Title page:

Efficient Prediction of Water Inundation Frequency using Long Short-Term Memory in comparison with Radial Basis Function Neural Network

Kovi Sai Ganesh[1], S. Kalaiarasi [2]

Kovi Sai Ganesh [1]
Research Scholar,
Department of Computer Science and Engineering,
Saveetha School of Engineering,
Saveetha Institute of Medical and Technical Sciences,
Saveetha University, Chennai, Tamil Nadu, India, Pincode: 602105.
192124028.sse@saveetha.com

S. Kalaiarasi [2]
Project Guide, Corresponding Author,
Department of Computer Science and Engineering,
Saveetha School of Engineering,
Saveetha Institute of Medical and Technical Sciences,
Saveetha University, Chennai, Tamil Nadu, India, Pincode: 602105.
kalaiarasis.sse@saveetha.com

Keywords : Deep Learning, Flood forecasting, Long Short-Term Memory,Machine Learning, Natural Catastrophes, Neural Networks, Radial Basis Function Neural Network,Water Inundation Frequency.

ABSTRACT

Aim:The main objective of this research is to determine the extent to which Long Short-Term Memory (LSTM) and Radial Basis Function Neural Networks (RBFNN) estimate the frequency of flooding, which is one of the most dangerous natural catastrophes that can cause significant damage to property and infrastructure. **Materials and Methods:** There are two sets of proposed algorithms. Radial Basis Function Neural Networks (RBFNN) are compared with Long Short-Term Memory (LSTM). Both algorithms are capable of accurately forecasting the frequency of a water inundation. The frequency of water inundations is examined for accuracy. Regarding the purpose of accurately simulating the intricate dynamics of water inundation, the long-term dependency handling capabilities of LSTM and the chaotic relationship-capturing capabilities of RBFNN are essential. The approach builds a model based on the characteristics of the training set of data, which it then applies to determine the value for the newly collected data.The algorithms utilized to identify the frequency of water inundation are LSTM and RBFNN with a sample size of N=20. Using SPSS, the significance value of the data set was estimated with a G-power value more than 80%. **Results:**The accuracy of 93.0095% that LSTM has achieved is relatively greater than that of 88.4810% that RBFNN has achieved. A significance value of 0.324 indicates that there is a substantial difference between the two groups.**Conclusion:**In conclusion, these results show that the LSTM has a higher predictive capacity for estimating the frequency of flooding.For applications requiring the capture of long-term dependencies in the data, LSTM performs better than RBFNN.

Keywords : Deep Learning, Floods, Long Short-Term Memory,Machine Learning, Natural Catastrophes, Neural Networks, Radial Basis Function Neural Network, Water Inundation Frequency.

INTRODUCTION

Flooding, also known as water inundation, is a serious environmental hazard that jeopardizes human life, property, and infrastructure. The top 10 countries, which include Bangladesh, India, and China, are currently at risk of flooding and may get worse if nothing is done, according to(Gude, Corns, and Long 2020) . It is true that recovering from flood damage can be extremely expensive. Thus, floods can happen in any nation.(Sankaranarayanan et al. 2020) states that the heterogeneity of the basins included in the regional analysis, the fit of the probability distribution, and the quality of the data are the main causes of uncertainty in the analysis of flood frequency. Lately, deep learning has emerged as one of the most trustworthy methods for time series forecasting (Kumar et al. 2023). Time series have shown to be an effective tool for flood forecasting, with many applications using(Sit and Demir 2019) .The 20 years from 2000 to 2019 saw a rise in the quantity and frequency of natural disaster-related damage as compared to the 20 years from 1980 to 1999. More damage was specifically caused by storms and floods than by other natural catastrophes.

The amount of damage produced by natural catastrophes over the years 1980 to 1999 and 2000 to 2019 is contrasted in Figure 1(Cho et al. 2022).Flooding, another name for water inundation, is a natural occurrence that happens when water transcends its intended boundaries and submerges normally dry land(Vineeth et al. n.d.2021).Flood risk assessment and mitigation, disaster preparedness and management, and water resource management are all dependent on the accuracy of water inundation predictions. Numerous industries, including urban planning, agriculture, and insurance, use water inundation prediction(Nayak, Das, and Senapati 2022).A thorough scan of the Scopus database showed that 2,801 publications in Science Direct and over 17,700 articles in Google Scholar have been published in the previous five years on the subject of water inundation prediction. Numerous methods have been used in these investigations, such as hydrological modeling, machine learning and deep learning techniques, and statistical methods. The current research focuses on creating novel machine learning-based techniques for water inundation analysis and natural catastrophes in the area of modeling and predicting water resources.Even with the progress made in the field of water inundation prediction, there are still a number of gaps in the literature, the most significant of which are related to the incorporation of uncertainty and real-time prediction. Because of the nonlinearity of the flood process and the unpredictable nature of the projected outcomes, the decision-making process for flood management is vulnerable to risks(Hayder et al. 2023). The Deep Learning model is trained only on publicly available datasets; it is not fair to utilize scientific concepts or guidelines during this process as this could lead to forecasts that are not based on logic(Bentivoglio et al. 2022). The prediction of stream flow is a crucial area of research to minimize harm to the environment and urban areas by natural catastrophes.The inundation of floodplains and wetlands is also necessary for the Earth system to function.

The ability of floodplain water storage to significantly lessen flood waves controls the global dynamics of the movement of energy, nutrients, and water to the coastal ocean(Nguyen and Chen 2020)By contrasting the performance of Long Short-Term Memory with Radial Basis Function Neural Network, this research work seeks to create a reliable and effective prediction model that will ultimately improve our capacity to predict water inundation caused by natural catastrophes events with increased accuracy and dependability.

MATERIALS AND METHODS

The data analytics lab of the Saveetha Institute of Medical and Technical Sciences was used to conduct this research. In the lab, a highly adjustable system allows for precise results and in-depth research. The study comprised ten respondents in all, who were divided into two groups: Group 2 employed RBFNN (Radial Basis Function Neural Networks) and Group 1 used the LSTM (Long Short-Term Memory) approach. The study's statistical power was sustained at a significance level of 0.05 (alpha) and a power level of 0.8 (beta) with an 80 percent G-Power value. Furthermore, a 95% confidence interval was upheld to methodically calculate and assess the variations between the two groups.

The "Kerala Floods" dataset(Devakumar 2019), which is arranged as comma-separated values (CSV), is used in this water inundation frequency prediction.The dataset incorporates variables like year, months, and annual rainfall to improve the accuracy of frequency determination. To facilitate training and testing, the dataset has been sanitized. The data has undergone normalization, missing values have been removed, and averages or medians have been used in place of null values. The LSTM algorithm receives the preprocessed dataset that includes features.

The technique proposed in this research was applied to a specific dataset, and the outcomes were contrasted with the well-known RBFNN algorithm. The dataset is set up in this case as a comma-separated value (CSV) file, which holds a range of data values. To perform an exhaustive examination and comparison, the dataset was subjected to a comprehensive analysis of the records using the algorithms presented in this research. The gear setup for this work was an Intel dual-core processor with 8 GB of RAM. Utilizing Jupyter Notebook, Python, and a MySQL database, the software configuration offered a stable platform for executing the algorithm and doing the comparative analysis.

Long Short-Term Memory

An algorithm type called Long Short-Term Memory is ideal for modeling sequential data, such time series data.A common issue with Recurrent Neural Network’s that can hinder them from learning long-term dependencies is the vanishing gradient problem, which LSTM networks are specifically made to address long-term dependencies(Li, Kiaghadi, and Dawson 2020). Because long-term factors like climate change and human activities are expected to have an impact on inundation frequency, LSTM deep learning networks are a promising option for simulating water inundation frequency.

Radial Basis Function Neural Networks

A nonlinear function type that can depict complex relationships between data points is the radial basis function. RBFNNs are highly known for their simplicity of training, quick learning and generalization rates, and ease of use(Panigrahi, Nath, and Senapati 2019). RBFNNs are a good choice for simulating the frequency of water inundations since it is likely that there is a complex and nonlinear relationship between the elements that generate floods and the frequency of inundation.

Statistical Analysis

The analysis was performed using IBM SPSS version 26. The collected accuracy values were examined in SPSS with a sample size of 20 for the algorithms Long Short-Term Memory(LSTM) and Radial Basis Function Neural Network (RBFNN). The LSTM and RBFNN denote Groups 1 and 2. The dataset's features are used to train the model. The model accuracy values for every epoch were used to compute the T-test. The performance comparison between the two algorithms is made easier by the t-test.

RESULTS

The accuracy raw data table for both Radial Basis Function Neural Networks (RBFN) and Long Short-Term Memory (LSTM) is displayed in Table 3. The accuracy values are computed using two different neural networks with sample sizes of 20, namely the Radial Basis Function Neural Network (RBFNN) and the Long Short-Term Memory (LSTM). The Radial Basis Function Neural Network (RBFNN) algorithm has an accuracy value of 88.4810%, while the Long Short-Term Memory (LSTM) approach has a mean accuracy value of 93.0095%. Table 4, which displays these accuracy numbers. This suggests that the recommended algorithm is more accurate than the others. This study used the Long Short-Term Memory (LSTM) approach as opposed to the Radial Basis Function Neural Network (RBFNN). Table 5 displays the "F" value as 0.998 and the "Sig" value as 0.324 with equal variances assumed. The 95% confidence intervals for the difference for equal variances assumed and not assumed are 6.22125 and 6.22580 for the upper case and 2.83575 and 2.83120 for the lower case (Table 5). The bar graph displayed in Fig. 2 is the result of plotting groups on the X-axis and mean accuracy on the Y-axis. The graph clearly shows that LSTM (Long Short-Term Memory) is more accurate than RBFNN (Radial Basis Function Neural Network).

DISCUSSION:

The study in question yielded a significant value of 0.324 (two-tailed, $p>0.05$), suggesting that Long Short-Term Memory (LSTM) is potentially superior to Regions using Radial Basis Function Neural Networks (RBFNs). While the mean accuracy of Regions using an RBFNN (Radial Basis Function Neural Network) classifier is 88.4810%, the mean accuracy analysis of the LSTM (Long Short-Term Memory) technique is 93.0095%. In comparison to RBFNN, this suggests that LSTM performs better.

Long Short-Term Memory (LSTM) and Radial Basis Function Neural Networks (RBFNNs) are powerful machine learning techniques, each having specific benefits and drawbacks. Applications such as time series forecasting and natural language processing profit immensely from LSTM's ability to accurately predict long-term dependencies for natural catastrophes in sequential data. Its gating system allows selective memory, which enhances its ability to recall complex temporal connections. Nevertheless, LSTM can be computationally expensive and challenging to train. However, because of its computational efficiency and ability to handle nonlinear interactions, RBFNN is more suitable for function approximation and pattern recognition(Kim, Han, and Lee 2020). Its relatively simple training method makes it interesting for smaller or simpler datasets(Zhang and Hao 2021).RBFNN's propensity for overfitting and its limited ability to identify long-term dependencies must be considered, though.

The size, diversity, and feature selection of the dataset, generalizability, and unproven algorithmic choices are among the potential limits of the study. To get over these limitations and progress the area, additional research should focus on expanding datasets that forecast increased frequency of flooding from various causes. Possible avenues to improve forecast precision and the understanding of Water Inundation Frequency are to highlight ensemble models, interpretability, and practical applications in environmental research and emergency response. Further investigation into alternative machine learning techniques(Ha, Liu, and Mu 2021), DL approaches, parametric optimization, and real-world validation may yield more promising predictive models.

CONCLUSION:

Water Inundation frequency prediction using the LSTM (Long Short-Term Memory) algorithm got the mean accuracy of 93.0095 % whereas the mean accuracy of Regions with RBFNN (Radial Basis Function Neural Network) is 88.4810% in this investigation. The study investigated the efficiency of LSTM and RBFNN models in predicting water inundation frequency. The results indicated that both models performed well, with LSTM demonstrating superior accuracy compared to RBFNN. LSTM's ability to capture long-term dependencies in time series data proved advantageous in predicting water inundation frequency, which is influenced by historical hydrological patterns and natural catastrophes. While RBFNN demonstrated satisfactory performance, it was less effective in capturing complex temporal relationships for natural catastrophes. The study concluded that LSTM is a more suitable choice for predicting water inundation frequency due to its superior ability to handle sequential data for natural catastrophes and long-term dependencies.

DECLARATION

Conflict of Interests

This paper contains no apparent conflicts of interest. To maintain our commitment to professional authenticity and prevent any inadvertent participation with concerns related to academic dishonesty, we carefully checked that our work is original.

Acknowledgment

In the absence of the invaluable support and supplies generously provided by the highly regarded Saveetha Institute of Medical and Technical Sciences, this research would not have been feasible. The authors would like to convey their profound appreciation for this respectable organization's priceless assistance.

Authors Contribution

Author KSG had an active role in the collection of data, its analysis, and composition. Conversely, Author SKA contributed to the research idea effectively, validated the data, and provided insightful feedback during the manuscript review process.

Funding

The following institutions and organizations generously offered crucial financial assistance that allowed us to successfully complete our research. We are truly appreciative of their significant contribution to our scientific pursuits.University of Saveetha

1. Saveetha Institute of Medical and Technical Sciences
2. Saveetha School of Engineering
3. Saveetha University

REFERENCES

Bentivoglio, Roberto, Elvin Isufi, Sebastian Nicolaas Jonkman, and Riccardo Taormina. 2022. "Deep Learning Methods for Flood Mapping: A Review of Existing Applications and Future Research Directions." *Hydrology and Earth System Sciences* 26 (16): 4345–78.

Cho, Minwoo, Changsu Kim, Kwanyoung Jung, and Hoekyung Jung. 2022. "Water Level Prediction Model Applying a Long Short-Term Memory (LSTM)–Gated Recurrent Unit (GRU) Method for Flood Prediction." *WATER* 14 (14): 2221.

Devakumar, K. P. 2019. "Kerala Floods 2018." https://www.kaggle.com/imdevskp/kerala-floods-2018.

Gude, Vinayaka, Steven Corns, and Suzanna Long. 2020. "Flood Prediction and Uncertainty Estimation Using Deep Learning." *WATER* 12 (3): 884.

Ha, Si, Darong Liu, and Lin Mu. 2021. "Prediction of Yangtze River Streamflow Based on Deep Learning Neural Network with El Niño–Southern Oscillation." *Scientific Reports* 11 (1): 1–23.

Hayder, Israa M., Taief Alaa Al-Amiedy, Wad Ghaban, Faisal Saeed, Maged Nasser, Ghazwan Abdulnabi Al-Ali, and Hussain A. Younis. 2023. "An Intelligent Early Flood Forecasting and Prediction Leveraging Machine and Deep Learning Algorithms with Advanced Alert System." *Processes* 11 (2): 481.

Kim, Hyun Il, Kun Yeun Han, and Jae Yeong Lee. 2020. "Prediction of Urban Flood Extent by LSTM Model and Logistic Regression." *KSCE Journal of Civil and Environmental Engineering Research* 40 (3): 273–83.

Kumar, Vijendra, Hazi Azamathulla, Kul Vaibhav Sharma, Darshan J. Mehta, and Kiran Tota Maharaj. 2023. "The State of the Art in Deep Learning Applications, Challenges, and Future Prospects: A Comprehensive Review of Flood Forecasting and Management." *Sustainability: Science Practice and Policy* 15 (13): 10543.

Li, Wei, Amin Kiaghadi, and Clint Dawson. 2020. "Exploring the Best Sequence LSTM Modeling Architecture for Flood Prediction." *Neural Computing & Applications* 33 (11): 5571–80.

Nayak, Monalisa, Soumya Das, and Manas Ranjan Senapati. 2022. "Improving Flood Prediction with Deep Learning Methods." *Journal of The Institution of Engineers (India): Series B* 103 (4): 1189–1205.

Nguyen, Dinh Ty, and Shien-Tsung Chen. 2020. "Real-Time Probabilistic Flood Forecasting Using Multiple Machine Learning Methods." *WATER* 12 (3): 787.

Panigrahi, Binaya Kumar, Tushar Kumar Nath, and Manas Ranjan Senapati. 2019. "An Application of Local Linear Radial Basis Function Neural Network for Flood Prediction." *Journal of Management Analytics*, January. https://doi.org/10.1080/23270012.2019.1566033.

Sankaranarayanan, Suresh, Malavika Prabhakar, Sreesta Satish, Prerna Jain, Anjali Ramprasad, and Aiswarya Krishnan. 2020. "Flood Prediction Based on Weather Parameters Using Deep Learning." *Journal of Water and Climate Change* 11 (4): 1766–83.

Sit, Muhammed, and Ibrahim Demir. 2019. "Decentralized Flood Forecasting Using Deep Neural Networks." http://arxiv.org/abs/1902.02308.

Vineeth, V., and E. A. Neeba. n.d. "Flood Detection Using Deep Learning." Accessed January 10, 2024. https://ieeexplore.ieee.org/abstract/document/9708240.

Zhang, Yu, and Yonghe Hao. 2021. "Loss Prediction of Mountain Flood Disaster in Villages and Towns Based on Rough Set RBF Neural Network." *Neural Computing & Applications* 34 (4): 2513–24.

TABLES AND FIGURES

Table 1. Representing LSTM (Long Short-Term Memory) algorithm pseudocode**.**

Input: Kerala floods data
Output: Prediction of Water Inundation Frequency
Phase 1: Compiling and Getting Ready Data Input: Raw data from Kerala floods dataset. Output: Processed data specific to Kerala floods. Description: Collect and preprocess the Kerala floods dataset, handling missing values, and converting the data into a format suitable for time-series analysis. This may involve cleaning satellite imagery data, integrating weather records, and ensuring the data covers the relevant time periods of the floods.
Phase 2: Selecting Elements Input: Processed Kerala floods data. Output: Subset of relevant features for LSTM modeling. Description: Identify features crucial for predicting flooding in Kerala. This may include rainfall data, water levels, land cover information, and historical flood occurrences.
Phase 3: Data Division Input: Selected features from Kerala floods data. Output: Training and testing datasets. Description: Split the processed Kerala floods data into training and testing sets, considering the temporal aspect to ensure that the model generalizes well to future flood occurrences.
Phase 4: Selecting a Model Input: LSTM Output: LSTM Model for Kerala floods. Description: Choose the LSTM model for its capability to capture sequential dependencies in

time-series data, which is relevant for predicting the frequency of floods in Kerala over time.

Phase 5: Model Instruction

Input: Training data and selected features from Kerala floods.

Output: Trained LSTM model for Kerala floods.

Description: Train the LSTM model using the selected features and historical data specific to Kerala floods. Adjust LSTM parameters to effectively capture temporal patterns in the dataset.

Phase 6: Evaluate the Model

Input: Trained LSTM model and testing data for Kerala floods.

Output: LSTM Model performance metrics for Kerala floods.

Description: Evaluate the LSTM model's performance on the testing dataset from the Kerala floods, considering factors such as prediction accuracy, sensitivity to flood events, and robustness in capturing temporal patterns.

Phase 7: Realizing

Input: Trained LSTM model for Kerala floods.

Output: Insights into LSTM model behavior for Kerala floods.

Description: Explore the interpretability of the LSTM model specifically in the context of Kerala floods. Understand how the model utilizes historical information to predict flooding events in the region.

Phase 8: Bring into Performance

Input: Trained LSTM model and selected features for Kerala floods.

Output: Deployed LSTM model for predicting floods in Kerala.

Description: Deploy the trained LSTM model for predicting floods in Kerala in a real-world setting. This may involve integrating the model into a system that can provide timely flood predictions for decision-makers in the region.

Table 2. Representing RBFNN (Radial Basis Function Neural Network) algorithm pseudocode.

Input: Kerala floods data
Output: Prediction of Water Inundation Frequency
Phase 1: Compiling and Getting Ready Data Input: Raw data from Kerala floods dataset. Output: Processed data specific to Kerala floods. Description: Collect and preprocess the Kerala floods dataset, handling missing values, and converting the data into a format suitable for time-series analysis. This may involve cleaning satellite imagery data, integrating weather records, and ensuring the data covers the relevant time periods of the floods.
Phase 2: Selecting Elements Input: Processed Kerala floods data. Output: Subset of relevant features for LSTM modeling. Description: Identify features crucial for predicting flooding in Kerala. This may include rainfall data, water levels, land cover information, and historical flood occurrences.
Phase 3: Data Division Input: Selected features from Kerala floods data. Output: Training and testing datasets. Description: Split the processed Kerala floods data into training and testing sets, considering the temporal aspect to ensure that the model generalizes well to future flood occurrences.
Phase 4: Selecting a Model Input: RBFNN Output: RBFNN Model for Kerala floods. Description: Opt for the RBFNN model due to its ability to handle non-linear relationships, which may be present in the complex patterns of Kerala floods data.

Phase 5: Model Instruction Input: Training data and selected features from Kerala floods. Output: Trained RBFNN model for Kerala floods. Description: Train the RBFNN model using the selected features and training data specific to Kerala floods. Adjust RBFNN parameters to effectively capture the non-linear relationships within the dataset.
Phase 6: Evaluate the Model Input: Trained RBFNN model and testing data for Kerala floods. Output: RBFNN Model performance metrics for Kerala floods. Description: Evaluate the RBFNN model's performance on the testing dataset from the Kerala floods. Assess its ability to approximate complex functions and capture non-linear dependencies in the flood data.
Phase 7: Realizing Input: Trained RBFNN model for Kerala floods. Output: Insights into RBFNN model behavior for Kerala floods. Description: Examine the interpretability of the RBFNN model specifically in the context of Kerala floods. Gain insights into how the model represents and processes information to predict flooding events.
Phase 8: Bring into Performance Input: Trained RBFNN model and selected features for Kerala floods. Output: Deployed RBFNN model for predicting floods in Kerala. Description: Deploy the trained RBFNN model for predicting floods in Kerala in a real-world setting. Ensure a seamless integration that allows the model to provide accurate predictions based on the unique characteristics of the Kerala floods dataset.

Table 3. Representing the Accuracy(93.005%) of LSTM & Accuracy(88.481%) of RBFNN Algorithms. The RBFNN Algorithm is less accurate than the LSTM Algorithm

SAMPLE NO	LSTM(%)	RBFNN(%)
1	96.81	92.45
2	96.25	92.21
3	95.89	91.29
4	95.74	91.25
5	94.78	90.67
6	94.68	90.34
7	94.45	90.31
8	94.32	89.88
9	92.56	89.39
10	92.55	89.34
11	92.41	89.26
12	91.37	88.92
13	92.31	88.26
14	91.63	87.93
15	91.49	87.77
16	91.34	86.34
17	91.14	84.93
18	90.85	84.25
19	89.67	82.68
20	88.95	82.15
AVERAGE	93.0095	88.481

Table 4. Group Statistics Results-LSTM has a mean accuracy (93.0095%), St. Deviation (2.23460), whereas RBFNN has a mean accuracy (88.4810%), St. Deviation (2.99839).

	Group	N	Mean	Std. Deviation	Std. Error Mean
Accuracy	LSTM	20	93.0095	2.23460	0.49967
	RBFNN	20	88.4810	2.99839	0.67046

Table 5. The independent sample t-test for Equality of Means is given with the Equal variances assumed and also Equal variances not assumed.

		Levene's Test for Equality of Variances		T-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Accuracy	Equal variances assumed	0.998	0.324	5.416	38	0.000	4.52850	0.83617	2.83575	6.22125
	Equal variances not assumed			5.416	35.130	0.000	4.52850	0.83617	2.83120	6.22580

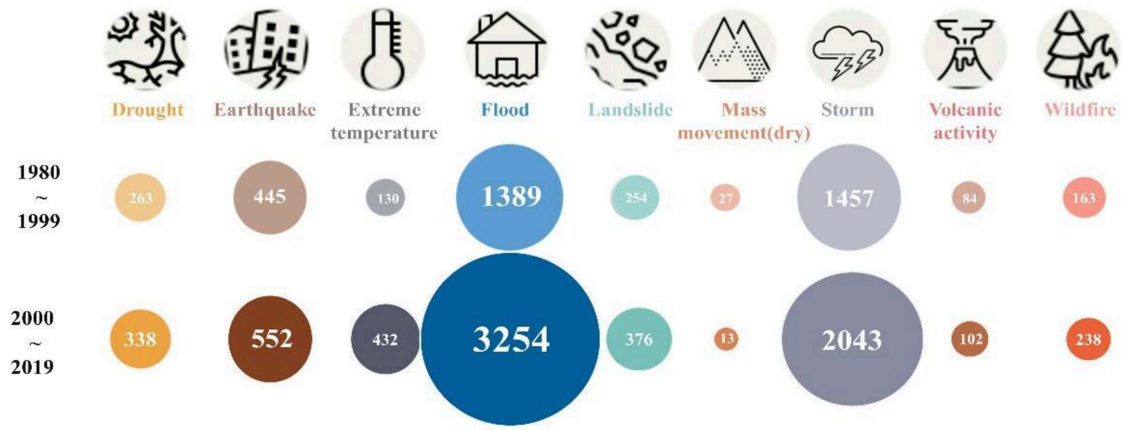

Figure 1. The categories of weather-related events : 1980–1999 vs. 2000–2019 from (Cho et al. 2022)**.**

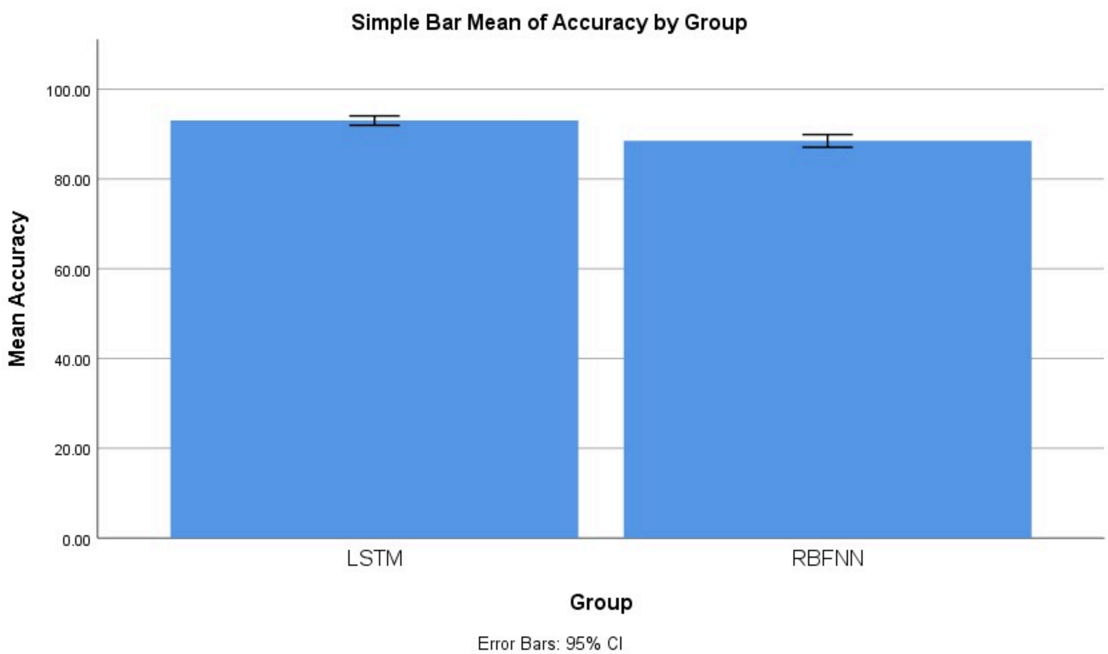

Figure 2: Mean accuracy comparison between RBFNN and LSTM algorithm. Compared to the RBFNN algorithm's mean accuracy of 88.4810, the LSTM's mean accuracy of 93.0095 is superior. Long Short-Term Memory (LSTM) vs. Radial Basis Function Neural Network (RBFN) on the X-axis. Y-axis: Precision(Accuracy).