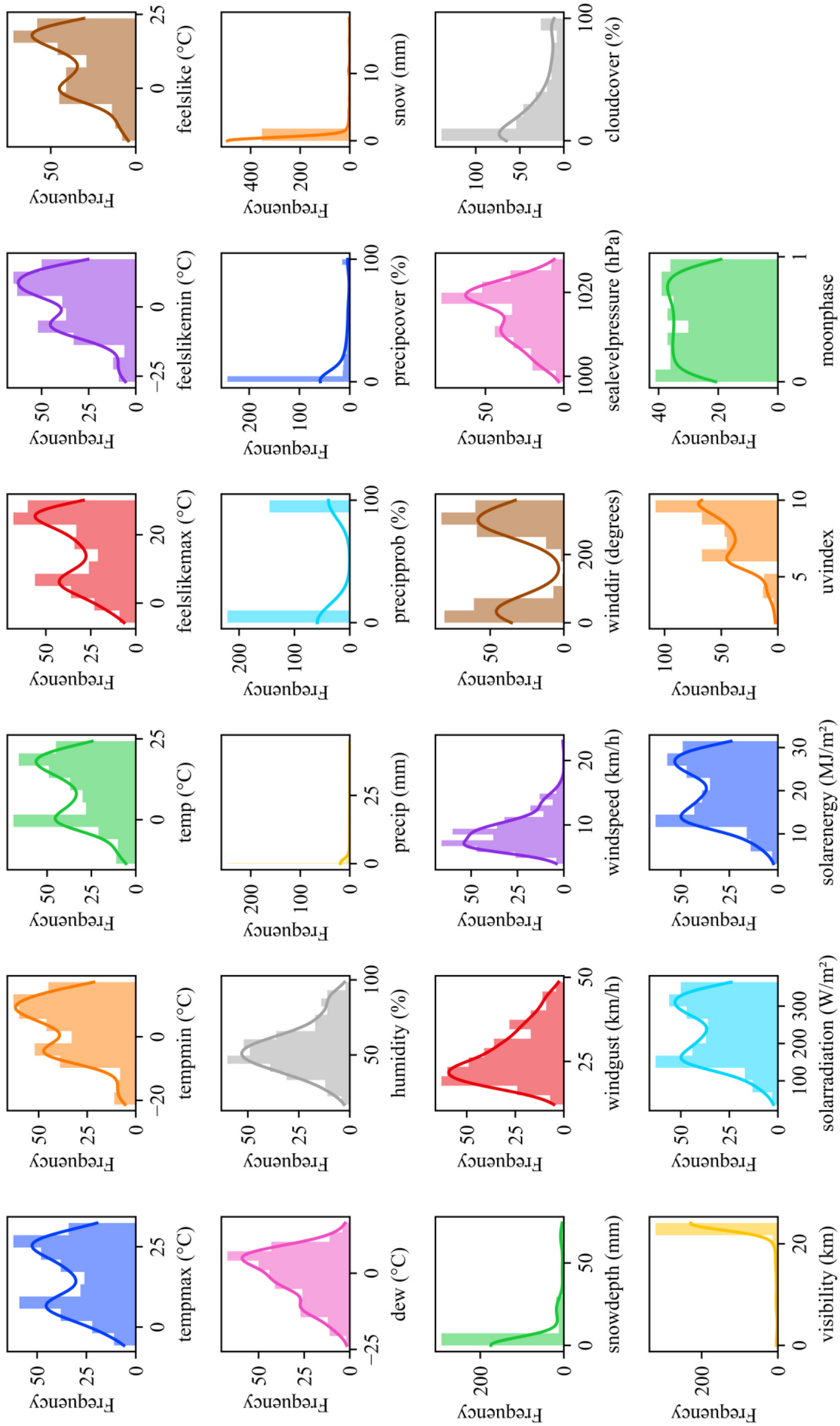


Fig. 3. Frequency analysis of key meteorological parameters.

Table 5. Hyperparameters and structural configuration of ML models.

Model	Hyperparameter	Selected Setting	Alternative Options	Description
Decision Tree Classifier (DTC)	Criterion	Gini	Gini, Entropy	Function to measure the quality of splits
	Splitter	Best	Best, Random	Strategy used to choose the split at each node
	Max Depth	5	None, 1, 2, 3, ...	Maximum depth of the tree
	Min Samples Split	3	2, 3, 4, ...	Minimum samples required to split an internal node
	Min Samples Leaf	3	1, 2, 3, ...	Minimum samples required at a leaf node
	Random State	938	Any integer	Ensures reproducibility of the results
Extra Trees Classifier (ETC)	Number of Estimators (Trees)	5	10, 50, 100, ...	Number of trees in the ensemble
	Max Features	sqrt	auto, sqrt, log2	Number of features considered at each split
	Max Depth	6	None, 1, 2, 3, ...	Maximum allowed depth of each tree
	Min Samples Split	2	2, 3, 4, ...	Minimum samples required to split an internal node
	Min Samples Leaf	1	1, 2, 3, ...	Minimum samples required at a leaf node
	Random State	938	Any integer	Ensures reproducibility of the results

The DTC method involved first randomizing the dataset and splitting it into training 70% and testing 30% subsets. The model's hyperparameters were manually optimized through iterative trials. The final hyperparameters that yielded the highest accuracy were criterion (Gini), splitter (Best), max depth (5), min samples split (3), min samples leaf (3), and random state (938). After training and evaluating the model's performance using standard classification metrics, the optimized model was saved and used for further feature importance analysis.

2.4.2. Extra trees classifier (ETC)

The ETC is an ensemble learning technique that builds on the principles of random forests by randomly selecting data points and features at each split to create a multitude of decision trees [41, 42]. This approach enhances model robustness and reduces overfitting by averaging multiple deep decision trees, each constructed from different random subsets of the data. The method operates by calculating the Gini index for each feature at every possible split within a randomly selected subset of the dataset [43]. The formula for the Gini index is [44]:

$$\text{Gini}(D) = 1 - \sum_{i=1}^{J} p_i^2 \tag{4}$$

Where, D is a dataset in a node, J represents the number of classes, and p_i is the proportion of class i instances within the dataset D. The splits that yield the lowest Gini index are selected to ensure that the trees are as pure as possible, thereby effectively increasing the prediction accuracy. This classifier is particularly effective in dealing with complex datasets like weather data, where high dimensionality and feature interactions can significantly affect prediction outcomes.

The ETC method began with randomizing the dataset and dividing it into training 70% and testing 30% subsets. The model's hyperparameters were iteratively optimized through manual trials, ultimately selecting the following parameters that yielded the highest accuracy: number of estimators (5), max features (sqrt), max depth (6), min samples split (2), min samples leaf (1), and random state (938). The ETC model was then trained, evaluated using standard classification metrics, and saved for subsequent feature importance analysis.

2.5. Performance evaluation metrics

In this study, a comprehensive set of performance evaluation metrics is employed to assess the effectiveness of the machine learning models used. These metrics provide insights into the accuracy and reliability of the classification outcomes across different weather conditions [45, 46].

Precision: The proportion of true positive predictions relative to the total predicted positives:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{5}$$

Recall (Sensitivity): Measures the proportion of true positives accurately identified out of all actual positive cases:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{6}$$

F1-score: Harmonic mean of Precision and Recall, providing a balance between these two metrics:

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

Overall Accuracy: The ratio of correctly predicted observations (both positive and negative) to the total number of predictions:

$$\text{Overall Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

Weighted Average (Precision, Recall, F1-score): Unlike macro averaging, the weighted-average calculation accounts for the number of instances in each class. Thus, classes with more instances influence the metrics proportionally more:

$$\text{Weighted avg Precision} = \sum \left(\frac{\text{number of instances in each class}}{\text{total instances}} \times \text{Precision of each class} \right) \tag{9}$$

$$\text{Weighted avg Recall} = \sum \left(\frac{\text{number of instances in each class}}{\text{total instances}} \times \text{Recall of each class} \right) \tag{10}$$

$$\text{Weighted avg F1-score} = \sum \left(\frac{\text{number of instances in each class}}{\text{total instances}} \times \text{F1-score of each class} \right) \tag{11}$$

Macro Average (Precision, Recall, F1-score): This calculation treats all classes equally, averaging the metrics across all classes regardless of the number of instances per class. It helps indicate the model's overall ability to handle various classes, particularly useful for datasets with imbalanced categories.

$$\text{Macro avg F1-score} = \frac{\sum (\text{F1-score of each class})}{\text{number of classes}} \tag{12}$$

$$\text{Macro Avg Precision} = \frac{\sum (\text{Precision of each class})}{\text{number of classes}} \tag{13}$$

$$\text{Macro Avg Recall} = \frac{\sum (\text{Recall of each class})}{\text{number of classes}} \tag{14}$$

Where, TP (True Positives) is the number of correctly predicted positive observations. FP (False Positives) is the number of incorrect predictions where the model predicted positive, but the truth was negative [47]. TN (True Negatives) indicates the number of correctly predicted negative observations. FN (False Negatives) is the number of incorrect predictions where the model predicted negative, but the truth was positive [48]. Sensitivity (Recall) demonstrates the metric indicating the model's ability to correctly identify positives. These metrics collectively offer a nuanced view of the models' performance, capturing both their accuracy and their ability to handle different classes in the dataset, essential for validating the robustness of weather prediction models.

3. Results representations

3.1. Performance evaluation of DTC and ETC models using confusion matrices

For the DTC model, the training phase results indicated that Class A (clear day) had 116 correct classifications, Class C (partly cloudy day) had 37, Class D (rainy day) had 80, and Class E (snowy day) had 21, with one miss classifications for Class B (cloudy day). The testing set for the DTC model indicates 50 correct classifications for Class A, 16 for Class C, 34 for Class D, and 9 for Class E, with a misclassification of 1 in Class B (Fig. 4). As the datapoints for Class B in whole year were 2 datapoints, this indicates low datapoints have low accuracy, where those of huge datapoints classified correctly. Conversely, the ETC model demonstrated a slightly different performance. In the training phase, Class A (clear day) and Class D (rainy day) were perfectly classified with 116 and 80 correct classifications respectively, Class C had 31, Class E had 19, and Class B had 1 misclassification. In the testing phase, Class A maintained a high accuracy with 50 correct predictions, but Class D showed 32 correct classified and 2 misclassified, Class C 7 correct classified and 9 misclassified, and Class E to 8 correct and 1 missed, with only 1 misclassification in Class B. From this analysis, the DTC model appears to slightly outperform the ETC model in terms of overall accuracy and consistency across

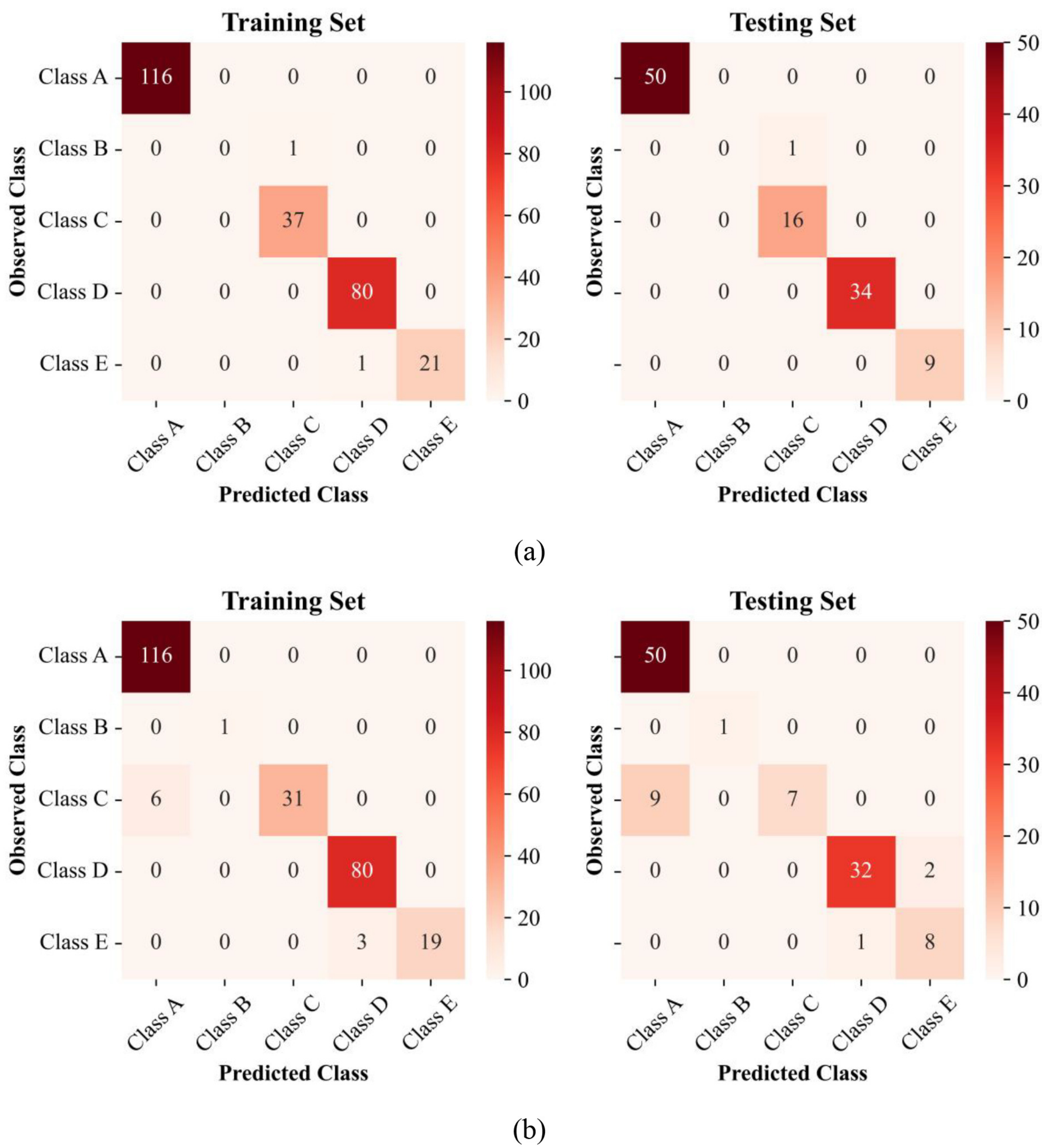
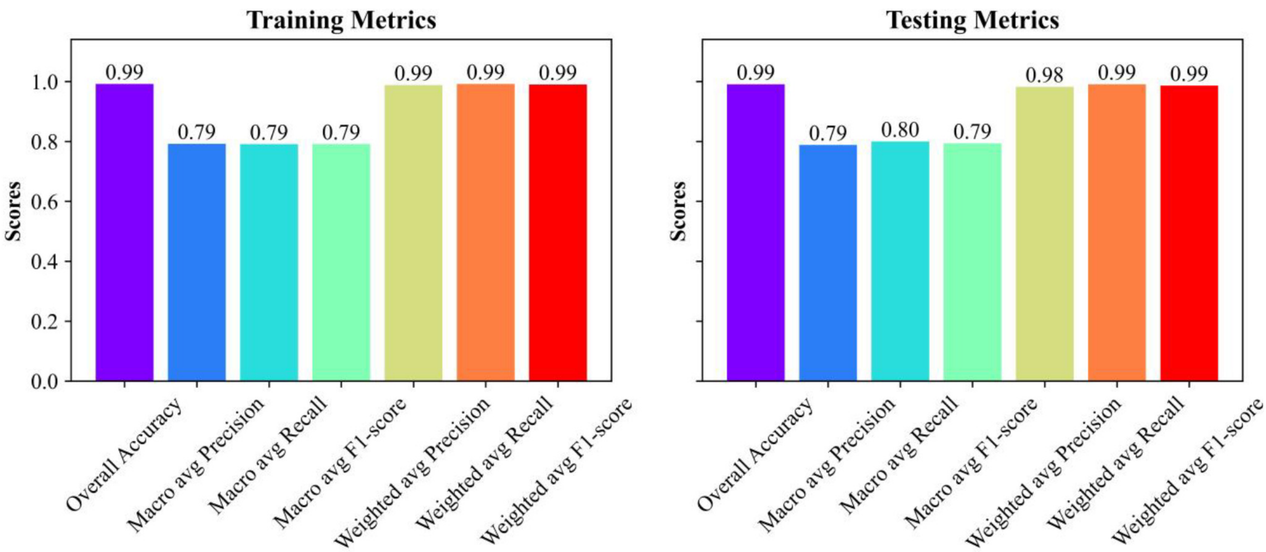


Fig. 4. Confusion matrix analysis for multi-class classification on training and testing phases; (a) DTC model, (b) ETC model.

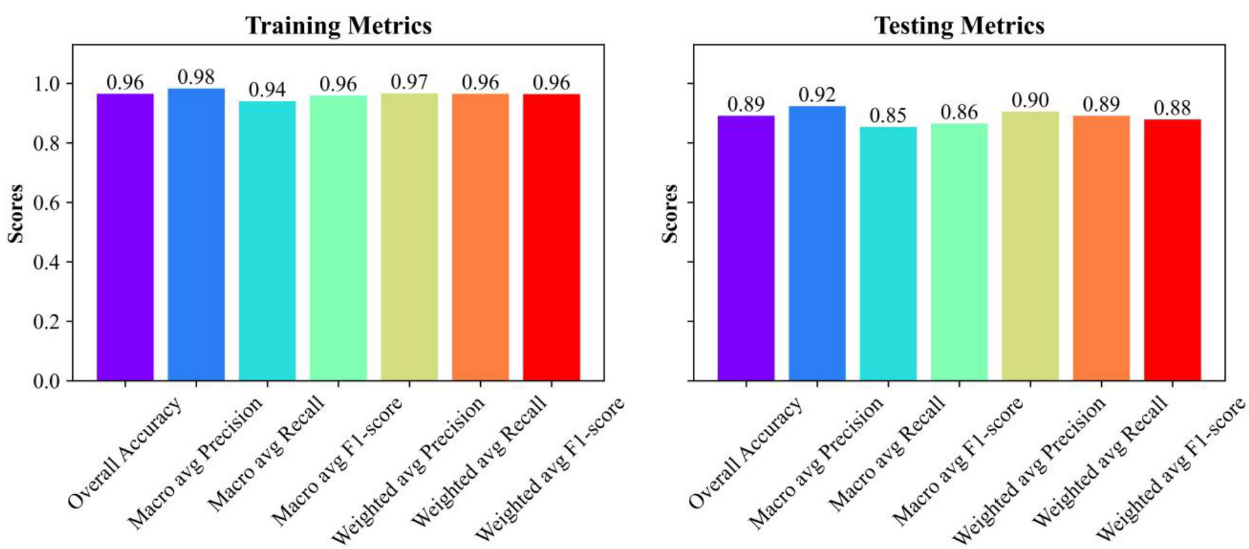
both the training and testing sets, suggesting it as the best model for handling multi-class classification for weather conditions.

3.2. Detailed performance metrics for DTC and ETC models

The detailed overall performance metrics for the DTC and ETC models are visualized in Fig. 5. Where, during the training phase the DTC model achieved an overall accuracy of 0.99, with macro avg and weighted avg for precision, recall, and F1-scores obtained high values of 0.79 and 0.99 respectively. In contrast, the ETC model during training phase recorded slightly lower overall accuracy at 0.96 but exhibited strong macro avg precision of 0.98 and macro avg recall of 0.94, with both macro avg and weighted avg for F1-scores at 0.96. During testing phase, the DTC model maintained high overall accuracy at 0.99, with macro avg for precision and recall at 0.79 and 0.80 respectively, and similarly high weighted metrics. The ETC model, however, showed



(a)



(b)

Fig. 5. Comparative analysis of model performance metrics across training and testing phases; (a) DTC model, (b) ETC model.

a decrease in testing metrics with an overall accuracy of 0.89, macro avg precision of 0.92, and macro avg recall of 0.85, along with macro avg and weighted avg for F1-scores of 0.86 and 0.88 respectively.

Building upon the detailed above evaluation, Fig. 6 further delineates class-specific performance metrics for the DTC and ETC models, demonstrating distinct strengths and weaknesses across various classes. For the DTC model, precision, recall, and F1-score in the training phase are perfect (1.00) for Class A, while Class B notably underperforms with zeros across all metrics. Consistently the issue with class B stands for low number of datapoints. Class C shows strong precision (0.97) and perfect recall (1.00), resulting in an F1-score of 0.99. Classes D and E also exhibit high performance with respective F1-scores of 0.99 and 0.98. In the testing phase, the DTC model retains high metrics for Classes A, D, and E, all achieving perfect scores (1.00) across precision, recall, and F1-Score. Class C, despite a lower precision of 0.94, achieves perfect recall and an F1-score of 0.97. On the other hand, the ETC model shows a varied performance spectrum. In training, Class A's metrics slightly lag behind with a precision of 0.95 but perfect recall, and Class C's lower recall of 0.84 brings its F1-score down to 0.91. However, Classes B and E excel with

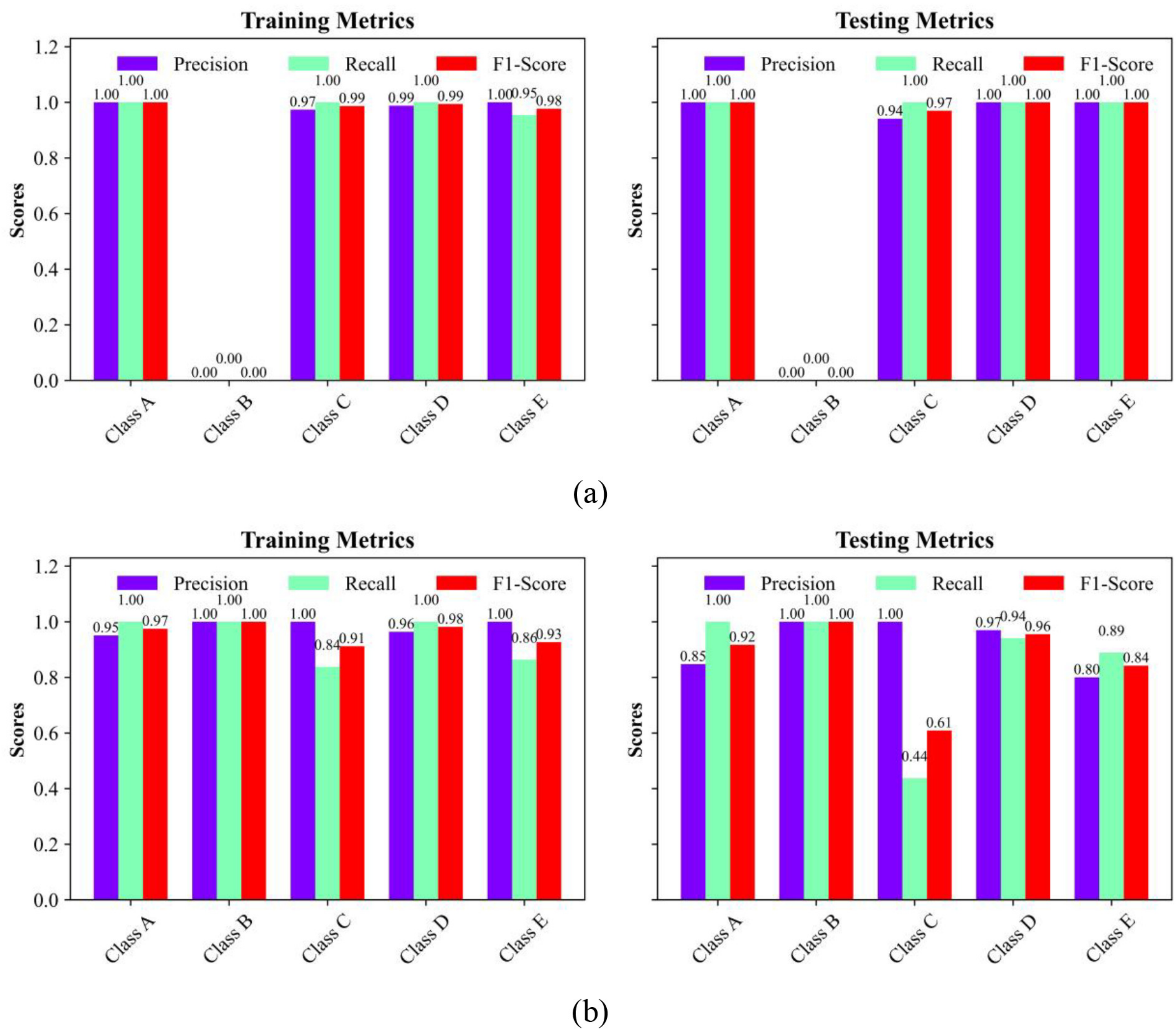
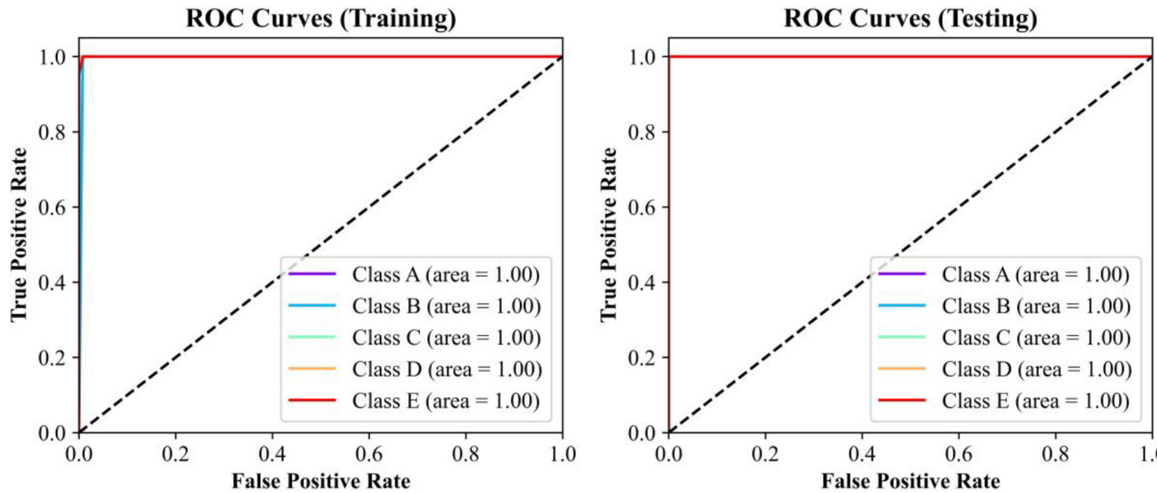


Fig. 6. Class-wise performance metrics in training and testing for model's evaluation; (a) DTC model, (b) ETC model.

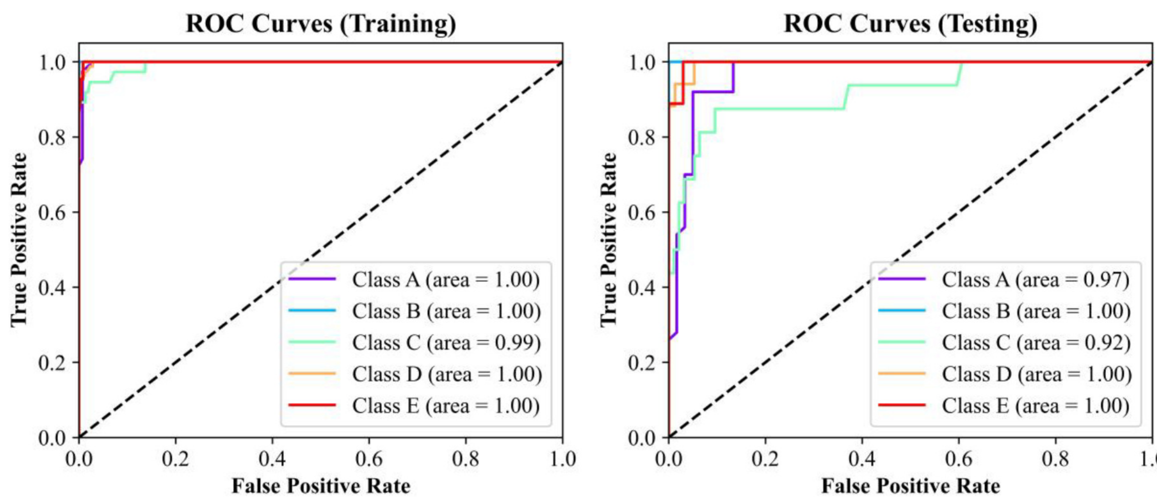
perfect precision and recall, leading to an F1-score of 1.00 and 0.93, respectively. The testing phase reveals some weaknesses, with Class C only achieving a precision of 1.00 but a lower recall of 0.44, resulting in a decreased F1-score of 0.61. Conversely, Class B maintains perfect scores, underscoring its robustness in both training and testing scenarios. Conclusively, while the ETC model shows exceptional results in certain classes, the DTC model demonstrates a more consistent and superior performance across a broader range of metrics, establishing it as the best performing model for this multi-class classification task.

3.3. Enhancing model evaluation with ROC curve analysis

The Receiver Operating Characteristic (ROC) curve, a graphical plot that illustrates the diagnostic ability of a binary classifier system, varies the discrimination threshold and plots the True Positive Rate (TPR) against the False Positive Rate (FPR). This method provides a comprehensive measure of model performance at different threshold settings. For the DTC model, the training and testing phases achieved perfect Area Under Curve (AUC) scores of 1.00 for all classes (A through E), indicating flawless classification without any false positives or negatives (Fig. 7a). Similarly, the ETC model showed nearly perfect AUC scores in training, with Class A, B, D, and E with scores of 1.00 and Class C slightly lower at 0.99. Testing metrics for the ETC model revealed minor variations, with Class A scoring 0.97, Class B, D, and E maintaining a perfect score of 1.00, and Class C at 0.92, indicating excellent but not flawless performance (Fig. 7b). These ROC curve results strongly support the findings of current study, demonstrating that both models exhibit robust classification capabilities, with the DTC model showing



(a)



(b)

Fig. 7. ROC curve analysis by multi-class classification for ML models performance evaluation; (a) DTC model, (b) ETC model.

significantly superior performance in maintaining consistent excellence across all scenarios.

3.4. Sensitivity analysis

Feature analysis and importance on how the DTC and ETC models utilized different subsets of features to obtain high predictive accuracy were carried out. The findings indicated that the DTC model highlights a focused approach, utilizing only five significant features: Precipitation Probability (40%), Cloud Cover (31%), Snow (18%), temperature Feels Like Max (8%), and Solar Radiation (3%) (Fig. 8a). This indicates that fewer features are required for data collection, making the DTC model efficient and effective for scenarios where limited data inputs are available. Conversely, the ETC model leverages a broader range of inputs, showing significant importance in nine features including Precipitation Probability (33%), Cloud Cover (16%), Dew (9%), Temperature Max (8%), Pressure (7%), temperature Feels Like Min (6%), Solar Energy (5%), Wind Max (5%), and Solar Radiation (4%) (Fig. 8b). This diversity suggests the ETC model's ability to incorporate and effectively utilize a wide variety of datapoints, enhancing its robustness and adaptability in more complex scenarios where more detailed environmental data is accessible. Conclusively, both models have their strengths, the DTC model is ideal for applications requiring fewer data inputs and where ease of interpretation is crucial. In contrast, the ETC model is preferable in situations where comprehensive data collection is

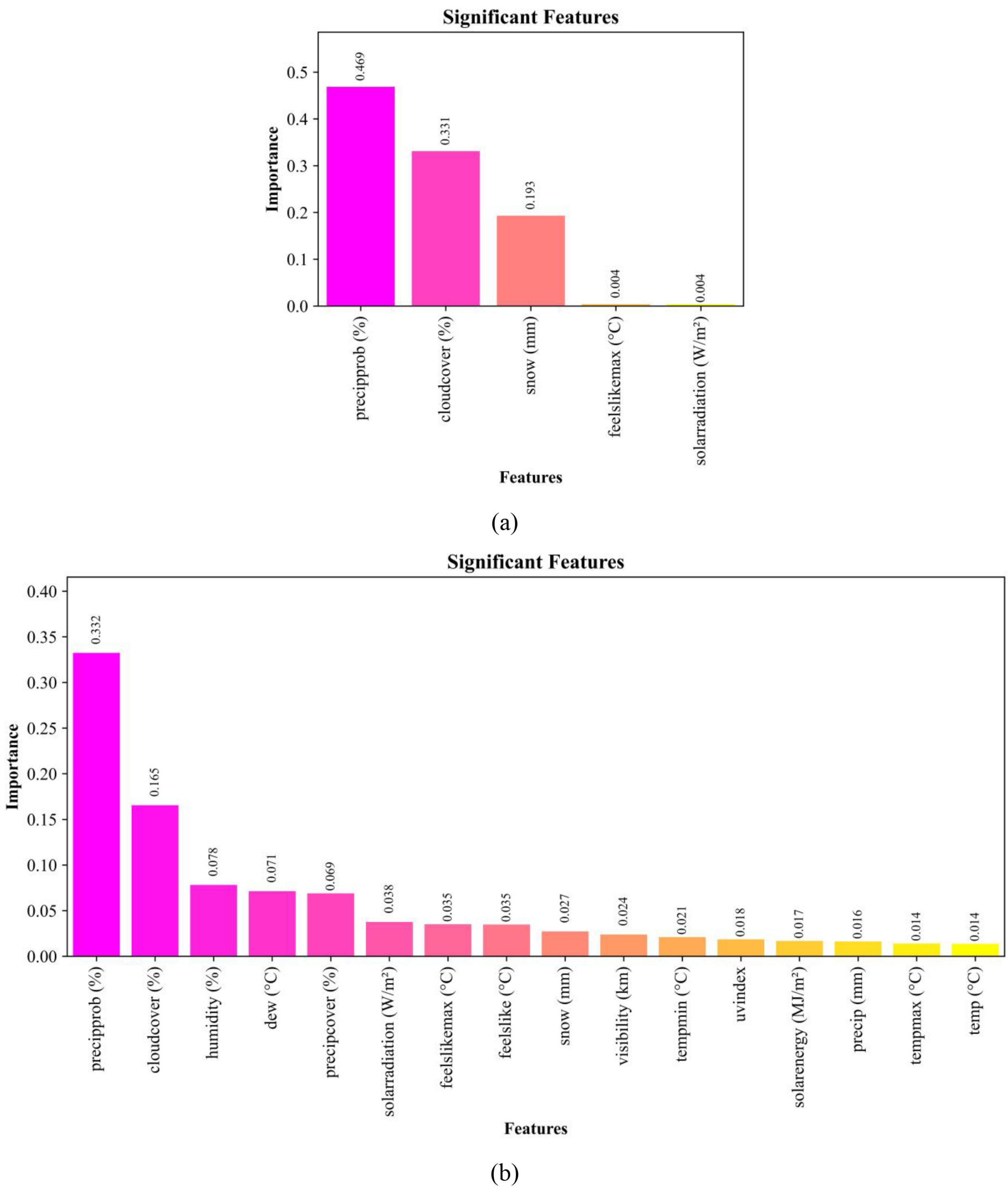


Fig. 8. Analysis of feature importance in predictive modeling; (a) DTC model, (b) ETC model.

feasible, and a more detailed analysis is beneficial. Ultimately, the choice between using the DTC or ETC model based on feature importance should be based on the specific needs and limitations of the application, as both offer valuable capabilities tailored to different operational preferences and conditions.

4. Discussion

The findings from the current study offers valuable insights into the efficacy of ML techniques in the classification of weather conditions, specifically within the challenging meteorological landscape of study region (Kabul, Afghanistan). This study differs significantly from other studies discussed below in its focus on a region with unique geographical and climatic challenges. The incorporation of specific models, namely the DTC and ETC, tailored to handle the complex, non-linear interactions of meteorological variables in Kabul, marks a distinct approach compared to other studies that have often employed more generic ML models across broader geographic areas.

In comparison, studies like [49, 50], explored weather classification using deep learning (DL) methods applied to image data, which contrasts with the current study's reliance on numerical weather data and structured decision tree methodologies. These differences highlight the varied applications of ML in meteorology, depending on the nature of the data and the specific requirements of the study area. Similarly, [51] used of convolutional neural networks for classifying weather images shows a different methodological approach, emphasizing visual data over numerical weather data used in the current study.

The findings of this study align closely with those of [52–54], where ML models demonstrated high accuracy in weather prediction. For example, the overall accuracy achieved by the DTC model in this study (99% in testing) is comparable to the high accuracy levels noted in [52] as 89.71% using ANN for visibility classification, and the study [53] obtained 97% accuracy in weather image classification using SVM. This similarity underscores the effectiveness of ML in accurately predicting weather conditions across different regions and datasets.

However, there is a contrast with studies like [55, 56], where the focus was on synthetic data and solar forecasting, respectively. These studies used ML for different applications within the meteorology field, indicating the versatility of ML techniques but also their specific adaptability to distinct study objectives and data types. The current study's used of real, ground-based meteorological data provides a direct analysis of weather patterns without the need for synthetic augmentation or the specific prediction of solar outputs.

The use of decision tree-based models in this study provides a clear, interpretable framework for classification, which is particularly important in operational settings such as agricultural planning and disaster management in the region. This approach aligns with the methodologies seen in [57, 58], where clear, actionable outputs from ML models are critical for real-time decision-making. On the other hand, the high dimensionality and feature interactions explored in studies like [59, 60], which use more complex algorithms such as boosted trees and hybrid models, offer a different perspective on handling meteorological data, emphasizing the trade-offs between model complexity and interpretability.

In summary, the current study not only advances our understanding of ML applications in meteorology for a specific regional context but also contributes to the broader discourse on the best practices and methodologies for weather classification across diverse environments. Its alignment with several previous studies highlights the robustness of ML in this field, while its distinctions underscore the importance of tailored approaches to meet specific regional challenges. This balance between adaptation to local conditions and alignment with global ML practices in meteorology forms a valuable part of the ongoing evolution of technological applications in environmental science.

5. Limitations, practical implications, and generalizability

Despite demonstrating high accuracy and effectiveness in classifying daily weather conditions, this study has certain limitations that should be acknowledged. The model's performance heavily relies on the quality, quantity, and representativeness of the local meteorological data collected specifically from Kabul, potentially affecting its direct applicability to other regions with significantly different climatic or topographical conditions. However, the practical implications remain substantial, as local authorities, urban planners, and agricultural stakeholders in Kabul can directly benefit from improved, reliable, and interpretable daily weather forecasts provided by these models. Regarding generalizability, while the specific results are most applicable to Kabul and similarly structured regions, the decision tree-based modeling framework, methodology, and systematic process developed herein offer broad applicability and can serve as a valuable blueprint for adapting and implementing similar ML-based forecasting systems in other regions facing comparable meteorological and environmental complexities.

6. Conclusion

Weather condition classification through advanced ML techniques represents a crucial advancement in meteorological sciences, allowing for more accurate and timely predictions critical for effective resource management and emergency response. This study objectives concentrated to implement and evaluate DTC and ETC advanced ML models for weather prediction in Kabul province of Afghanistan, assessing the models' effectiveness in different weather conditions, and explore the applicability of these models for local and global meteorological challenges. Where, a novel methodology involved a detailed analysis of data collected over a complete year (1.1.2024 to 31.12.2024 with 366 records) from a central station in the region, followed by the application of ML algorithms. The models were carefully trained and tested, ensuring comprehensive coverage of different atmospheric conditions prevalent in the region. Results indicated that the DTC model obtained an overall accuracy of 99% in both training and testing phases, with average precision, average recall, and average F1-scores of 0.79, 0.80 and 0.99 respectively in the testing phase. In contrast, the ETC model, while powerful in the training phase with an overall accuracy of 96% and macro average precision of 0.98, showed decreased performance in the testing phase with an overall accuracy of 89%, macro average precision of 0.92, and macro average F1-score of 0.86. These findings signify the outperforming of the DTC model over ETC model in handling different weather conditions efficiently, suggesting its potential for broader implementation in similar climatic regions. Given the results, it is recommended that further validation with larger datasets be conducted to confirm the models' reliability and adaptability. In addition, consideration should be given to integrating real-time data feeds to enhance predictive capabilities and model responsiveness. Future research should focus on exploring ensemble methods that combine the strengths of different ML techniques to improve prediction accuracy. Investigating the impact of extreme weather events on model performance and extending the application of these models to other regions with different climatic conditions are also critical areas for further study.

Abbreviations

ML: Machine Learning
DTC: Decision Tree Classifier
ETC: Extra Trees Classifier
AE: Autoencoders
TPR: True Positive Rate
TNR: True Negative Rate
CNN: Convolutional Neural Networks
DL: Deep Learning
ROC-AUC: Receiver Operating Characteristic - Area Under Curve
RF: Random Forest
SVM: Support Vector Machine
ANN: Artificial Neural Networks
K-NN: K-Nearest Neighbors
YOLOv3: You Only Look Once version 3
HSV: Hue-Saturation-Value
ROC: Receiver Operating Characteristic
AUC: Area Under Curve

References

1. F. Nebeker, "Calculating the weather: Meteorology in the 20th century," Elsevier, 1995.
2. M. Fathi, M. Haghi Kashani, S. M. Jameii, and E. Mahdipour, "Big data analytics in weather forecasting: A systematic review," *Arch. Comput. Methods Eng.*, vol. 29, no. 2, pp. 1247–1275, 2022.
3. S.-P. Xie *et al.*, "Towards predictive understanding of regional climate change," *Nat. Clim. Chang.*, vol. 5, no. 10, pp. 921–930, 2015.
4. G. L. Hammer *et al.*, "Advances in application of climate prediction in agriculture," *Agric. Syst.*, vol. 70, no. 2–3, pp. 515–553, 2001.
5. Z. H. Doost, L. Goliatt, M. S. Aldlemy, M. Ali, and B. da S. Macêdo, "Enhancing predictive accuracy of compressive strength in recycled concrete using advanced machine learning techniques with k-means clustering," *AUIQ Tech. Eng. Sci.*, vol. 1, no. 1, p. 10, 2024.
6. H. Ghouse, U. C. Ishmael, E. D. Obando-Paredes, H. S. Shakir, and A. Al-Bayaty, "Field canals improvement projects duration prediction: A comparative analysis of machine learning models," *AUIQ Tech. Eng. Sci.*, vol. 1, no. 2, p. 3, 2024.
7. R. Yang *et al.*, "Interpretable machine learning for weather and climate prediction: A review," *Atmos. Environ.*, p. 120797, 2024.
8. A. S. Albahri *et al.*, "A systematic review of trustworthy artificial intelligence applications in natural disasters," *Comput. Electr. Eng.*, vol. 118, p. 109409, 2024.
9. Y. Varshney and N. Ram, "Weather forecasting: A systematic review using ai approaches," *Weather*, vol. 4, no. 10, 2024.
10. Z. H. Doost and Z. M. Yaseen, "The impact of land use and land cover on groundwater fluctuations using remote sensing and geographical information system: Representative case study in Afghanistan," *Environ. Dev. Sustain.*, pp. 1–24, 2023, doi: 10.1007/s10668-023-04253-2.
11. Z. H. Doost and Z. M. Yaseen, "Allocation of reservoirs sites for runoff management towards sustainable water resources: Case study of Harirud River Basin, Afghanistan," *J. Hydrol.*, vol. 634, p. 131042, May 2024, doi: 10.1016/j.jhydrol.2024.131042.
12. L. Chen, B. Han, X. Wang, J. Zhao, W. Yang, and Z. Yang, "Machine learning methods in weather and climate applications: A survey," *Appl. Sci.*, vol. 13, no. 21, p. 12019, 2023.
13. S. Verma, K. Srivastava, A. Tiwari, and S. Verma, "Deep learning techniques in extreme weather events: A review," arXiv Prepr. arXiv2308.10995, 2023.
14. P. Zhang, X. Liu, and K. Pu, "Classification of dry and wet periods using commercial microwave links: A one-class classification machine learning approach based on autoencoders," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, pp. 1–13, 2023.
15. Y. Wang and Y. Li, "Research on multi-class weather classification algorithm based on multi-model fusion," in *2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, 2020, vol. 1, pp. 2251–2255.
16. X. Zhao and C. Wu, "Weather classification based on convolutional neural networks," in *2021 International Conference on Wireless Communications and Smart Grid (ICWCSG)*, 2021, pp. 293–296.
17. Y. Fu *et al.*, "A machine-learning-based study on all-day cloud classification using Himawari-8 infrared data," *Remote Sens.*, vol. 15, no. 24, p. 5630, 2023.

18. M. Elhoseiny, S. Huang, and A. Elgammal, "Weather classification with deep convolutional neural networks," in *2015 IEEE international conference on image processing (ICIP)*, 2015, pp. 3349–3353.
19. J. C. V. Guerra, Z. Khanam, S. Ehsan, R. Stolkin, and K. McDonald-Maier, "Weather classification: A new multi-class dataset, data augmentation approach and comprehensive evaluations of convolutional neural networks," in *2018 NASA/ESA Conference on Adaptive Hardware and Systems (AHS)*, 2018, pp. 305–310.
20. A. W. Fazil, M. Hakimi, R. Akbari, M. M. Quchi, and K. Q. Khaliqyar, "Comparative analysis of machine learning models for data classification: An in-depth exploration," *J. Comput. Sci. Technol. Stud.*, vol. 5, no. 4, pp. 160–168, 2023.
21. B. L. Pavuluri, R. S. Vejendla, P. Jithendra, T. Deepika, and S. Bano, "Forecasting meteorological analysis using machine learning algorithms," in *2020 International conference on smart electronics and communication (ICOSEC)*, 2020, pp. 456–461.
22. M. Ghasedi, M. Sarfjoo, and I. Bargegol, "Prediction and analysis of the severity and number of suburban accidents using logit model, factor analysis and machine learning: a case study in a developing country," *SN Appl. Sci.*, vol. 3, no. 1, p. 13, 2021.
23. X.-Z. Chen, C.-M. Chang, C.-W. Yu, and Y.-L. Chen, "A real-time vehicle detection system under various bad weather conditions based on a deep learning model without retraining," *Sensors*, vol. 20, no. 20, p. 5731, 2020.
24. W. Wafa, M. H. Hairan, and H. Waizy, "The impacts of urbanization on Kabul city's groundwater quality," *Int. J. Adv. Sci. Technol.*, vol. 29, no. 4, pp. 10796–10809, 2020.
25. S. Das, "Challenges in predicting extreme weather events over the South Asian region," in *Extreme Natural Events: Sustainable Solutions for Developing Countries, Springer*, 2022, pp. 51–106.
26. A. R. Noori and S. K. Singh, "Status of groundwater resource potential and its quality at Kabul, Afghanistan: a review," vol. 80, p. 654, 2021, doi: 10.1007/s12665-021-09954-3.
27. D. K. Sinha, "Natural disaster reduction: South East Asian realities, risk perception and global strategies," *Anthem press*, 2007.
28. B. Sivakumar, "Global climate change and its impacts on water resources planning and management: Assessment and challenges," *Stoch. Environ. Res. Risk Assess.*, vol. 25, pp. 583–600, 2011.
29. M. Salimi and S. G. Al-Ghamdi, "Climate change impacts on critical urban infrastructure and urban resiliency strategies for the Middle East," *Sustain. Cities Soc.*, vol. 54, p. 101948, 2020.
30. Q. Mahdawi, J. Sagin, M. Absametov, and A. Zaryab, "Water recharges suitability in Kabul aquifer system within the upper Indus Basin," *Water*, vol. 14, no. 15, p. 2390, 2022.
31. Visual Crossing, "Timeline weather API," *Visual Crossing*, 2025.
32. A. Nasir, V. Gurupur, and X. Liu, "A new paradigm to analyze data completeness of patient data," *Appl. Clin. Inform.*, vol. 7, no. 03, pp. 745–764, 2016.
33. P. Ahlgren, B. Jarneving, and R. Rousseau, "Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient," *J. Am. Soc. Inf. Sci. Technol.*, vol. 54, no. 6, pp. 550–560, 2003.
34. T. Mahmood *et al.*, "Enhancing coronary artery disease prognosis: A novel dual-class boosted decision trees strategy for robust optimization," *IEEE Access*, 2024.
35. T. Thomas, A. P. Vijayaraghavan, S. Emmanuel, T. Thomas, A. P. Vijayaraghavan, and S. Emmanuel, "Applications of decision trees," *Mach. Learn. approaches cyber Secur. Anal.*, pp. 157–184, 2020.
36. B. Gupta, A. Rawat, A. Jain, A. Arora, and N. Dhami, "Analysis of various decision tree algorithms for classification in data mining," *Int. J. Comput. Appl.*, vol. 163, no. 8, pp. 15–19, 2017.
37. S. Samadianfard, F. Mikaeili, and R. Prasad, "Evaluation of classification and decision trees in predicting daily precipitation occurrences," *Water Supply*, vol. 22, no. 4, pp. 3879–3895, 2022.
38. B. Charbuty and A. Abdulazeez, "Classification based on decision tree algorithm for machine learning," *J. Appl. Sci. Technol. Trends*, vol. 2, no. 01, pp. 20–28, 2021.
39. F. Baig, L. Ali, M. A. Faiz, H. Chen, and M. Sherif, "How accurate are the machine learning models in improving monthly rainfall prediction in hyper arid environment?," *J. Hydrol.*, vol. 633, p. 131040, 2024.
40. V. Kumar, N. Kedam, O. Kisi, S. Alsulamy, K. M. Khedher, and M. A. Salem, "A comparative study of machine learning models for daily and weekly rainfall forecasting," *Water Resour. Manag.*, pp. 1–20, 2024.
41. S. González, S. García, J. Del Ser, L. Rokach, and F. Herrera, "A practical tutorial on bagging and boosting based ensembles for machine learning: Algorithms, software tools, performance study, practical perspectives and opportunities," *Inf. Fusion*, vol. 64, pp. 205–237, 2020.
42. A. Berrouachedi, R. Jaziri, and G. Bernard, "Convolutional, extra-trees and multi layer perceptron," in *2022 IEEE/ACS 19th International Conference on Computer Systems and Applications (AICCSA)*, 2022, pp. 1–8.
43. T. E. Mathew, "An optimized extremely randomized tree model for breast cancer classification," *J. Theor. Appl. Inf. Technol*, vol. 31, pp. 5234–5246, 2022.
44. D. Baby, S. J. Devaraj, and J. Hemanth, "Leukocyte classification based on feature selection using extra trees classifier: Atransfer learning approach," *Turkish J. Electr. Eng. Comput. Sci.*, vol. 29, no. 8, pp. 2742–2757, 2021.
45. N. K. A. Appiah-Badu, Y. M. Missah, L. K. Amekudzi, N. Ussiph, T. Frimpong, and E. Ahene, "Rainfall prediction using machine learning algorithms for the various ecological zones of Ghana," *IEEE Access*, vol. 10, pp. 5069–5082, 2021.
46. A. Gumilar, S. S. Prasetiyowati, and Y. Sibaroni, "Performance analysis of hybrid machine learning methods on imbalanced data (rainfall classification)," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 6, no. 3, pp. 481–490, 2022.
47. A. Sarasa-Cabezuelo, "Prediction of rainfall in Australia using machine learning," *Information*, vol. 13, no. 4, p. 163, 2022.
48. G. Singh and D. Kumar, "Hybrid prediction models for rainfall forecasting," in *2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, 2019, pp. 392–396.
49. M. Roser and F. Moosmann, "Classification of weather situations on single color images," in *2008 IEEE Intelligent Vehicles Symposium*, 2008, pp. 798–803.
50. S. Goel, S. Markanday, and S. Mohanty, "Analysis of multi-class weather classification using deep learning models and machine learning classifiers," in *2022 OITS International Conference on Information Technology (OCIT)*, 2022, pp. 223–227.
51. J. An, Y. Chen, and H. Shin, "Weather classification using convolutional neural networks," in *2018 International SoC Design Conference (ISOCC)*, 2018, pp. 245–246.
52. L. Ortega, L. D. Otero, and C. Otero, "Application of machine learning algorithms for visibility classification," in *2019 IEEE International Systems Conference (SysCon)*, 2019, pp. 1–5.
53. E. Ship, E. Spivak, S. Agarwal, R. Birman, and O. Hadar, "Real-time weather image classification with SVM," arXiv Prepr. arXiv2409.00821, 2024.
54. A. Sharma and Z. S. Ismail, "Weather classification model performance: using CNN, Keras-Tensor Flow," in *ITM Web of Conferences*, 2022, vol. 42, p. 1006.

55. S. Minhas, Z. Khanam, S. Ehsan, K. McDonald-Maier, and A. Hernández-Sabaté, "Weather classification by utilizing synthetic data," *Sensors*, vol. 22, no. 9, p. 3193, 2022.
56. Z. Liu and Z. Zhang, "Solar forecasting by K-nearest neighbors method with weather classification and physical model," in *2016 North American power symposium (NAPS)*, 2016, pp. 1–6.
57. S. Dalal, B. Seth, M. Radulescu, T. F. Cilan, and L. Serbanescu, "Optimized deep learning with learning without forgetting (LwF) for weather classification for sustainable transportation and traffic safety," *Sustainability*, vol. 15, no. 7, p. 6070, 2023.
58. E. M. Ali and M. M. Ahmed, "Employment of instrumented vehicles to identify real-time snowy weather conditions on freeways using supervised machine learning techniques–A naturalistic driving study," *IATSS Res.*, vol. 46, no. 4, pp. 525–536, 2022.
59. K. Purwandari, J. W. C. Sigalingging, T. W. Cenggoro, and B. Pardamean, "Multi-class weather forecasting from twitter using machine learning aprroaches," *Procedia Comput. Sci.*, vol. 179, pp. 47–54, 2021.
60. M. S. Ghaleb, H. Moushier, H. Shedeed, and M. Tolba, "Weather classification using fusion of convolutional neural networks and traditional classification methods," *Int. J. Intell. Comput. Inf. Sci.*, vol. 22, no. 2, pp. 84–96, 2022.