Enhancing Flood Monitoring and Prevention Using Machine Learning and IoT Integration

By:

Syed Asad Shabbir Bukhari

(Registration No.: MS-CE-21-360454)

<u>Supervisor:</u>

Dr. Imran Shafi

DEPARTMENT OF COMPUTER& SOFTWARE ENGINEERING,

COLLEGE OF ELECTRICAL & MECHANICAL ENGINEERING,

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY,

ISLAMABAD

June 26, 2024

Annex A

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS/MPhil thesis written by NS **Syed Asad Shabbir Bukhari** Registration No. **00000360454**, of College of E&ME has been vetted by undersigned, found complete in all respects as per NUST Statutes/Regulations, is free of plagiarism, errors and mistakes and is accepted as partial fulfillment for award of MS/MPhil degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the thesis.

Signature: ______________

Name of Supervisor: Col Dr Imran Shafi

Date: 26-06-2024

Signature of HOD: ______________
(Dr Usman Qamar)
Date: 26-06-2024

Signature of Dean: Rashid
(Brig Dr Nasir Rashid)
Date: 26-06-2024

Enhancing Flood Monitoring and Prevention Using Machine Learning and IoT Integration

By:

Syed Asad Shabbir Bukhari

(Registration No.: MS-CE-21-360454)

A thesis submitted to National University of Sciences and Technology, Islamabad

in partial fulfillment of the requirements for the degree of

Master of Sciences in Computer Engineering

<u>Supervisor:</u>

Dr. Imran Shafi

DEPARTMENT OF COMPUTER& SOFTWARE ENGINEERING,

COLLEGE OF ELECTRICAL & MECHANICAL ENGINEERING,

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY,

ISLAMABAD

June 26, 2024

Dedicated to my parents, whose tremendous continuous support and endless prayers led me to this accomplishment.

Acknowledgment

First and foremost, I would like to express my gratitude to Allah Almighty, the most merciful and the kindest, for bestowing His blessings upon me and granting me the strength to complete this work.

I am deeply indebted to my supervisor, Dr. Imran Shafi, for his unwavering intellectual support and valuable ideas throughout this research journey. He has been an inspirational instructor, providing me with technical guidance and moral encouragement. His insightful feedback has honed my skills and critical thinking, enabling me to achieve my research goals. Without his continuous support, completing this dissertation would not have been possible. I am immensely grateful for the motivation he has instilled in me throughout this challenging expedition. Concluding my research study under his supervision has been a great honor.

I would also like to express my appreciation to my GC-Members, Dr.Ali Hassan and Dr. Hammad Tanveer Butt, for his valuable suggestions and guidance. Also for their assistance and direction in helping me achieve my research objectives.

I dedicate this work to my supervisor, mentor, and teacher, Dr. Imran Shafi, and my beloved parents, Syed Shabbir Hussain Shah and Mrs. Shabbir, who have supported and stood by me unwaveringly throughout this journey. Without their love and encouragement, I would not have been able to accomplish this work. To my dear"Ami and Abu," thank you for everything. I also extend my gratitude to my sisters and brothers for their support.

Abstract

Floods require good monitoring and preventive measures; they pose risk factors to humankind, property, and infrastructure. This thesis proposes a new approach to enhancing the integration of Internet of Things (IoT) infrastructure through utilizing Machine Learning (ML) methods for improved flood management and prevention. The system comprises three stations, namely: the water station, repeater station, and siren station. The water station is implanted with a radar sensor to monitor the water level continuously. Repeater stations actuate smooth communication and data flow by transmitting data between the stations. The siren station is anchored with a suite of environmental sensors that includes wind speed, wind direction, humidity, air pressure, atmospheric temperature, and rain gauges to present an overall view conducive to flood events.

It is the data from this gathered sensor that becomes the basis for machine learning data set development. Then, one scenario is tested to determine how predictive the suggested method would be. Rainfall is the output variable in the scenario, which holds wind speed, wind direction, humidity, atmospheric pressure, atmospheric temperature, humidity, and water level as input features. In this dataset, preprocessing techniques are applied to remove the outliers, noise level, and missing values, assured with analysis at every step of the accuracy and reliability of the input data for further research on the study.

The collected sensor data is utilized to predict flood episodes with the help of machine learning models, including 1D Convolutional Neural Networks (CNN) and Multivariate Long Short-Term Memory (LSTM) networks. The 1D-CNN models include the spatial relationships among the input characteristics in the case of the Multivariate LSTM models; it makes use of its capacity to capture the temporal dependencies in multivariate time series data. The models were evaluated through standard measures, such as Mean Square Error (MSE), and informed about their generalization and prediction accuracy.

The implications are further important for flood monitoring and prevention efforts. The combination of IoT technology and machine learning methodologies will enable the authorities to preempt better and prevent the incidence of floods. The combined approach of environmental and radar sensors offers the most comprehensive approach to flood monitoring, designed to consider both meteorological and hydrological parameters. Future lines of research can include the development of flood management and catastrophe response decision support systems, the investigation of highly advanced machine learning algorithms, and the incorporation of increased sensor data.

Key Words: IoT, Machine Learning, Flood, LSTM Model, CNN Model

Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Motivations

Natural disasters often lead to long-term disturbances engendered by different components of the socio-economic system. For instance, a single large-scale catastrophic event, such as a flood, can cause tremendous damage to complex infrastructural networks, leading to long-term failures and profound socio-economic effects that undermine development in general. Flash floods or rapid increases in river water levels are tremendous hazards, especially to human lives in mountain regions, where they cause considerable deaths. Mountainous flash floods are different from a regular flood in that they strike suddenly, and very limited early warnings are offered. Such events continue to increase due to the ranges of the combined effect of climatic change, Glacier Bursting, and human activities, with mountainous flash floods proving more destruction than their urban counterparts. Intense rainfall and Glacier Bursting are often associated with thunderstorms or the passage of a typhoon and have the potential to increase river flow and stage in mountainous regions dramatically. This makes for their rapid response to high rainfall rates, with the unique characteristics of mountainous watersheds being steep slopes and quasicircular morphologies. In addition, they are particularly prone to extreme rain in either the total volume or intensity. This means that under mountainous terrains, peak flow and water levels in rivers may reach their maximum within hours, making it unpredictable; hence, very little or no warning for effective prevention of damage from flooding is granted.

It was estimated that by the end of 2020, the Internet of Things (IoT) technology would connect approximately 50 billion devices [5]. The effect of IoT is varied and widespread: industry, agriculture, health, automobiles, etc. This technology automated not just one's home but many other things, too. The various applications allowed by IoT have considerably sped up the transfer of data and information [6]. IoT is critical in bringing together various intelligent applications and devices in real-life scenarios. Examples include the innovative IoT flood monitoring system, which plays an essential role and serves as assistance in the proper management of flood situations by governments and societies. At last, this reduces disasters' impacts on victims[7][8].

Floods are recognized as one of the most damaging natural disasters, creating massive destruction in cities, agriculture, and infrastructure globally. With increasing climate change, which intensifies occurrences and vulnerabilities of extreme weather conditions, monitoring and control systems would now require more intelligence. Traditional flood monitoring systems have a delay both in flood detection and response time because they often involve manually collected data and smaller sensor networks. New technologies such as machine learning and the Internet of Things offer plausible means to improve flood monitoring and forecast capacities[9].

The ML algorithms, combined with the currently available IoT sensors, have created a significant opportunity for flood monitoring and prevention capacities. Machine learning approaches include predictive modeling and data analytics, which help in the examination of considerable datasets to detect patterns and trends of occurrences of floods [7]. ML algorithms can enhance the prediction of crises such as floods by utilizing historical information about floods, meteorological forecasts, and sensor readings produced by IoT devices to achieve accurate forecasting. Internet-of-things devices, such as water level monitors, rainfall gauges, and weather stations, provide continuous surveillance of environmental variables to enable early detection of potential flood hazards [11]. Combining the two technologies of Machine Learning and the Internet of Things provides decision-makers with practical information to take required steps beforehand and reduce adversity from flooding effects on communities and structures.

Floods are also considered an essential problem, further exacerbated by rapidly increasing urbanization and demographic growth. The denser communities' setup in flood-prone areas makes the region increasingly vulnerable to flood-related disasters. Rapid urbanization leads to encroachments over floodplains in most cases; this circumstance, together with natural drainage deterioration in the areas, enhances vulnerability to floods. In the case of dense urban areas, the impacts of flooding go beyond mere property losses and include transport, public amenities, as well as disruptions to health services. As a result, there is an urgent need for innovative ways to apply advanced technologies, such as machine learning and the Internet of Things, to build resilient cities against floods altogether while focusing on disaster preparedness. Embedded into the planning and building process for cities, such technologies enable urban planners to implement measures that lower the probability of such floods and protect the welfare of the people living in them [12].

1.2 Objectives

One of the major concerns of this study is to develop an integrated infrastructure of IoT for monitoring and preventing flooding as efficiently as possible. First, to avoid irregular flooding through smooth data collection and communication, a robust Internet of Things network with water stations, repeater stations, and siren stations is currently being constructed. In addition to the use of radar sensors in water stations to monitor river and other water body levels continuously, there will be early warning of possible occurrence of floods. Incorporation of environmental sensors, including humidity, air pressure, temperature, wind direction, wind speed, and rain gauges in siren stations, will collect information on meteorological parameters that are prone to flood disasters. The main goal is to construct machine models, for instance, 1D Convolutional Neural Networks and Multivariate Long Short-Term Memory Networks, used to evaluate sensor data in making accurate predictions on flooding occurrences.

This will be done by testing the models for performance and generalization capability with such metrics as Mean Square Error (MSE). Further study on optimizing the data pretreatment methods will also be carried out to ensure that the integrity and reliability of the dataset are safeguarded during machine learning. A comparison and contrast of various machine learning methodologies based on their effectiveness in predicting flood disasters, considering the contextual elements, will be highlighted. Finally, this research will look into how IoT integration impacts flood monitoring, considering flood management decision support systems and identifying future research avenues based on holistic solutions for flood management. This research aims at studying the socio-technical aspects of a flood monitoring approach for flood mitigation, over and above the development of machine learning models and an integrated IOT infrastructure.

Community participation, stakeholder collaboration, and public awareness campaigns play an important role in contributing toward the effectiveness of strategies in controlling floods. This research will look into the perceptions, attitudes, and behaviors toward flood-monitoring technologies based on the Internet of Things by different stakeholders: government agencies, emergency responders, and grassroots citizens. The study will apply social science research to develop effective strategies through which the acceptance and implementation of IoT technology in the context of flooding areas can be encouraged. The ultimate goal is to enhance resilience and readiness in communities facing disasters.

1.3 Scope and Organization

One of the chief objectives is to design a well-integrated IoT infrastructure to enable effective flood monitoring and prevention. An elaborate Internet of Things network level, consisting of water stations, repeater stations, and siren stations, will be created for smooth data gathering and communication. Due to the radar sensors attached to the infrastructure at the water stations, constant monitoring of the levels of rivers and other bodies of water shall be realized, so early warnings on possible occurrences of flood can be reported. Environmental sensors such as humidity, air pressure, temperature, wind direction, wind speed, and rain gauges shall also be installed at siren stations to gather meteorological parameters that favor flood disasters. The main goal is to create machine learning models, like 1D Convolutional Neural Networks (CNN) This research aims at developing and executing an integrated flood monitoring and prevention system by using machine learning and Internet of Things technologies.

Most importantly, it facilitates data collection and communication by developing and installing end-to-end Internet of Things, including water stations, repeater stations, and siren stations. The water stations are installed with radar sensors to monitor water levels in rivers and other water sources. In contrast, the siren stations have environmental sensors to gather data on weather conditions that may lead to flooding. The study shall build and test machine learning models for flood forecasts based on sensor data, 1D-CNN, and Multivariate Long Short-Term Memory networks. The optimization techniques will ensure dependability and accuracy for machine learning of the dataset. This study will take a structured format with chapters, including methodology, results and analysis, discussion, and conclusion. The framework to be followed is systematic. Each chapter would contribute toward attaining the research objectives mentioned above and yield a holistic understanding of ways to improve monitoring and preventive strategies for floods using machine learning and IoT.

1.4 Significance of Research

This discovery has the potential to revolutionize flood monitoring and prevention procedures, which may significantly reduce the adverse effects of floods on infrastructure, people, and the environment [13]. Therefore, the objective of the current study is to apply machine learning and IoT technology to existing flood monitoring systems for improved accuracy, timeliness, and effectiveness of flood prediction and early warning systems. This would permit officials to preplan, predict, and directly mitigate the effects of flood calamities, along with strongly affecting disaster preparedness and response activities. Moreover, the new approach proposed in this study could also be adequately extrapolated and tailored to different geographical localities and climatic conditions, thus being versatile and very flexible in addressing the universal problem of flood management. Ultimately, the findings from this work can be put to use in safeguarding vital infrastructure, saving human lives, and protecting the sustainability of livelihoods. The implications for these results are dynamic, with dramatic improvements for communities worldwide.

1.5 Summary

This is an imperative piece of research since it introduces new ways to monitor and also manages floods using both machine learning and IoT technology. Such works have a significant potential impact on lessening the effects of floods on the structures, people, and the environment. It would be possible to better the accuracy and efficiency of flood forecasting and early warning systems. The study proposes innovative ways to enhance the prediction, planning, and prevention of flood events. With these concepts put in place, authorities can predict, prepare, and reduce the consequences of a flood event better. In addition, the scalability of the capability of the proposed solutions makes them suitable for application in other environmental situations and geographic locales, enhancing their reach into global flood control interventions.

Chapter 2: Related Work

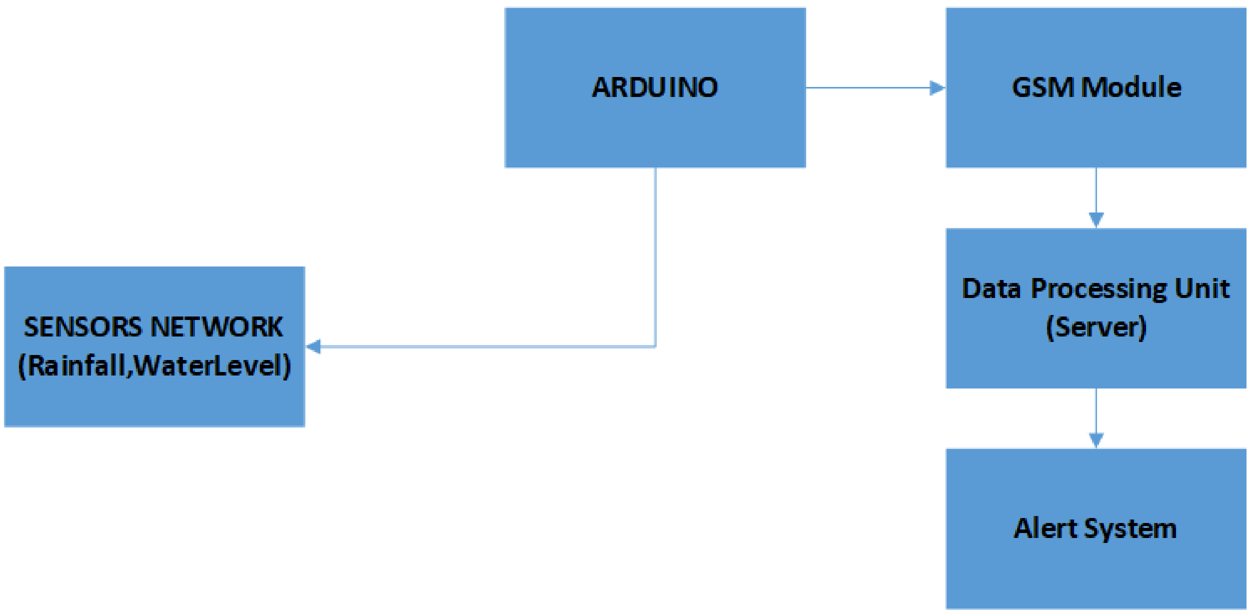
This section briefly looks at some of the critical scholarly research studies that are directly related to our project. D'Addabo et al. used Bayesian network methodology for flood detection by developing a model with historical datasets rather than real-time measurements. LiDAR technology was used in data collection in their research. Wu and Wang carried out research on the development of network sensors for implementing a portable luminescence flood detection system. Sensors monitored roads to send alerts to drivers in case of detected flooding [15]. Other sensor and machine learning works include using ML techniques for flood detection. In this study, a sensor network was employed to establish water levels and give alerts via SMS in incidences of flooding. The authors randomly selected ML algorithms for showing the applicability of the use of architectural time series in ML algorithms [16].

Numerous techniques can be used for flood risk assessment and management. In order to estimate flood risk, any method must first identify and assess potential risks and vulnerabilities [13]. This entails assessing any potential weak points or vulnerabilities in the neighborhood or community as well as estimating the likelihood that a specific area will experience flooding and estimating the likely intensity of a flood event [38]. The next step is to assess every potential consequence of a flood, including potential property damage, impacts on important infrastructure, like roads and bridges, and potential fatalities [14]. After assessing the risks, vulnerabilities, and potential outcomes, the next step is to create a risk management plan[15].

To lessen the risk of flooding, this may entail constructing levees or flood walls, moving vital infrastructure out of flood-prone areas[16] or creating early warning systems[17]. In the event of flooding, the risk management plan should address recovery and emergency response measures. Evacuating impacted areas, providing emergency shelter and supplies, and initiating a recovery effort to restore infrastructure and essential services are examples of possible actions. Finally, the risk analysis methodology should incorporate provisions for ongoing monitoring and evaluation in order to ensure that the risk management plan is up to date and functional [18].

Hydrologic modeling is an essential part of managing and assessing the risk of flooding. Hydrologic modeling is a computational and mathematical process that analyzes the water flow in a river or stream system [19].Gathering data is required for topographical features such as land usage and elevation, as well as features like soil moisture, evaporation, and precipitation. The data is then used to model the water flow rate, volume, and timing at various points across the river or stream system[20].Hydrologic modeling provides data for hydraulic modeling, which simulates the behavior of water during a flood event [21].

This involves evaluating the flow of water through a system of rivers or streams as well as the landscape around them, considering parameters such as channel geometry, roughness, and water velocity. Hydraulic modeling can provide more information into the depth and scope of flooding in each area, as well as the location and timing of peak flows[22].



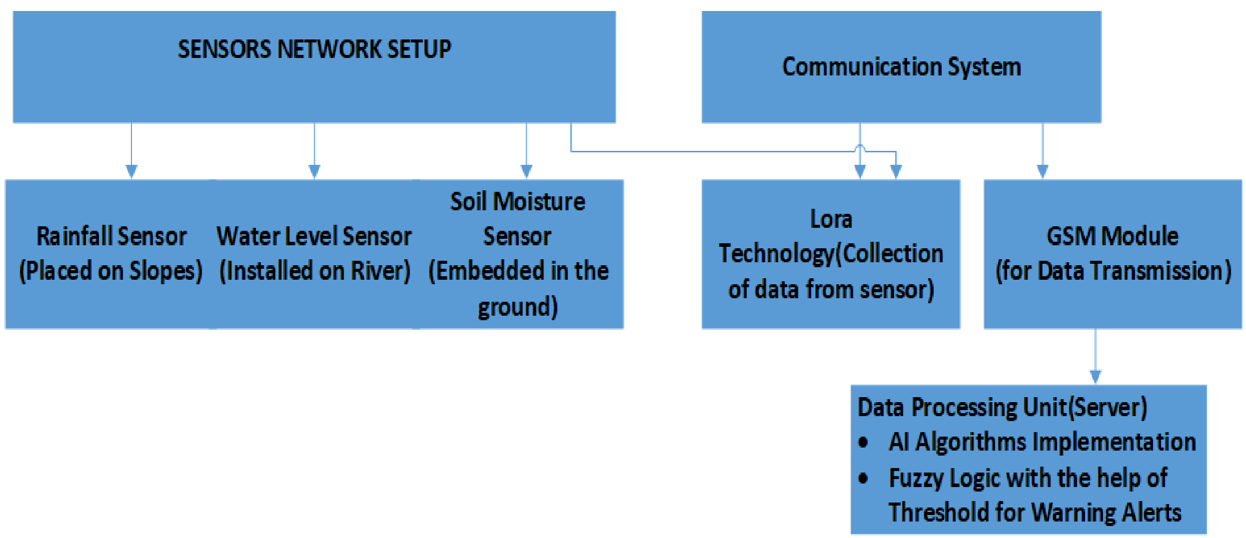
2.1 Real-Time Early Warning System Design

Hydrologic Modeling is necessary for both comprehending flood risk and developing effective mitigation strategies[23].It is possible to identify areas that are very vulnerable to flooding as well as the potential consequences of flooding, such as damage to infrastructure and buildings and fatalities, by monitoring the flow of water during a flood event[24].

Strategies for reducing flood risk, both structural and non-structural, are created using this data, and emergency response planning is also influenced[25].It is important to remember that the quality and accessibility of the data, together with the assumptions and limitations of the models used, all have an impact on how accurate hydrologic modeling is[26].

Computerized simulations of flood behavior that utilize computational methods and mathematical skills are known as numerical flood models[33].These models often employ numerical methods to solve mathematical equations describing the flow of water in a river or stream while taking into account variables like precipitation, runoff, channel geometry, and roughness of the riverbank. Numerical flood models can be used to simulate the effects of various flood scenarios and assess the efficacy of recommended flood mitigation strategies [34]. Additionally, they are employed to simulate how flooded areas will behave in response to modifications in the climate, land use, and other factors.

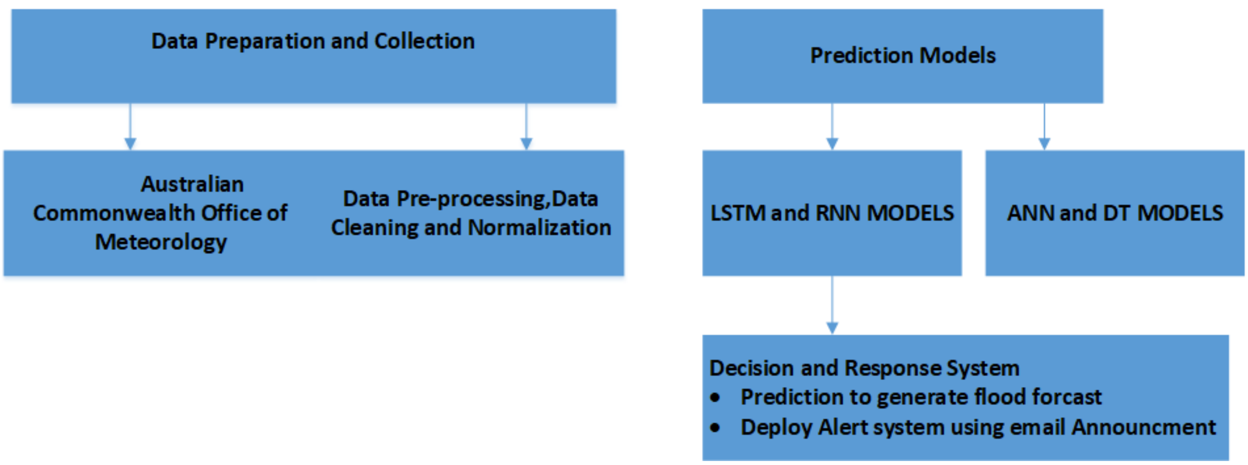
However, deep learning-based models such as Long Short-term Memory have been optimized using graphical processing units and high optimization techniques compared to artificial neural networks with computational constraints on both the processing units and a number of layers. LSTM can capture time series and memorize long-term association with the inclusions of the forget gate, making it useful for sophisticated, longtime lag applications. The LSTM model has been used by Kratzert et al. in predicting daily stream flow and by Hu et al. for predicting hourly stream flow. Widiasari et al. have applied an LSTM model for river water level forecasts in the Semarang region. Mousavi et al. have proposed an IoT-based flood early detection system, and various ML and DL algorithms are used for monitoring.



2.2 Early warning of impending flash food based on AIoT

To predict water tables in agricultural zones, Zhang et al. [21] adopted LSTM; meanwhile, a model for the hourly stream flow prediction approach was presented by Xiang et al. [22] based on LSTM with the seq2seq structure since Damavandi et al. [23] proposed an approach to stream flow forecasting using supplemental information from digital elevation models that has been appended to historical observed data with LSTM layers. Dong et al. [24] introduce a dynamic sliding window mechanism with LSTM for urban flooding forecasting. Won et al. [25] Introduce an urban flooding forecasting model by employing ANN, LSTM, biLSTM, and StackLSTM, where biLSTM displayed superiority in forecasting high water-level locations.

Other researchers developed a model to predict floods in the Red River of the North using various ML and DL models in the water level forecasting model. Kunverji et al. developed models with high accuracy for flood prediction by developing ML algorithms, including Decision Trees, Random Forests, and Gradient Boosts [26].

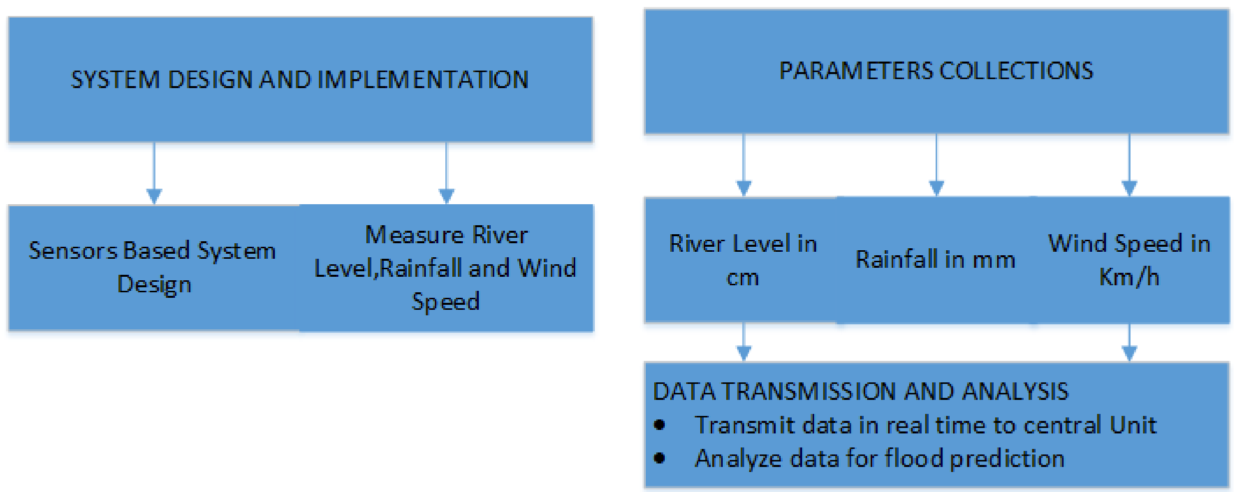


2.3 An Intelligent Early Flood Forecasting and Prediction

The model for urban flood prediction was developed by Chen et al. [27] based on LSTM and a numerical model, presenting good predictive accuracy and rapid detection for a daily flood event with a fast response time. Over the years, combining machine learning (ML) and deep learning (DL) techniques has primarily revolutionized the earlier detection and prediction systems, providing greater accuracy and efficiency.

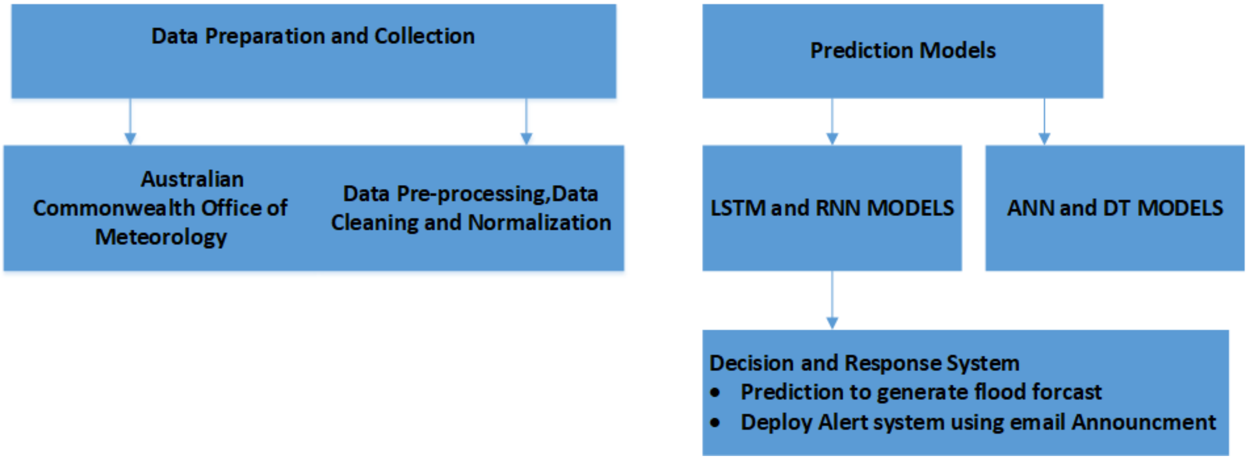
One such model that proves successful in capturing temporal dependencies and long-run patterns in time series data is the Long Short-term Memory (LSTM) model, belonging to the group of

Recurrent Neural Networks. Kratzert et al. and Hu et al. for this critical reason, different implemented LSTM models have been used in various studies related to hydrological forecasting. Estimation of the level of water in rivers is also carried out within the area of Semarang by Widiasari et al. [30] using an LSTM-based procedure. The methodology proved validated in showing high reliability through forecasting floods. These studies will further build the accuracy and timeliness of flood prediction to make early warning systems stronger. Moreover, Internet of Things (IoT) technology has dramatically improved systems for monitoring and early detection of flooding using deep learning models.



2.4 A Multi-Modal Wireless Sensor System for River Monitoring

Mousavi et al. put forth an early warning system for floods that is based on IoT-enabled sensors to deploy various ML and DL methods applied for the continuous monitoring of locations prone to flooding. Therefore, this system, based on Internet of Things (IoT) gadgets that include sensors, collects real-time information regarding the water level, rainfall, and environmental conditions. For instance, this supports the early detection of floods, where timely warnings are issued to relevant stakeholders. Zhang et al.[21] has developed LSTM models for water table forecasting in agricultural areas, hence continuing to show the potential of using Deep Learning techniques with IoT sensor data in fine agriculture and water resource management. Xiang et al.[15] developed an LSTM and seq2seq-based model to forecast hourly stream flow. They focused on the potential of IoT-enabled DL models in improving the accuracy of flood forecasting at highly short time scales.



2.5 An Intelligent Early Flood Forecasting and Prediction Leveraging Machine and Deep Learning Algorithms

Similarly, researchers have also worked on integrating machine learning algorithms with numerical models for better accuracy and speed in predicting urban flooding. Wang et al. [31] proposed a new urban flooding prediction model that combined LSTM with a numerical model, showing accurate predictions in significantly reduced times, where daily floods can be rapidly discovered with rapid response. By integrating machine learning approaches with numerical simulations, these models can better understand complex urban hydrologic processes and actionable insights for flood prevention and disaster response. These studies highlight the necessity of interdisciplinary approaches in research on flood prediction through the capitalization of experience from hydrology, meteorology, and computer science to develop robust and reliable flood forecasting systems.

Moreover, the use of ensemble learning approaches has been prevalent among flood prediction studies, which are aimed at providing more enhanced predictive accuracy and resilience. The technique combines weak learners, such as decision trees, neural networks, and support vector machines, to bring up a much stronger predictive model. Won et al. [27] developed an urban flood forecasting model using an ensemble learning technique composed of an Artificial Neural Network (ANN), a Long Short Term Memory (LSTM), biLSTM, and StackLSTM. Here, the high water levels characteristic of the urban flood event are well represented, implying that the complex dynamics involved were well captured by the performance of the ensemble approach model in surpassing the individual models. Similarly, [26] developed machine learning algorithms that included Decision Trees, Random Forests, and Gradient Boosts in forecasting floods. They achieved high classification accuracy by using a variety of different learners. This provides a feasible way to develop better ways in which the system of flood prediction can be enhanced since ensemble learning techniques combine the capabilities of several models to overcome the constraints of individual algorithms and improve the performance of the collective forecast.

Several studies have been carried out on the utilization of GIS data with remote sensing techniques in the field of flood prediction and monitoring using classical ML and DL methods. Damavandi et al. [23] proposed a method to predict stream flow by adding digital terrain model (DTM) data to past observed data. Spatial information can be included in this model to enhance the level of prediction accordingly. GIS-equipped DL models will effectively delineate spatial diversity of environmental elements and terrain features that would further help improve the prediction accuracy in complex environments. Dong et al.[32] applied a dynamic sliding window method using LSTM to predict urban floods. This methodology further uses satellite imagery and GIS data to capture the spatiotemporal patterns of flood occurrences. Geographic Information System (GIS) combined with remote sensing data may use machine and deep learning techniques to have better flood forecasting skills potentially. Integration like this allows for more precision in early warnings, which would reduce the adverse effects on populations and infrastructure.

Chapter 3: Material and Methods

3.1 Description of Study Area

The study was conducted in Arkari Chitral, part of the Chitral District in the Khyber Pakhtunkhwa province, Pakistan. It lies in the northwestern part of Pakistan and is surrounded by the magnificent Hindu Kush mountain range. The area is characterized by rugged topography, steep hills, and tremendous scenic beauty [33].

The Arkari Glacier, located near Arkari Chitral, supplies the Chitral River. The temperatures are getting warmer and warmer, a factor that is quickening the process of glacial melting, known as glacier bursting; this is a massive threat in the area. Huge masses of water, ice, and debris might break off from the glacier suddenly, and within no time, the flow is upon the valley floor as concentrated debris flows and flash floods downstream [34].This region is characterized by high susceptibility to flash floods and glacier breaches, with narrow valleys choked by dense forests and alpine meadows. The meandering Chitral River across the landscape originates from the glaciers of the Hindu Kush Mountains, defining all the hydrological characteristics.

The sites for data collection were randomly selected throughout the Arkari Chitral to represent an array of environmental conditions and flood-prone areas. The basis of site selection was accessibility, elevation above sea level, proximity to water bodies, historical flood data, and the possibility of glacier breakage events. The research area shows apparent variations seasonally. Summers are hot, with on-and-off rains attributed to the monsoon, while the winters are cold, further marked by snowfall [36]. Therefore, effectiveness in monitoring and managing floods in Arkari Chitral lies in an understanding of the complex linkages between many other factors within the environment, such as glacier dynamics, patterns of precipitation, and changes in land structures. For example, this research will improve the already developed techniques to mitigate the impacts of floods and glacier outbursts on the local population and infrastructure in Arkari Chitral. This is because the study involves an analysis of the hydrological and climatic parameters that act as precursors for the events.

It is essential to mention some primary socio-economic conditions and livelihoods adopted by the indigenous people inhabiting Arkari Chitral. The region has several small settlements and villages whose single sustenance is basically from agricultural and animal activities. This further poses a significant threat to the lives and livelihoods of these communities. Hence glacier busting and flash flooding resulting in the economic activities of the communities living in the area. Additionally, the area's infrastructure is highly inadequate, and its terrain is very rugged for emergency response and people evacuation. The people in the region are greatly affected by the floods since they deter access to health services and education, thus increasing vulnerability. Arkari Chitral is also unique in its own cultural and ecological nature. The mountains are home to a wide variety of diversified plant and animal species, adapted to thrive in the tough alpine conditions.

The Chitral River system and its tributaries represent significant habitats for a great variety of fish and wildlife, which underpin the overall ecological diversity of the region. Furthermore, Arkari Chitral holds great cultural significance to the local population living in the area since antiquity. These people have many heritage and ritual sites attached to their culture. Preservation of the cultural and ecological integrity of Arkari Chitral can, therefore, only be achieved through sustainable development and improvements in the well-being of the present and future generations. Consequently, flood monitoring and prevention strategies have to be undertaken in the region, bearing in mind and incorporating socio-cultural and ecological values of the local dwellers. This would ensure they are actively involved and take ownership of programs designed to develop resilience.

3.2 Data Collection Methods

A wide-ranging data collection process was initiated at Arkari Chitral, on the river masses of Arkari River, to estimate flood risk and environmental dynamics. A strategically situated sensor for measuring the water level was present at the most critical locations along the river to know about the variations in water levels and the velocity of water flow. The data collected by the sensors were able to detect the dynamics in rivers continuously, hence assessing if there was a flood event that was about to take place. This would greatly assist the response agencies to prepare and prevent the damage caused by the flood. Another critical segment was an environmental monitoring system that was integrated, bringing in several sensors that measured different environmental variables; In addition to the sensors above for monitoring soil moisture, land cover, and topographic features, the weather stations recorded temperature, humidity, wind speed, and wind direction.

The integration of environmental monitoring system sensors and water level measurement sensors generated broad information on the hydrological and meteorological factors that cause flood events in Arkari Chitral. Such information has been of paramount importance in designing flood management systems, forecasting models, and assessment of risks occurring from floods in the area. The strategic location of sensors to measure environmental parameters and water level was critical in helping us understand the dynamics of the environment and flooding risk factors in Arkari Chitral. Changes in river flow could be seen if considered indicators of potential impending flood events by monitoring the level of water throughout the day.

The environmental monitoring system has also generated data on climatic conditions, soil moisture content, and changes in land cover that could affect flood susceptibility. High-resolution information about the collection of climate, topography, and hydrology gives a detailed understanding of complex interactions in the study area.

3.3 Instrumentation and Sensors

3.3.1 Water Level Measurement Sensor

The Rikka Company employed a measuring radar sensor. This sensor, with a measuring range of up to 10 meters, is highly accurate in monitoring water levels together with the Arkari River and its tributaries.

The measurement sensor of the water level sensor can ideally detect the changes in water levels and flow rates, thus helping determine floods before they take place.

Manufacturer and Model: Rikka Company RKL-02 10m Measuring Radar Sensor.

Deployment Sites: The river stage measuring sensor was installed at selected sites on the Arkari River and its tributaries to track variations of river flow dynamics.

Figure 3.1 Radar Sensor

3.3.2 Wind Speed Sensor

The anemometer by Rikka Company was used to measure the wind speed. The sensor is helpful in making important decisions regarding the evaluated parameters related to flood dynamics.

The wind speed sensor measures the speed of wind effectively, providing key information towards determining wind patterns and their impact on flood hazard.

The wind speed sensor is of Rikka Company; its make and model is: RKL-100-02.

Deployment sites: For the purpose of recording proper measurements and fluctuations in wind speed and direction, the wind speed sensor was definitely placed at various strategic points around the study area.

Figure 3.2 Wind Speed Sensor

3.3.3 Wind Direction Sensor

In the measurement of the wind direction in the region under research, a wind direction sensor from Rikka Company was used. This sensor has the capability of detecting the flow direction of the wind with great precision, hence improving the data super numerated by the wind speed sensor.

The wind direction sensor provides precise measurements regarding the direction of wind, which comes to evaluate meteorological variables that significantly influence flood events.

The wind direction sensor refers to the model number RKL-100-01, manufactured by Rikka Company.

Deployment sites: The anemometer was deployed with the wind speed sensor at many places around the project area.

Figure 3.3 Wind Direction Sensor

3.3.4 Rain Gauge

The study area made use of the rain gauge to measure the intensity of rainfall through the quantity of water accumulated over a unit of time; hence providing precious data for forecast and flood analysis.

The rain gauge is able; therefore, to realize the intensity of rainfall accurately and be able to quantify rain events with quite a good preciseness level that affects flood dynamics.

Manufacturer and model: Rikka Company's RK400-01 Precipitation Gauge.

Locations for deployment: A good number of rain gauges were placed at different locations, which could help record the spatial variations of rainfall intensity over the study area.

Figure 3.4 Rain Gauge Sensor

3.3.5 Atmospheric Temperature and Humidity Sensor

In the research area, Rikka Company sensors were installed to detect atmospheric pressure and humidity. The sensors are designed to collect relevant data regarding the atmospheric environment. Thus, they assist in predicting the weather conditions, including those that cause the occurrence of floods.

The sensors for pressure in the atmosphere and relative humidity give exact measurements of the atmospheric pressure and the humidity level, thus adding value to the understanding of climatic features that predispose to occurrences of floods.

Manufacturer and Model: Rikka Company RKL-330-01 Atmospheric Pressure and Humidity Sensors.

The atmospheric pressure and humidity sensors were deliberately distributed around the test site so that spatial variations in the atmospheric conditions could be captured.

Figure 3.5 Atmospheric Temperature and Humidity Sensor

3.3.6 Barometric Pressure Sensor

The measurement of air pressure differences was done using a barometric pressure sensor provided by Rikka Company in the study location. This forms essential information regarding changes in atmospheric pressure, which could be expected to affect changes in climate and, therefore, flooding patterns. This barometric pressure sensor can determine atmospheric pressure variation sensitively, hence enabling the understanding of the climatic parameters that affect flooding.

Manufacturer and Model: Rikka Company RKL-330-01 Barometric Pressure Sensor.

To capture spatial variations in barometric pressure, the barometric pressure sensor was intentionally collocated with other sensors at several sites across the study area.

Figure 3.6 Barometric Pressure Sensor

3.4 Methods

This research aims to assess the performance of machine learning algorithms in conjunction with flood prediction, as presented in this work, showing an alternative to classical forecasting methods. For this, the following set of algorithms will be employed: 1D-CNN and M-V-LSTM.

One commonly used neural network architecture that is applied in processing and interpreting sequential data, such as time series, audio signals, and text, is a 1D Convolutional Neural Network (1D CNN). Although 1D CNNs are designed for 1D sequence, standard CNNs are built primarily to learn from image data with 2D structures. The 1D CNN model is constructed for Rainfall Level Prediction using the Tensor Flow and Keras libraries to predict the amount of rainfall levels by taking environmental parameters obtained from the Arkari region. This implementation adheres to industry standards in developing and fine-tuning a 1D CNN model for a regression task.

3.4.1 Data Preparation

The target variable(y) and features (X) are separated of the dataset. A second division into training and test sets is done using the sci-kit-learn train_test_split function.'Arkari_Water_Level','Arkari_AirTC','Arkari_Humidity','Arkari_BP_mbar','Arkari_Wind_Speed', and’Arkari_Wind_Direction' are among the features. The target variable (y) and features (X) are separated out of the dataset. It is implemented through standardization by the use of StandardScaler.

3.4.2 Model's Architecture

3.4.2 1D CNN Model

Convolutional Layer: Uses the ReLU activation function in extracting features from the input sequences, to which a convolutional layer with 64 filters, each of size 3, is applied.

It is through a pool size of 2 that the feature maps are down-sampled.

Flattening Layer: In some sense, it flattens the 3D volumes into a 1D vector.

Dense Layers: Two dense layers in which the output features and target predictions are calculated using ReLU activation functions with 50 and 1 neurons, respectively.

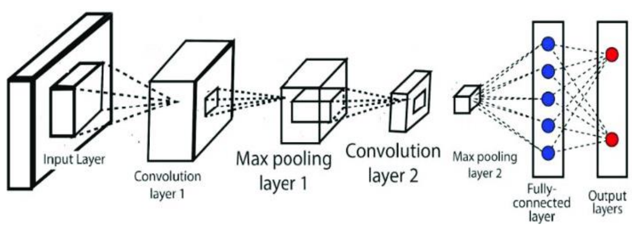


Figure 3.7 1D CNN Model Architecture

3.4.2.1 Collection and Training of Models

The model is compiled, using Adam as the optimizer and mean squared error as the loss function. Model fitting is achieved by effectively training the function with training data (X_train_reshaped, y_train) over a given batch size and number of epochs. In the training process, data for validation (X_test_reshaped, y_test) is used for monitoring model performance.

3.4.2.2 Prediction

The model predicts the feature test data's rainfall levels following training (X_test).

3.4.2.3 Evaluation

The measurement of error and generalization performance with the training is done through metrics such as mean squared error, mean absolute error and R-squared score.

3.4.3 Multivariate LSTM

LSTM provided an essential innovative addition to RNN[34, 35] and fixed some shortcomings of earlier versions. Hoch Reiter and Schmidhuber invented the LSTM in the year 1997. In LSTM, more units were added that helped it in learning very long-term dependencies and remembering it for a more extended period. The sequential shape of the model was retained, but its recurrent unit was significantly changed. As observed, LSTM differs from the usual RNNs, which are unrolled and composed of four network layers connected through a specific communication method. It has three gate modules: input, which holds the information happening right now; output, which provides the information; and forgets, which decides to keep or discard the information [18]. This current study uses multivariate LSTM model to model temporal dependencies in the multivariate time series data collected from various sensors installed across the Arkari region. The section below describes the methodology to predict future environmental parameters, including model training and development and clearly describing how the data was prepared [38].

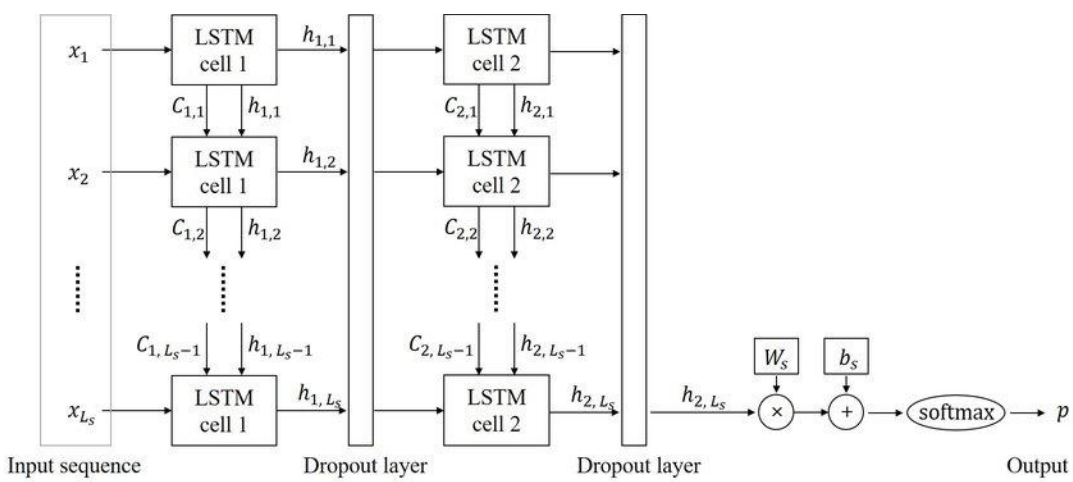


Figure 3.8 Multi-LSTM Model Architecture

3.4.3.1 Data Preparation

A multivariate time series dataset is restructured into input-output pairs to easily make the LSTM model learn past observations and future predictions. Input refers to a previous observation with fixed-length windows, and output refers to a future observation.

The dataset is further divided into separate subsets into which testing and training will be done to look for how well the model performs on untried data.

3.4.3.2 Model Architecture

The architecture of the LSTM model is designed with great attention to detail, ensuring that it manages complex information associated with multivariate time series forecasting.

Here, we use stacked Long Short-Term Memory Layers (LSTM) to capture the intricate temporal relationships present in the input sequences. Generally, an increase in memory units that can be added per layer allows the user to give more ability to the LSTM layer to remember data over long time intervals [39].

The LSTM layers are supercharged using a dense (fully connected) layer to take the features learned by the LSTM and output a prediction in the desired output space. This layer should have the same number of neurons as the number of output time steps.

Non-linearity and regression suitability are explicitly introduced at the output and LSTM layers [40], through specific activation functions.

3.4.3.3 Model Training

The optimization method, with a well-formulated loss function, plays an essential role in updating the parameters of a model in a way that minimizes the difference between its expected and observed values during the training process.

The model parameters are optimized based on the loss of mean squared error (MSE) and the famous Adam optimizer, acclaimed for its flexible learning rate capability [42].

The training procedure involves feeding data into the model iteratively in batches so that it is exposed to the model in a way that allows the extraction of very complex interactions and patterns embedded in the data.

Validation data is inserted purposely during training to prevent over fitting and enable monitoring of model performance.

Chapter 4: Dataset

4.1 Dataset Description

This dataset had 40,030 observations with eight characteristics, of which 7 were the input parameters and the other one was the target parameter, representing each row as a unique observation from their fieldwork dataset in Arkari Chitral. It has detailed information about flood dynamics in the study area, including various environmental and meteorological variables that might influence them.

	A	B	C	D	E	F	G	H
1	TimeStamp	Arkari_Water_Level	Arkari_AirTC	Arkari_Humidity	Arkari_BP_mbar	Arkari_Rain_mm_Tot	Arkari_Wind_Speed	Arkari_Wind_Direction
2	2022-04-20 11:15:00	0.379	6.59	100	744	0.4	1.232639	223.9788
3	2022-12-11 09:45:00	0.581	0.86	68.6	745.3	0	0	0
4	2021-12-30 11:00:00	0	1.38	32.7	745.3	0	1.454014	0
5	2022-04-23 04:30:00	0.36	5.81	100	742.1	0	0.730389	175.2958
6	2022-11-28 14:30:00	0.555	11.19	27.6	740.6	0	2.802194	217.4463
7	2022-10-27 20:45:00	0.586	11.2	18.3	744.5	0	1.818403	127.7851
8	2022-11-25 18:45:00	0.567	7.7	45.7	745.1	0	0	0
9	2023-07-07 12:15:00	0.657	21.45	49.8	735.9	0	7.61382	206.3639
10	2022-01-14 05:15:00	0	-6.2	63.4	740.8	0	1.652056	0
11	2023-04-09 21:45:00	0	4.48	100	747	0.6	0	0
12	2021-12-14 05:15:00	0.391	-1.32	59.3	737.8	0	0.8798195	0
13	2022-12-16 08:45:00	0.556	-0.26	40.5	743.4	0	0	0
14	2023-11-18 13:00:00	0.767	11.65	33.6	743.3	0	0.9970694	233.7885
15	2022-04-03 20:00:00	0.367	12.12	39.2	742.6	0	1.9095	175.6242
16	2022-03-13 10:30:00	0	10.42	59.5	746.3	0	1.266764	0
17	2022-03-30 09:30:00	0.354	9.85	24.5	742.8	0	1.412597	74.91104
18	2022-04-23 23:30:00	0.344	10.95	63.5	743.2	0	0.5012777	113.9104
19	2021-12-06 09:15:00	0.398	1.13	72.9	743.4	0	0.8591111	0
20	2023-07-16 13:15:00	0.68	24.92	58	737	0	1.334139	226.4623
21	2022-01-23 17:30:00	0	-0.84	84.1	731.5	0	1.426694	0
22	2022-12-14 06:00:00	0.557	-1.22	54.1	744.1	0	0	0
23	2022-05-21 12:00:00	0.363	18.3	35.8	735.3	0	2.223236	79.7707
24	2022-09-05 11:15:00	0.648	22.63	59.5	738.4	0	3.248056	219.638

Figure 4.1 Dataset

4.2 Input Features

Timestamp: Date and time stamp of each observation documented while in the field.

Water Level: obtained from the sensor of the Rikka Company installed in the study area.

Wind velocity: This is determined in the study area. The sensor used for the wind speed measurement is from Rikka Company.

Wind Direction: Directional wind data registered in the flow from the wind direction sensor made by Rikka Company.

Rainfall intensity: Data of precipitation taken with the aid of a rain gauge used in ascertaining the precipitation's event intensity.

Atmospheric Pressure: Atmospheric pressure readings were measured using the Rikka Company atmospheric pressure sensors already installed across the study area.

Humidity: Humidity level recorded on the Rikka Company sensors to give an idea of atmospheric content moisture.

Barometric Pressure: A type of atmospheric pressure variation recorded at Rikka Company using a barometric pressure sensor.

Target Attribute - Scenario 1 (Maximum Intensity of Rainfall):

Rainfall Intensity: The data extracted from the rain gauge portrays continuous measurements of rainfall intensity, which forms the target variable in this first scenario. In this first scenario, the task is to forecast rainfall intensity based on meteorological and environmental variables.

Target Feature - Scenario 2 (Water Level):

Water Level: Measurements related to water levels in the Arkari River and its network are done using a 10 m measuring radar sensor of Rikka Company. This becomes the target variable for the second scenario, where we aim to predict water levels based on meteorological and environmental variables.

4.3 Data Preparation

The data preparation procedure for this study comprised the following steps in the series:

4.3.1 Handling Missing Data

Fig below shows that there is no missing sample in 08 variables/features.

```
missing_values = data.isnull().sum()
print("Missing values in the dataset:")
print(missing_values)
```

```
Missing values in the dataset:
TimeStamp                 0
Arkari_Water_Level        0
Arkari_AirTC              0
Arkari_Humidity           0
Arkari_BP_mbar            0
Arkari_Rain_mm_Tot        0
Arkari_Wind_Speed         0
Arkari_Wind_Direction     0
```

Figure 4.2 Data Samples

4.3.2 Categorical Variables to Numeric

The shared snippet below shows only one feature, that is Timestamp as Non-Numeric Value.

```
# Check for non-numeric values in each column
non_numeric_values = data.applymap(lambda x: not isinstance(x, (int, float)))
print("Columns with non-numeric values:")
print(non_numeric_values.any())
```

```
Columns with non-numeric values:
TimeStamp                 True
Arkari_Water_Level       False
Arkari_AirTC             False
Arkari_Humidity          False
Arkari_BP_mbar           False
Arkari_Rain_mm_Tot       False
Arkari_Wind_Speed        False
Arkari_Wind_Direction    False
```

Figure 4.3 Categorical Variables to Numeric

Below Snippet shows Conversion of Timestamp Column to Date time Column.

```
# Convert timestamp column to datetime format
data['TimeStamp'] = pd.to_datetime(data['TimeStamp'])

# Print the first few rows of the DataFrame to verify the changes
print(data.head())
```

Figure 4.4 Date time Column

4.3.3 Correlation Analysis

Descriptive analysis on Variables after the data preprocessing had been conducted, the next thing was to look into the variables descriptively to get a feel of the data that was going to be analyzed. In particular, an assessment was made of the degree of correlation between the variable "Arkari_Rain_mm_Tot" and the rest of the variables. Below, the correlations are ordered by their absolute values in descending order.

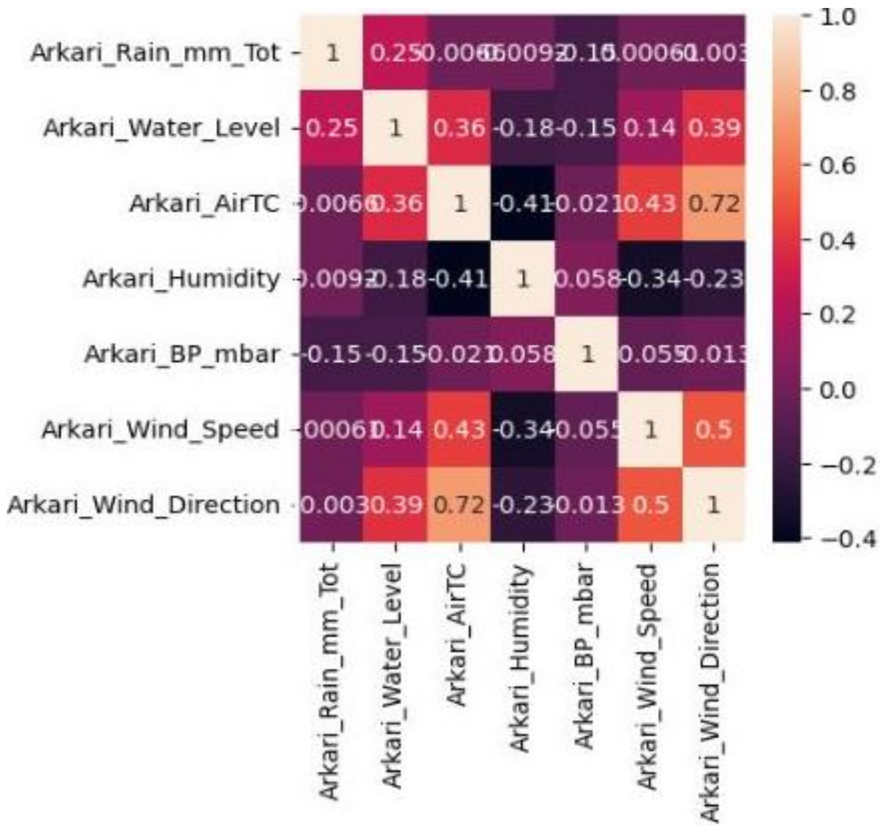


Figure 4.5 Correlation Analysis

Chapter 5: Proposed Prototype

The proposed IoT-enabled system comprises three units: a Water Station, a Repeater Station, and a Siren Station, which are all part of the monitoring and alert system.

5.1 Water Station

The Water Station harbors a Water Level Sensor: Radar-Based, a Micro-Controller, a Solar Charge Controller, and a Lora Wireless Module into the IoT Ecosystem. This water level sensor can monitor the Height of Water. The data is sent to the microcontroller, which processes the signal picked up by the sensor; it then communicates using LoRa technology. It regulates the power coming from the solar panel, which makes the station functional continuously.

Figure 5.1 Installed Radar Sensor

Figure 5.2 Water Station operating Equipment's

5.2 Repeater Station

The Repeater Station is already a complete IoT gateway, with a solar charge controller and Lora module. The solar charge controller assures the availability of power from its solar panel, while the Lora module supports the relay of data between the Water Station and Siren Station within this larger IOT framework. Data from these stations are transmitted and controlled by a Thing-speak server.

Figure 5.3 Repeater Station Operating Equipment's

5.3 Siren Station

For the design of the Siren Station, the following components are used: micro-controller, solar charge controller, GSM module, and siren sound system. The microcontroller is supplied with data from the Repeater Station, processes it to generate an alarm message, and then triggers the GSM module into dialing and sending an alert message. The siren sound system is simultaneously triggered to give an alarm. This is enhanced with a solar charge controller that will ensure continuous power from the solar panel.

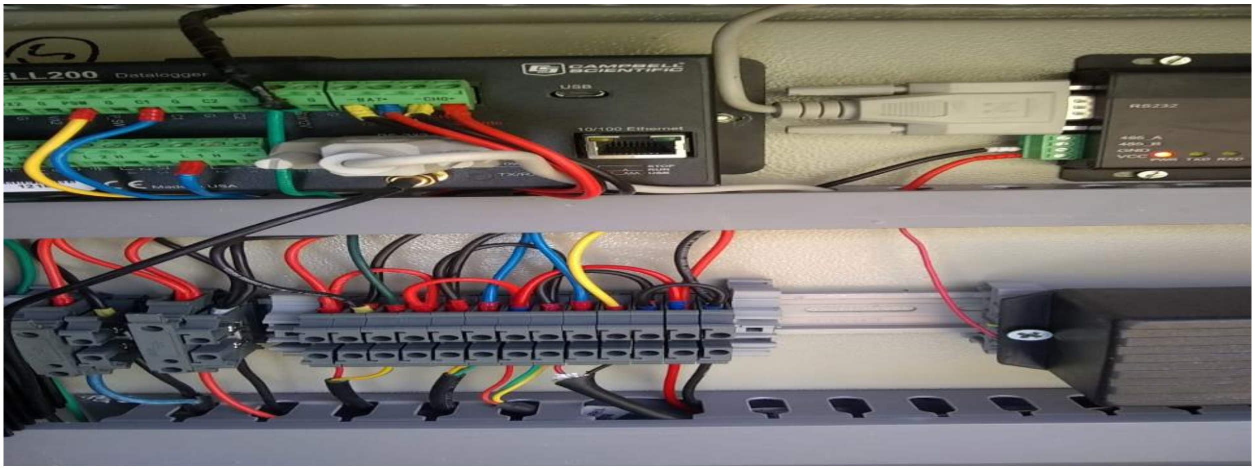


Figure 5.4 Siren Station Operating Environment

Figure 5.5 Solar System and Enclosure

5.4 Data Transmission

The Water Station sends the information about the water level and water content data from time to time to the Repeater Station with the help of the Lora module. Having an infrastructure based on the Internet of Things, the Repeater Station forwards the information to the Siren Station through the Lora network. The entire data is sent and stored in the Think Speak server for higher-level analysis within the broader frame of the Internet of Things.

This integration with Thing Speak enhances the capability of the system by offering a centralized platform for the storage and analysis of data.

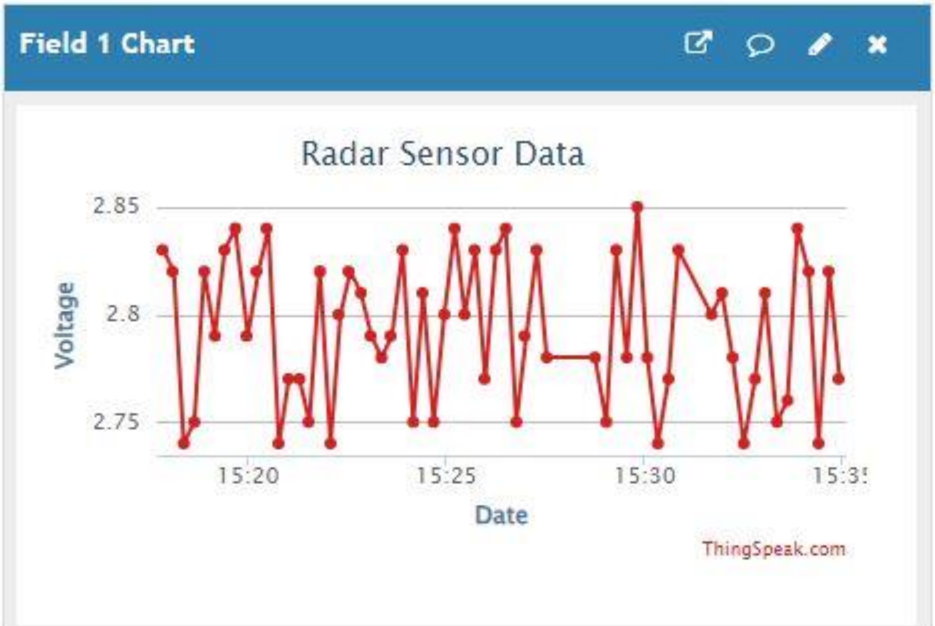


Figure 5.6 Graphical Visualization of Radar Sensor Data

Thing-speak utilize the MQTT (Message Queuing Telemetry Transport) protocol as a solid method to transport data from the attached devices to its server. MQTT, famous for its lightness and low bandwidth, supports seamless communication between IoT things and the Thing-speak platform [44]. It is very critical to ensure reliable information exchange; for instance, the Water Station, Repeater Station, and Siren Station in the proposed monitoring system send information regarding water level, content in the water, and alarm triggers, respectively, to the Thing-speak server. By using MQTT, responsiveness and reliability in data transfer are improved and hence contribute accordingly to the realization of the IoT ecosystem that is put in place within this water monitoring solution [45].

Chapter 6: Results and Discussion

The results chapter presents the work carried out for improved flood monitoring and prevention through machine learning and IoT integration. This chapter summarizes the results obtained from the analysis with sensory data and performance assessment of the models developed.

Key performance assessment metrics that were used to evaluate the machine learning models, especially 1D Convolutional Neural Networks (CNN) and Multivariate Long Short-Term Memory (LSTM) networks, are mean squared error (MSE). In this respect, the water level was the target variable for both the multivariate LSTM model and CNN. Moreover, for both the 1D CNN and the Multivariate LSTM models, the variable rain_mm_total was used as an added target variable.

6.1 1D Convolutional Neural Network (CNN) Model

6.1.1 Water Level as Goal

Performance evaluation for the 1D CNN model was done using a mean squared error loss curve. The dataset contained environmental features of the Arkari River and the goal variable, Arkari Water Level, trained over 50 epochs with a batch size of 32.

The model kept improving its predictions with the training process, thereby reflecting the respective decrease in validation and training MSE loss values in the following epochs. In epoch 1, for example, the value of validation MSE was 0.0247; however, the training loss was first observed at 0.1743. The same pattern continued upon variations in the following epochs, showing how the model adapted to the training set.

The MSE of the validation was stabilized at 0.0226 after the 50th epoch, while the final training MSE converged to 0.1722. The values of the training set and validation set MSE are close to each other, which indicates no over fitting to the training data and, hence, generalization well on new data. Also, low values of MSE add to the evidence that the model can capture the complicated correlations between features and the Arkari water level. In general, from the MSE loss curve, a 1D CNN structure model could be applied further in future water level prediction of the Arkari River together with environmental inputs; this would open insights into model training dynamics and performance.

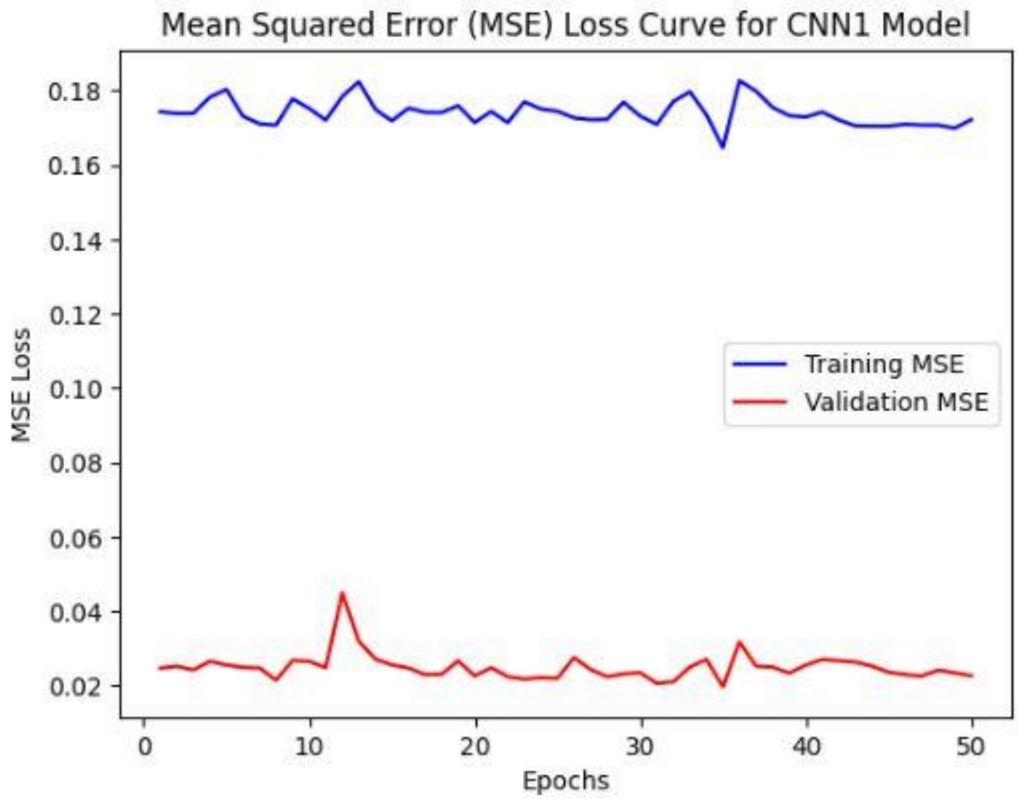


Figure 6.1 MSE Loss Curve for CNN1 Model

6.1.2 Rain_mm_total as Goal

The MSE loss curve was used to know how well the model performed for the 1D CNN. The training was done with the dataset having features regarding the environmental characteristics of Arkari River and a goal variable, Arkari Rain_mm_total, for 50 epochs and a batch size of 32.

The model improved its predictive power during training, showing a decreasing trend for the MSE loss values between validation and training in the subsequent epochs. At epoch 1, for example, the validation MSE was 0.0042, while the training MSE first showed 17.2178. This trend was maintained in the subsequent epochs, and the fluctuations indicate how the model adapts to the training set.

Here, the validation MSE becomes stable at 0.0166 after 50 epochs, and the final value of the training MSE converges at 5.5035. The obtained MSE values are relatively low, showing the model's ability to capture the subtle correlations among the input features and with the target variable, Arkari Rain_mm_total.

In short, the MSE loss curve indicates that the model of the 1D CNN is appropriate to predict Rain_mm_total in Arkari River based on environmental inputs and gives an insight into the training dynamics and performance of the model.

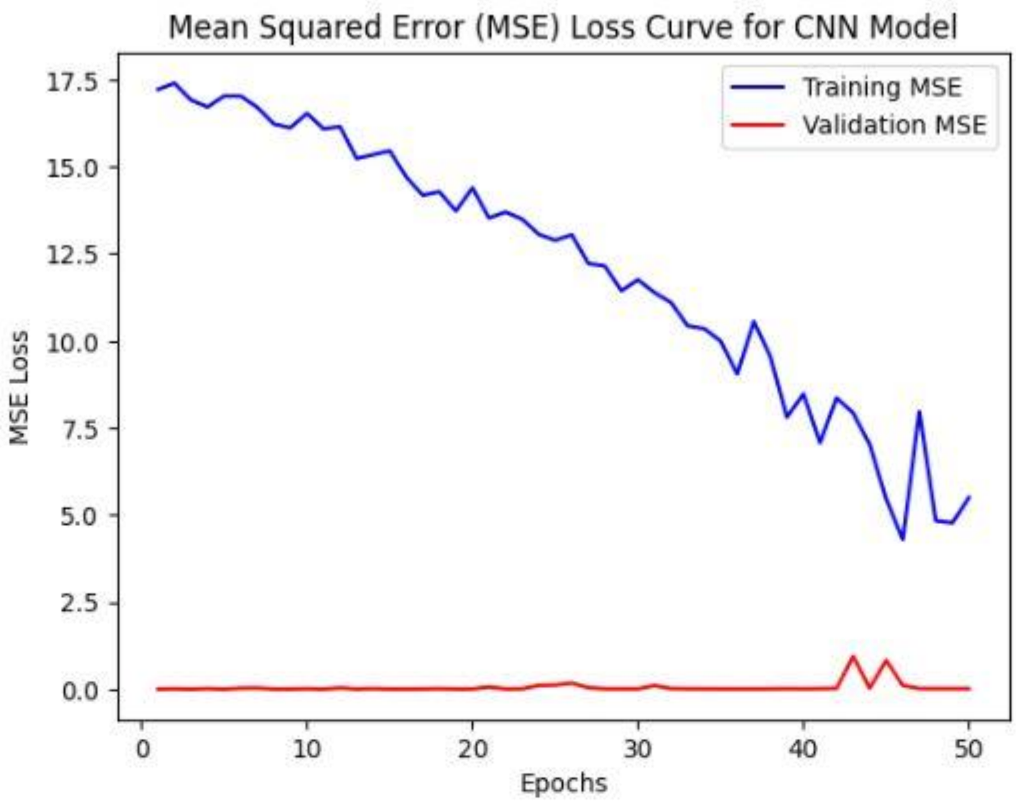


Figure 6.2 MSE Loss Curve for CNN Model

6.2 Multivariate LSTM Model

6.2.1 Water Level as Goal

These scores were evaluated using the mean-squared error (MSE) loss curve to the multivariate LSTM model. The Arkari River environmental characteristic-related feature dataset and the goal variable, which is predicting the Arkari water level, were used to train the model over 60 epochs with a batch size of 32.

For instance, it was indicated by the model improving its predictive abilities with a drop in the validation and training MSE loss values through successive or subsequent epochs in the training process. For instance, at epoch 1, the validation MSE was 6.7080e-04. In contrast, the training MSE was first observed to be 1.9338e-04, after which other epochs followed suit through these changes in variations to show how the model updated the one-pass training play.

After 60 epochs, it attained stability at 1.1029e-04 in validation MSE and reached a final convergence of 9.0664e-06 for training MSE. This further asserts the fact that due to shallow MSE values obtained, signifying that this model can capture the intricate correlations between the input features and the target variable—Arkari Rain_mm_total.

In general, the MSE loss curve shows that the multivariate LSTM model is good at making time-series predictions of Rain_mm_total in Arkari River based on environmental inputs and gives insight into the training dynamics and performance of this model.

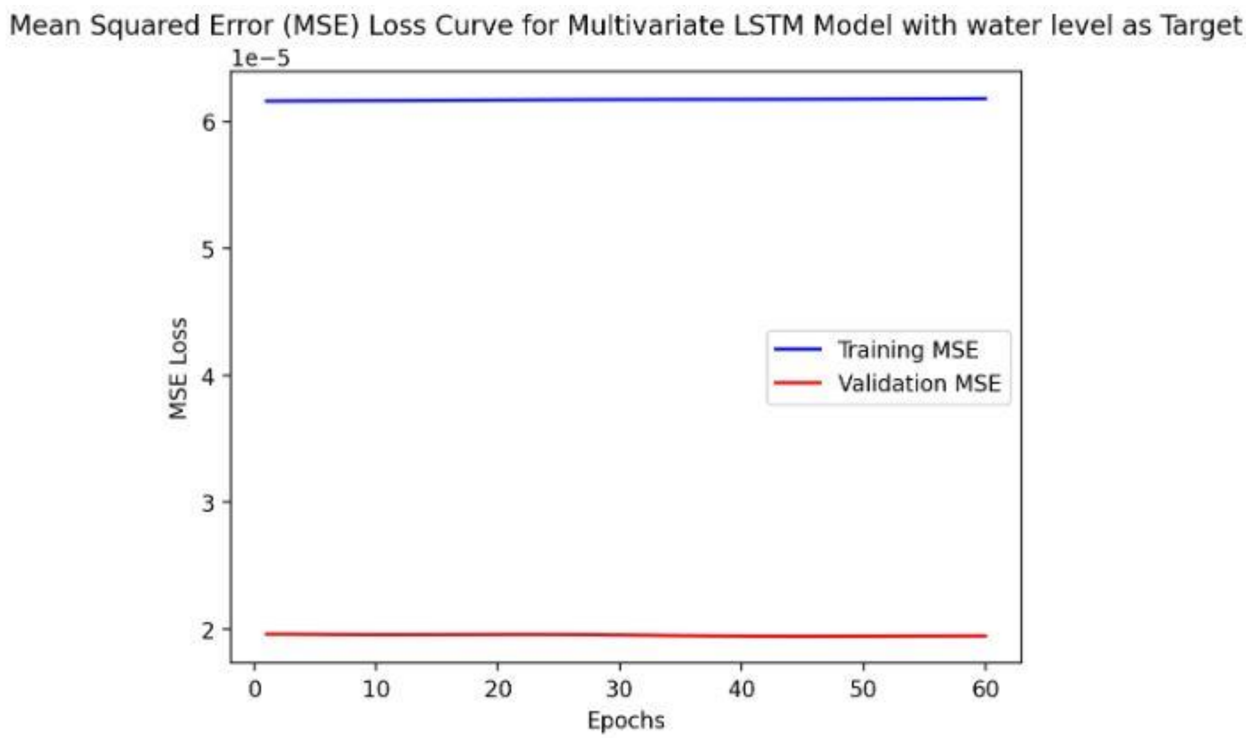


Figure 6.3 MSE Loss Curve for Multi-LSTM

6.2.2 Rain_mm_total as Goal

The MSE loss curve was taken as an evaluation strategy for the Multivariate LSTM model. The dataset to train the model contained the environmental characteristics related to the Arkari River and the goal variable, which was the Arkari Rain_mm_total due to 60 epochs and a batch size 32.

The model developed its training toward an improvement in the predictive process: i.e., a trend of decreasing validation values and training MSE loss with increasing epoch. For instance, at epoch 1, the validation MSE presented 6.7080e-04, while the training MSE was first observed to be 0.0142. The following epochs maintained this pattern; the variations showed the behavior adapting to the set.

For validation, the model reached an MSE of 1.1029e-04 for the final 60 epochs, with training MSE finishing at 9.0664e-06. These shallow values of MSE prove that the model can capture the intricate correlations between features and target variables of the Arkari Rain_mm_total input features.

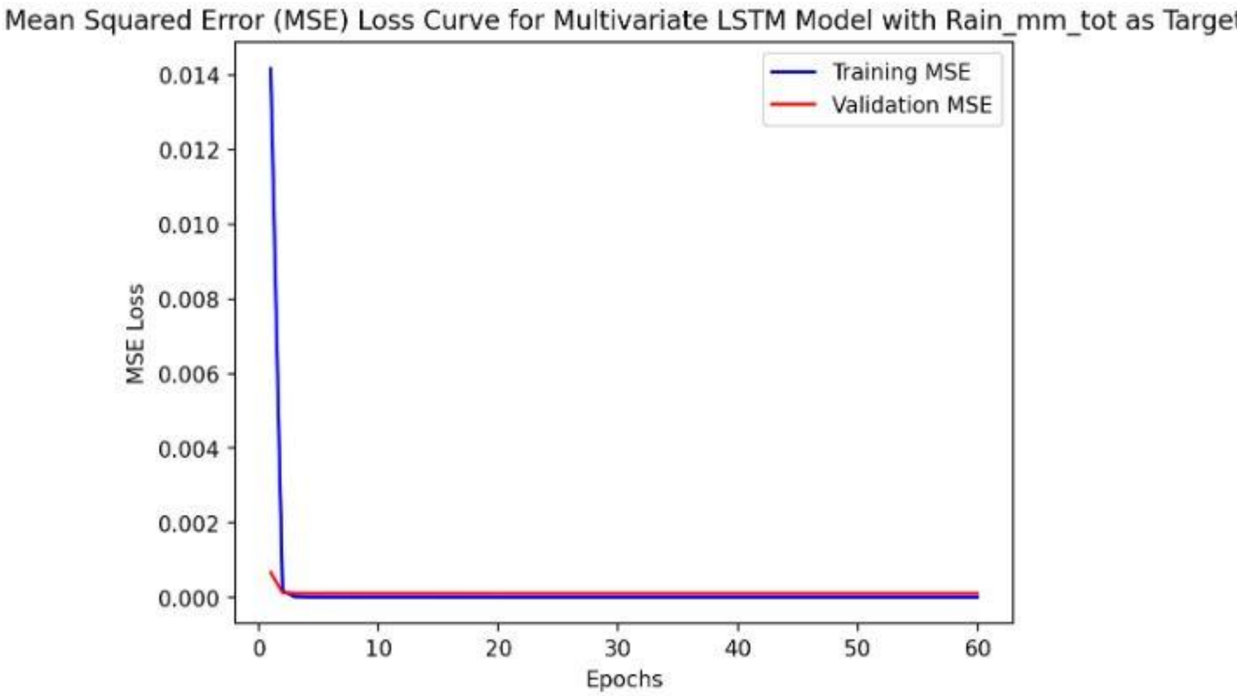


Figure 6.4 MSE Loss Curve for Multi-LSTM with Rain mm as Target

6.3 Comparative Analysis

Table 1: Comparative Analysis of Proposed Study with Existing Study

Aspect	Proposed Study	Existing Studies	Key Differences
Predictive Accuracy	1D CNN and Multivariate LSTM, Used	Kratzert et al. (2018) used LSTM[46] Xiang et al. (2018) used LSTM only[22]	Combination of CNN and LSTM Improved spatial and temporal feature extraction
Data Integration	Environmental sensors (wind, humidity, etc.) Water level radar sensors	Mousavi et al. (2020) used diverse environmental data[47] Damavandi et al. (2018) focused on limited data sources[48]	Comprehensive data integration Real-time water level data included
Real-time Monitoring	IoT infrastructure with water, repeater, and siren stations Real-time data transmission	Nguyen et al. (2019) explored IoT for real-time monitoring	Robust communication network Timely alerts and predictive analytics
Decision Support	Actionable insights for decision-makers Predictive analytics	Silva et al. (2021) focused on early warning systems[49] Singh et al. (2018) provided basic decision support[50]	Detailed and actionable decision support Enhanced early warning systems
Implications for Flood Management	Improved early warning Informed decision-making Scalable and adaptable	Similar findings in enhanced early warning (Silva et al., 2021)	More reliable and accurate systems Comprehensive decision support
Future Research Directions	Explore advanced ML models (e.g., Transformers) Integration with other disaster management systems Focus on cost-effective IoT solutions	Kaur et al. (2020) suggested integration with other systems[51]	New avenues for model enhancement Comprehensive disaster management integration

6.4 Discussion

6.4.1 Overview of Main Discoveries

This thesis offers a new way of monitoring floods and how the ML technique will be used in mitigation with IoT technology. The key findings, as based on the study, are given below:

Using 1D Convolutional Neural Networks (CNN) and Multivariate Long Short-Term Memory (LSTM) networks is very accurate in predicting disaster floods. Both models demonstrated their effectiveness with reliable predictions using the Mean Square Error (MSE) metric, although the performance of the second model was superlative since it captures temporal dependencies.

Comprehensive Data Integration: All the data parameters are fully integrated with the combination of environmental sensors—wind speed, wind direction, humidity, air pressure, temperature, and rainfall—with water-level radar sensors. The aggregation of many sensors allows the prediction of floods to be more accurate and timely.

The use of a resilient framework for the Internet of Things, including water stations, repeater stations, and siren stations for instant data sending and communicating, ensures a continuous process of surveillance with instant notifications, therefore protecting and providing flood control.

The Decision Support System (DSS) provides enormous information to decision-makers by enhancing their ability to forecast, plan for, and mitigate the impacts of flooding disasters—most notably those that occur suddenly and at high intensity.

6.4.2 Comparative Analysis of Previous Studies

The findings are consistent with and expand important domains not only from previous studies:

Machine Learning for Flood Prediction: It was already found in one of the past studies that ML techniques, for example, LSTM and ANN, work well in the forecasting of flood events (Kratzert et al., 2019; Xiang et al., 2018). Our research supports these results and emphasizes the advantage heaped by the CNN features that capture spatial correlations within the data.

The application of IoT in environmental sensing has been extensively studied (Mousavi et al. 2020). Our work adds to this literature by showing an applied example of IoT in the context of flood sensing, in which a dense network of sensors and rapid communication of sensed information are essential.

The use of machine learning, the Internet of Things, and hydrological knowledge in a combined manner is a recommended conduct of flood prediction, as shown by Damavandi et al., 2018 and Dong et al., 2019; this has also proven to be reliable yet effective in this research. This study re-confirms the effectiveness of an interdisciplinary approach toward enhancing the accuracy and hence specializes in predicting floods.

6.4.3 Implications for Flood Management

The significance of this study is that the developed approach shall enhance better early warning systems by giving accurate and timely predictions of floods, hence reducing loss of life and property.

Building Community Resilience: Predictive insights can guide authorities on predictive decisions regarding evacuation notices, resource allocations, and infrastructure protections. All these build the community ability to withstand and recover from adverse events.

The system that was developed showed scalability and adaptability, allowing it to work efficiently in different geographical and climatic conditions, hence making it a perfect solution to solve global flood-management challenges.

Community Engagement: It is essential that human elements, such as community acceptance and stakeholder participation, be understood and addressed. This study identifies public awareness creation and stakeholder participation as critical factors during the setup of flood monitoring systems based on the Internet of Things.

6.4.4 Constraints and Prospects for Further Investigation

Notwithstanding its encouraging outcomes, this study is subject to many constraints:

Data Limitations: Forecasts' precision is grossly affected by the quality and quantity of data collected. Future research should prioritize increasing the data set by including more of the environmental variables and expanding the periods.

Provide Model Enhancements: Although the results provided by both the CNN and LSTM models are already effective, a further search for more advanced ML algorithms, such as Transformer models, may help improve predicting accuracy.

Integration into Other Systems: Future endeavors should be conducted to verify whether this system will be integrated with other systems developed for disaster management, such as earthquake monitoring and wildfire alerts, to constitute a complete disaster response system.

Expense and Maintenance: The development of IoT infrastructure requires enormous investment and maintenance work. Thus, low-maintenance sensor systems are often sought after.

Chapter 7 Conclusion and Future Directions

In conclusion, the proposed system deals with the vital need for an effective way of flood prediction and monitoring using a combined approach of Internet of Things technologies and machine learning algorithms. This study brought out the increased occurrences of flash flooding in mountainous areas, which poses more challenges since they can strike at any place at any time, with short warning lead times. Integrating these systems with IoT-based flood monitoring systems and machine learning algorithms brings an excellent promise for reducing the adverse impact of these disasters.

The study made use of several machine learning algorithms up to Convolutional Neural Networks and Long Short-Term Memory. The dataset used in the implementation was obtained from Kaggle.

The data pre-processing steps have been very carefully carried out: from treating missing data and changing the variable nature from categorical to numeric to reducing the number of variables, normalization, and detecting outliers. Correlation analysis provides the relationship between different kinds of variables that help in selecting important features for flood prediction.

The used evaluation matrices like Mean Square Error (MSE) are very important in studying the machine learning model performance. The research article displayed a comparative analysis of different algorithms, which mentioned their power and limitation in flood prediction scenarios.

The proposed IoT-enabled prototype introduced three key stations: a Water Station, a Repeater Station, and a Siren Station. Each was important in monitoring and alerting to create a robust and interconnected system. The technology of water level monitoring using IoT proved its practicability based on data transmission and alarm triggering.

In the end, this study has demonstrated that the application of new technologies, such as the Internet of Things and machine learning, has a potential contribution to flood prediction and monitoring. The system also serves as a foundation for future advancements in disaster management and specifically points out that proactive measures can be taken to prevent the socioeconomic consequences of natural disasters

References

[1] W. Botzen, O. Deschenes, and M. Sanders, "The Economic Impacts of Natural Disasters: A Review of Models and Empirical Studies," *Rev. Environ. Econ. Policy*, vol. 13, pp. 167–188, Jul. 2019, doi: 10.1093/reep/rez004.

[2] M. Stoffel, B. Wyżga, T. Niedźwiedź, V. Ruiz-Villanueva, J. A. Ballesteros-Cánovas, and Z. W. Kundzewicz, "Floods in mountain basins," *Flood Risk Up. Vistula Basin*, pp. 23–37, 2016.

[3] M. Kelsch, "Hydrometeorological Characteristics of Flash Floods," *Coping with Flash Floods*, pp. 181–193, Jan. 2001, doi: 10.1007/978-94-010-0918-8_18.

[4] M. Aslam, R. Hayat, N. Pari, A. Sameen, and M. Ahmed, "Climate Change, Flash Floods and Its Consequences: A Case Study of Gilgit-Baltistan," 2023, pp. 293–310. doi: 10.1007/978-981-99-1763-1_14.

[5] T. Alam, "A reliable communication framework and its use in internet of things (IoT)," *CSEIT1835111| Receiv.*, vol. 10, pp. 450–456, 2018.

[6] S. Kumar, P. Tiwari, and M. Zymbler, "Internet of Things is a revolutionary approach for future technology enhancement: a review," *J. Big Data*, vol. 6, no. 1, p. 111, 2019, doi: 10.1186/s40537-019-0268-2.

[7] S. Zahir *et al.*, "Smart IoT Flood Monitoring System," *J. Phys. Conf. Ser.*, vol. 1339, p. 12043, Dec. 2019, doi: 10.1088/1742-6596/1339/1/012043.

[8] N. H. D. M. R. H. Abhishek Raj, "IOT based Flood Monitoring and Alerting System," 2022.

[9] K. Furtak and A. Wolińska, "The impact of extreme weather events as a consequence of climate change on the soil moisture and on the quality of the soil environment and agriculture – A review," *CATENA*, vol. 231, p. 107378, 2023, doi: https://doi.org/10.1016/j.catena.2023.107378.

[10] D. S. Rani, G. N. Jayalakshmi, and V. P. Baligar, "Low cost IoT based flood monitoring system using machine learning and neural networks: flood alerting and rainfall prediction," in *2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*, IEEE, 2020, pp. 261–267.

[11] M. Wajid, M. Abid, A. Raza, M. Haroon, and A. Mudasar, "Flood Prediction System Using IOT & Artificial Neural Network," *VFAST Trans. Softw. Eng.*, vol. 12, pp. 210–224, Mar. 2024, doi: 10.21015/vtse.v12i1.1603.

[12] B. Feng, Y. Zhang, and R. Bourke, "Urbanization impacts on flood risks based on urban growth data and coupled flood models," *Nat. Hazards*, vol. 106, Mar. 2021, doi: 10.1007/s11069-020-04480-0.

[13] H. S. Munawar, A. W. A. Hammad, and S. T. Waller, "Remote sensing methods for flood

prediction: A review," *Sensors*, vol. 22, no. 3, p. 960, 2022.

[14] A. D'Addabbo, A. Refice, G. Pasquariello, F. Bovenga, M. Chiaradia, and D. O. Nitti, *A Bayesian Network for Flood Detection*. 2014. doi: 10.1109/IGARSS.2014.6947260.

[15] I. M. Hayder *et al.*, "An Intelligent Early Flood Forecasting and Prediction Leveraging Machine and Deep Learning Algorithms with Advanced Alert System," *Processes*, vol. 11, no. 2, p. 481, 2023.

[16] M. Khalaf, *Flood Detection using Sensor Network and Notification via SMS and Public Network*. 2011.

[17] A. Vasileiadis, E. Alexandrou, L. Paschalidou, M. Chrysanthou, and M. Hadjichristoforou, *Artificial Neural Network and Its Applications*. 2024.

[18] G. Van Houdt, C. Mosquera, and G. Nápoles, "A Review on the Long Short-Term Memory Model," *Artif. Intell. Rev.*, vol. 53, Dec. 2020, doi: 10.1007/s10462-020-09838-1.

[19] W. Xu, Y. Jiang, X. Zhang, Y. Li, R. Zhang, and G. Fu, "Using long short-term memory networks for river flow prediction," *Hydrol. Res.*, vol. 51, Oct. 2020, doi: 10.2166/nh.2020.026.

[20] I. R. Widiasari, L. E. Nugoho, and R. Efendi, "Context-based hydrology time series data for a flood prediction model using LSTM," in *2018 5th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*, IEEE, 2018, pp. 385–390.

[21] J. Zhang, Y. Zhu, X. Zhang, M. Ye, and J. Yang, "Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas," *J. Hydrol.*, vol. 561, pp. 918–929, 2018.

[22] Z. Xiang, J. Yan, and I. Demir, "A rainfall-runoff model with LSTM-based sequence-to-sequence learning," *Water Resour. Res.*, vol. 56, no. 1, p. e2019WR025326, 2020.

[23] H. G. Damavandi, R. Shah, D. Stampoulis, Y. Wei, D. Boscovic, and J. Sabo, "Accurate prediction of streamflow using long short-term memory network: a case study in the Brazos River Basin in Texas," *Int. J. Environ. Sci. Dev.*, vol. 10, no. 10, pp. 294–300, 2019.

[24] L. Dong *et al.*, "Prediction of streamflow based on dynamic sliding window LSTM," *Water*, vol. 12, no. 11, p. 3032, 2020.

[25] Y.-M. Won, J.-H. Lee, H.-T. Moon, and Y.-I. Moon, "Development and application of an urban flood forecasting and warning process to reduce urban flood damage: A case study of Dorim River basin, Seoul," *Water*, vol. 14, no. 2, p. 187, 2022.

[26] K. Kunverji, K. Shah, and N. Shah, "A flood prediction system developed using various machine learning algorithms," in *Proceedings of the 4th International Conference on Advances in Science & Technology (ICAST2021)*, 2021.

[27] J. Chen, Y. Li, C. Zhang, Y. Tian, and Z. Guo, “Urban Flooding Prediction Method Based on the Combination of LSTM Neural Network and Numerical Model,” *Int. J. Environ. Res. Public Health*, vol. 20, no. 2, p. 1043, 2023.

[28] F. Ostadkalayeh, S. Moradi, A. Asadi, A. Moghaddam Nia, and S. Taheri, “Performance Improvement of LSTM-based Deep Learning Model for Streamflow Forecasting Using Kalman Filtering,” *Water Resour. Manag.*, vol. 37, pp. 1–17, Mar. 2023, doi: 10.1007/s11269-023-03492-2.

[29] M. Cheng, F. Fang, T. Kinouchi, I. M. Navon, and C. C. Pain, “Long lead-time daily and monthly streamflow forecasting using machine learning methods,” *J. Hydrol.*, vol. 590, p. 125376, 2020, doi: https://doi.org/10.1016/j.jhydrol.2020.125376.

[30] Q. Ye, X. Yang, C. Chen, and J. Wang, *River Water Quality Parameters Prediction Method Based on LSTM-RNN Model*. 2019. doi: 10.1109/CCDC.2019.8832885.

[31] H. Wang, S. Xu, H. Xu, Z. Wu, T. Wang, and C. Ma, “Rapid prediction of urban flood based on disaster-breeding environment clustering and Bayesian optimized deep learning model in the coastal city,” *Sustain. Cities Soc.*, vol. 99, p. 104898, 2023, doi: https://doi.org/10.1016/j.scs.2023.104898.

[32] H. Farahmand, X. Liu, S. Dong, A. Mostafavi, and J. Gao, “A Network Observability Framework for Sensor Placement in Flood Control Networks to Improve Flood Situational Awareness and Risk Management,” *Reliab. Eng. Syst. Saf.*, vol. 221, p. 108366, Feb. 2022, doi: 10.1016/j.ress.2022.108366.

[33] R. Khatoon, I. Hussain, M. Anwar, and M. A. Nawaz, “Diet selection of snow leopard (Panthera uncia) in Chitral, Pakistan,” *Turkish J. Zool.*, vol. 41, no. 5, pp. 914–923, 2017.

[34] U. Naeem, H. Hashmi, H.-U.-R. Mughal, and A. Shakir, “Flow trends in river Chitral due to different scenarios of glaciated extent,” *KSCE J. Civ. Eng.*, vol. 17, Jan. 2013, doi: 10.1007/s12205-013-1978-1.

[35] Z. Syed *et al.*, “Hydroclimatology of the Chitral River in the Indus Basin under Changing Climate,” *Atmosphere (Basel).*, vol. 13, p. 295, Feb. 2022, doi: 10.3390/atmos13020295.

[36] H. Waqas *et al.*, “Flash flood susceptibility assessment and zonation using an integrating analytic hierarchy process and frequency ratio model for the Chitral District, Khyber Pakhtunkhwa, Pakistan,” *Water*, vol. 13, no. 12, p. 1650, 2021.

[37] S. Hochreiter and J. Schmidhuber, “Long Short-term Memory,” *Neural Comput.*, vol. 9, pp. 1735–1780, Dec. 1997, doi: 10.1162/neco.1997.9.8.1735.

[38] J. Charco, T. Roque-Colt, K. Egas-Arizala, C. Perez Espinoza, and C. C. Angelica María, *Using Multivariate Time Series Data via Long-Short Term Memory Network for Temperature Forecasting*. 2020. doi: 10.1007/978-3-030-59194-6_4.

[39] A. Salman, Y. Heryadi, E. Abdurahman, and W. Suparta, “Single Layer & Multi-layer Long Short-Term Memory (LSTM) Model with Intermediate Variables for Weather Forecasting,” *Procedia Comput. Sci.*, vol. 135, pp. 89–98, Jan. 2018, doi:

10.1016/j.procs.2018.08.153.

[40] X. Zhao, L. Wang, Y. Zhang, X. Han, M. Deveci, and M. Parmar, “A review of convolutional neural networks in computer vision,” *Artif. Intell. Rev.*, vol. 57, no. 4, p. 99, 2024, doi: 10.1007/s10462-024-10721-6.

[41] I. H. Sarker, “Machine Learning: Algorithms, Real-World Applications and Research Directions,” *SN Comput. Sci.*, vol. 2, no. 3, p. 160, 2021, doi: 10.1007/s42979-021-00592-x.

[42] Z. Zhang, *Improved Adam Optimizer for Deep Neural Networks*. 2018. doi: 10.1109/IWQoS.2018.8624183.

[43] H. Allamy and R. Z. Khan, *Methods to Avoid Over-Fitting and Under-Fitting in Supervised Machine Learning (Comparative Study)*. 2014. doi: 10.3850/978-981-09-5247-1_017.

[44] M. Bani Yassein, M. Shatnawi, S. Aljawarneh, and R. Al-Hatmi, *Internet of Things: Survey and open issues of MQTT Protocol*. 2017. doi: 10.1109/ICEMIS.2017.8273112.

[45] N. S. Alotaibi, H. I. Sayed Ahmed, S. O. M. Kamel, and G. F. ElKabbany, “Secure Enhancement for MQTT Protocol Using Distributed Machine Learning Framework,” *Sensors*, vol. 24, no. 5, p. 1638, 2024.

[46] L. Windheuser, R. Karanjit, R. Pally, S. Samadi, and N. C. Hubig, “An end-to-end flood stage prediction system using deep neural networks,” *Earth Sp. Sci.*, vol. 10, no. 1, p. e2022EA002385, 2023.

[47] N. Mohd Nawi, M. Makhtar, M. Salikon, and Z. Afip, “A comparative analysis of classification techniques on predicting flood risk,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 18, p. 1342, Jun. 2020, doi: 10.11591/ijeecs.v18.i3.pp1342-1350.

[48] J. Wang *et al.*, “Disaster mapping and assessment of Pakistan’s 2022 mega-flood based on multi-source data-driven approach,” *Nat. Hazards*, vol. 120, Dec. 2023, doi: 10.1007/s11069-023-06337-8.

[49] D. Perera, J. Agnihotri, O. Seidou, and R. Djalante, “Identifying Societal Challenges in Flood Early Warning Systems,” *Int. J. Disaster Risk Reduct.*, vol. 51, p. 101794, Aug. 2020, doi: 10.1016/j.ijdrr.2020.101794.

[50] S. Ahmad and S. Simonovic, “An Intelligent Decision Support System for Management of Floods,” *Water Resour. Manag.*, vol. 20, pp. 391–410, Jun. 2006, doi: 10.1007/s11269-006-0326-3.

[51] M. Truu, I. Annus, J. Roosimägi, N. Kändler, A. Vassiljev, and K. Kaur, “Integrated decision support system for pluvial flood-resilient spatial planning in urban areas,” *Water*, vol. 13, no. 23, p. 3340, 2021.

[52] L. Mdegela, Y. De Bock, E. Municio, E. Luhanga, J. Leo, and E. Mannens, “A Multi-Modal Wireless Sensor System for River Monitoring: A Case for Kikuletwa River Floods

in Tanzania," *Sensors*, vol. 23, no. 8, p. 4055, 2023.