Title page:

Improved Accuracy in Prediction of Water Inundation Frequency using Long Short-Term Memory compared over Temporal Convolutional Networks .

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, Floods, Long Short-Term Memory,Machine Learning, Natural Disasters, Neural Networks, Time series,Temporal Convolutional Networks , Water Inundation Frequency.

ABSTRACT

Aim:The primary objective of the research is to assess how well Long Short-Term Memory and Temporal Convolutional Networks anticipate the frequency of flooding, which is one of the most dangerous catastrophic events which can cause significant damage to buildings and facilities. **Materials and Methods:** There are two distinct groups of suggested algorithms. Temporal Convolutional Network is compared with Long Short-Term Memory. Both systems are capable of accurately forecasting the frequency of a water inundation. The frequency of water inundations is examined for accuracy. Long-term dependencies can be handled by LSTMs, while TCNs can compute more quickly, have less gradient vanishing, have flexible receptive fields, and have better stability. The algorithm builds a model based on the characteristics of the training set of data, which it then applies to determine the value of fresh data.For the sample size (N=20), the LSTM and TCN algorithms are both employed to identify the frequency of water inundation. Using SPSS, the significance value of the data set was estimated with a G-power value more than 80%. **Results:** Compared to TCN, which has an accuracy of 86.7730%, LSTM has achieved a greater accuracy of 93.0095%. A significant difference ($p<0.05$) is seen between the two groups, with a significance value of 0.046.**Conclusion:** Ultimately, these results show that the LSTM has a higher predictive capacity for estimating the frequency of flooding.LSTM outperforms TCN for applications that need to capture long-term dependencies in the data.

Keywords : Deep Learning, Floods, Long Short-Term Memory,Machine Learning, Natural Disasters, Neural Networks, Time series, Temporal Convolutional Networks , Water Inundation Frequency.

INTRODUCTION

"Water inundation," the term for the either periodic or ongoing flooding of land by water, is a serious global concern. Changes in land use, urbanization, and climate change are interacting in a way that is making floods more frequent and intense(Y. Xu et al. 2021). Both coastal and interior areas are impacted by these phenomena; the latter are affected by storm surges and tsunamis, while the former are affected by riverine overflows and flash floods(Yao et al. 2023). Accurately predicting the frequency of these floods events has become crucial for disaster preparedness, resource allocation, and vulnerable population protection(K. Lin et al. 2020).Let's now explore the field of deep learning, where methods like long short-term memory and recurrent neural networks hold enormous promise for understanding the intricate temporal correlations observed in water inundation dynamics(Shao et al. 2023).

An extensive statistical examination of flood incidents demonstrates the gravity of the issue. A graphic representation of the extent of the damage caused by natural catastrophes may be found in Figure 1(Cho et al. 2022). It highlights the growing impacts that have been demonstrated throughout the course of two notable decades, from 2000 to 2019 and 1980 to 1999. During our research, we scanned the Scopus database extensively. What we found was that, over the preceding five years, there were 2,801 publications on Science Direct and over 17,700 papers on Google Scholar. The intricacy of the subject at hand as well as the expanding importance of water inundation prediction are shown by this growth in scholarly attention.Beyond the ecology of the area, floods and other types of flooding pose a complex challenge(Chen et al. 2022).In addition to the risk of direct physical harm, there is also the potential for interruptions to agricultural productivity, public safety risks, and potential consequences on the integrity of the socioeconomic system as a whole(H. Xu 2023). Because of the scope of this issue, innovative and workable solutions are needed.Water inundation events are becoming more frequent and unpredictable, which emphasizes how important it is to have accurate forecast models in order to reduce possible damage(Bayat and Tavakkoli 2022).

The primary objective of this research is to conduct a thorough and effective evaluation of the performance of Long Short-Term Memory (LSTM) and Recurrent Neural Network (RNN) models for forecasting the frequency of flooding. Our main objective is to improve the accuracy and dependability of these models using state-of-the-art deep learning (DL) techniques, with the potential to greatly enhance early warning systems and disaster preparedness initiatives(Tao et al. 2023). Beyond merely honing prediction skills, our research intends to explore the intricate dynamics surrounding water inundation, accounting for significant factors such as land-use patterns, urbanization, and climate change(Dijkstra 2019).

MATERIALS AND METHODS

The data analytics lab of the Saveetha Institute of Medical and Technical Sciences was used to conduct this investigation. In the lab, a highly adjustable system allows for precise results and in-depth research. Ten respondents in all were included in the study, and they were separated into two groups: Group 1 utilized the Long Short-Term Memory approach and Group 2 used the Temporal Convolutional Networks . 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.

A comma-separated values dataset named "Kerala Floods"(Devakumar 2019) is used in this water inundation frequency estimate.To improve the precision of frequency determination, the dataset incorporates variables like year, month, and annual rainfall. The dataset has been cleansed for testing and training. Normalization of the data has been performed, along with the removal of missing values and the replacement of null values with means or medians. The preprocessed feature-rich dataset is fed into the LSTM algorithm.

The method presented in this study was applied to a specific dataset, and the outcomes were contrasted with the well-known RNN algorithm. The dataset is set up in this case as a comma-separated value 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

The algorithm type called Long Short-Term Memory is ideal for modeling sequential data, such time series data.A common issue with RNNs that can hinder them from learning long-term dependencies is the vanishing gradient problem, which Long Short-Term Memory networks are specifically made to address(Faruq et al. 2020). This makes Long Short-Term Memory networks a viable candidate for predicting water inundation frequency, as the frequency of inundation is expected to be influenced by long-term causes such as climate change and human activities(Kimura et al. 2019).

Temporal Convolutional Networks

A kind of convolutional neural network called a temporal convolutional network is made especially for modeling sequential data, such time series data(Q. Lin et al. 2023). Although TCNs may identify temporal patterns and correlations in sequential data, CNNs are usually employed for picture classification and other tasks using spatial data.Dilated causal convolutional layers, a kind of convolutional layer that guarantees that the output of the layer only depends on values that are earlier in the input sequence, make up TCNs(Ogunjo, Olusola, and Olusegun 2023).Long-distance dependencies in sequential data can also be captured by TCNs. This is so that they may integrate data from a wide range of time steps since the dilated convolutional layers can have broad receptive fields**.**

Statistical Analysis

The analysis was performed using IBM SPSS version 26. The acquired accuracy values were examined in SPSS using a sample size of 20 for the algorithms Long Short-Term Memory and Temporal Convolutional Networks . Long Short-Term Memory and Temporal Convolutional Networks are used to identify Groups 1 and 2. The model is trained using the dataset's features. The model accuracy values for each epoch were used to generate the T-Test. The performance comparison between the two methods is made easier using T-Test.

RESULTS

The correctness of the raw data table for temporal convolutional networks and long short-term memory networks is displayed in Table 1. The 20-sample sample sizes of the Temporal Convolutional Network and the Long Short-Term Memory are used to compute the accuracy values. The Temporal Convolutional Networks technique has an accuracy value of 87.2605%, whereas Long Short-Term Memory has a mean accuracy value of 93.0095%. This accuracy data is displayed in Table 2. This suggests that the recommended algorithm is more accurate than the others. This study used the LSTM (Long Short-Term Memory) approach instead of Temporal Convolutional Networks.

Table 3 displays the "F" value as 4.749 and the "Sig" value as 0.036 with equal variances assumed. For equal variances assumed and not assumed, the 95% confidence intervals for the difference are 3.59931 and 3.57718 for the lowercase and 7.89869 and 7.92082 for the upper case(shown in table 3). Plotting mean accuracy on the Y-axis and groups on the X-axis results in the bar graph displayed in Figure 1. The graph clearly shows that LSTM (Long Short-Term Memory) is more accurate than Temporal Convolutional Networks (TCN).

DISCUSSION

The research's significant value, which is 0.019 (two-tailed, $p>0.05$), suggests that regions with temporal convolutional networks are not as good as long-term memory.While the mean accuracy of the Regions with Temporal Convolutional Networks classifier is 87.2605%, the Long Short-Term Memory approach's mean accuracy analysis is 93.0005%. which suggests that in comparison to TCN, LSTM performs better.

Both Long Short-Term Memory (LSTM) and Temporal Convolutional Networks (TCN) are effective techniques, each having certain 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 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 computing speed and ability to handle nonlinear interactions, TCN is perfect for modeling sequential data, such as time series data. Its relatively simple training method makes it interesting for smaller or simpler datasets. But it's crucial to consider TCN's tendency toward overfitting and its limited ability to identify long-term dependencies.

The size, diversity, and feature selection of the dataset, generalizability, and unproven algorithmic choices are among the potential limits of the research(Zhang et al. 2022). 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. Potential approaches for boosting prediction accuracy and the comprehension of Water Inundation Frequency include focusing ensemble models, interpretability, and real-world applications in environmental science and crisis management(Z. Xu et al. 2022). Further investigation into alternative machine learning techniques, DL approaches, parametric optimization, and real-world validation may yield more promising predictive models.

CONCLUSION

In this examination, the mean accuracy of water inundation frequency prediction using the Long Short-Term Memory (LSTM) algorithm was 93.0005%, whereas the mean accuracy of regions utilizing Temporal Convolutional Networks (TCN) was 87.2605%. The effectiveness of Long Short-Term Memory and Temporal Convolutional Networks models in forecasting the frequency of water inundations was examined in this study. The outcomes showed that both models functioned effectively, with Long Short-Term Memory outperforming Temporal Convolutional

Networks in terms of accuracy. The application of Long Short-Term Memory in time series data to identify long-term dependencies was beneficial in forecasting the frequency of water inundations, which is impacted by past hydrological trends. Although Temporal Convolutional Networks performed satisfactorily, it was not as good at capturing intricate temporal correlations. The study came to the conclusion that because Long Short-Term Memory can handle sequential data better than other models, it is a better option for forecasting the frequency of water inundations.Given that long-term factors like climate change and human activity are predicted to have an impact on inundation frequency, long short-term memory networks are a good option for forecasting water inundation frequency.

DECLARATION

Conflict of Interests

This article contains no declared conflicts of interest. To maintain our commitment to academic integrity 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 materials generously provided by the highly regarded Saveetha Institute of Medical and Technical Sciences, this study would not have been possible. The authors would like to convey their profound appreciation for this respectable organization's priceless assistance.

Authors Contribution

Author KSG participated fully in the data creation, analysis, and gathering. On the other hand, Author SK made a significant contribution to the research proposal, verified the data, and offered critical criticism throughout the manuscript review procedure.

Funding

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

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

REFERENCES

Bayat, Mehrsa, and Omid Tavakkoli. 2022. "Application of Machine Learning in Flood Forecasting." *Future Technology* 1 (1): 01–06.

Chen, Delong, Ruizhi Zhou, Yanling Pan, and Fan Liu. 2022. "A Simple Baseline for Adversarial Domain Adaptation-Based Unsupervised Flood Forecasting." http://arxiv.org/abs/2206.08105.

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.

Dijkstra, Rick. 2019. "Predicting Water Levels in the Rhine River Using Temporal Convolutional Networks," July. https://theses.ubn.ru.nl/handle/123456789/10693.

Faruq, A., S. S. Abdullah, A. Marto, M. A. A. Bakar, Samin, and A. Mubin. 2020. "River Water Level Forecasting for Flood Warning System Using Deep Learning Long Short-Term Memory Network." *IOP Conference Series: Materials Science and Engineering* 821 (1): 012026.

Kimura, Nobuaki, Ikuo Yoshinaga, Kenji Sekijima, Issaku Azechi, and Daichi Baba. 2019. "Convolutional Neural Network Coupled with a Transfer-Learning Approach for Time-Series Flood Predictions." *WATER* 12 (1): 96.

Lin, Kangling, Sheng Sheng, Yanlai Zhou, Feng Liu, Zhiyu Li, Hua Chen, Chong-Yu Xu, Jie Chen, and Shenglian Guo. 2020. "The Exploration of a Temporal Convolutional Network Combined with Encoder-Decoder Framework for Runoff Forecasting." *Hydrology Research* 51 (5): 1136–49.

Lin, Qinyue, Zeping Yang, Jie Huang, Ju Deng, Li Chen, and Yiru Zhang. 2023. "A Landslide Displacement Prediction Model Based on the ICEEMDAN Method and the TCN–BiLSTM Combined Neural Network." *WATER* 15 (24): 4247.

Ogunjo, Samuel, Adeyemi Olusola, and Christiana Olusegun. 2023. "Predicting River Discharge in the Niger River Basin: A Deep Learning Approach." *NATO Advanced Science Institutes Series E: Applied Sciences* 14 (1): 12.

Shao, Pingping, Jun Feng, Yirui Wu, Wenpeng Wang, and Jiamin Lu. 2023. "Data-Driven and Knowledge-Guided Heterogeneous Graphs and Temporal Convolution Networks for Flood Forecasting." *NATO Advanced Science Institutes Series E: Applied Sciences* 13 (12): 7191.

Tao, Lizhi, Xinguang He, Jiajia Li, and Dong Yang. 2023. "A Multilevel Temporal Convolutional Network Model with Wavelet Decomposition and Boruta Selection for Forecasting Monthly Precipitation." *Journal of Hydrometeorology* 24 (11): 1991–2005.

Xu, Hongyi. 2023. "Multistep Prediction of Multiple Small Hydropower Stations' Total Power in a Watershed within a Day Considering the Distributed Discharge TCN-LSTM Model." *Journal of Physics. Conference Series* 2625 (1): 012067.

Xu, Yuanhao, Caihong Hu, Qiang Wu, Zhichao Li, Shengqi Jian, and Youqian Chen. 2021. "Application of Temporal Convolutional Network for Flood Forecasting." *Hydrology Research* 52 (6): 1455–68.

Xu, Zhihao, Zhiqiang Lv, Jianbo Li, and Anshuo Shi. 2022. "A Novel Approach for Predicting Water Demand with Complex Patterns Based on Ensemble Learning." *Water Resources Management* 36 (11): 4293–4312.

Yao, Jinliang, Zhipeng Cai, Zheng Qian, and Bing Yang. 2023. "A Noval Approach Based on TCN-LSTM Network for Predicting Waterlogging Depth with Waterlogging Monitoring Station." *PloS One* 18 (10): e0286821.

Zhang, Di, Kai Wei, Yi Yao, Jiacheng Yang, Guolong Zheng, and Qing Li. 2022. "Capture and Prediction of Rainfall-Induced Landslide Warning Signals Using an Attention-Based Temporal Convolutional Neural Network and Entropy Weight Methods." *Sensors* 22 (16): 6240.

TABLES AND FIGURES

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

Input: Kerala floods data
Output: Prediction of Water Inundation Frequency
Step 1: Data Collection and Preprocessing 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.
Step 2: Feature Selection: 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.
Step 3:Data Splitting: 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.
Step 4: Model Selection: 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.
Step 5: Model Training: 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.
Step 6:Model Evaluation: 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.
Step 7: Interpretability: 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.
Step 8: Deployment: 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 TCN(Temporal Convolutional Network) algorithm pseudocode.

Input: Kerala floods data
Output: Prediction of Water Inundation Frequency
Step 1: Data Collection and Preprocessing 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.
Step 2: Feature Selection: 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.
Step 3:Data Splitting: 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.
Step 4: Model Selection: Input: TCN Output: TCN Model. Description: Choose the TCN model for its ability to capture temporal dependencies using convolutional operations. TCN is known for its parallel processing capabilities and has shown effectiveness in time-series forecasting tasks.
Step 5: Model Training: Input: Training data and selected features. Output: Trained TCN model.

Description: Train the TCN model using the selected features and historical data. Adjust TCN parameters to effectively capture temporal patterns associated with water inundation events.
Step 6:Model Evaluation: Input: Trained TCN model and testing data. Output: Performance metrics for the TCN Model. Description: Evaluate the TCN model's performance on the testing dataset, considering metrics such as MAE, RMSE, and any other relevant metrics for accuracy assessment. Compare the performance with that of the LSTM model.
Step 7: Interpretability: Input: Trained TCN model. Output: Insights into TCN model behavior. Description: Explore the interpretability of the TCN model to understand how it captures temporal dependencies and leverages selected features for predicting water inundation frequency.
Step 8: Deployment: Input: Trained TCN model and selected features. Output: Deployed TCN model for predicting water inundation frequency. Description: Deploy the trained TCN model in a real-world setting to provide efficient predictions of water inundation frequency. Ensure smooth integration for practical applications.

Table 3. Representing the Accuracy(93.005%) of LSTM & Accuracy(87.2605%) of TCN Algorithms. The TCN Algorithm is less accurate than the LSTM Algorithm

SAMPLE NO	LSTM(%)	TCN(%)
1	96.81	91.83
2	96.25	91.67
3	95.89	91.21
4	95.74	90.99
5	94.78	90.83
6	94.68	90.23
7	94.45	89.93
8	94.32	89.12
9	92.56	88.92
10	92.55	88.82
11	92.41	88.56
12	91.37	87.35
13	92.31	87.34
14	91.63	86.23
15	91.49	85.43
16	91.34	84.98
17	91.14	83.33
18	90.85	82.63
19	89.67	79.17
20	88.95	76.64
AVERAGE	93.0095	87.2605

Table 4. Group Statistics Results-LSTM has a mean accuracy (93.0095%), St. Deviation (2.23460), whereas TCN has a mean accuracy (87.2605%), St. Deviation (4.19033).

	Group	N	Mean	Std. Deviation	Std. Error Mean
Accuracy	LSTM	20	93.0095	2.23460	0.49967
	TCN	20	87.2605	4.19033	0.93699

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	4.749	0.036	5.414	38	0.000	5.74900	1.06189	3.59931	7.89869
	Equal variances not assumed			5.414	28.998	0.000	5.74900	1.06189	3.57718	7.92082

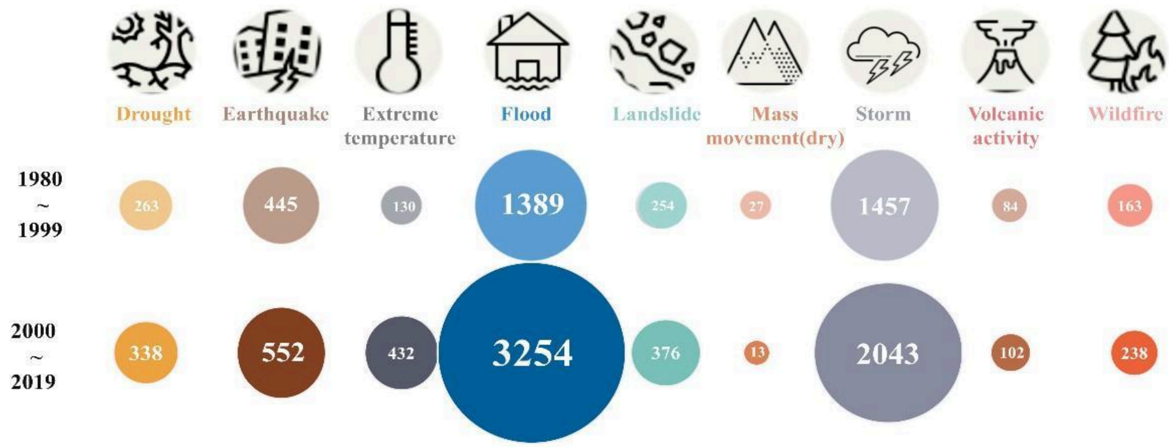

Figure 1. The Natural Hazards events types : 1980–1999 vs. 2000–2019 from (Cho et al. 2022)

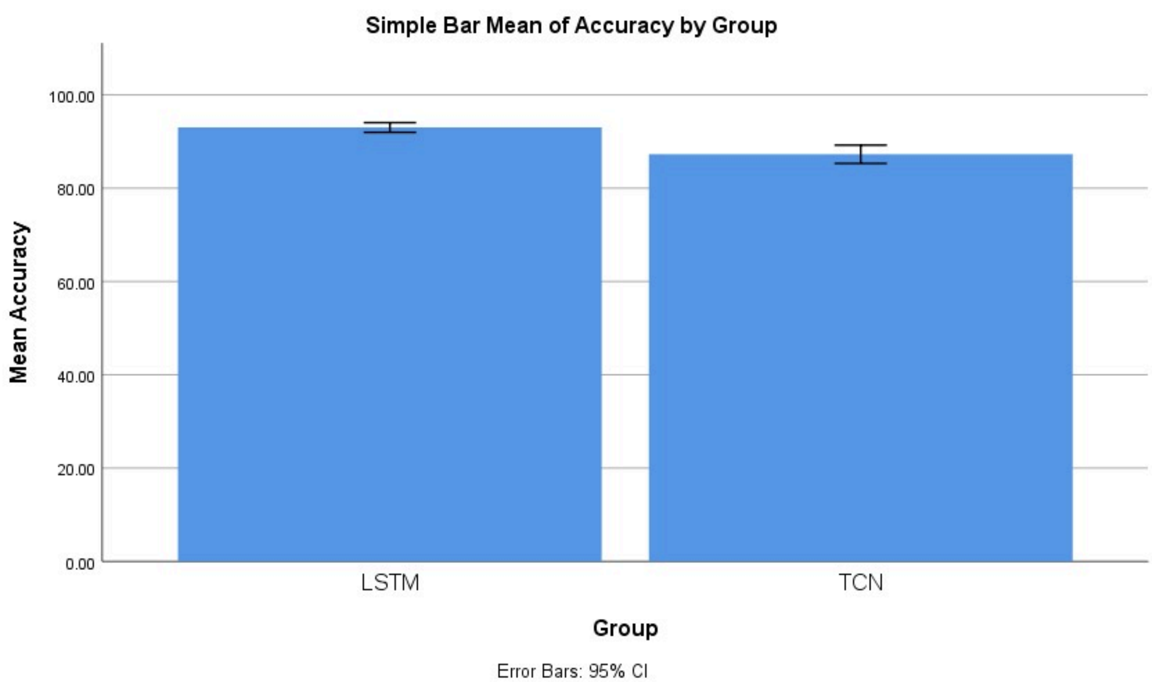

Figure 2. Comparison of LSTM Algorithm and TCN in terms of mean accuracy. The mean accuracy of the LSTM(93.0095) is better; than the mean accuracy of TCN algorithm(87.2605). X-axis: LSTM(Long Short-Term Memory) VS TCN(Temporal Convolutional Network). Y-axis: Mean accuracy.