

# Forecasting of Flood in the Non-Tidal River of Northern Regions of Bangladesh Using Machine Learning-Based Approach

Md. Khalid Hasan<sup>1\*</sup>, Md. Mofizul Islam<sup>2</sup>, Maisha Fahmida<sup>3</sup>

<sup>1-3</sup>Department of Agricultural & Industrial Engineering, Hajee Mohammad Danesh Science and Technology University, Dinajpur-5200, Bangladesh

## Abstract

Floods are among the most devastating natural disasters, causing extensive damage to property and posing a threat to human lives. However, significant progress has been made in mitigating their impact through the development of effective early warning systems. Over the past two decades, advances in machine learning (ML) technology have played a crucial role in enhancing the predictive capabilities of these systems. A recent study focused on predicting floods in non-tidal rivers by proposing various machine-learning models. The research findings indicate that the Random Forest algorithm emerges as the most effective, offering an accuracy of 87% with high precision, recall, and F1 scores, using an 80:20 training and testing data ratio. These findings provide valuable insights for hydrologists and make a significant contribution to flood forecasting and mitigation efforts. The study has significant implications for flood understanding and management, offering a better understanding of machine learning model performance in predicting floods in non-tidal rivers. This research provides a solid foundation for the development of more efficient early warning systems. The information gleaned from this study can be utilized by hydrologists, climate scientists, and other related practitioners to develop more accurate and reliable forecasting strategies in the face of flood threats. Thus, this research is not only a valuable scientific contribution but also a practical tool for future flood disaster risk mitigation efforts.

**Keywords:** Flood Forecasting; River; Machine Learning Algorithm; Data Analysis; Model Development.

Received: 31 December 2023

Revised: 29 March 2024

Accepted: 29 April 2024

## Introduction

Floods are widely recognized as one of the most devastating natural disasters, with impacts that can persist for years (McGlade et al., 2019). They are caused by a variety of factors, as seen in different parts of the world. For instance, the Queensland floods resulted from the soil's inability to absorb water, leading to excess water flowing through river channels and causing flooding (Chinchor, 1992; Tiggeloven et al., 2021). Rivers can also flood when the volume of water exceeds the channel's capacity, particularly at bends and meanders. Floods pose significant environmental challenges globally (Khan et al., 2024; Green et al., 2024). While preventing floods is crucial, early forecasting is seen as a lifesaving measure, especially for people living in river basin areas, as it can reduce damage to property.

Bangladesh, often referred to as the land of rivers, is home to approximately 700 rivers, including tributaries, with a combined length of 24,140 kilometers, encompassing all rivers, streams, creeks, and channels (Rahman et al., 2015; Alam et al., 2022). Some of these rivers rank among the world's largest based on criteria such as catchment size, length, and discharge volume (Hirpa et al., 2013). However, annually, about 26,000 square kilometers of land in Bangladesh are flooded, resulting in the loss of 5,000 lives and the destruction of over seven million homes (Price, 2020). Rivers in northern Bangladesh are non-tidal and generally have low water levels except during the rainy season, exhibiting different characteristics compared to rivers in the southern part of the country.

All rivers studied in this research originate from Bangladesh and India (Baxter, 2018), with many originating from the Himalayan Mountains, and their water flow is directly affected by rainfall in the Indian states that share the rivers' basins (Dewan, 2015; Zheng et al., 2021). The flood situation and water levels in northern Bangladesh's rivers are directly influenced by flooding and water levels in the common river portions of India (Roy et al., 2019; Ali et al., 2019; Becker et al., 2020). In April 2017, heavy rains caused flooding in several parts of Bangladesh, damaging pre-harvested crops. Flood inundation maps generated using Sentinel-1 data for March, April, June, and August 2017 showed the

\*Corresponding author.

E-mail address: [khalidhstu1996@gmail.com](mailto:khalidhstu1996@gmail.com) (Md. Khalid Hasan)



presence of perennial water bodies covering 5.03% of Bangladesh's total land area in March 2017. The most significant flooding occurred in croplands in April (1.51%), followed by rural settlements and homestead orchards (0.21%), and other areas (0.29%), with a total flooded area of 2.01%. Larger areas were submerged during the catastrophic months of June and August, with inundation covering 4.53% and 7.01% of the land, respectively.

In 2017, South Asian countries, including Bangladesh, India, and Nepal, experienced a devastating flood, affecting approximately 41 million people and resulting in 1200 deaths and 950,000 damaged houses across these nations (George, 2017; Siddique, 2017). The flood originated in India before affecting downstream Bangladesh (Mondal et al., 2020; Chakraborty et al., 2021; Palash et al., 2020). Particularly, the northern region (Aziz et al., 2022). Lack of early warnings left people with insufficient time to evacuate, exacerbating the impact (Kumar et al., 2022; Erdianto, 2023). This tragedy inspired the selection of a study area in northern Bangladesh to enhance flood prediction accuracy (Sarkar et al., 2022; Hossain, 2024).

Various analysis techniques have been employed to forecast floods and assess high water levels (Munawar et al., 2022; Umgiesser et al., 2020). Previous researchers have utilized hydrological, data-driven, statistical, machine learning, and deep-learning methods for flood prediction (Bentivoglio et al., 2022; Speight et al., 2021), mainly focusing on major rivers, with some exceptions involving different channels (Puttinaovarat & Horkaew, 2020). Physical models, such as global circulation, have also been proposed for flood forecasting based on environmental time series data. While these models have shown effectiveness in anticipating various flooding scenarios (Riza & Nuryadi, 2023; Wheeler, 2002; Mosavi et al., 2018; Brunner et al., 2021), they often require extensive hydro-geomorphological monitoring datasets, leading to high computational costs and limiting short-term predictions (Bates, 2022; Guo et al., 2021; Zahura et al., 2020). The construction of physically based models also demands profound knowledge and expertise in hydrological factors, which can be challenging (Kim et al., 2015).

(Valipour et al., 2012) introduced physical and numerical models, while data-driven models have a long history in flood modeling. These approaches have recently gained popularity, particularly autoregressive moving average (ARIMA) models. (Adamowski et al., 2012) developed multiple linear regression (MLR) models, which have been widely used for flood forecasting. Machine learning, neural networks, and other technologies have also been developed for flood forecasting and time series analysis. Researchers successfully predicted Tarbela reservoir inflow using fusion models of regression and neural networks. (Cheng et al., 2015) proposed a neural network-based model with particle swarm optimization to predict daily river flows. They applied the SVM-BP model for flood forecasting in the Changhua River, showing that this combination outperformed other methods. (Rasel et al., 2018) forecasted tidal levels in the Karnafuli River, crucial for the daily activities of Chittagong residents such as fishing, waterway traffic, and port activity regulation. They developed a machine learning model using ten years of historical Karnafuli River datasets (2007–2017) and compared SVM, BP-ANN, and DNN algorithms, with DNN achieving 99% accuracy in water level prediction.

This study focuses on advanced techniques like Machine Learning (ML), proposing a hybrid model to enhance prediction accuracy. The objectives include predicting floods in the non-tidal rivers of northern Bangladesh using a machine learning approach, designing a methodology to predict floods more accurately than existing models, and conducting a comparative study of algorithms to select the best model for flood prediction in the region.

## Method

Case study area and data collection. Bangladesh is a riverine country with a total area of 14.4 million hectares. The northern part of Bangladesh, particularly Kurigram, is a high-altitude area with non-tidal rivers. It is noteworthy that many rivers, such as the Tista and Brahmaputra, flow into Bangladesh from India, and the Kurigram district is located in the northern region of Bangladesh near the Indian border. Heavy rainfall in India often causes floods in Kurigram. Most rivers in the northern region of Bangladesh are connected to those in West Bengal and Assam states of India. Flood data for Bangladesh was obtained from the Bangladesh Water Development Board, while Indian data was sourced from Kaggle.com. The variables used in this study are detailed in Table 1.

### 1. Dataset Description

Table 1. Name of the variables used in this analysis with a short description

Variable Name	Description of Variable
Date	Date to collect data from different stations.
Water Level	Measured the water level of different stations of the rivers.
Rainfall of Bangladesh	Daily recorded rainfall by Bangladesh Meteorological Department.
Discharge	The volume of water that flows in a given time.
Temperature	Temperature of the region.
Humidity	Relative Humidity of the Study Area.
Wind speed	It indicated the speed of wind in the Kurigram district.
Cloud Coverage	The optimal coverage of cloud in the area.
Bright Sunshine	The sunshine of the river study area.
Indian Rainfall	The Assam rainfall that associates with the Bangladeshi data.
Flood	1 if the water level is greater than a threshold value, otherwise 0.

### 2. Overview of Proposed Methodology

This study aimed to forecast the flood of the rivers in the Kurigram district using Machine Learning (ML) models, as depicted in Figure 1.

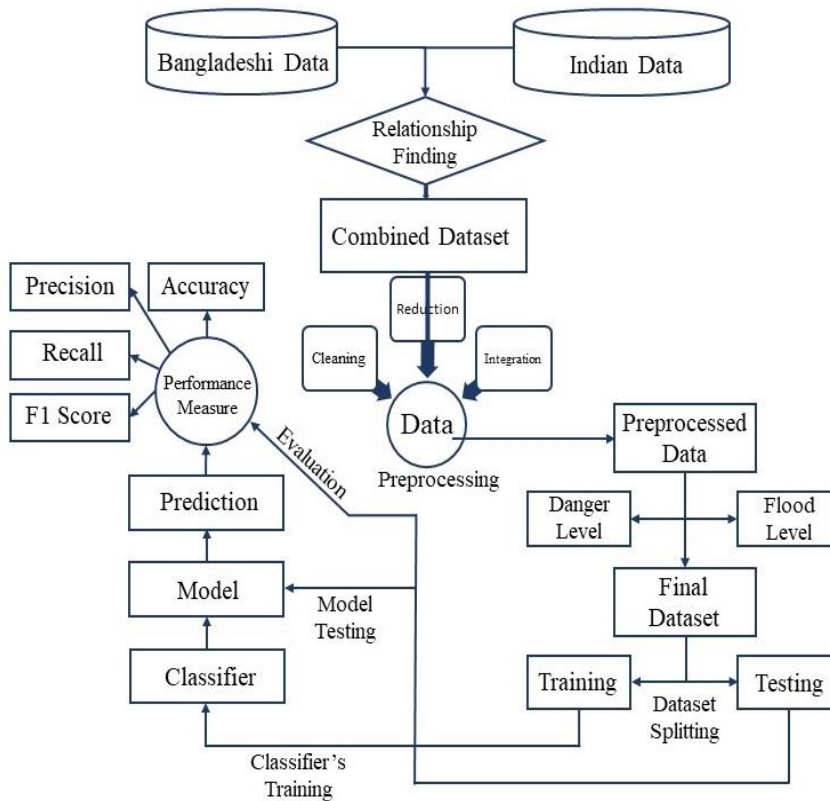


Figure 1. Proposed methodology of this study

To make the dataset more suitable for analysis, long-term pre-processing was conducted, including handling missing values, normalization of data, integration of data, and more. The correlation among variables was also calculated for a better understanding of the dataset. As per the standard practice for analysis using machine learning techniques, the dataset was divided into two parts, i) training and ii) testing data, in ratios of 80:20, 70:30, and 50:50.

3. Data Pre-processing

The necessary variables were collected from the Bangladesh Water Development Board. The main rivers in Kurigram were selected for analysis. After collecting the data, a relationship with the Indian rainfall was established and the dataset was pre-processed to make it suitable for ML training.

4. Description of the ML algorithms

Traditional machine learning algorithms including Deep learning algorithms were used to predict the flood as well as the danger level of water to the rivers of Kurigram District.

5. Performance Measure techniques

To measure the performance of the algorithms, we use evaluation techniques that are called classification reports.

- a. Accuracy: To quantify how well a system performs, we look at how many examples of data were correctly categorized as a percentage of all occurrences of data. In technical words, accuracy is defined as the ratio of true positives (TP) and true negatives (TN) to false positives (FP) and false negatives (FN) that are represented in equation (xiii).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

- b. Precision: To what degree of precision do the optimistic projections hold true? The formal definition of precision for binary classification is the ratio of true positives (TP) to false positives (FP). When the goal is to reduce FP in an unbalanced dataset, precision might be used to achieve this. Equation (xiv) showed the measurement mathematically.

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

- c. Recall: The percentage of happy memories can be calculated by thinking back on all the nice times. It is also known as True Positive Rate (TPR) or sensitivity. The percentage of correct responses (TP) is divided by the total number of correct responses (TP and FN) is equation (xv). Recall is appropriate when the goal is to reduce FN in an unbalanced dataset.

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

- d. F-1 score: Precision and Recall are averaged symmetrically to create the F-1 score. There is no need to consider a model's intelligence if only accuracy is used as a criterion for success. The mathematical representation in equation (xvi).

$$F - 1 \text{ Score} = \frac{2PR}{P + R} \quad (8)$$

**Results and Discussion**

**Result**

The data was trained and tested into 80:20, 70:30, and 50:50 ratios to show the stability of the machine learning models. Before the machine learning result, a short tour was used on the dataset to find the correlation and statistical properties of the dataset.

Table 2. Statistical properties of the variables

Variable	Mean	Std	Min	Max
WL	19.88	0.342	19.50	21.15
Discharge	59896.09	15818.241	24143.38	115024.11
Rainfall	33.26	3.599	23.60	41.00
Max_Temp	33.05	3.474	24.90	39.80
Min_Temp	20.24	5.317	9.10	27.00
Humidity	81.12	6.127	56.00	91.00
Wind Speed	1.45	0.672	0.40	5.20
Cloud Coverage	3.55	2.142	0.30	7.20
Bright Sunshine	6.34	1.606	1.70	9.40
Rainfall Indian	2405.59	837.021	735.20	4257.00

Table 2 represents the statistical properties of the variables. The mean water level was 19.88 feet, indicating that the river had sufficient water throughout the year, which was significantly different from the other rivers in Bangladesh. The standard deviation was also 0.342, showing that the water level varied only slightly during the season and remained constant most of the time. It had a strong correlation with discharge, as depicted in the heatmap. The discharge was consistently high. The minimum discharge of the river was 24143.38 m3/day, and the maximum was 115024.11 m3/day, resulting in a mean of 59896.09 m3/day, which was considered good for a non-tidal river. Table 2 also indicated that the rainfall in India was higher than in Bangladesh, which was one of the primary causes of flooding in rivers that entered Bangladesh from India.

The heatmap in Figure 2 depicted the correlation of the variables using Pearson's correlation, a parametric correlation procedure that showed the negative and positive correlations of the variables. In the heatmap, red represented a neutral correlation, white represented a strong positive correlation, and black represented a strong negative correlation of the variables. The heatmap indicated that most of the variables were not strongly correlated with each other, and that most of the environmental variables were strongly correlated with each other. Cloud coverage had an 87% positive correlation with minimum temperature, 57% with humidity, 66% with maximum temperature, and 41% with wind speed. Additionally, the rainfall in Bangladesh had a 46% positive correlation with humidity. The last column indicates the association of the variables with the flood.

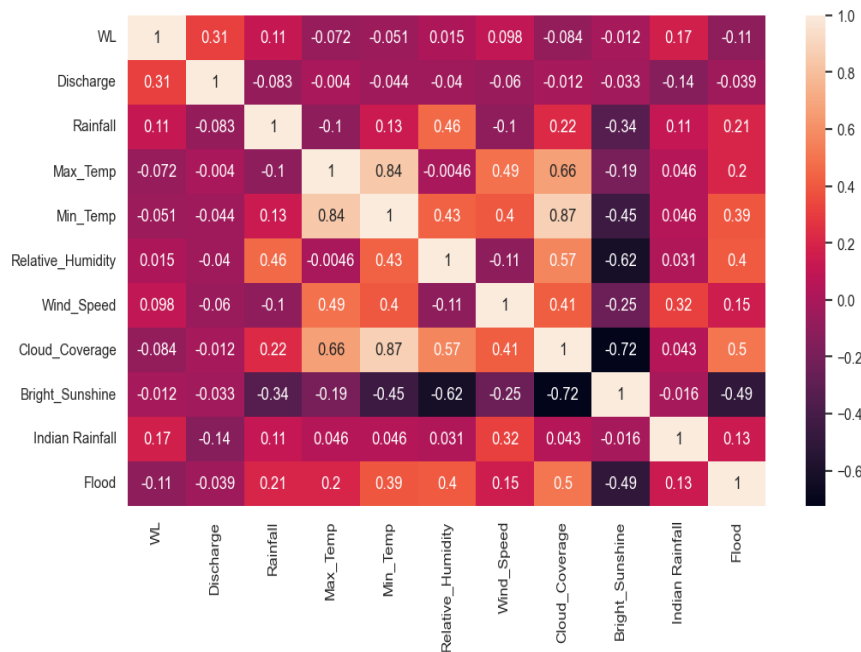


Figure 2. Correlation of the variables

1. Obtain results in an 80:20 ratio of training testing

Different machine learning algorithms were used to forecast the flood from the independent variables. The data was divided into a training and testing ratio of 80:20 to build and evaluate the model's performance. The Random Forest (RF) algorithm showed 87% accuracy, with 87% precision, recall, and f1 score, respectively. Except for the Naive Bayes (NB) algorithm, all other algorithms performed well with this ratio. The Logistic Regression (LR) algorithm also showed 87% accuracy, precision, recall, and f1 score. The Support Vector Machine (SVC) and K-Nearest Neighbor (KNN) algorithms both showed 81% accuracy, but the other performance measures fluctuated. The SVC showed 65% precision, 81% recall, and 72% f1 score, while the KNN achieved 75% precision, 81% recall, and 74% f1 score, respectively. The NB showed the lowest performance with 49% accuracy, 69% precision, 49% recall, and

54% f1 score for this ratio of data. The performance of each algorithm was visualized in Figure 3, and in all cases, LR and RF showed a superior prediction.

Table 3. Performance of the algorithm to forecast the model in 80:20 ratio

Model	Accuracy	Precision	Recall	F1 Score
SVC	0.81	0.65	0.81	0.72
LR	0.87	0.87	0.87	0.87
NB	0.49	0.69	0.49	0.54
KNN	0.81	0.75	0.81	0.74
RF	0.87	0.87	0.87	0.87

The ROC curve in figure 4 represented the classification performance of each algorithm in terms of true positive rate and false positive rate. SVC, KNN and NB showed average results, as the lines of the algorithms were close to the middle average line. However, the performance of RF and LR was superior to the three algorithms, as indicated by the lines of the two algorithms in the ROC.

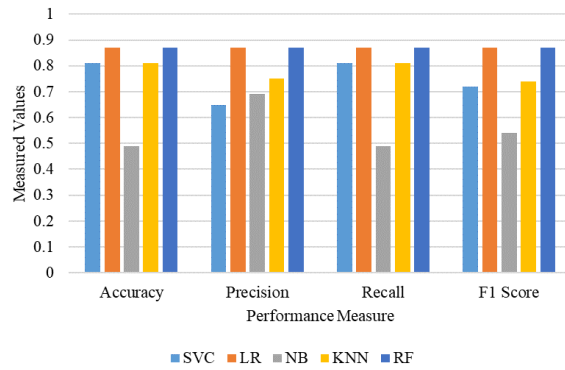


Figure 3. Performance of the ML algorithms in 80:20 ratio

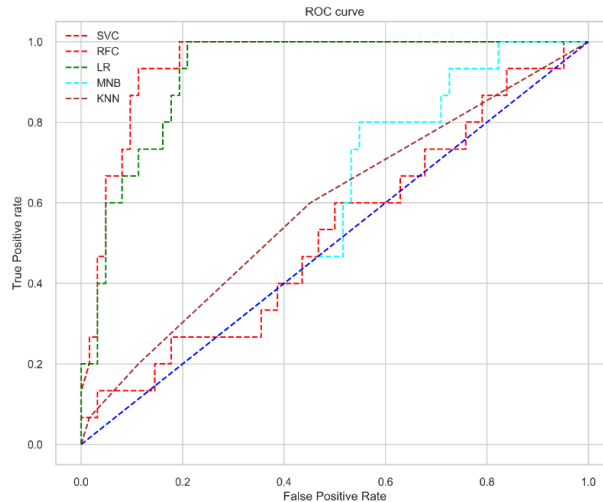


Figure 4. ROC curve of the ML algorithm in an 80:20 ratio

2. Obtain results in a 70:30 ratio of training testing.

To evaluate the performance of the algorithms, a 70:30 training and testing ratio of the dataset was used to build and test the models. The results are shown in Table 4. In this segment, RF also demonstrated the maximum 90% accuracy with 90% precision, recall, and f1 score. LR showed the second-highest result, achieving 88% accuracy with 87% precision, 88% recall, and 88% f1 score. KNN and SVC showed similar performance in this segment.

KNN had 83% accuracy with 78% precision, 83% recall, and 77% f1 score, which demonstrated the instability of this algorithm in this forecasting methodology. Similarly, SVC's performance was not stable either and showed 82% accuracy, 68% precision, 83% recall, and 75% f1 score. NB had the lowest performance in this case, with 52% accuracy and recall, 73% precision, and 58% f1 score, respectively. The performance of the models was visualized in Figure 5, which clearly indicated the superiority of the RF model in forecasting river floods in the study area.

Table 4. Performance of the algorithm to forecast the model in a 70:30 ratio

Model	Accuracy	Precision	Recall	F1 Score
SVC	0.82	0.68	0.83	0.75
LR	0.88	0.87	0.88	0.88
NB	0.52	0.73	0.52	0.58
KNN	0.83	0.78	0.83	0.77
RF	0.90	0.90	0.90	0.90

To demonstrate the classification performance of the model in terms of the true positive rate and false positive rate, a ROC curve was created in Figure 6 that showed the performance of the models in individual lines. The performance of SVC was below average, while NB and KNN were very close to average performance.

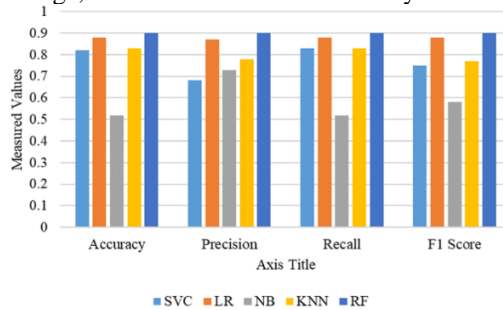


Figure 5. Performance of the ML algorithms in 70:30 ratio

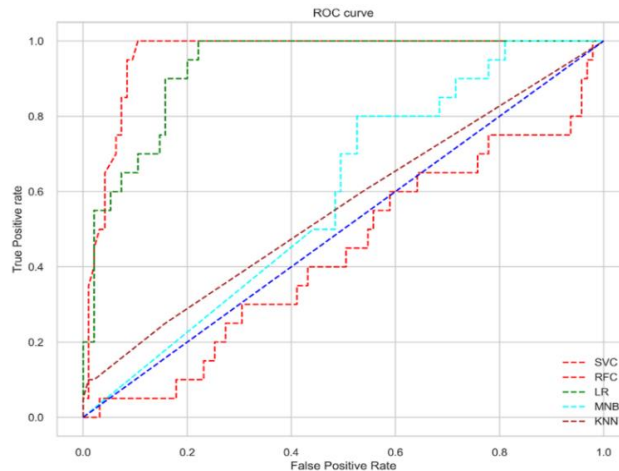


Figure 6. ROC curve of the ML algorithm in 70:30 ratio

3. Obtain result in a 50:50 ratio of training testing

Finally, the data was divided into a 50:50 ratio of training and testing, and the model was built and tested to determine the performance of the models. Similar to the previous segments, RF showed the maximum accuracy in this case as well. RF had 92% accuracy with 92% precision, recall, and f1 score.

It was proven that the performance of RF was stable in all cases. LR and SVC had 87% accuracy, but LR was superior to SVC in terms of precision. SVC had 76% precision, 87% recall, and 81% f1 score, while LR had 90%

precision, 87% recall, and 88% f1 score, respectively. Like the previous ratio, NB had the lowest result, showing 53% accuracy, 79% precision, 53% recall, and 60% f1 score. All results were tabulated in Table 5. The performance of the algorithms was visualized in a bar chart in Figure 7, where RF showed its superiority over other algorithms in this ratio of training and testing.

Table 5. Performance of the algorithm to forecast the model in a 50:50 ratio

Model	Accuracy	Precision	Recall	F1 Score
SVC	0.87	0.76	0.87	0.81
LR	0.87	0.90	0.87	0.88
NB	0.53	0.79	0.53	0.60
KNN	0.86	0.81	0.86	0.82
RF	0.92	0.92	0.92	0.92

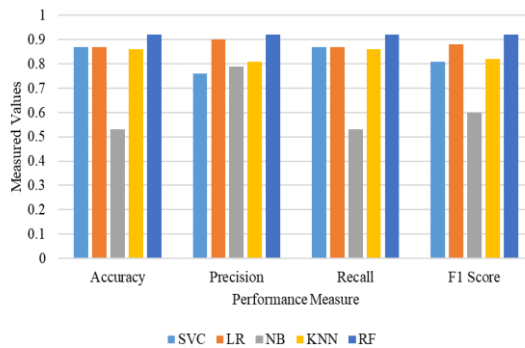


Figure 7. Performance of the ML algorithms in 50:50 ratio

This study also demonstrated the performance of the algorithm using a ROC chart that represented the true positive rate and false positive rate. Except for LR and RF, the performance of the remaining algorithms was close to the average value. Again, RF demonstrated its superiority in forecasting the Flood. The ROC curve is shown in Fig. 8.

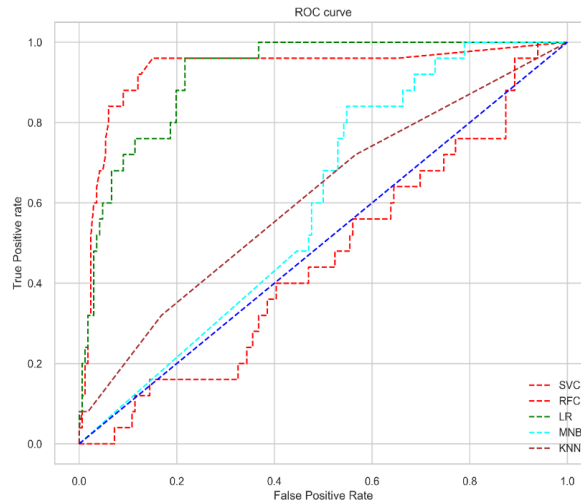


Figure 8. ROC curve of the ML algorithm in 50:50 ratio

#### 4. Forecasting phase by the best model

The developed model in this study can be predicted next 30 days outcome. It can be improved by adding more data in further study. The Figure 9 showed the possibility of flood. It indicated that in this winter season, the possibility of flood is almost zero. But the 20<sup>th</sup> day have a possibility that depends on the other independent variables. After

adding current data to the dataset, the model can predict the next 30 days outcome at 87% accurately that was already discussed in the result section.

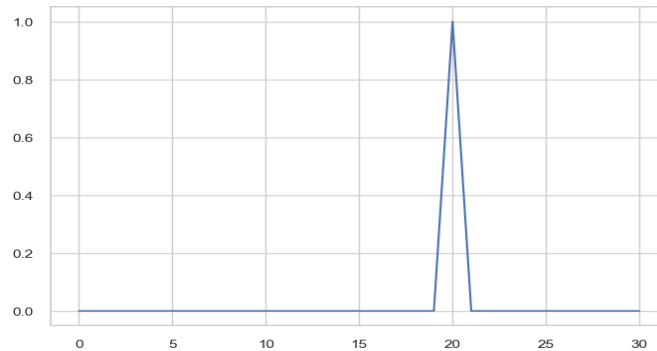


Figure 9. The forecasting of a flood of the next 30 days by the best model

## Discussion

This study aimed to predict floods in the Kurigram district using various datasets. Machine learning algorithms were employed to assess their prediction performance. The datasets were divided into ratios of 80:20, 70:30, and 50:50 for model building and algorithm performance evaluation. Significance testing was conducted to determine the algorithm's superiority. The Random Forest algorithm demonstrated superior performance and stability across all ratios, suggesting it as the best algorithm for flood forecasting in this region. The proposed methodology can forecast floods more accurately with limited data. Researchers and policymakers can utilize this methodology to proactively manage floods in specific regions through data forecasting, potentially saving lives and resources.

## Conclusions and Suggestions

### Conclusions

Floods have been identified as significant natural disasters globally. Early prediction of these events has been instrumental in minimizing damage and loss of life. This study focused on predicting floods in Bangladesh's Kurigram district by collecting and analyzing data from its main rivers. The urgency of this prediction was highlighted to aid government and authorities in making informed decisions based on flood forecasts, which are crucial for the planning of long- and short-term projects. Various machine learning models were employed in this study to forecast floods in Kurigram, with the Random Forest model demonstrating superior performance over others. The model was validated using different training and testing ratios, with the RF model consistently producing better results. This model also showed promise in forecasting other datasets in different domains. Researchers in this field can utilize this model for improved flood predictions in diverse regions and study areas. The proposed model can assist authorities in making informed decisions for both long-term infrastructure development and providing early warnings to mitigate property damage and save lives.

### Suggestions

This study only analyzed data from the rivers of the Kurigram district and used data from non-tidal rivers. There is still scope for analysis of tidal rivers in Bangladesh. In the future, the study area can be expanded to cover the entire country and consider both tidal and non-tidal rivers.

## References

Adamowski, J., Fung Chan, H., Prasher, S. O., Ozga-Zielinski, B., & Sliusarieva, A. (2012). Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban water demand forecasting in Montreal, Canada. *Water Resources Research*, 48(1). <https://doi.org/10.1029/2010WR009945>

- Alam, S., Jahan, S., & Noor, F. (2022). The surface water system, flood and water resources management of Bangladesh. *Bangladesh Geosciences and Resources Potential*, 467–546.
- Ali, M. H., Bhattacharya, B., Islam, A., Islam, G. M. T., Hossain, M. S., & Khan, A. S. (2019). Challenges for flood risk management in flood-prone Sirajganj region of Bangladesh. *Journal of Flood Risk Management*, 12(1), e12450. <https://doi.org/https://doi.org/10.1111/jfr3.12450>
- Aziz, M. A., Moniruzzaman, M., Tripathi, A., Hossain, M. I., Ahmed, S., Rahaman, K. R., Rahman, F., & Ahmed, R. (2022). Delineating flood zones upon employing synthetic aperture data for the 2020 flood in Bangladesh. *Earth Systems and Environment*, 6(3), 733–743. <https://doi.org/https://doi.org/10.1007/s41748-022-00295-0>
- Bates, P. D. (2022). Flood inundation prediction. *Annual Review of Fluid Mechanics*, 54, 287–315. <https://doi.org/https://doi.org/10.1146/annurev-fluid-030121-113138>
- Baxter, C. (2018). Bangladesh: From a nation to a state. *Routledge*.
- Becker, M., Papa, F., Karpytchev, M., Delebecque, C., Krien, Y., Khan, J. U., Ballu, V., Durand, F., Le Cozannet, G., Islam, A. K. M. S., & others. (2020). Water level changes, subsidence, and sea level rise in the Ganges Brahmaputra Meghna delta. *Proceedings of the National Academy of Sciences*, 117(4), 1867–1876. <https://doi.org/https://doi.org/10.1073/pnas.1912921117>
- Bentivoglio, R., Isufi, E., Jonkman, S. N., & Taormina, R. (2022). Deep learning methods for flood mapping: a review of existing applications and future research directions. *Hydrology and Earth System Sciences*, 26(16), 4345–4378. <https://doi.org/https://doi.org/10.5194/hess-26-4345-2022>
- Brunner, M. I., Slater, L., Tallaksen, L. M., & Clark, M. (2021). Challenges in modeling and predicting floods and droughts: A review. *Wiley Interdisciplinary Reviews: Water*, 8(3), e1520. <https://doi.org/https://doi.org/10.1002/wat2.1520>
- Chakraborty, D., Mondal, K. P., Islam, S. T., & Roy, J. (2021). 2017 flash flood in Bangladesh: Lessons learnt. *Disaster Resilience and Sustainability*, 591–610. <https://doi.org/https://doi.org/10.1016/B978-0-323-85195-4.00007-X>
- Cheng, C., Niu, W., Feng, Z., Shen, J., & Chau, K. (2015). Daily reservoir runoff forecasting method using artificial neural network based on quantum-behaved particle swarm optimization. *Water*, 7(8), 4232–4246. <https://doi.org/https://doi.org/10.3390/w7084232>
- Chinchor, N. (1992). MUC-4 evaluation metrics in Proc. of the Fourth Message Understanding Conference. *Morgan Kaufmann*. <https://doi.org/https://doi.org/10.3115/1072064.1072067>
- Dewan, T. H. (2015). Societal impacts and vulnerability to floods in Bangladesh and Nepal. *Weather and Climate Extremes*, 7, 36–42. <https://doi.org/https://doi.org/10.1016/j.wace.2014.11.001>
- Erdianto, M. A. (2023). Perancangan Model Peramalan Jangka Pendek Harga Komoditas Pertanian di Indonesia Menggunakan Machine Learning. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 3(4), 338–346.
- George, S. (2017). A third of Bangladesh under water as flood devastation widens. *Cable News Network*. Retrieved September, 22, 2017.
- Green, J., Haigh, I. D., Quinn, N., Neal, J., Wahl, T., Wood, M., Eilander, D., de Ruiter, M., Ward, P., & Camus, P. (2024). A Comprehensive Review of Coastal Compound Flooding Literature. *ArXiv Preprint ArXiv:2404.01321*. <https://doi.org/https://doi.org/10.5194/egusphere-egu24-1669>
- Guo, K., Guan, M., & Yu, D. (2021). Urban surface water flood modelling--a comprehensive review of current models and future challenges. *Hydrology and Earth System Sciences*, 25(5), 2843–2860. <https://doi.org/https://doi.org/10.5194/hess-25-2843-2021>
- Hirpa, F. A., Hopson, T. M., De Groeve, T., Brakenridge, G. R., Gebremichael, M., & Restrepo, P. J. (2013). Upstream satellite remote sensing for river discharge forecasting: Application to major rivers in South Asia. *Remote Sensing of Environment*, 131, 140–151. <https://doi.org/https://doi.org/10.1016/j.rse.2012.11.013>
- Hossain, M. S. (2024). Assessing the viability of the non-monetary flood insurance market for Bangladeshi smallholder farmers. *Natural Hazards*, 1–22. <https://doi.org/https://doi.org/10.1007/s11069-024-06454-y>
- Khan, N. S., Roy, S. K., Talukdar, S., Billah, M., Iqbal, A., Zzaman, R. U., Chowdhury, A., Mahtab, S. B., & Mallick, J. (2024). Empowering real-time flood impact assessment through the integration of machine learning and Google Earth Engine: a comprehensive approach. *Environmental Science and Pollution Research*, 1–16. <https://doi.org/https://doi.org/10.1007/s11356-024-33090-7>
- Kim, B., Sanders, B. F., Famiglietti, J. S., & Guinot, V. (2015). Urban flood modeling with porous shallow-water equations: A case study of model errors in the presence of anisotropic porosity. *Journal of Hydrology*, 523, 680–692. <https://doi.org/https://doi.org/10.1016/j.jhydrol.2015.01.059>

- Kumar, A., Mondal, S., & Lal, P. (2022). Analysing frequent extreme flood incidences in Brahmaputra basin, South Asia. *Plos One*, 17(8), e0273384. <https://doi.org/https://doi.org/10.1371/journal.pone.0273384>
- McGlade, J., Bankoff, G., Abrahams, J., Cooper-Knock, S. J., Cotecchia, F., Desanker, P., Erian, W., Gencer, E., Gibson, L., Girgin, S., & others. (2019). Global assessment report on disaster risk reduction 2019. *UN Office for Disaster Risk Reduction*.
- Mondal, M. S. H., Murayama, T., & Nishikizawa, S. (2020). Assessing the flood risk of riverine households: A case study from the right bank of the Teesta River, Bangladesh. *International Journal of Disaster Risk Reduction*, 51, 101758. <https://doi.org/https://doi.org/10.1016/j.ijdrr.2020.101758>
- Mosavi, A., Ozturk, P., & Chau, K. (2018). Flood prediction using machine learning models: Literature review. *Water*, 10(11), 1536. <https://doi.org/https://doi.org/10.3390/w10111536>
- Munawar, H. S., Hammad, A. W. A., & Waller, S. T. (2022). Remote sensing methods for flood prediction: A review. *Sensors*, 22(3), 960. <https://doi.org/https://doi.org/10.3390/s22030960>
- Palash, W., Akanda, A. S., & Islam, S. (2020). The record 2017 flood in South Asia: State of prediction and performance of a data-driven requisitely simple forecast model. *Journal of Hydrology*, 589, 125190. <https://doi.org/https://doi.org/10.1016/j.jhydrol.2020.125190>
- Price, S. (2020). Floods compound COVID-19 emergency. *Green Left Weekly*, 1276, 16.
- Puttinaovararat, S., & Horkaew, P. (2020). Flood forecasting system based on integrated big and crowdsource data by using machine learning techniques. *IEEE Access*, 8, 5885–5905. <https://doi.org/https://doi.org/10.1109/ACCESS.2019.2963819>
- Rahman, M. A., Mondal, M. N., Hannan, M. A., & Habib, K. A. (2015). Present status of fish biodiversity in Talma River at Northern Part of Bangladesh. *International Journal of Fisheries and Aquatic Studies*, 3(1), 341–348.
- Rasel, R. I., Uddin, M. N., Islam, F., & Haroon, A. (2018). Application of deep neural network for predicting river tide level. *International Conference on Innovations in Science, Engineering and Technology (ICISSET)*, 311–316. <https://doi.org/https://doi.org/10.1109/ICISSET.2018.8745593>
- Riza, O. S., & Nuryadi, A. (2023). Bibliometric Study: Rainfall Classification-Prediction using Machine Learning Methods. *Digital Zone: Jurnal Teknologi Informasi Dan Komunikasi*, 14(2), 206–218. <https://doi.org/https://doi.org/10.31849/digitalzone.v14i2.16618>
- Roy, B., Islam, A. K. M. S., Islam, G. M. T., Khan, M. J. U., Bhattacharya, B., Ali, M. H., Khan, A. S., Hossain, M. S., Sarker, G. C., & Pieu, N. M. (2019). Frequency analysis of flash floods for establishing new danger levels for the rivers in the northeast Haor region of Bangladesh. *Journal of Hydrologic Engineering*, 24(4), 5019004.
- Sarkar, S. K., Ansar, S. Bin, Ekram, K. M. M., Khan, M. H., Talukdar, S., Naikoo, M. W., Islam, A. R. T., Rahman, A., & Mosavi, A. (2022). Developing robust flood susceptibility model with small numbers of parameters in highly fertile regions of Northwest Bangladesh for sustainable flood and agriculture management. *Sustainability*, 14(7), 3982. <https://doi.org/https://doi.org/10.3390/su14073982>
- Siddique, H. (2017). South Asia floods kill 1,200 and shut 1.8 million children out of school. *The Guardian*, 31.
- Speight, L. J., Cranston, M. D., White, C. J., & Kelly, L. (2021). Operational and emerging capabilities for surface water flood forecasting. *Wiley Interdisciplinary Reviews: Water*, 8(3), e1517. <https://doi.org/https://doi.org/10.1002/wat2.1517>
- Tiggeloven, T., Couasnon, A., van Straaten, C., Muis, S., & Ward, P. J. (2021). Exploring deep learning capabilities for surge predictions in coastal areas. *Scientific Reports*, 11(1), 17224. <https://doi.org/https://doi.org/10.1038/s41598-021-96674-0>
- Umgiesser, G., Bajo, M., Ferrarin, C., Cucco, A., Lionello, P., Zanchettin, D., Papa, A., Tosoni, A., Ferla, M., Coraci, E., & others. (2020). The prediction of floods in Venice: methods, models and uncertainty. *Natural Hazards and Earth System Sciences Discussions*, 2020, 1–47. <https://doi.org/https://doi.org/10.5194/nhess-2020-361>
- Valipour, M., Banihabib, M. E., & Behbahani, S. M. R. (2012). Parameters estimate of autoregressive moving average and autoregressive integrated moving average models and compare their ability for inflow forecasting. *J Math Stat*, 8(3), 330–338. <https://doi.org/https://doi.org/10.3844/jmssp.2012.330.338>
- Wheater, H. S. (2002). Progress in and prospects for fluvial flood modelling. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 360(1796), 1409–1431. <https://doi.org/https://doi.org/10.1098/rsta.2002.1007>
- Zahura, F. T., Goodall, J. L., Sadler, J. M., Shen, Y., Morsy, M. M., & Behl, M. (2020). Training machine learning surrogate models from a high-fidelity physics-based model: Application for real-time street-scale flood prediction

in an urban coastal community. *Water Resources Research*, 56(10), e2019WR027038.

Zheng, Z., Wang, D., Gong, F., He, X., & Bai, Y. (2021). A Study on the Flux of Total Suspended Matter in the Padma River in Bangladesh Based on Remote-Sensing Data. *Water*, 13(17), 2373. <https://doi.org/https://doi.org/10.3390/w13172373>