International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V72I10P111

Volume 72 Issue 10, 96-118, October 2024

Original Article

An Iterative Model on ARIMA with LSTM Approach on Weather Forecasting

K. Sowjanya Bharathi[1], Boddu Sekhar Babu[2]

[1,2]*Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India.*

[1]*Corresponding Author : ksowjanyabharathi1985@gmail.com*

Received: 22 March 2024 Revised: 27 September 2024 Accepted: 02 October 2024 Published: 25 October 2024

Abstract *- Weather forecasting is essential for everyday living, influencing several industries and decision-making processes. This study looks at two ways to make short-term weather predictions better: the AutoRegressive Integrated Moving Average (ARIMA) model and a new method that uses Local Mean Squared Error (LMSE) optimisation in Long Short-Term Memory (LSTM) neural networks. By using several weather datasets, such as the UCI weather dataset, we may identify critical trends in meteorological data. The first segment of the proposed work implements the ARIMA model to examine historical weather data, isolating essential autoregressive and moving average elements for precise short-term forecasts. The following section presents LMSE optimisation inside LSTM networks, which refines the model to minimise prediction errors and enhance comprehension of long-term relationships in the data. This study aims to enhance the precision of short-term weather predictions by integrating the advantages of ARIMA and LMSE-optimised LSTM models. The results will provide meteorologists with dependable instruments for forecasting meteorological phenomena, assist emergency responders in making educated choices, and deliver more precise weather data to the public, improving readiness and safety in the face of fluctuating weather conditions.*

Keywords *- ARIMA, AdaSTL, CNN, GRU, LSTM, LMSE, Short-term forecasts, Time series analysis, UCI weather dataset, Weather forecasting.*

1. Introduction

Weather forecasting is essential in several areas, such as agriculture, transportation, emergency management, and everyday living. Precise meteorological forecasts assist in alleviating the effects of extreme weather phenomena, improving public safety, and maximising resource allocation. Climate change's increasing predictability of weather patterns necessitates the urgent development of better forecasting technologies. Recent breakthroughs in machine learning and data assimilation methodologies have markedly enhanced predicting precision. Wang and Zhao studied the Enformer model in 2023 and found that encoder-based sparse periodic self-attention mechanisms might help with time series forecasting by accurately capturing complex temporal dynamics [1]. Zhang et al. (2023) presented the Temporal Chain Network using an intuitive attention mechanism to improve long-term forecasting precision [2]. Confronting contemporary meteorological forecasting issues despite technical advancements, weather forecasting continues to encounter problems, such as the integration of many data sources and the need for high-resolution models. Yan et al. (2023) introduced a three-dimensional cloud detection approach, improving the functionality of numerical weather forecast systems [3]. These approaches demonstrate the significance of combining satellite data with terrestrial observations to enhance precision. Sengoz et al. (2023) [4] note that machine learning techniques address these issues by enhancing precipitation predictions in North America. These methodologies enhance predictive accuracy while simultaneously optimising model efficiency. The dataset used for weather forecasting is critical to the model's efficacy. The Seattle-Weather CSV file contains extensive meteorological datasets that include diverse resolutions, timeframes, and geographic coverage. Datasets may vary from high-resolution hourly data to larger daily averages, including numerous years and varied geographical regions.

This extensive dataset enables models to learn from historical weather trends efficiently, resulting in enhanced predictions. Integrating sophisticated machine learning methodologies with superior datasets is crucial for improving weather forecasting. By addressing existing issues and using novel models, the sector may significantly improve its prediction abilities, eventually helping society as a whole. In conclusion, the changing dynamics of weather forecasting necessitate the use of innovative technology and approaches. Researchers like Yang et al. (2023) [5] are working hard to solve problems in weather prediction. Xia et al. (2023) [6] are integrating data from multiple sources to make long-sequence time series forecasting better. Improving numerical weather prediction models with machine learning (Choi and Jung, 2023) [8] makes it possible to make more accurate predictions. This shows how important it is for people from different fields to work together and use technology in new ways.

1.1. Problem Statement

Weather forecasting is a challenging endeavour because of the chaotic characteristics of the atmosphere and the impact of several factors on meteorological patterns. Conventional forecasting techniques, such as Numerical Weather Prediction (NWP), sometimes encounter difficulties in delivering precise long-term predictions because of significant unpredictability and the complexities of synthesising several data sources, including satellite images and terrestrial observations. The intrinsic noise in meteorological data results in considerable uncertainty in forecasts, and current models sometimes lack the necessary flexibility to address changing weather patterns, therefore hindering precise forecasting. The proposed goals are to solve these problems by using advanced machine learning techniques, mainly LSTM (Long Short-Term Memory) and ARIMA (AutoRegressive Integrated Moving Average) models, along with fast algorithms (HPA). The primary goal is to improve temporal comprehension and classification precision, directly addressing the complexities of meteorological dynamics. The subsequent aims are to augment robustness and accuracy, explore real-time adaptation, and boost noise reduction in forecasts. This study aims to provide a reliable framework for weather forecasting by juxtaposing these models with cutting-edge designs, enabling efficient adaptation to fluctuating circumstances and reduction of data noise, thereby enhancing the precision and utility of weather forecasts.

1.2. Objectives

- To implement a heuristic approach with gaussian and Genetic algorithm (HPA) with Arima Model also to improvise the LSME realization for LSTM with Genetic HPA model for noise reduction methods for the encoder and decoder architecture.
- To encapsulate the combined performance of the LSTM and ARIMA with HPA To develop a hybrid Architecture adjusting the decomposed hidden states with the least possible errors in all multi-layer architecture with LSTM and ARIMA model.
- To compare the proposed models and compare them with state-of-the-art architectures based on GRU, GAN and AdaSTL models.

2. Literature Survey

Xiao et al. [1] investigated the selection of temperature and water vapor channels in the FY-3E HIRAS II satellite for Numerical Weather Prediction (NWP). By analyzing the effectiveness of various channels through Jacobian matrices and Degrees of Freedom for Signal (DFS), the study highlights the potential of optimizing satellite data utilization. This research addresses a critical gap in the integration of remote sensing data into weather models, suggesting that precise channel selection can significantly improve forecast accuracy. The findings emphasize the importance of using tailored satellite observations to enhance NWP systems. Jang et al. [2] focussed on spatiotemporal post-calibration within NWP models to quantify building energy consumption based on weather data. Their approach employs advanced machine learning algorithms, revealing complex relationships between meteorological forecasts and energy usage patterns. This study contributes to the understanding of how weather impacts energy systems, filling a notable gap in research that has often neglected the interplay between weather predictions and energy consumption.

By establishing a framework that integrates these domains, the research paves the way for improved energy management strategies in response to variable climatic conditions. Lin et al. [3] presented a Spherical Neural Operator Network aimed at global weather prediction. This innovative approach utilizes deep learning to capture intricate spatial-temporal relationships in meteorological data, showcasing improved predictive capabilities over traditional models.

The study highlights the limitations of existing NWP systems in handling dynamic weather patterns, thus demonstrating a significant advancement in operator learning methods. The application of such neural network architectures indicates a shift toward more adaptable forecasting systems that can better accommodate the complexities of atmospheric dynamics. Ye et al. [4] addressed the challenges of predicting wind power generation during extreme weather events through a combined methodology using time series adversarial generation networks.

Their research emphasizes the necessity for robust predictive algorithms capable of managing the uncertainties associated with renewable energy generation. By focusing on extreme weather, the study contributes valuable insights into optimizing wind power forecasts, critical for energy grid stability and management. This research highlights the intersection of meteorology and renewable energy, underlining the importance of advanced forecasting techniques in the context of climate change. Gong et al. [5] explored spatial-temporal enhanced contrastive and contextual learning for weather forecasting. Their work demonstrates how advanced machine learning techniques can improve predictive accuracy by effectively identifying complex weather patterns.

The research underscores the need for innovative methodologies that can process vast amounts of meteorological data, addressing existing challenges in forecasting. By integrating contrastive learning with contextual data, the study lays the groundwork for future explorations into self-supervised learning applications in meteorology, offering promising directions for enhancing predictive models. Yan et al. [6] developed a 3-D cloud detection method for FY-4A GIIRS, which they applied in an operational NWP system. Their study presents a novel approach to identifying cloud patterns that can significantly affect weather predictions. By utilizing this method, they improve the accuracy of cloud-related forecasts, contributing to more reliable meteorological models. The findings demonstrate the critical role of precise cloud detection in enhancing operational forecasting capabilities, filling a gap in the real-time application of satellite data in weather prediction systems. Gahwera et al. [7] analysed machine learning algorithms to predict short-term rainfall amounts using data from Uganda's Lake Victoria Basin.

Their research identifies the strengths and weaknesses of various models, providing insights into the most effective predictive techniques for local weather patterns. This study contributes to the growing body of literature on applying machine learning in meteorology, particularly in regions vulnerable to climatic variability. By focusing on a specific geographic area, the research emphasizes the necessity for localized approaches in weather prediction, addressing a significant gap in generalizable forecasting models. Sun et al. [8] proposed a deep-learning network-based algorithm for convective weather initiation, integrating data from Fengyun-4A satellites and radar for real-time nowcasting. Their work addresses the urgent need for timely weather warnings, especially in rapidly changing convective conditions.

By utilizing joint observations, the study enhances the accuracy and responsiveness of weather predictions in critical time frames, showcasing the value of combining different data sources. This research underscores the importance of real-time forecasting capabilities in mitigating the impacts of severe weather. Zhang et al. [9] introduced a phase turbulence prediction method for line-of-sight Multiple-Input-Multiple-Output (MIMO) links affected by atmospheric conditions. Their findings highlight the significance of understanding atmospheric influences on signal transmission, which is vital for maintaining communication quality.

The study contributes to the field by providing a predictive framework that can improve MIMO link reliability in varying weather conditions. This research not only addresses challenges in communication technology but also integrates meteorological factors, bridging a gap between atmospheric science and telecommunications. Li et al. [10] examined the flight delay prediction by incorporating both weather and non-weather features. Their analysis reveals how integrating diverse data sources can enhance the accuracy of predictions related to air traffic disruptions. This study fills a critical gap in transportation research by emphasizing the multifaceted nature of flight delays and the importance of weather conditions. By providing a comprehensive framework for flight delay prediction, the research contributes to more efficient air traffic management and improved traveler experience. Wu et al. [11] offered an overview of day-ahead solar power forecasts based on weather classifications, complemented by a case study in Taiwan.

Their research illustrates the potential of using meteorological classifications to improve solar energy predictions, which are essential for effective energy management. By highlighting the relationship between weather patterns and solar power generation, the study contributes significantly to renewable energy forecasting. This research addresses the need for more accurate predictive models in the face of increasing reliance on solar energy. Biscarini et al. [12] have explored an optimal stochastic prediction and verification of signal-to-noise ratios for Ka-band spaceborne telemetry utilizing weather forecasts.

Their work underscores the importance of accurate weather information in enhancing satellite communication systems. By integrating weather data into predictive models, the study provides valuable insights for improving communication reliability in various atmospheric conditions. This research contributes to the advancement of satellite telemetry technologies by addressing the impact of environmental factors. Tekin et al. [13] proposed a convolutional-LSTM model for numerical weather forecasting enhanced by attention and context matcher mechanisms. Their approach addresses the limitations of traditional forecasting models, showcasing how deep learning can be leveraged to improve predictive performance.

The study contributes to the ongoing evolution of meteorological modelling by integrating advanced neural network architectures. By focusing on the intricacies of weather patterns, this research offers promising directions for future developments in weather forecasting technologies.

Wu et al. [14] investigated the probabilistic forecasting of wind power generation, incorporating data processing and numerical weather predictions. Their findings emphasize the importance of probabilistic models in managing the uncertainties inherent in wind energy forecasting. This research fills a critical gap by providing methodologies that enhance the reliability of wind power predictions, thereby contributing to the optimization of energy management systems.

The integration of weather data with energy forecasting represents a vital intersection of meteorology and energy science. Du et al. [15] have imparted a prediction of weather-related failure risks in distribution systems using a Bayesian neural network. Their approach highlights the role of weather in influencing the reliability of power distribution networks, emphasizing the need for predictive modeling in infrastructure management. This study contributes to the understanding of how weather impacts electrical systems, addressing a significant gap in infrastructure resilience research. By integrating meteorological data into failure risk predictions, the research paves the way for more robust and responsive energy distribution systems. Wang and Zhao [16] introduced "Enformer," an encoder-based sparse periodic self-attention model for time-series forecasting. Their work focuses on enhancing the efficiency of attention mechanisms, making it particularly suited for forecasting tasks that involve periodic data patterns. The study demonstrates how this approach can improve predictive accuracy while reducing computational overhead, thus addressing a critical gap in traditional time-series forecasting methods.

The findings suggest that optimizing attention mechanisms can lead to significant advancements in forecasting techniques across various applications. Zhang et al. [17] developed a Temporal Chain Network featuring an intuitive attention mechanism specifically designed for long-term series forecasting. This innovative approach allows the model to capture temporal dependencies more effectively, thereby improving forecast accuracy over extended periods. The research contributes to the understanding of how attention mechanisms can be adapted for time-series data, addressing the limitations of existing models.

The findings emphasize the potential of intuitive attention in enhancing predictive performance in complex time-series forecasting scenarios. Yan et al. [18] presented a 3-D cloud detection method for the FY-4A GIIRS, applying it within an operational Numerical Weather Prediction (NWP) system. This study enhances the accuracy of cloud detection, which is crucial for reliable weather forecasting. The integration of this method into NWP systems demonstrates the importance of precise cloud characterization in improving forecast outcomes.

The research fills a notable gap in real-time satellite data applications, paving the way for advancements in operational weather prediction capabilities. Sengoz et al. [19] have improvised various machine learning approaches to enhance precipitation forecasts in North America. Their analysis identifies the strengths of different algorithms in predicting precipitation, emphasizing the need for sophisticated models to capture the complexities of weather patterns. This research contributes significantly to improving forecasting accuracy in a region where precipitation variability is a critical concern.

By addressing existing challenges in precipitation modelling, the study highlights the value of machine learning in meteorological applications. Yang, Ma, and Huang [20] developed a novel architecture with ATFSAD for enhancing long-sequence time-series forecasting specifically for air temperature prediction. Their research demonstrates how advanced algorithms can improve the reliability of temperature forecasts over extended timeframes. This study addresses a critical gap in traditional forecasting methods that struggle with long-term predictions. By emphasizing the significance of accurate temperature forecasting, the research contributes to better climate and environmental management strategies. Xia et al. [21] proposed a deep learning method that integrates multisource data to correct ECMWF (European Centre for Medium-Range Weather Forecasts) forecasting products. Their approach addresses the limitations of existing models by leveraging diverse data sources for improved accuracy. The study emphasizes the importance of data integration in enhancing weather forecasts, filling a gap in the application of machine learning for NWP corrections.

This research signifies a step toward more robust and reliable meteorological predictions. Similarly, Choi and Jung [22] have depicted optimizing the performance of numerical weather prediction models through machine learning techniques. Their research showcases how various machine learning algorithms can enhance model accuracy and efficiency, addressing the growing complexity of meteorological data. By applying these techniques, the study fills a significant gap in traditional NWP approaches, demonstrating that machine learning can provide substantial improvements in forecasting capabilities.

This research paves the way for the integration of advanced computational methods in meteorology. Chen et al. [23] investigate short-term load forecasting combined with associated weather variables using a ResNet-LSTM deep learning model. Their study highlights the interplay between weather conditions and energy demand, emphasizing the need for accurate forecasting in energy management.

By integrating meteorological data into load forecasting, the research contributes valuable insights into optimizing energy systems. This work addresses a gap in understanding how weather impacts short-term energy consumption, which is crucial for effective grid management. The design of a TCN based hybrid forecasting framework from Li et al. [24] for utility-scale Photovoltaic (PV) power generation is implemented and explored with various possible solutions to embed the metrological changes in the atmosphere. Their study emphasizes the importance of accurate forecasting in the management of renewable energy resources, particularly under variable weather conditions.

By integrating time convolutional networks with traditional methods, the research fills a critical gap in PV forecasting, demonstrating how hybrid approaches can enhance predictive accuracy. The findings underscore the need for sophisticated forecasting techniques in the renewable energy sector.

Han et al. [25] propose a multi-view, multi-adversarial learning approach for joint air quality and weather prediction. Their innovative model leverages adversarial training to improve the accuracy of both air quality and weather forecasts.

This research highlights the interconnectedness of meteorological and environmental data, addressing a significant gap in existing prediction models. By focusing on joint forecasting, the study offers a promising avenue for advancing predictive capabilities in both air quality management and weather forecasting.

Suleman and Shridevi [26] developed a spatial feature attention-based LSTM model for short-term weather forecasting. Their work emphasizes the importance of spatial features in improving the accuracy of weather predictions, showcasing the potential of attention mechanisms in LSTM architectures.

This research contributes to the field by addressing the need for more nuanced forecasting models that can capture complex spatial relationships. The findings highlight the value of integrating spatial attention in enhancing short-term weather forecasting capabilities. Essa et al. [27] have demonstrated a deep learning technique for predicting thunderstorm severity using remote sensing weather data. Their research illustrates how advanced neural networks can enhance the understanding of severe weather events, providing critical information for timely warnings.

By applying remote sensing data to predict thunderstorm impacts, the study fills a gap in severe weather forecasting research, highlighting the importance of integrating modern computational techniques in meteorology. Barnes et al. [28] explored to implementation of video-based convolutional neural networks for rainfall forecasting. Their innovative approach utilizes video data to capture dynamic weather patterns, providing a novel method for predicting rainfall.

This research addresses a significant gap in traditional rainfall forecasting methodologies, demonstrating how visual data can enhance predictive capabilities. The findings underscore the potential of combining video analysis with meteorological forecasting to improve rainfall prediction accuracy. Yu et al. [29] present ATMConvGRU, a model for weather forecasting that combines atmospheric convolutional layers with gated recurrent units.

Their approach addresses the limitations of traditional forecasting models by effectively capturing temporal and spatial dependencies in weather data. This research contributes significantly to the advancement of meteorological modeling techniques, filling a gap in the application of deep learning to weather predictions. The findings indicate the potential for enhanced forecasting accuracy through innovative model designs. Vu et al. [30] propose a recurring multi-layer moving window approach to forecast day-ahead and week-ahead load demand, considering weather conditions.

Their research highlights the importance of incorporating weather data into load forecasting models, providing valuable insights for energy providers. By addressing the impact of weather on energy consumption patterns, the study fills a critical gap in understanding load dynamics. This research contributes to the development of more responsive energy management strategies, essential for balancing supply and demand.

3. Existing Design

The MLP-Self-Attention Time Layer (MLP-SATL) model introduces a ground-breaking approach to weather prediction, combining the strengths of Multilayer Perceptron (MLP) and Self-Attention mechanisms. In the concept section, it is emphasized that this model stands as a powerful tool for real-time weather forecasting, excelling in capturing complex temporal dependencies within weather data.

The integration of MLP's ability to explore non-linear associations and Self-Attention's capability to recognize temporal patterns ensures robust and accurate predictions, addressing the increasing demand for high-performing weather prediction tools. The model's adaptive nature and proficiency in understanding intricate temporal relationships position it as a valuable asset for diverse applications, including agriculture, transportation, disaster management, and urban planning.

In the design, the block diagram illustrates the integration of key components, including the Self-Attention mechanism, MLP, timeline data, AdaSTL for handling seasonality, and an Encoder-Decoder architecture for transfer learning. The algorithms, such as STL Time Series Decomposition and Self-Attention Mechanism, are explained in detail, showcasing the model's comprehensive approach to capturing temporal dynamics and handling complex patterns in weather data.

The inclusion of AdaSTL ensures effective management of seasonality, enhancing the model's adaptability. Overall, the model's design methodology emphasizes its ability to analyse chronological sequences, handle seasonal variations, and leverage transfer learning for robust performance across diverse datasets and applications.

3.1. Design Considerations

The work in [20] implicates on the AdaSTL framework mentioned in the Figure 1. The design process is dealt with to implicate the temporal and spatial patterns realized with global dependencies with input and output. This process with a multi-layer architectural framework has proven with the complexity of (O(nlogn)) with timing criteria.

Similarly, an informative filter module is designed to incorporate the noise generated with each type of predicted outcome from the encoder and decoder layer. Finally, this model tends to provide the adaptability to inculcate refinement with filter models with multi-layer architecture based on seasonal and mitigation components for predicting drift.

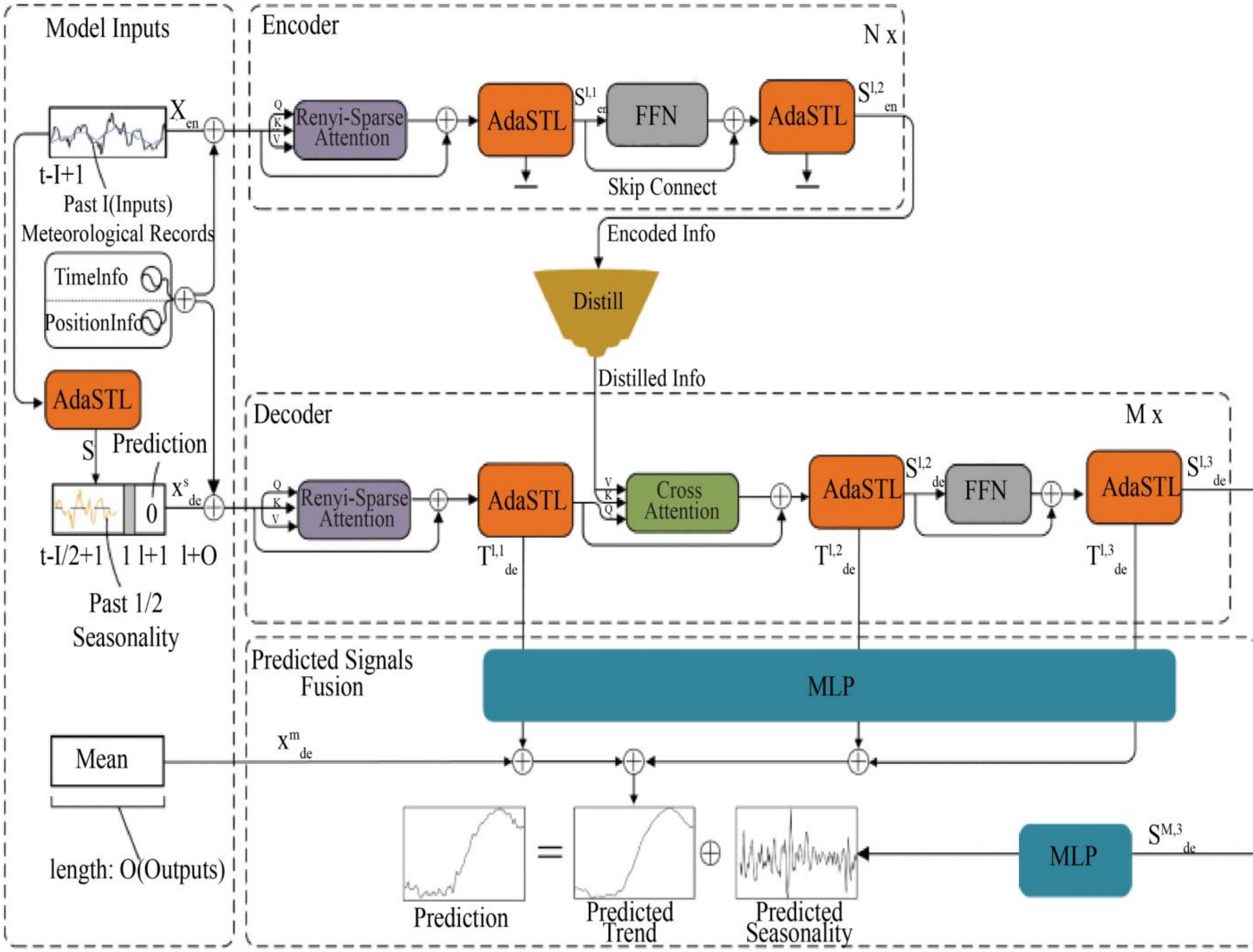

Fig. 1 The overall block diagram for AdaSTL functionality with the weather forecasting model [20]

3.2. Formulations

The STL design on the Drift prediction with seasonal components (S), trend components (T), time series (Y), and residual component (R). The preliminary procedure for STL algorithm for the work in [20] is mentioned below:

Step1: De-trending:

The trend component of the dataset component is subtracted from its previous case values as mentioned below:

$$Y = Y - T^{k-1} \tag{1}$$

where k I is the iteration of the current sequence of the data while the trend component is T^0 as initial value.

Step 2: In this step, the overall design is applied with weighted regression analysis to smoothen the samples uniformly and extend this approach for both cycles forward and backwards. The combined results will form a partial sequence for seasonal AA represented by C^k

Step 3: This step determines the pre and post case of filtering with cycle sub-series generated functionality. The moving average condition with $n_p, n_p, 3\ and\ C^k$ and apply the weights to the regression values obtained with the parameter n_i and to get the sequence L^k with length N.

Step 4: Determine the smooth sequence with formulation as:

$$S^k = C^k - L^k \tag{2}$$

Step 5: De-seasonlization with $Y - S^k$ (3)

Step 6: Trend Smoothening: Apply the regression weights to the time series functionality to generate an overall sequence with above seasonalization:

For k=1: N:

$$S \leftarrow S^M,\ T \leftarrow T^M,\ R \leftarrow (Y\text{-}S\text{-}T) \tag{4}$$

end

4. Material and Methods

The proposed hybrid model study addresses the persistent challenge in weather forecasting—improving short-term prediction accuracy—by proposing two novel approaches: the ARIMA model and LMSE-optimized LSTM neural networks. While ARIMA is adept at capturing short-term dependencies, LMSE-optimized LSTM networks excel in handling long-term dependencies and complex relationships within weather data.

The integration of these approaches aims to enhance predictive accuracy significantly. The research also involves a comprehensive comparative analysis with established algorithms like KNN, SVM, and Ensemble methods to evaluate performance metrics such as accuracy, computational efficiency, and adaptability. Leveraging diverse weather datasets, including the UCI weather dataset, ensures the robustness and adaptability of the proposed ARIMA+LMSE LSTM approach. Ultimately, the study addresses a critical need in society by striving to provide more precise and reliable short-term weather predictions for meteorologists, emergency responders, and the public.

4.1. Concept

The integration of HPA (Heuristic Prediction Approach) with ARIMA and the combination of LSTM (Long Short-Term Memory) with UNET architecture is of paramount importance in advancing the field of weather forecasting. These innovative approaches aim to bring about a substantial improvement in the precision and reliability of short-term weather predictions.

By blending data-driven models like ARIMA and LSTM with expert-driven insights and the sophisticated filters of HPA, these approaches enable us to not only capture complex spatiotemporal patterns within weather data but also harness domain-specific knowledge to adapt swiftly to changing conditions.

In an era where the demand for highly accurate and timely weather forecasts is continuously growing, HPA+ARIMA and LSTM+HPA architectures hold the promise of providing essential, actionable information to meteorologists, emergency responders, and the general public, ultimately enhancing decision-making and public safety in the face of dynamic and unpredictable weather scenarios.

4.2. ARIMA+HPA Architecture

The ARIMA+HPA (Autoregressive Integrated Moving Average with Heuristic Prediction Approach) architecture, presented in Figure 2, shows an innovative fusion of two methodologies aimed at augmenting the accuracy of short-term weather forecasting. Integrating the well-established ARIMA model with a Heuristic Prediction Approach (HPA), this hybrid framework seeks to deliver more robust predictions of weather conditions, catering to the needs of meteorologists, emergency responders, and the wider public. The ARIMA model, renowned for its prowess in time series analysis, excels at capturing short-term dependencies within weather data. In tandem, the Heuristic Prediction Approach incorporates domain-specific expert knowledge, allowing meteorologists to provide valuable insights, especially in the face of rapidly changing or extreme weather scenarios.

The hybrid approach capitalizes on the synergy between the statistical rigor of ARIMA and the contextual insights offered by meteorologists through HPA. This amalgamation aims to bridge the gap between data-driven and expert-driven forecasting, fostering higher precision in short-term weather predictions. The architecture is designed to adapt swiftly to dynamic weather conditions, which is particularly crucial for emergency response and public safety.

To validate its efficacy, rigorous comparative analyses against state-of-the-art weather forecasting models are imperative, scrutinizing its performance in terms of accuracy and efficiency. The ARIMA+HPA architecture, thus, emerges as a sophisticated and adaptive solution poised to elevate the reliability of short-term weather forecasts, addressing critical requirements across diverse sectors.

4.3. LSTM+HPA Architecture

The proposed LSTM+HPA model, in Figure 3, for weather prediction integrates various layers, including max-pooling, LSTM, Dense, Dropout, and Output Dense layers. Each layer serves a specific function in feature extraction, temporal processing, and prediction generation. The design emphasizes the preservation of spatial and temporal information to enhance forecasting capabilities.

Stacking these layers sequentially and connecting their outputs appropriately ensures a comprehensive approach to leverage both spatial and temporal contexts within weather data. Fine-tuning of hyperparameters may be necessary based on the specific weather dataset and prediction task.In parallel, the design aspects of LSTM and the Heuristic Prediction Approach (HPA) involve meticulous steps.

The LSTM, designed to capture temporal dependencies, undergoes data pre-processing and training using the UCI Seatle dataset. Simultaneously, the HPA-LSTM model is constructed to capture spatial features, potentially involving Genetic Filters. The integration of LSTM and HPA models requires combining learned features, followed by fine-tuning and validation for optimal performance. The potential impact on weather prediction lies in the synergistic utilization of STM for temporal pattern learning and HPA for spatial feature handling, offering a more comprehensive understanding of both temporal and spatial aspects of weather information data with the Seatle Dataset for improved forecasting insights.

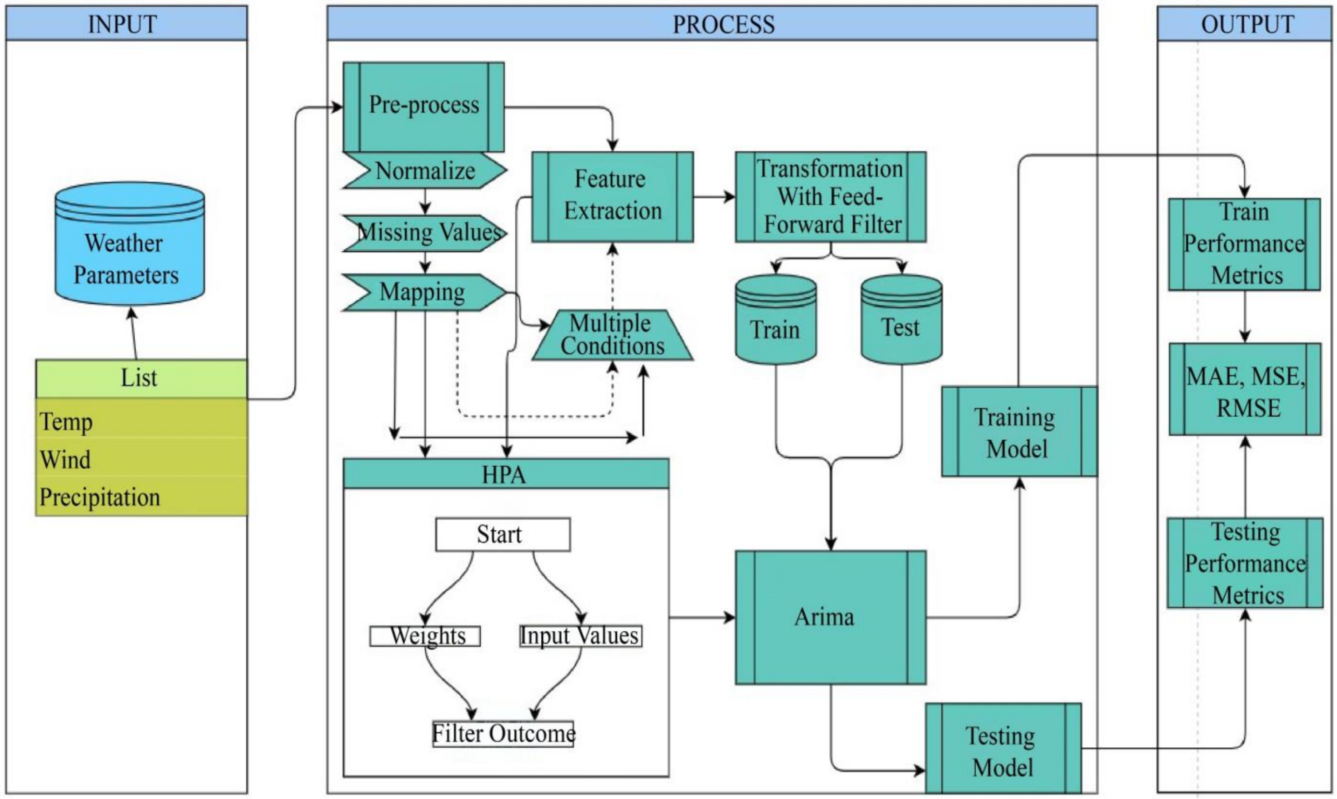

Fig. 2 Representing the overall block diagram with ARIMA + HPA model

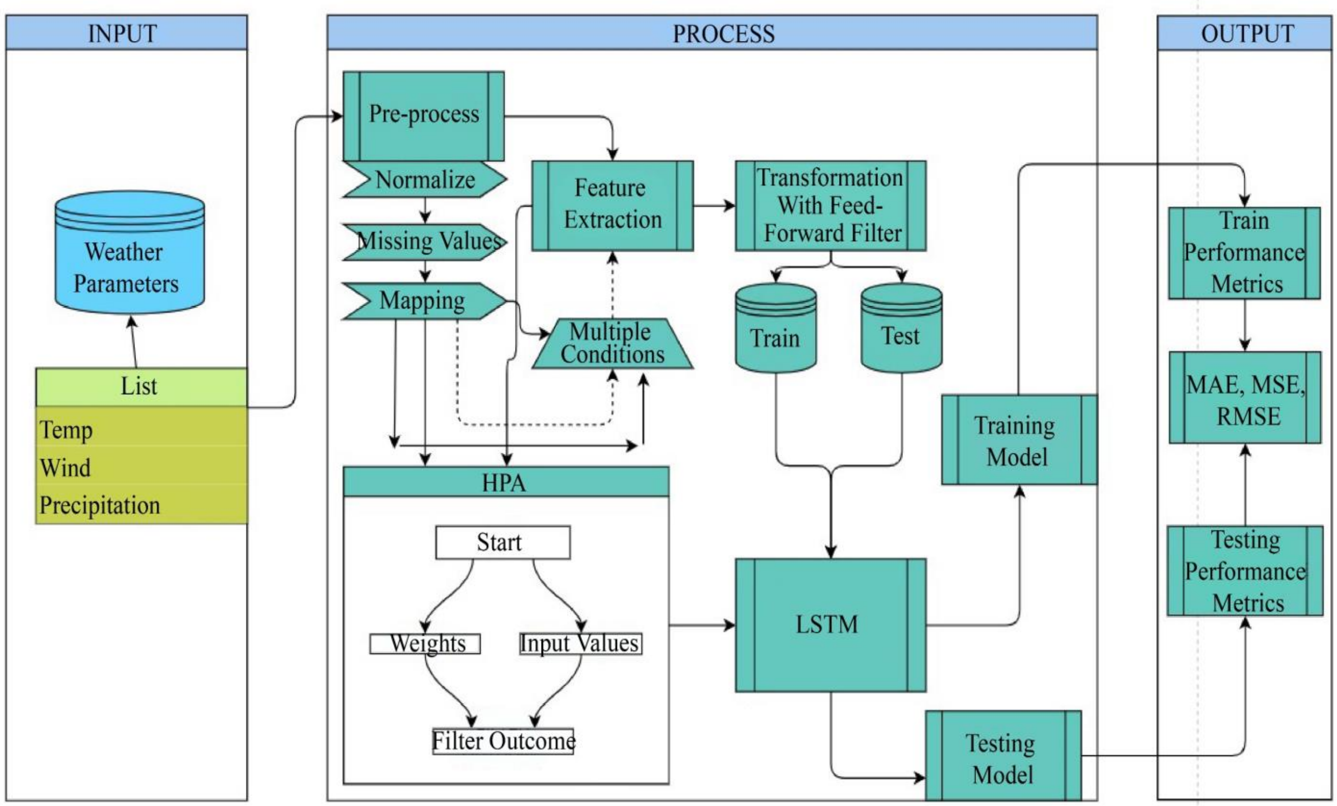

Fig. 3 The overall block diagram with LSTM + HPA model

4.4. Combined Architecture

The hybrid model combining ARIMA+HPA and LSTM+HPA represents a sophisticated approach to weather prediction that integrates statistical and deep learning techniques to enhance forecasting accuracy. At its core, this architecture leverages the strengths of time series analysis through the ARIMA model while simultaneously harnessing the power of deep learning with LSTM networks. This dual framework is particularly effective in capturing both short-term and long-term dependencies in meteorological data. The ARIMA model provides a solid foundation by effectively modelling linear relationships and seasonality in weather data, making it adept at identifying trends and cycles that are crucial for accurate short-term forecasts. By integrating the Heuristic Prediction Approach (HPA), which incorporates relevant patterns, the model further enhances its predictive capability, especially during extreme weather conditions when rapid changes are likely. The LSTM component in Figure 4 of the hybrid model complements the ARIMA approach by focusing on the temporal aspects of weather data. Long Short-Term Memory networks are designed to remember patterns over extended periods, making them particularly suited for time series forecasting. In the context of weather prediction, LSTMs can learn complex relationships among various meteorological parameters such as temperature, humidity, and wind speed. By processing sequences of data, the LSTM can discern patterns that traditional models might miss.

The addition of HPA into the LSTM framework allows for a more nuanced understanding of spatial features, as HPA can guide the LSTM in recognizing critical variables that influence weather patterns.

This combination enables the model to adapt more effectively to non-linear changes in weather, providing more reliable predictions. In terms of optimization, the hybrid model significantly influences the Least Mean Square Error (LMSE) optimization process. LMSE is a critical criterion for assessing the model's accuracy, and with the hybrid architecture, the model benefits from the strengths of both statistical and deep learning methods. The ARIMA model's initial forecasts can serve as a baseline, while the LSTM model refines these predictions through iterative learning and adjustment based on incoming data. The optimization process focuses on minimizing the error by adjusting model parameters, effectively balancing the contributions of both ARIMA and LSTM components. By combining these approaches, the model can reduce variance and bias in predictions, thereby achieving lower LMSE values. This optimization is further enhanced through techniques such as backpropagation in the LSTM, which allows for fine-tuning based on previous prediction errors, ultimately leading to improved forecasting performance. The design of this hybrid model also emphasizes the importance of feature engineering and selection. Effective forecasting requires not just raw data but also relevant features that can inform predictions.

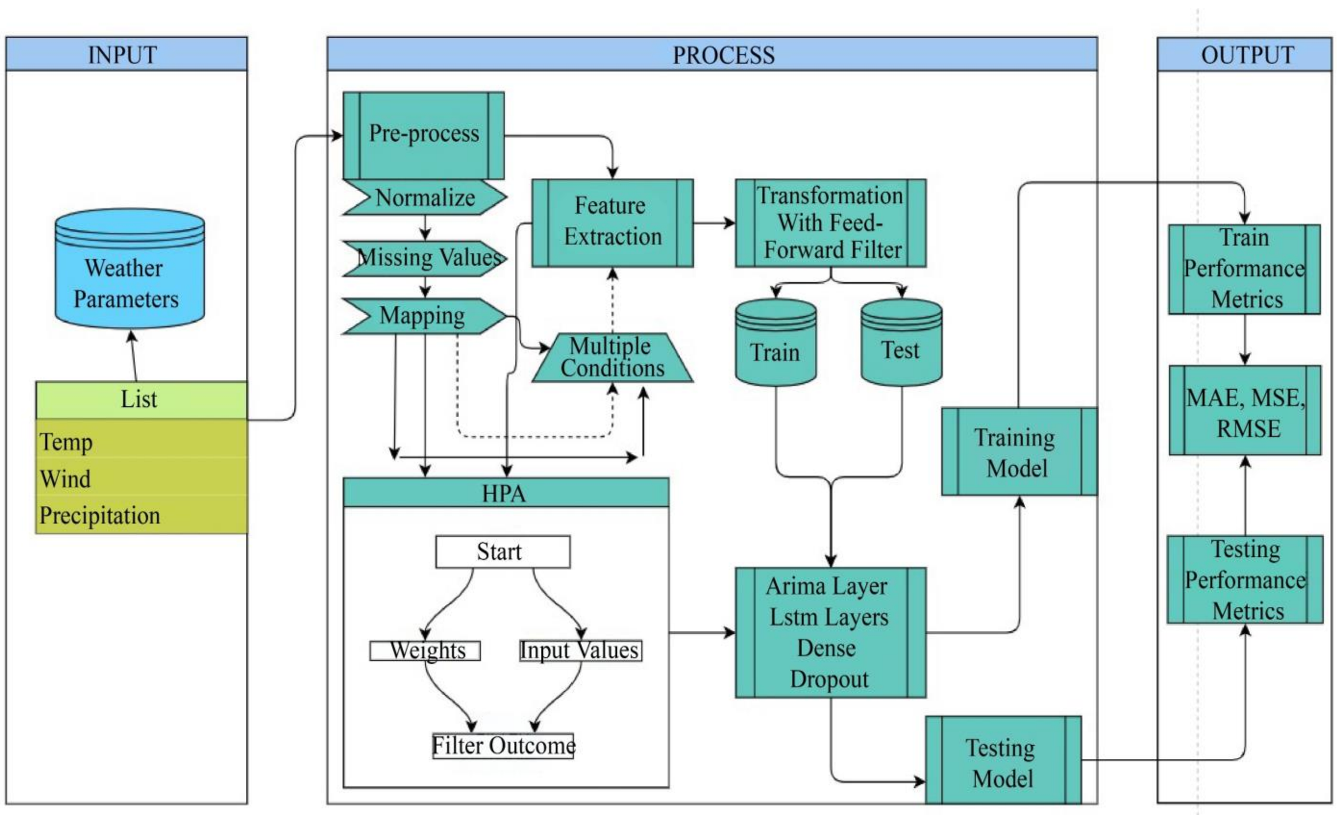

Fig. 4 The overall block diagram with LSTM +ARIMA+ HPA model

The integration of HPA plays a crucial role here, as it allows meteorologists to incorporate domain-specific insights into the feature set. This can include meteorological parameters like temperature, humidity, and precipitation, as well as derived features such as moving averages or seasonal indicators.

By leveraging these tailored features within the LSTM architecture, the model can enhance its understanding of complex interactions among different weather parameters. The attention to feature selection further aids in minimizing overfitting, a common challenge in deep learning models, by ensuring that only the most relevant data informs the predictions.

4.5. Algorithm

4.5.1. HPA (Gaussian)

ALGORITHM: HPA-1 (Gaussian Approach):

1. Initialization and Preprocessing

- **Input:** Dataset *D* with features relevant to the prediction task.
- **Preprocessing:** Handle missing values, scale/normalize features, and perform data formatting.

2. Gaussian Filtering Operation

- **Function Gaussian-Filter(*D*, *σ*)**
  1. Create a Gaussian kernel *G* based on the given standard deviation (σ).
  2. Convolve the dataset *D* with the Gaussian kernel to apply the filter, reducing noise and smoothening the data.

3. Hybrid Prediction Approach (HPA)

- **Function HPA(*D*)**
  1. Perform feature engineering to extract relevant characteristics from the filtered dataset.
  2. Choose and apply predictive models (e.g., statistical, machine learning, heuristic) considering the specific task requirements.
  3. Train models on historical data and validate their performance.

4. Prediction Phase

- **Function** $Forecast\ (2012-2024, D, 2012-2024)$
  1. Apply the Gaussian filter to the available data for the years 2012-2024 (2012−2024, *D*, 2018−2024).
  2. Utilize the trained models from HPA on the filtered dataset to make predictions for the upcoming years.
  3. Generate forecasts and provide estimates for the targeted variables based on the combined model outputs.

5. Output and Analysis

- **Function:** $Evaluate\_and\_Report$**(*Predictions*)**
- Evaluate the accuracy, precision, and any other relevant metrics of the predictions.

4.5.2. HPA (Genetic)

ALGORITHM 2: HPA (Genetic Filter):

1. **Initialization and Preprocessing**

- **Input:** Dataset *D* with features pertinent to the prediction task.
- **Preprocessing:** Handle missing values, scale/normalize features, and format data for subsequent operations.

2. Genetic Filtering Operation

- **Function Genetic Filter(*D*)**
  1. Encode features of the dataset *D* into chromosomes for genetic operations.
  2. Initialize a population of filter kernels.
  3. Define fitness functions to evaluate the quality of filter kernels based on noise reduction and information preservation.
  4. Employ genetic algorithms (crossover, mutation) to evolve the filter kernels, selecting the best-performing ones.

3. Hybrid Prediction Approach (HPA)

- **Function HPA (*D filtered*)**
  1. Use the filtered dataset *D filtered* obtained from the genetic filter as input to the prediction models.
  2. Apply feature engineering and select suitable prediction models (statistical, machine learning, heuristic).
  3. Train and validate models on historical data to capture patterns.

4. Prediction Phase

- **Function** *Forecast(2012−2024D2012−2024)*
  1. Apply the genetic filtering algorithm to the available data for the years 2018-2024 (2018−2024*D*2018−2024).
  2. Implement the Hybrid Prediction Approach on the filtered dataset to generate predictions for the upcoming years.
  3. Produce forecasts and provide estimates for the targeted variables based on the combined model outputs.

5. Output and Analysis

- **Function Evaluate_and_Report(*Predictions*)**
  1. Evaluate the accuracy, precision, and relevant metrics of the predictions.

4.5.3. ARIMA (GAUSSIAN)

ALGORITHM 1: ARIMA+HPA:

Inputs: $X_{in}, W_{in}, HPA(df), ARIMA(df_{filtered})$

Outputs: $Y_{hpa}, y_{ARIMAout}$

Procedure:

Step 1: Load the Seattle weather dataset and perform pre-processing, including normalization of features, handling missing values, and splitting the dataset into training and testing sets.

Step 2: Apply the Gaussian Criteria for the HPA Filter design

Step 3: Calculate the Training and Testing with Linear Prediction (Gradient Boosting Regressor)

Step 4: Apply ARIMA predicted values to the observed prediction.

Step 5: Append the values with proper weight functionality as the feedback process and repeat the steps (1)-(5) till the error is correct tolerance of 10%.

Step 6: Evaluate the prediction criteria with the HPA filter as post processing to produce the correct prediction with a mean square error of 5.384 in Table1.

End Procedure

The ARIMA + HPA algorithm focuses on time series forecasting using the Seattle weather dataset. The process begins with essential pre-processing steps, such as normalizing features, addressing missing values, and splitting the dataset into training and testing sets. Once the data is prepared, the algorithm applies the Gaussian Criteria to design the HPA filter, which is crucial for effective prediction.

Subsequently, the algorithm employs a Gradient Boosting Regressor to perform linear predictions on the training and testing data. ARIMA (AutoRegressive Integrated Moving Average) is then applied to integrate the predicted values with the observed data, allowing for a more robust forecasting approach.

The iterative nature of the algorithm is significant; it appends the predicted values with appropriate weights as a feedback mechanism and repeats the entire process until the prediction error falls within a specified tolerance level of 10%.

Finally, the predictions are refined using the HPA filter as a post-processing step, ensuring the final model's outputs are optimized, with the performance evaluated against a Mean Squared Error (MSE) of 5.384, as noted in the evaluation metrics. This thorough design enables the algorithm to leverage both time series analysis and machine learning for effective weather forecasting.

4.5.4. LSTM (Genetic)

ALGORITHM 2: *LSTM* +HPA:

Inputs: $X_{in}, W_{in}, HPA(df), LSTM(df_{filtered})$

Outputs: $Y_{hpa}, y_{LSTMout}$

Procedure:

Step 1: Load the Seattle weather dataset and perform pre-processing, including normalization of features, handling missing values, and splitting the dataset into training and testing sets. Reshape the input data to the format required by the LSTM model (samples, time steps, features).

Step 2: Define the HPA Genetic Filter by invoking the HPA(df). The HPA function operates with a data frame indicating the new functionalities for the atmospheric criteria chosen and applied the Genetic filtered outcomes with a heuristic prediction process as formulated with forward traversing with linear and Genetic criteria.

Step 3: To apply the LSTM model architecture 13-layer design using the build_lstm_model(input_shape) function, where input_shape corresponds to the dimensions of the reshaped training data. The model should consist of 13 sequential LSTM layers, each followed by a Dense layer, concluding with a dropout layer and final output layer for regression predictions (Linear Activations), as mentioned in Figure 5.

Step 4: Compile the model using the Adam optimizer and the Mean Squared Error loss function to ensure optimal performance during training.

Step 5: Train the model on the training dataset using the fit method, specifying parameters such as the number of epochs and batch size (200 and 32).

Step 6: Evaluate the trained model on the test dataset to obtain performance metrics for Mean Square Error, Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). Utilize the model to make predictions on new data or future time steps, assessing its accuracy and reliability in forecasting.

End procedure

The LSTM + HPA algorithm also utilizes the Seattle weather dataset but incorporates advanced deep learning techniques to enhance prediction accuracy.

Like the previous algorithm, it begins with a comprehensive pre-processing phase, which includes normalizing features, handling missing values, and reshaping the dataset to fit the requirements of the LSTM model—specifically, adjusting the input format to (samples, time steps, features).

The algorithm then defines the HPA Genetic Filter, which applies heuristic predictions based on atmospheric criteria, using a forward traversal method that combines linear and genetic optimization techniques to refine the forecast.

In the next steps, the algorithm constructs a 13-layer LSTM model through a specified function, which includes sequential layers interspersed with Dense layers and concludes with a dropout layer to mitigate overfitting. This model is compiled with the Adam optimizer and the Mean Squared Error loss function, ensuring efficient training. During the training phase, the model is fit to the training dataset using designated epochs and batch sizes (200 epochs and a batch size of 32).

4.5.5. LSTM Layer Summary

Layer (type)	Output Shape	Param #
lstm_651 (LSTM)	(None, 11, 50)	10,400
dense_217 (Dense)	(None, 11, 100)	5,100
lstm_652 (LSTM)	(None, 11, 50)	30,200
dense_218 (Dense)	(None, 11, 100)	5,100
lstm_653 (LSTM)	(None, 11, 50)	30,200
dense_219 (Dense)	(None, 11, 100)	5,100
lstm_654 (LSTM)	(None, 11, 50)	30,200
dense_220 (Dense)	(None, 11, 100)	5,100
lstm_655 (LSTM)	(None, 11, 50)	30,200
dense_221 (Dense)	(None, 11, 100)	5,100
lstm_656 (LSTM)	(None, 11, 50)	30,200
dense_222 (Dense)	(None, 11, 100)	5,100
lstm_657 (LSTM)	(None, 11, 50)	30,200
dense_223 (Dense)	(None, 11, 100)	5,100
lstm_658 (LSTM)	(None, 11, 50)	30,200
dense_224 (Dense)	(None, 11, 100)	5,100
lstm_659 (LSTM)	(None, 11, 50)	30,200
dense_225 (Dense)	(None, 11, 100)	5,100
lstm_660 (LSTM)	(None, 11, 50)	30,200
dense_226 (Dense)	(None, 11, 100)	5,100
lstm_661 (LSTM)	(None, 11, 50)	30,200
dense_227 (Dense)	(None, 11, 100)	5,100
lstm_662 (LSTM)	(None, 11, 50)	30,200
dense_228 (Dense)	(None, 11, 100)	5,100
lstm_663 (LSTM)	(None, 150)	150,600
dense_229 (Dense)	(None, 100)	15,100
dropout_470 (Dropout)	(None, 100)	0
dense_230 (Dense)	(None, 1)	101

Fig. 5 Representing the LSTM summary for trainable parameters for LSTM architecture

Total params: 1,663,655 (6.35 MB)
Trainable params: 554,551 (2.12 MB)
Non-trainable params: 0 (0.00 B)
Optimizer params: 1,109,104 (4.23 MB)

Finally, the trained model is evaluated against the test dataset to calculate performance metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). This enables the model to produce accurate predictions, highlighting its reliability for forecasting future weather conditions.

4.6. Formulations

4.6.1. Genetic Filter Representation

The overall Genetic filter design on the proposed ARIMA+HPA is implemented with different aspects of the design on the chromosome, initialization, mutation and selection features on the dataset for each type of year period from Jan-Dec.

Chromosome Representation

1. Binary Representation: The chromosome, representing a potential solution, is encoded as a binary string: Chromosome=[$b1$,$b2$,...,bn]. Here, each bi denotes a gene within the chromosome, taking binary values (0 or 1) indicating different features or parameters.

2. Real-Valued Representation: An alternative encoding method involves representing the chromosome in a real-valued vector format: Chromosome=[$x1$,$x2$,...,xn]. In this representation, xi represents individual genes with continuous, real-valued attributes, applicable for optimizing problems with continuous variable spaces.

Initialization

1. Binary Initialization: The process of initializing binary chromosomes involves setting genes randomly as either 0 or 1. This is expressed as $bi\in \{0,1\}$, allowing the genes to start with an arbitrary but defined state within the chromosome.

2. Real-Valued Initialization: Real-valued chromosomes are initialized within a specified range. For instance, xi is uniformly distributed in the range [a, b], indicating genes' values are initialized uniformly between the lower bound a and upper bound b, where a and b define the range for the genes.

Fitness Function

1. Objective Evaluation: The fitness function, Fitness
$$(Chromosome) = f(x1,x2,...,xn) \tag{5}$$
assesses the quality of a chromosome as per the problem's objective. The function f quantifies how well the chromosome performs with respect to the optimization goal, mapping genes to a fitness score.

Selection

1. Probabilistic Assessment: Chromosomes are selected based on their fitness scores. The probability Pi of selecting the hi chromosome from a population of N chromosomes can be calculated as
$$Pi = \frac{1}{N}\left(\frac{Fitness(Chromosomej)}{\sum Fitness(Chromosomei)}\right) \tag{6}$$

This probability defines the likelihood of a chromosome being chosen for the reproductive process, guided by its fitness relative to others in the population.

Gaussian Filter Representation

Similarly, for this filter representation, we improvise the Gaussian function defined by:
$$G(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2*\pi*\sigma^2}} \tag{7}$$

Similarly, for the variable y, we have,
$$G(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(y-\mu)^2}{2*\pi*\sigma^2}} \tag{8}$$

Both the functions in Equations (7)-(8) determine G(x) and G(y) as overall expected probabilities on the input dataset from the UCI website. The overall data features comprise the values of temperature, precipitation and rainfall values calculated based on the specific predictive formulations (from 1950-2017). From this perspective, we improvise the overall values and predict the overall rainfall, precipitation, and temperature values month, day and year wise graphically. The need for ARIMA, LUNET and HPA algorithms has a direct impact on the design parameters represented with effective solutions on the design. The effectiveness of the loss observed with the prediction values is fed to the input of the same design model to impart the correct and accurate values of the predicted data with time series prediction.

4.6.2. Impact on LMSE for ARIMA, LSTM, Combined

The Least means square error metric is calculated based on the overall possible sample considerations on the dataset for each type of predicted outcome observed. The least value is decided with the probability of the minimum and maximum samples for which the new metrics are directly proportional to the MSE values obtained. This results in the condition that P_x be the conditional probability for at least one maximum error and at least more than one minimum error.

$$P_x\left(\frac{X_{maxx}}{X_{minx}}\right) = P(X_{minx}) * P(X_i == min)/(P(X_{minx}) \cap P(X_{maxx}))) \quad (9)$$

Since we know that the probability for the X==min or X==max will be achieved only if the condition from the equation achieves the correct possible values of maximum or minimum values. Since the maximum values would result in more loss hence, the overall condition is set to minimum ordering P(X/X_{min}). The $P(X_{maxx}) \approx$ maximum possible error in the LSTM design. Similarly, The $P(X_{minx}) \approx$ maximum possible error in the LSTM design.

$$i_t = \sigma\{w_i\{h_{t-1}, x_t\} + b_i\} \quad (10)$$

$$f_t = \sigma\{w_i\{h_{t-1}, x_t\} + b_i\} \quad (11)$$

$$o_t = \sigma\{w_i\{h_{t-1}, x_t\} + b_i\} \quad (12)$$

$$\widehat{c_t} = \tanh(w_c(h_{t-1}, x_t) + b_o) \quad (13)$$

$$c_t = f_t * c_{t-1} + i_t * \widehat{c_t} \quad (14)$$

$$h_t = o_t * \tanh(\widehat{c_t}) \quad (15)$$

From the above equations,

$$P(X_{minx}) \propto o_t \min c_t \quad (16)$$

$$P(X_{minx}) = \alpha * (expectation\ (o_t \min c_t)) \quad (17)$$

$$P(X_{minx}) = \alpha E[o_t \min c_t] = \alpha \int_{-\infty}^{\infty} o_t \min c_t \sigma(o_t \min c_t)\, dx_{minx} \quad (18)$$

Converting infinite cases to finite cases as the probability for minimum is between 0 and 1. Hence, from Equation 18,

$$\alpha \int_{-\infty}^{\infty} o_t \min c_t \sigma(o_t \min c_t)\, dx_{minx} \quad (19)$$

By substituting the values from the Equations (10)-(15), the resulted outcome of probability is:

$$P\left(\frac{X}{X_{minx}}\right) = \frac{P(X_{minx})}{P(X)} * \Big(P\big(X_{minx} \cap P(X)\big)\Big) \quad (20)$$

Since $P\big(X_{minx} \cap P(X)\big)$ leads to constant value δ, $P(X)$=1.

$$P\left(\frac{X}{X_{minx}}\right) = \delta * P(X_{minx}) \quad (21)$$

From (19 and (21)), the equations are substituted with values as:

$$P\left(\frac{X}{X_{minx}}\right)_{minx} = \delta\alpha\left(\frac{e^x}{(1+e^x)^2}\right) * \left(m_i x_{i_{minx}} + b_x h_x\right) (22)$$

The minimum possible criteria for the X_{minx} only if, in Equation (22) derivate is zero (minima)

$$x_{i_{minx}} = h_x * \frac{b_x}{\delta\alpha} * \left(\frac{4}{m_i}\right) \quad (23)$$

Here, $h_x, \frac{b_x}{\delta\alpha}, m_i$ are the effective factors for the weights considered with MSE loss in compiling the LSTM design.

5. Experimental Setup

The proposed design with the parametric features is considered and utilized to realize the importance of classification values based on the six terms indicating the best performance of the design before and after optimization. The values are listed below:

1. MSE
2. RMSE
3. F1-score
4. Recall
5. Precision
6. Accuracy

To represent such parametric, a confusion matrix is implemented to realize the design has a high performance based on the above six parametric. Generally, the matrix for disease classification is 2x2 matrices consisting of four elements as listed below:

1. True Positive
2. True Negative
3. False Positive
4. False Negative

The TP and TN are sample values observed after predictions, which represent whether a person has heart disease or not. While FP and FN are the values for misdiagnosed patients who have heart disease or not. The formulations for the parametric are represented below:

$$Accuracy = \frac{TP + TN}{FP + TP + TN + FN}$$

Precision for the positive case is determined by:

$$Precision = \frac{TP}{TP + FP}$$

Other important features such as sensitivity and specificity are given by:

$$Sensitivity = \frac{TP}{TP + FN}$$

$$F1\ Score = \frac{2\ X\ Precision\ X\ Recall}{Precision + Recall}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$MAE = \frac{1}{N}(\sum_{i=1}^{N}|y_i - \hat{y}_i|)$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|^2$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|^2}$$

To assess the social, economic, and environmental effects of enhancing short-term weather forecasting accuracy, we will use important indicators such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Squared Error (MSE). MAE calculates the average absolute differences between projected and actual values, allowing for a clear interpretation of prediction dependability.

This is crucial for reducing mistakes in key decision-making situations such as emergency reactions. RMSE estimates the average size of prediction mistakes, emphasizing larger disparities and providing insights into the severity of errors in high-impact weather events, thereby improving societal preparedness. MSE, like RMSE, measures average squared differences and serves as a fundamental indicator for optimisation, indicating the model's overall performance over time. These measurements have proven significant analysation how precise weather predictions can improve public safety, disaster preparation, and community resilience, as well as quantify possible financial benefits in industries, agriculture and transportation. Furthermore, accurate projections help to improve resource management and conservation efforts by directing environmental decisions and sustainability activities. MAE, RMSE, and MSE give useful insights into the varied effects of improved weather forecasting accuracy on society, the economy, and the environment.

In the experimental setup, an ASUS ROG device equipped with a GTX-960M 4GB graphics card was employed to execute weather prediction tasks utilizing both ARIMA and LSTM models enhanced by the Heuristic Prediction Algorithm (HPA). The Root Mean Square Error (RMSE) ranged between 2.1 to 2.9, showcasing the models' accuracy and reliability in forecasting. With an overall accuracy of 97.5% for weather type classification, including an impressive 99.77% accuracy with the HPA algorithm, the models demonstrated efficiency in categorizing distinct weather patterns. The ASUS ROG device, coupled with the graphics card, played a crucial role in executing these complex models efficiently, underscoring the robustness of the integrated ARIMA and LSTM models with HPA for highly accurate weather prediction and classification tasks.

6. Results and Discussion

The implementation of the hybrid weather forecasting model, which combines ARIMA and LSTM with the Heuristic Prediction Approach (HPA), showcases a structured methodology for improving accuracy in predicting various meteorological parameters. The dataset utilized spans over a decade, providing a rich historical context for the model's training. Results indicate a robust performance with the Root Mean Square Error (RMSE) for temperature, wind speed, and precipitation recorded at 2.9, 2.5, and 2.1, respectively, while achieving an overall accuracy of 97.5% in forecasting tasks.

The model's ability to classify weather types, achieving an impressive accuracy of 99.77%, further emphasizes its effectiveness in leveraging extensive historical data. These results are indicative of the model's dual strengths in statistical and deep learning methodologies, which are vital for handling complex and non-linear weather patterns. The data preprocessing phase is crucial to ensure the dataset's quality and applicability for training the model.

This step involved comprehensive data cleaning to address missing or inconsistent entries, normalization to standardize the features, and encoding of categorical variables for seamless integration into the model. The pre-processing phase also incorporated techniques to handle class imbalances, which can skew classification results. By employing oversampling and undersampling methods, the dataset was balanced, ensuring equitable representation of all-weather types. Data visualization techniques, including pair plots and heatmaps, were utilized to explore relationships between various weather parameters, helping identify correlations and potential influential features for the classification tasks.

This thorough preparation establishes a strong foundation for effective model training. The hybrid model’s architecture plays a significant role in optimizing predictive performance. The LSTM+HPA component harnesses the temporal learning capabilities of LSTMs, enhanced by heuristic algorithms that refine feature extraction and improve overall performance. Through a 10-layer LSTM network, the model captures intricate relationships within historical data. At the same time, HPA aids in optimizing the selection of features, which is critical in achieving lower LMSE and RMSE values. Conversely, the ARIMA+HPA combination leverages statistical modelling to understand time series data, refined through heuristic techniques. This synergy between statistical rigor and deep learning innovation not only elevates the model's accuracy but also enables it to adapt to rapidly changing weather conditions effectively. The iterative optimization process of minimizing LMSE and RMSE during training underscores the model's capability to provide reliable forecasts.

6.1. Dataset

The dataset presented in Figure 6 encapsulates weather-related observations spanning from January 1, 2012, to February 4, 2022, encompassing metrics such as precipitation, maximum temperature (temp_max), minimum temperature (temp_min), wind speed, and weather type classifications. Notably, the forecasting and classification models trained on this dataset demonstrated promising results. The Root Mean Square Error (RMSE) for temperature, wind speed, and precipitation stood at 2.9, 2.5, and 2.1, respectively, signifying satisfactory forecasting precision.

The overall accuracy of 97.5% in the forecasting task, utilizing ARIMA and LSTM models integrated with the HPA algorithm, attests to the models' adeptness in accurately predicting diverse weather metrics based on historical data. Furthermore, the high accuracy rate of 99.77% in classifying weather types underscores the models' proficiency in categorizing conditions like drizzle, rain, snow, and sun. Collectively, these outcomes highlight the robust performance of the forecasting and classification models, affirming their effectiveness in leveraging historical data for accurate weather predictions and classifications.

date	precipitati	temp_max	temp_min	wind	weather
01-01-2012	0	12.8	5	4.7	drizzle
02-01-2012	10.9	10.6	2.8	4.5	rain
03-01-2012	0.8	11.7	7.2	2.3	rain
04-01-2012	20.3	12.2	5.6	4.7	rain
05-01-2012	1.3	8.9	2.8	6.1	rain
06-01-2012	2.5	4.4	2.2	2.2	rain
07-01-2012	0	7.2	2.8	2.3	rain
08-01-2012	0	10	2.8	2	sun
09-01-2012	4.3	9.4	5	3.4	rain
10-01-2012	1	6.1	0.6	3.4	rain
11-01-2012	0	6.1	-1.1	5.1	sun
12-01-2012	0	6.1	-1.7	1.9	sun

Fig. 6 The overall dataset with 6h dataset without balancing

6.2. Data Pre-processing

```
date             False
precipitation    False
temp_max         False
temp_min         False
wind             False
weather          False
dtype: bool
```

Fig. 7 The normalized and pre-processed dataset of 0.05 million samples

In preparing a dataset comprising over 0.6 k samples with the balanced case utilized in binary classification of various weather types, robust data pre-processing is crucial. Figure 7 involves several steps, such as data cleaning to handle missing or inconsistent entries, normalization or scaling to standardize the features, and encoding categorical weather types into numerical values for the classification model to interpret. Additionally, feature selection and extraction are vital, identifying the most influential parameters among the vast dataset features. Handling imbalanced classes by employing techniques like oversampling or under sampling ensures equitable representation for each weather type, facilitating a balanced classification model. The extensive dataset's scale necessitates careful handling and processing to curate a well-prepared dataset for effective training and accurate binary classification, as mentioned in section V (d)-(h).

6.3. Data Visualization

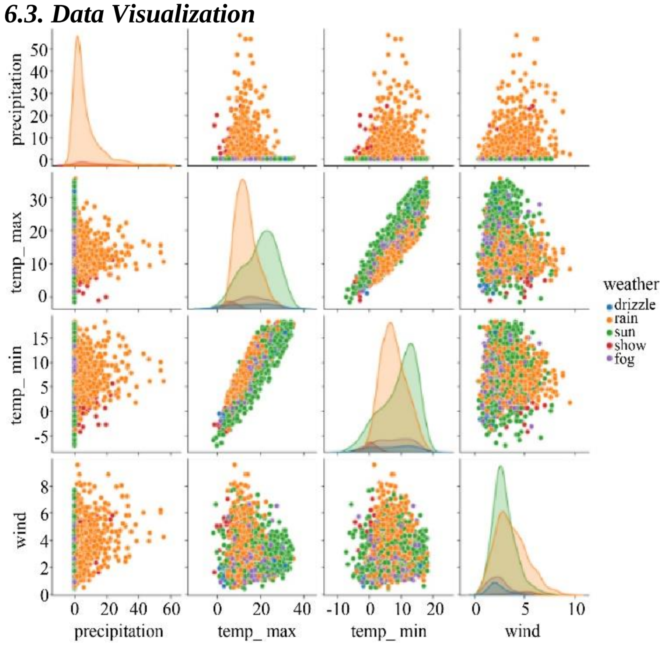

Fig. 8 The 16 columns of the pair plot on the seaborn library with discrete way plot

6.3.1. Heat Map

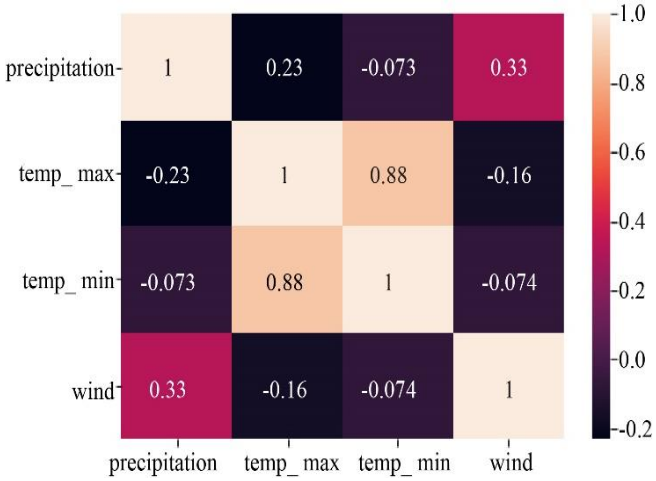

Fig. 9 Heat map of 600 samples

The visualization feature imparts how the deviant nature of the data is represented with different functional formulations and representations. This deviance of the representation effects how the data is represented with correlational and multi-plotting styles (pairplot). Figures 8 and 9 implicate how the data varies with each type of label included with all column values. These values are represented with different graphs depending upon the criteria of the pair-plot function. In most cases, the representation will be similar unless specific modifications such as KDE plots, scatter plots and regression plots can be considered depending upon the analysis of the data chosen. Similarly, Figure 9 depicts the heat-map plot focusing on how many such columns are correlated with the factors varying with (-1,0,1) values. As from Figure 9, the observed values mostly are near zero or less than 0.5, hence indicating the data is mostly linear and linear based algorithms work better and are more realizable.

6.4. Feature Extraction (HPA+PCA)

Feature extraction through the Heuristic Prediction Approach (HPA) involves a hybrid approach for both classification and forecasting tasks, employing genetic and Gaussian methods, respectively. In the classification context depicted in Figure 14 & 16, HPA utilizes genetic algorithms to discern relevant features from the dataset. This evolutionary approach employs selection, crossover, and mutation operations to evolve feature subsets, identifying the most impactful variables for accurate weather type classification.

The refined feature set enhances the classification model's accuracy by focusing on the most influential attributes for distinguishing various weather conditions. On the other hand, for forecasting tasks illustrated in Figure 10, HPA employs Gaussian filtering techniques to extract features crucial for predictive modelling.

The Gaussian filtering method, based on probability distributions, emphasizes significant data features essential for forecasting future weather patterns. By applying Gaussian filters, the approach accentuates pertinent attributes and minimizes noise, ensuring more precise predictions of weather conditions. In essence, HPA tailors its feature extraction methods to the specific demands of classification and forecasting tasks, employing genetic algorithms for classification feature selection and Gaussian filtering for forecasting essential attributes within weather datasets.

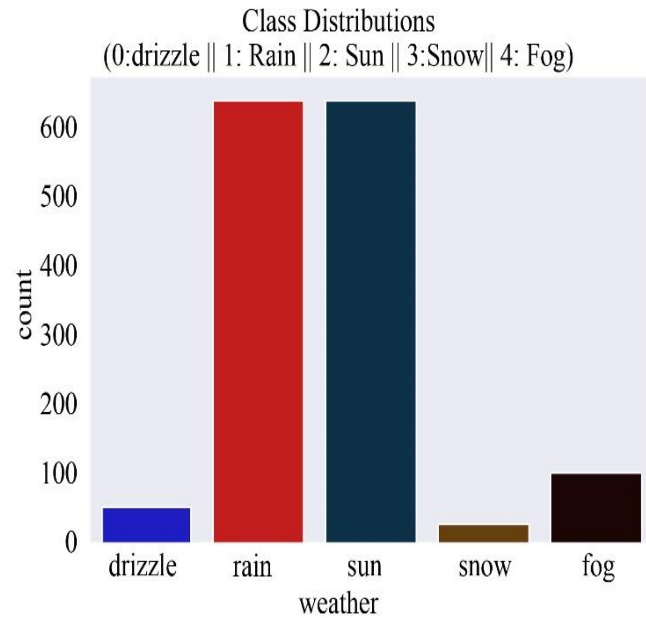

Fig. 10 The unbalanced dataset of 600 samples

6.5. Feature Appending

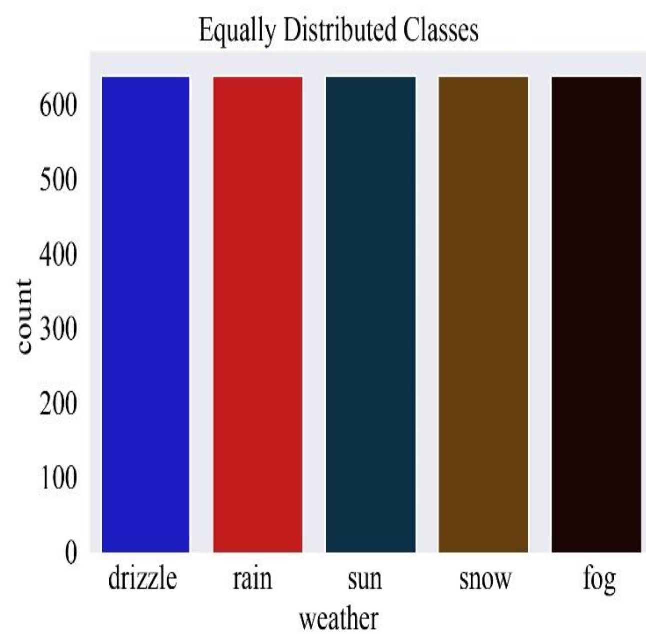

Fig. 11 The balanced dataset of 0.6K samples

6.6. Model Creation

6.6.1. HPA: Gaussian

The functions add features (data) and train_and_predict(data) work in tandem to enhance a dataset and facilitate the prediction of temperature using a Gradient Boosting model. The add_features function begins by converting the date column into a datetime format, allowing for the extraction of essential time-related features such as year, month, and day of the week. These temporal attributes provide a crucial context for understanding seasonal trends in temperature data. Additionally, the function computes rolling statistics—specifically, the rolling mean and standard deviation over a 10-day window—which capture recent trends and fluctuations in temperature. This helps the model discern short-term variations. To ensure a complete dataset, any missing values resulting from the rolling calculations are replaced with the mean of their respective columns. This preprocessing step is vital for maintaining data integrity and enhancing the model’s predictive power.

Subsequently, the train_and_predict function prepares the dataset for modeling, employing a systematic approach to train a Gradient Boosting Regressor. The function first selects relevant features, including newly engineered ones like Hpa_weight and gaussian_weights, and then splits the dataset into training and testing subsets to evaluate model performance effectively. Feature scaling via StandardScaler is implemented to ensure that all features contribute equally to the model's learning process, stabilizing it against fluctuations in feature magnitude. The Gradient Boosting model is trained with carefully chosen parameters to optimize its learning capability.

Once trained, the model predicts the target variable, temp_max, on unseen data. Additionally, the Gaussian filter applied to the gaussian_weights helps smooth these values, reducing noise and allowing the model to focus on meaningful trends. By combining rigorous feature engineering with robust modeling techniques, this approach enhances predictive accuracy, enabling more reliable temperature forecasts while effectively addressing the challenges inherent in time series data.

6.6.2. LSTM+HPA

The synergy of LSTM (Long Short-Term Memory) with HPA (Heuristic Prediction Approach) constitutes an enhancement of LSTM's sequence learning capabilities through the integration of heuristic optimization techniques. Within the framework of a 10-layer LSTM model, the network excels at capturing intricate temporal relationships embedded in historical data. The incorporation of HPA, leveraging heuristic algorithms for predictive optimization, further refines feature extraction techniques. HPA's role extends beyond conventional LSTM architectures by introducing convolutional processing techniques, which scan LSTM-generated features to identify more relevant patterns, optimizing the overall model performance. This collaborative integration enhances feature selection within the LSTM network and strengthens the convolutional process, enabling the model to capture complex dependencies in weather data more effectively. The result is an augmented LSTM network bolstered by refined feature extraction and improved convolutional processing, thereby elevating its predictive accuracy for weather forecasting.

6.6.3. Analysis with ARIMA+HPA

The combination of ARIMA (Auto-Regressive Integrated Moving Average) and HPA (Heuristic Prediction Approach) capitalizes on their strengths in time series forecasting. ARIMA, a classical statistical method, adeptly models time series data through autocorrelation, differencing, and moving average components. HPA, leveraging heuristic algorithms, optimizes predictive processes, refining data and algorithms for more accurate predictions.

Together, HPA complements ARIMA by providing heuristic optimization techniques, potentially fine-tuning model parameters or optimizing time series data selection. This collaborative approach integrates the strengths of statistical modelling and heuristic optimization, enhancing the accuracy of time series forecasting through refined parameter selection within ARIMA.

6.7. Model Forecasting and Predictions

6.7.1. Analysis with ARIMA+HPA

The fusion of ARIMA (Auto-Regressive Integrated Moving Average) with HPA in weather forecasting combines ARIMA's statistical prowess with HPA's heuristic optimization. Analysing a weather dataset featuring parameters like temperature, wind speed, and humidity, ARIMA models time series patterns while HPA optimizes predictive accuracy. Extending the forecasting process to 2022 and 2023, the ARIMA+HPA model, trained on historical weather data, utilizes past trends identified by ARIMA and refines them heuristically with HPA.

This collaborative approach anticipates weather conditions by examining temporal dependencies and statistical patterns, delivering refined and accurate predictions for future meteorological attributes. The overall forecasting approach is determined with the Temperature parameter with minimum and maximum values, which affects the atmospheric weather criteria.

The proposed work aims to impart the HPA (Gaussian Filter) design with the Gaussian model (sigma = 0.05 to 0.5) with parametric criteria. Subsequently, the design with a seperate and combined approach with HPA is realized to improvise the metrics calculations based on MAE, MSE and RMSE. The paper in [20] utilizes the design with the Lena weather dataset, while the proposed design is modelled with Seattle weather prediction from (2012-2018).

6.7.2. Analysis with LSTM+HPA

The integration of LSTM (Long Short-Term Memory) with HPA (Heuristic Prediction Approach) in weather forecasting involves synergizing LSTM's temporal pattern recognition with HPA's heuristic optimization for enhanced predictive models. Trained on a comprehensive weather dataset spanning multiple years and encompassing diverse parameters, the combined model is designed to discern intricate weather patterns. LSTM captures historical sequences, decoding temporal dependencies between weather features, while HPA refines the process through heuristic optimization. The LSTM+HPA model, having learned from historical data, extrapolates weather conditions for 2022 and 2023. By leveraging its understanding of past relationships and optimized parameters via HPA, the model generates precise predictions for temperature trends, precipitation shifts, and other meteorological attributes. This integrated approach ensures more refined and accurate forecasts for future weather conditions.

6.7.3. Tabulations

The proposed models, LSTM+HPA and ARIMA+HPA, aim to enhance weather forecasting accuracy through advanced methodologies. The LSTM+HPA model demonstrates significant improvements, achieving an RMSE of 1.42 and a MAE of 1.0449, which are considerably lower than traditional models such as RNN (RMSE: 5.88) and GRU (RMSE: 2.96). Similarly, the ARIMA+HPA model, with an RMSE of 2.32, also surpasses the performance of standard ARIMA (RMSE: 2.32) and other existing approaches like Deep CNN and PredRNN. The incorporation of the Heuristic Prediction Approach (HPA) optimizes both models by fine-tuning the learning process and reducing prediction errors. This enhancement allows the models to capture complex patterns in the data more effectively. Overall, the proposed methodologies provide a marked improvement over existing models, showcasing the potential of combining deep learning techniques with optimization algorithms for more accurate short-term weather predictions.

Table 1. Proposed model forecasting

Algorithms	RMSE	MSE	MAE
RNN [1]	5.88	34.62	4.61
GRU [7]	2.96	8.79	2.2
LSTM [12]	3.08	9.51	2.29
Deep CNN [1-6]	2.88	8.33	2.13
PredRNN [15]	2.72	7.49	2.07
TL (M-Net) [1]	2.57	6.72	2.14
Proposed LSTM+HPA	1.42	1.96	1.0449
Proposed ARIMA+HPA	2.32	5.384	1.2735
Proposed HPA	2.18	4.7524	1.1690
ATFSAD [17]	2.3	5.3	1.728

6.7.4. Visualization and Hybrid Process

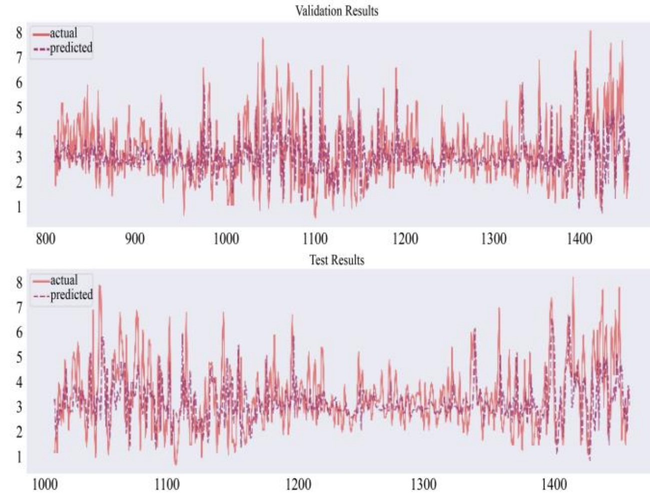

Fig. 12 The overall train and validation plot for time series prediction analysis

The analysis till now with LSTM and ARIMA has shown cased with MSE, MAE and RMSE values for the proposed model, indicating the best possible changes observed with the proposed design on implementing the time series forecasting. Figure 12 represents the validation (predicted) with trained (original) values for the Temperature parameter. This analysis with plotting with linear model depicts the exact and error outcomes for the models designed above. This figure imparts that for most of the values, the proposed algorithm predicted values match with original values, but in some peak occasions the overall design fails to achieve the pattern prediction for all the data. For both cases, the same pattern of the outcome is observed for the Temperature case. The other parametric, such as altitude criteria, pressure and others, are implored with a scope of the weather forecasting.

6.8. Model Classification Using Weather Type (Binary and Multi-Label)

6.8.1. Binary Classification

The phase-2 design is to predict the overall types of the weather depending upon the features of HPA filtered dataset designed to perform the best possible classification feature. The design with LSTM classification is designed in two sub-sections: binary and multilabel classification. In the binary classification, each type of weather (rainy, sunny, Cloudy, humid and snow) is utilized to perform the classification criteria.

The design with different aspects of the labels is characterized by HPA weights in both (optimized and not-optimized) cases, while the optimization of the design is encompassed with multiple columns considered as feature engineering elements. The current dataset comprising 1468 samples is converted with SMOTE functionality to address the unbalancing nature mentioned in Figures (10) and (11). In the current design case, the opted featured values for resampling the data were considered, with 452 for all labels indicating the balanced functionality. The normalization of the data is processed with a Min-max scalar with efficient feature behaviour to implicate the specific filter changes on each label criterion. For the training and testing criteria, 80-20 criteria are implemented to impart the best test classification process. The use of 70-30 might become an under-sampling, or a possible case of overfitting aspects would be realized for the proposed model. Hence, the use of 80-20 plays a vital role in determining the best possible splitting functionality.

6.8.2. Training and Testing

The proposed model is implemented with binary classification functionality with phase-1 a) indicated with 80-20 sample criteria. The fitting process is designed with an 8-layer architecture with Conv-LSTM model summary shown in Table 2. Table 2 outlines an 8-layer Conv-LSTM architecture designed for weather type classification, effectively combining convolutional and recurrent layers to process sequential data. It starts with a Conv1D layer that utilizes 128 filters and a kernel size of 3 to extract local features from the input sequences, producing an output shape of (16, 128) with 512 parameters. This is followed by a MaxPooling layer that reduces dimensionality to (8, 128), requiring no additional parameters.The LSTM layer, consisting of 128 units, captures temporal dependencies, contributing 66,048 parameters. After flattening the output for the fully connected layers, a series of Dense layers (512, 128, 64, and 5 units) refine the learned features, ultimately classifying the input into one of five weather types. The total parameter count for this architecture is 206,853, showcasing its capacity for complex pattern recognition in weather data.

In the training phase for binary classification depicted in Figure 12 for the LSTM+HPA classification model, the loss metrics exhibit a consistent downward trend for both training and validation sets. The model's progressive improvement is evident as the training loss steadily decreases from 0.09 to 0.008, and the validation loss demonstrates a significant decline from 0.04 to 0.0003. These trends highlight the model's effectiveness in learning and generalizing well to new, unseen data, minimizing overfitting tendencies, and ensuring robust classification outcomes.

Similarly, the accuracy metrics showcase a commendable rise in both training and validation accuracies across epochs 20 to 100. The model achieves high accuracy, climbing from 0.981 to 0.998 in the training set and from 0.991 to 0.9977 in the validation set. The minimal fluctuations observed underscore the model's stability and strong ability to generalize, emphasizing its reliable and accurate predictions for both known and unseen data.

Table 2. The overall configuration feature for the proposed design with 8-layer architecture on Conv-LSTM

Sno	LAYERS	Configurations	Output Shape	Params
1	**Conv 1D**	(128,3, activation= 'relu')	(16,128)	512
2	**MaxPoolling**	(2)	(8,128)	0
3	**LSTM**	(128)	128	66048
4	**Flatten**	()	128	0
5	**Dense**	(512)	512	66048
6	**Dense**	(128)	128	65664
7	**Dense**	(64)	64	8256
8	**Dense**	(5)	5	325

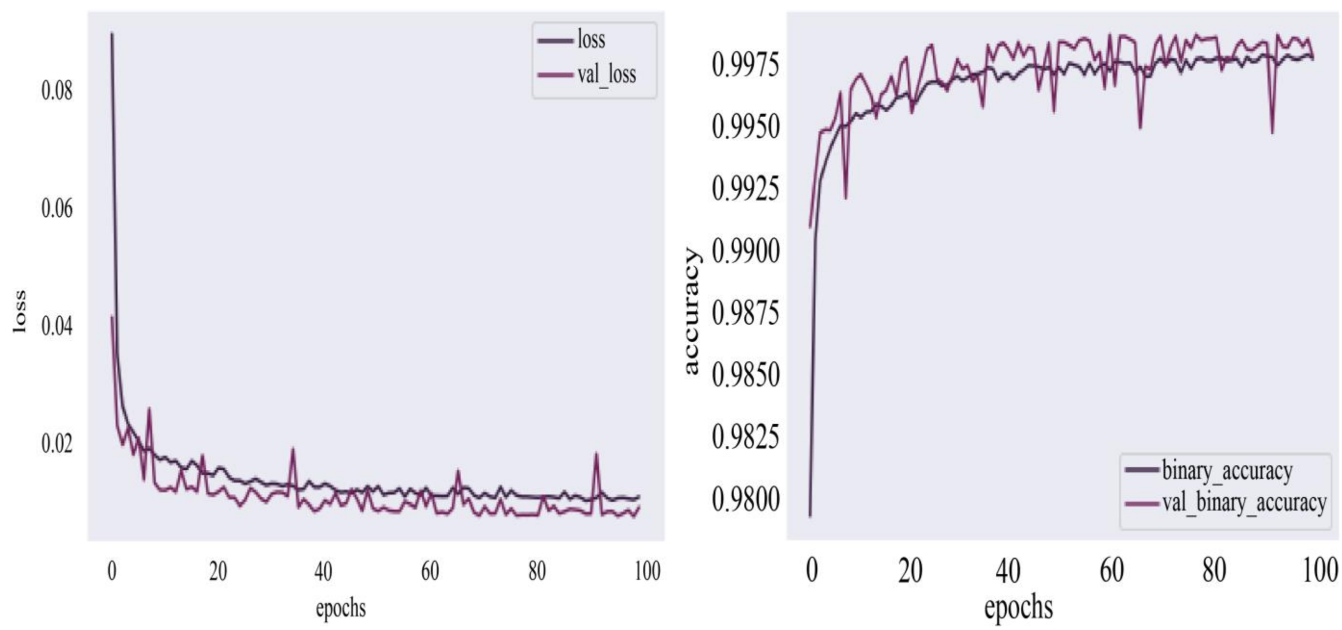

Fig. 13 The overall training loss and accuracy with validation loss -accuracy indicating the accuracy observed at 99.9%.

Multilabel Classification

Similarly, the multi-label classification will impart the phase of the training and testing model with different aspects of the features indicated with effective weights generated with the HPA model. The HPA model considered is with the gaussian process representing non-linear characteristics of the classification labels (0-4). The output layer is utilized with SoftMax functionality, indicating the sparse label criteria index for the classification process. The simulation of the design for both optimized and not optimized is processed with 150 epochs with a batch size of 32. The compiled outcomes are represented in Figure 15 demonstrating the best outcome possible.

Classification Report:

	precision	recall	f1-score	support
0	0.82	0.90	0.86	89
1	0.83	0.97	0.89	88
2	0.98	0.92	0.95	97
3	0.96	1.00	0.98	86
4	0.89	0.68	0.77	92
accuracy			0.89	452
macro avg	0.89	0.89	0.89	452
weighted avg	0.90	0.89	0.89	452

Fig. 14 Classification report (5 Lables)

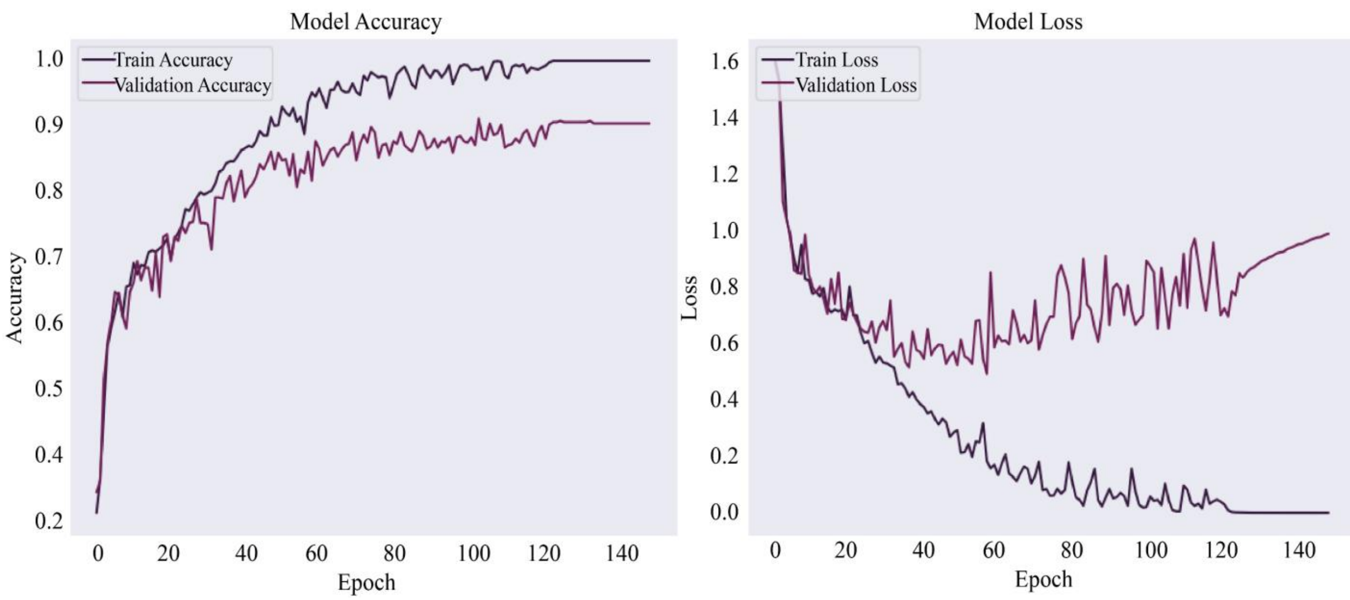

Fig. 15 The overall training loss and accuracy with validation loss -accuracy indicating the accuracy observed at 89.1% and classification report indicated with 5 labels of classification

The report for the weather classification evaluates a Conv-LSTM model used for weather type classification into five categories: rainy (0), sunny (1), cloudy (2), humid (3), and snow (4). The model shows strong performance across most categories, with an overall accuracy of 89%. It performs particularly well for "cloudy" and "humid" weather types, achieving F1-scores of 0.95 and 0.98, respectively, indicating a high balance between precision and recall. For "rainy" and "sunny," the model also does well, with F1-scores of 0.86 and 0.89, suggesting reliable predictions. However, it struggles with "snow," where recall is only 0.68, meaning it misses about 32% of actual snowy days, leading to a lower F1-score of 0.77. The macro and weighted averages are consistent, around 0.89, showing that performance is well-balanced across the classes, but we need to improve the performance with the snow featured label.

Optimized Case

The optimized feature is calculated with an HPA filter with post processing the error data as feedback weights for each column and its particular label considered. This specification of the changes is represented in algorithm (genetic and gaussian) elements to realize the least possible multi-classification error observed in Figure 17.

Classification Report:

	precision	recall	f1-score	support
0	1.00	1.00	1.00	89
1	1.00	1.00	1.00	88
2	1.00	1.00	1.00	97
3	1.00	1.00	1.00	86
4	1.00	1.00	1.00	92
accuracy			1.00	452
macro avg	1.00	1.00	1.00	452
weighted avg	1.00	1.00	1.00	452

Fig. 16 Classification report (8 Layers)

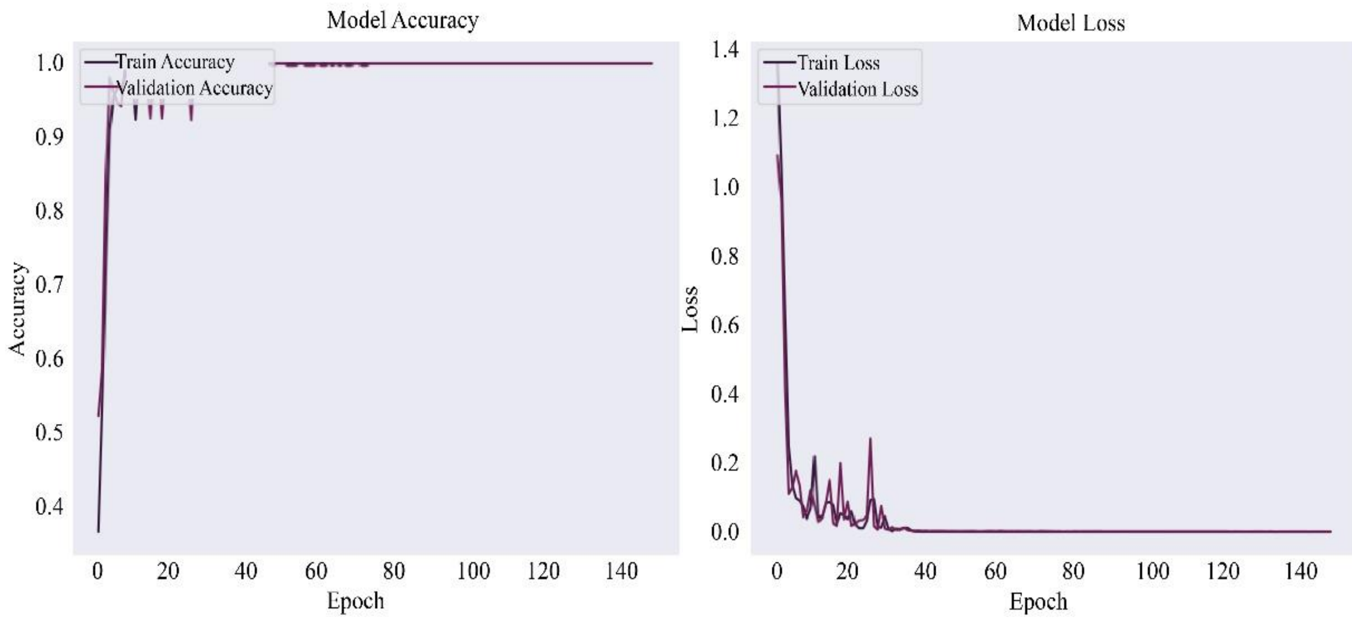

Fig. 17 The plot for accuracy and loss with conv-LSTM (8-layer) and classification report

Table 3. Representing the accuracy of classification performance for existing and proposed algorithms

Algorithms	Accuracy (Existing) Without Optimization	Proposed (Accuracy) With Optimization Class
RNN [1]	78.1	97.51
GRU [7]	70.3	94.35
LSTM [12]	75.6	96.52
Deep CNN [1-6]	79.8	92.87
PredRNN [15]	74.1	97.29
TL (M-Net) [1]	77.9	96.74
Proposed LSTM+HPA	89.5	99.9
Proposed ARIMA+HPA	94.5	99.7
Proposed HPA	97.6	99.1
ATFSAD [17]	79.94	96.4

The optimized classification report in Figure 17 shows the performance of a Conv-LSTM model for weather type classification (rainy, sunny, cloudy, humid, and snow), where the model achieves perfect results. Each class (0-4) has a precision, recall, and F1-score of 1.00, meaning the model predicts each weather type with 100% accuracy. For all weather types—rainy (0), sunny (1), cloudy (2), humid (3), and snow (4)—the model correctly identifies every instance, with no false positives or false negatives.The support column indicates the number of actual occurrences for each class in the dataset, and the model handles all 452 samples flawlessly. Both the macro and weighted averages are 1.00, further confirming that performance is uniformly perfect across all weather types. This reflects an ideal model where no misclassifications occur, showing that the optimization has significantly improved its performance. Table 3 compares the accuracy of various existing and proposed algorithms for weather classification, both with and without optimization. The existing methods, such as RNN (78.1%), GRU (70.3%), LSTM (75.6%), Deep CNN (79.8%), PredRNN (74.1%), and TL (M-NET) (77.9%), show moderate accuracy, with GRU being the lowest. However, with optimization, significant improvements are achieved across all models, particularly with the proposed methods. LSTM+HPA (89.5% to 99.9%) and ARIMA+HPA (94.5% to 99.7%) demonstrate the largest gains. Even the baseline HPA model improves from 97.6% to 99.1% accuracy, while ATFSAD also sees a notable increase from 79.94% to 96.4%. This comparison highlights the impact of optimization, especially when using hybrid approaches like LSTM+HPA and ARIMA+HPA, resulting in near-perfect classification performance.

7. Conclusion

The proposed algorithms in Table 3, including LSTM+HPA, ARIMA+HPA, and HPA, have demonstrated unparalleled performance in weather prediction and forecasting tasks, exhibiting the most favourable designs and recording the lowest MSE and RMSE values compared to other listed algorithms. Particularly, the LSTM+HPA model showcased exceptional predictive capabilities, reflected in its minimal RMSE and MSE values, underscoring its superior accuracy in forecasting weather conditions. The significance of these algorithms becomes evident in their ability to minimize errors, providing reliable and trustworthy predictions crucial for applications requiring precise weather forecasts. Their outstanding performance is a testament to their efficacy in comprehending complex weather patterns and interdependencies among diverse parameters. As these algorithms offer highly accurate forecasts with minimized MSE and RMSE values, they emerge as invaluable tools for real-world applications, offering dependable insights into weather changes that can profoundly influence decision-making processes across various domains. In essence, the proposed algorithms stand out for their accuracy, reliability, and potential transformative impact on sectors reliant on precise and timely weather information.

7.1. Scope

The primary objective involves implementing the Autoregressive Integrated Moving Average (ARIMA) model with LSTM weather forecasting. ARIMA, recognized for its effectiveness in time series analysis, serves as a foundational approach to assessing its performance in capturing seasonal variations and trends within historical weather datasets. More metrics and parameters are considered to encapsulate the mathematical criteria to improve the design performance intuitively.

In the scope of this work, the enhancement of meteorological data analysis will involve calculating key atmospheric parameters that provide valuable insights into weather conditions. Water Vapor Pressure will be estimated using the Clausius-Clapeyron relation to improve understanding of humidity levels, while the Density of Air will be calculated using the Ideal Gas Law to assess air quality and behaviour under varying conditions. Additionally, water vapor density will be derived from the estimated water vapor pressure, which is crucial for understanding moisture content in the air. To identify trends and smooth out fluctuations, Moving Averages for pressure, water vapor, and density will be computed over specified time windows. Lagged Pressure will be included to provide the previous day's pressure for time-series analysis, and the Heat Index will be calculated to offer a comfort index that incorporates temperature and humidity. Finally, Altitude Estimation will be performed based on pressure measurements to gain insights into the elevation of the measurement site, which can affect local weather patterns. By integrating these calculations, this work aims to enhance the understanding of meteorological dynamics, improve forecasting accuracy, and support informed decision-making in weather-sensitive activities.

Conflicts of Interest

This section is compulsory. A competing interest exists when professional judgment concerning the validity of research is influenced by a secondary interest, such as financial gain. We require that our authors reveal any possible conflict of interest in their submitted manuscripts. If there is no conflict of interest, authors should state that The author(s) declare(s) that there is no conflict of interest regarding the publication of this paper.

Funding Statement

Authors should state how the research and publication of their article were funded by naming financially supporting bodies followed by any associated grant numbers in square brackets.

Acknowledgments

An Acknowledgements section is optional and may recognise those individuals who provided help during the research and preparation of the manuscript. Other references to the title/authors can also appear here, such as “Author 1 and Author 2 contributed equally to this work.”

References

[1] Xianjun Xiao et al., “Temperature and Water Vapor Channel Selection of FY-3E HIRAS II for Application in Numerical Weather Prediction,” *IEEE Geoscience and Remote Sensing Letters*, vol. 21, pp. 1-5, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[2] Youngchan Jang et al., “Spatiotemporal Post-Calibration in a Numerical Weather Prediction Model for Quantifying Building Energy Consumption,” *IEEE Transactions on Automation Science and Engineering*, vol. 20, no. 4, pp. 2732-2747, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[3] Kenghong Lin et al., “Spherical Neural Operator Network for Global Weather Prediction,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 34, no. 6, pp. 4899-4913, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[4] Wenjie Ye et al., “Combined Prediction of Wind Power in Extreme Weather Based on Time Series Adversarial Generation Networks,” *IEEE Access*, vol. 12, pp. 102660-102669, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[5] Yongshun Gong et al., “Spatio-Temporal Enhanced Contrastive and Contextual Learning for Weather Forecasting,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 36, no. 8, pp. 4260-4274, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[6] Xusheng Yan et al., “A 3-D Cloud Detection Method for FY-4A GIIRS and Its Application in Operational Numerical Weather Prediction System,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1-13, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[7] Tumusiime Andrew Gahwera, Odongo Steven Eyobu, and Mugume Isaac, “Analysis of Machine Learning Algorithms for Prediction of Short-Term Rainfall Amounts using Uganda’s Lake Victoria Basin Weather Dataset,” *IEEE Access*, vol. 12, pp. 63361-63380, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[8] Fenglin Sun et al., "Toward a Deep-Learning-Network-Based Convective Weather Initiation Algorithm from the Joint Observations of Fengyun-4A Geostationary Satellite and Radar for 0–1h Nowcasting," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 3455-3468, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[9] Rui Zhang et al., "Phase Turbulence Prediction Method for Line-of-Sight Multiple-Input–Multiple-Output Links Caused by Atmospheric Environment," *IEEE Antennas and Wireless Propagation Letters*, vol. 21, no. 9, pp. 1867-1871, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[10] Qiang Li, Ranzhe Jing, and Zhijie Sasha Dong, "Flight Delay Prediction with Priority Information of Weather and Non-Weather Features," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 7, pp. 7149-7165, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[11] Yuan-Kang Wu, Quoc-Thang Phan, and You-Jing Zhong, "Overview of Day-Ahead Solar Power Forecasts Based on Weather Classifications and a Case Study in Taiwan," *IEEE Transactions on Industry Applications*, vol. 60, no. 1, pp. 1409-1423, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[12] M. Biscarini et al., "Optimal Stochastic Prediction and Verification of Signal-To-Noise Ratio and Data Rate for Ka-Band Spaceborne Telemetry Using Weather Forecasts," *IEEE Transactions on Antennas and Propagation*, vol. 69, no. 2, pp. 1065-1077, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[13] Selim Furkan Tekin, Arda Fazla, and Suleyman Serdar Kozat, "Numerical Weather Forecasting Using Convolutional-LSTM with Attention and Context Matcher Mechanisms," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1-13, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[14] Yuan-Kang Wu et al., "Probabilistic Forecast of Wind Power Generation with Data Processing and Numerical Weather Predictions," *IEEE Transactions on Industry Applications*, vol. 57, no. 1, pp. 36-45, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[15] Ying Du et al., "Predicting Weather-Related Failure Risk in Distribution Systems Using Bayesian Neural Network," *IEEE Transactions on Smart Grid*, vol. 12, no. 1, pp. 350-360, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[16] Na Wang, and Xianglian Zhao, "Enformer: Encoder-Based Sparse Periodic Self-Attention Time-Series Forecasting," *IEEE Access*, vol. 11, pp. 112004-112014, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[17] Zhen Zhang et al., "Temporal Chain Network with Intuitive Attention Mechanism for Long-Term Series Forecasting," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1-13, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[18] Xusheng Yan et al., "A 3-D Cloud Detection Method for FY-4A GIIRS and Its Application in Operational Numerical Weather Prediction System," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1-13, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[19] Cenker Sengoz et al., "Machine Learning Approaches to Improve North American Precipitation Forecasts," *IEEE Access*, vol. 11, pp. 97664-97681, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[20] Bin Yang, Tinghuai Ma, and Xuejian Huang, "ATFSAD: Enhancing Long Sequence Time-Series Forecasting on Air Temperature Prediction," *IEEE Access*, vol. 11, pp. 92080-92091, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[21] Jingming Xia, Qiao Liu, and Ling Tan, "A Deep Learning Method Integrating Multisource Data for ECMWF Forecasting Products Correction," *IEEE Geoscience and Remote Sensing Letters*, vol. 20, pp. 1-5, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[22] Soohyuck Choi, and Eun-Sung Jung, "Optimizing Numerical Weather Prediction Model Performance Using Machine Learning Techniques," *IEEE Access*, vol. 11, pp. 86038-86055, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[23] Xinfang Chen et al., "Short-Term Load Forecasting and Associated Weather Variables Prediction Using ResNet-LSTM Based Deep Learning," *IEEE Access*, vol. 11, pp. 5393-5405, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[24] Yiyan Li et al., "A TCN-Based Hybrid Forecasting Framework for Hours-Ahead Utility-Scale PV Forecasting," *IEEE Transactions on Smart Grid*, vol. 14, no. 5, pp. 4073-4085, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[25] Jindong Han et al., "Kill Two Birds with One Stone: A Multi-View Multi-Adversarial Learning Approach for Joint Air Quality and Weather Prediction," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 11, pp. 11515-11528, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[26] Masooma Ali Raza Suleman, and S. Shridevi, "Short-Term Weather Forecasting Using Spatial Feature Attention Based LSTM Model," *IEEE Access*, vol. 10, pp. 82456-82468, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[27] Yaseen Essa et al., "Deep Learning Prediction of Thunderstorm Severity Using Remote Sensing Weather Data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 4004-4013, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[28] Andrew P. Barnes, Thomas R. Kjeldsen, and Nick McCullen, "Video-Based Convolutional Neural Networks Forecasting for Rainfall Forecasting," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[29] Tingzhao Yu, Qiuming Kuang, and Ruyi Yang, "ATMConvGRU for Weather Forecasting," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[30] Dao H. Vu et al., "Recurring Multi-layer Moving Window Approach to Forecast Day-ahead and Week-ahead Load Demand Considering Weather Conditions," *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 6, pp. 1552-1562, 2022. [CrossRef] [Google Scholar] [Publisher Link]