EARLY WARNING FOR PRE AND POST FLOOD RISK MANAGEMENT

2021-124

Final Report

Vidura Yasiru Samarasiri

(IT18012620)

B.Sc. (Hons) Degree in Information Technology Specializing in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

13th October 2021

EARLY WARNING FOR PRE AND POST FLOOD RISK MANAGEMENT

2021-124

Vidura Yasiru Samarasiri

(IT18012620)

Dissertation submitted in partial fulfillment of the requirements for the Bachelor of Science

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

13th October 2021

1 DECLARATION

I declare that this is my own work, and this dissertation does not incorporate without acknowledgement any material previously submitted for a degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to Sri Lanka Institute of Information Technology, the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Name	Registration Number	Signature
V.Y Samarasiri	IT18012620	*vidurasamarasiri*

The above candidate is carrying out research for the undergraduate Dissertation under my supervision.

Signature of the Supervisor: Date:

2 ACKNOWLEDGEMENT

Firstly, I would convey my heartfelt appreciation and credits to our supervisor Mr. Samantha Rajapakse, who provided me the treasurable prospect to test myself at the uppermost level I could do. You have invigorated and helped me to grow up from a fresh graduate student to a mature researcher. I have learnt so much from your both educationally and personally which has bought me to a great place in my studentship times. Thank you for all the priceless comments and guidance on my research. I would also like to express gratitude my parents for all the backing they gave me with no anticipation of a compensation. Their constant help, support and tolerance have made my life full of success, and I am glad for everything they have done.

3 ABSTRACT

Flooding has been a treacherous situation in Sri Lanka. The studies have found potential risks and eradications caused by natural disasters are rising. These flooding and extreme weather conditions occurs as a result of changes of climates, inappropriate land use and impact of severe metrological conditions. Analyzing these factors, it is reasonable to expect property damages and human losses caused by floods will climb up more in future. Therefore, there is a need of modern tool for flood forecasting and it is important for taking further decisions in a flooding situation. Developing a structure to forecast extreme weather conditions will be a great aid for citizens who are affected from disasters. However, these approaches should be incorporated with tools to advance natural disaster consciousness and mitigate the risks exposed to public. Implementation of crowdsourcing methodologies with the help of volunteers will improve disaster governance. The information technology is evolving fast in now days and people are using technology to transform their living style to be more productive.

Crowdsourcing has been an innovational approach for solving problems by participating public crowd and with the help of information technology-based crowdsourcing, the main goal is to gather responses from public crowd through their devices such as mobile phone to facilitate disaster response authorities with appropriate weather information in case of flooding situations and extreme weather conditions. The crowdsourcing solution consists with a mobile application to engage with the community to gather response to mitigate flood risks by enhancing active participation of the public crowd. The overall outcome of the validated crowdsourcing data sets will be an effective mechanism to assist and support implementing strategies for decision making procedure before, during and after a crisis to promote disaster awareness and preparedness in the community.

Keywords: Crowdsourcing, Mobile application development, Awareness, Avoidance, flood risk, readiness

Table of Contents

4 List of Figures

5 List of Tables

6 List of Abbreviations

IoT	Internet of Things
API	Application Programming Interface
SDLC	Software Development Lifecycle

7 INTRODUCTION

7.1 Background Study

7.1.1 flood risk management systems

The incidence of floods is seen various localities across the world. Floods cover disasters arising from all sources of water overflows including incidence of Tsunami and tidal waves. As many people live on low land areas and close to waters such as sea, rivers, and reservoirs due to economic benefits, the exposure to flood risk among the people is considered high. This has begun from ancient civilizations that evolved around water resources due to economic benefits of water. Highlands also confront floods including earth slips as they are the catchment areas of rainfall and rivers and, therefore, people living in highland areas also confront flood risks. As such, flood is a known risk confronted by a large segment of people across the world.

From time to time, societies have attempted to control and manage flood risk through various systems. In primitive societies, people lived around rivers and seas and dispersed to highlands when floods were affecting. Once floods are over, they came back rebuilt settlements. This is just food risk confrontation or tolerance. After water engineering got developed, governments started flood protection systems to prevent or block floods. This includes building concrete barriers on the sides of cannels and rivers, blocking rivers and taking waters through cannels to agricultural areas and reservation of lowlands to hold water. However, now such absolute flood protection systems are considered to be unachievable and unsustainable due to high cost and inherent risks. Therefore, now societies are moving form flood protection systems to flood risk management systems.

The key steps generally seen in the flood risk management system are prediction of rainfalls, tsunami, etc., to detect possible incidence or build-up of floods, issuance of alerts to people living in risky areas, evacuation of people from flooding areas to common places such as schools, provision of basic living requirements by the

government and social organizations to evacuated people and resettlement and recovery from floods. This system is designed on science and technology and community engagement.

However, present flood risk management systems in Sri Lanka are not inclusive operations to minimize damages and losses to people and properties as the components of the present risk management systems are not integrated with the modern technology and community engagement. Although weather predictions are available, people disregard them due to most incidences of inaccuracy. Therefore, occurrence of floods has become an annual event in certain areas. The evacuation starts only at the onset of floods. Relief measures to evacuated people are very poor while recovery of settlement and earnings are not assisted by any social safety net system. The use of technology and community engagement in the system are hardly seen.

- **Environmental Risks**

Although floods are a result of a natural hazard due to sudden excessive rainfall, reasons caused by people have increased the risk of floods even due to normal rainfalls. Those reasons are mostly connected with various misuses of the land resources that are unfriendly or obstructing natural means of water flows to lows lands and sea. For example, increasing use of lowlands for residential purpose, blockages of banks of rivers and cannels, unmaintained drainage systems and deforestation cause frequent floods affecting the people. In addition, global warming mainly due to emission of greenhouse gas by human activities also has raised sea water levels and floods. The present global temperature has increased by about 1.2 Celsius degrees above the pre-industrial era average and is expected to rise to about 1.5 degrees by 2040. As a result, sea water level in the future is expected to increase by about 15 feet and some islands such as Maldives are expected to disappear from the surface.

- **Short-term and Long-term Strategies**

Therefore, the flood risk management involves in both a short-term strategy and long-term strategy. The long-term strategy requires globally agreed efforts to control human activities to protect the environment and control the global warming below 1.5 degrees

over the pre-industrial climate level and other human activities that pollute the environment.

The short-term flood risk management is primarily based on an analysis of flood hazard, exposure to flood hazard and vulnerability of people and property to danger. It is recommended that people learn to live with flood risks, gaining and promoting a clear understanding of flood risks, quantifying, and modifying the flood hazard, regulating exposure to the hazard, and reducing their vulnerability to danger. Therefore, it requires integration of government support, technology, and community engagement. The system is primarily to protect people from ongoing risks of floods. This strategy requires the techniques for prediction of floods and flood disaster management system to issue alerts, evacuate people urgently from flooding areas, look after them until returning to normalcy and a national insurance and safety net system to recover people and property from damages caused by floods. Media reports on recent flood events show that even short-term flood risk management in Sri Lanka is poor as disruptions to people due to floods have been severe and rising. The main problem confronted in flood risk management system is the information lag on fold risks.

7.1.2 crowdsourcing defined

Crowdsourcing is the strategy used to outsource identified operational lines of a project or scheme to a wider segment of people who directly benefit from the project. As such, crowdsourcing is the means of getting wider participation of the beneficiaries in a social project. Therefore, this strategy helps beneficiaries to become stakeholders of the project. The outsourcing is the term used to mean the outsource of operations of a business or a project to a set of external persons. In most businesses, marketing function is outsourced to a network of agents operating on various remuneration packages such as commissions and price margins. Further, certain clerical and accounting operations of institutions are outsourced to third parties in order to reduce permanent staff and cost being subjected to numerous labor laws and regulations. Therefore, outsourcing of functions or operations has been a popular business/human

resource management strategy during the past decade. Therefore, the crowdsourcing is the new mode of outsourcing of certain operations of a socially beneficial project to its beneficiary members of the society. Therefore, this is a strategy adopted to involve a wider community engagement/participation in a project.

Therefore, this research is to develop a crowdsourcing solution to bridge the gaps between information and actions among the relevant authorities and public through wider community engagement. This will help relevant parties/participants to get activated early and implement the disaster management system in a timely manner.

- **Coverage of Crowdsourcing**

Crowdsourcing is the element of operation of the flood risk management project/system with respect to exchange of information whereas the activities/operations with regard to risk assessment, risk aversion and risk mitigation/control are left with the designers and managers of the system.

The exchange of information is two-tier. First, gathering input information from the field or beneficiaries. This implies that the risk management software has a portal authorized to persons living in the areas under the flood risk. They are provided with the access to report classified information useful to assess the probable incidence or probability of flooding in the respective areas. That will cover the extent of rainfall, water levels of surrounding rivers, canals, tanks, ponds, drainage, etc., and water flows on lowlands and highlands. This will be used in addition to relevant information available from official sources such as Meteorology and Irrigation officials. However, since official information sources are subject to time lags, information provided by the participants in the crowdsource will be faster and online as and when risks build up at each stage such as start of the rainfall, duration of the rainfall, initial signs of flooding and gradual developments.

The second tier is the output information or release of risk and risk management information to the participants to get ready and prepare for risk handling. In this regard, risk handing system should have been designed as a part of the system. For example,

issuance of alerts, evacuation instructions and situation information to participants/beneficiaries should be defined and covered in the system. The distribution flood benefits also can be covered in the system. This will help faster handling of flood risks. Accordingly, crowdsourcing can be linked through smart phones and other relevant devices that are accessible to the crowd participants.

7.2 Literature Survey

The aim of this chapter is to give a complete review of existing vast literature which provides an understanding about the role of crowdsourcing can be applied to reinforce flood risk management. Incidence of floods is a known risk to several localities across the world. Although floods are a result of a natural hazard due to sudden excessive rainfall, reasons caused by people have increased the risk of floods even due to normal rainfalls. Those reasons are mostly connected with various misuses of the land resources that are unfriendly or obstructing natural means of water flows to lows lands and sea. For example, increasing use of lowlands for residential purpose, blockages of rivers and cannels, unmaintained drainage systems and deforestation cause frequent floods affecting the people.

In addition, global warming mainly due to emission of greenhouse gas by human activities also has raised sea water levels and floods. For many countries, the factors that effect is weak flood monitoring structure led to damage, loss of lives and economic losses when a flooding situation takes place because of weak and poor flood monitoring it can cause devastating effect for human and environment. The strategies of flood risk management is growing fast in today's world [1].

Flood risk management consists with both long term strategy and short-term strategy. The long-term strategy requires globally agreed efforts to control human activities to protect the environment and control the global warming below 1.5 degrees over the pre-industrial climate level. The short-term strategy is to protect people from ongoing risks of floods. This strategy requires the techniques for prediction of floods and flood disaster management system to issue alerts, evacuate people urgently from flooding areas, look after them until returning to normalcy and a national insurance system to protect people from damages caused by floods.

Flood risk management is the structure for managing flooding situations and reducing its impact to citizens. The main intention of flood risk management is to identify floods and to get activated early and prepare for the disaster management [2]. The strategies and frameworks that are executed in flood risk management will be helpful for

implementing strategies for weather forecasting and creating rehabilitation plans [3]. In a disastrous situation communication, availability of information helps to promote self-organization. Self-organizing is the procedure where different people will communicate information and those behaviors and information will be helpful in achieving the common goal of disaster risk reduction [4].

Crowdsourcing is the process where volunteer information is gathered during a disastrous situation and those information falls under disaster response. Engagement of the public crowd and citizens promotes efficient decision making through exchanging weather information. There are set of frameworks defined for engaging community in disaster response [5]. Conventional crowdsourcing refers to offline methods for engaging community [6].

The information technology is evolving fast in now days and people are using technology to transform their living style to be more productive. Information technology together with the Internet provides an effective online platform for community engagement. In disaster risk reduction, crowdsourcing has become a low-cost approach for gathering disaster responses [7]. Existing research emphasis on what are the techniques for gathering community data and techniques for analyze, classify the data sets as well as benefits of effective communication to support on identifying damages, losses and disaster decision making.

Crowdsourcing has become an innovational approach to encourage in involving active volunteers to obtain weather information [8-10]. Crowdsourcing is known as novel method for gathering responses from public crowd to improve in overcoming disastrous situations, availability of information and promotes the ability in public crowd to respond in a flexible way [10]. Lots of weather factors can be analyzed and gathered in the occurrences of floods in this matter. The framework for crowdsourcing is outlined in the bellow mentioned figure [1]. It illustrates three phases of community engagement that can provide information regarding flood risk management, and it is emphasis as citizen science, volunteer information and crowdsourcing,

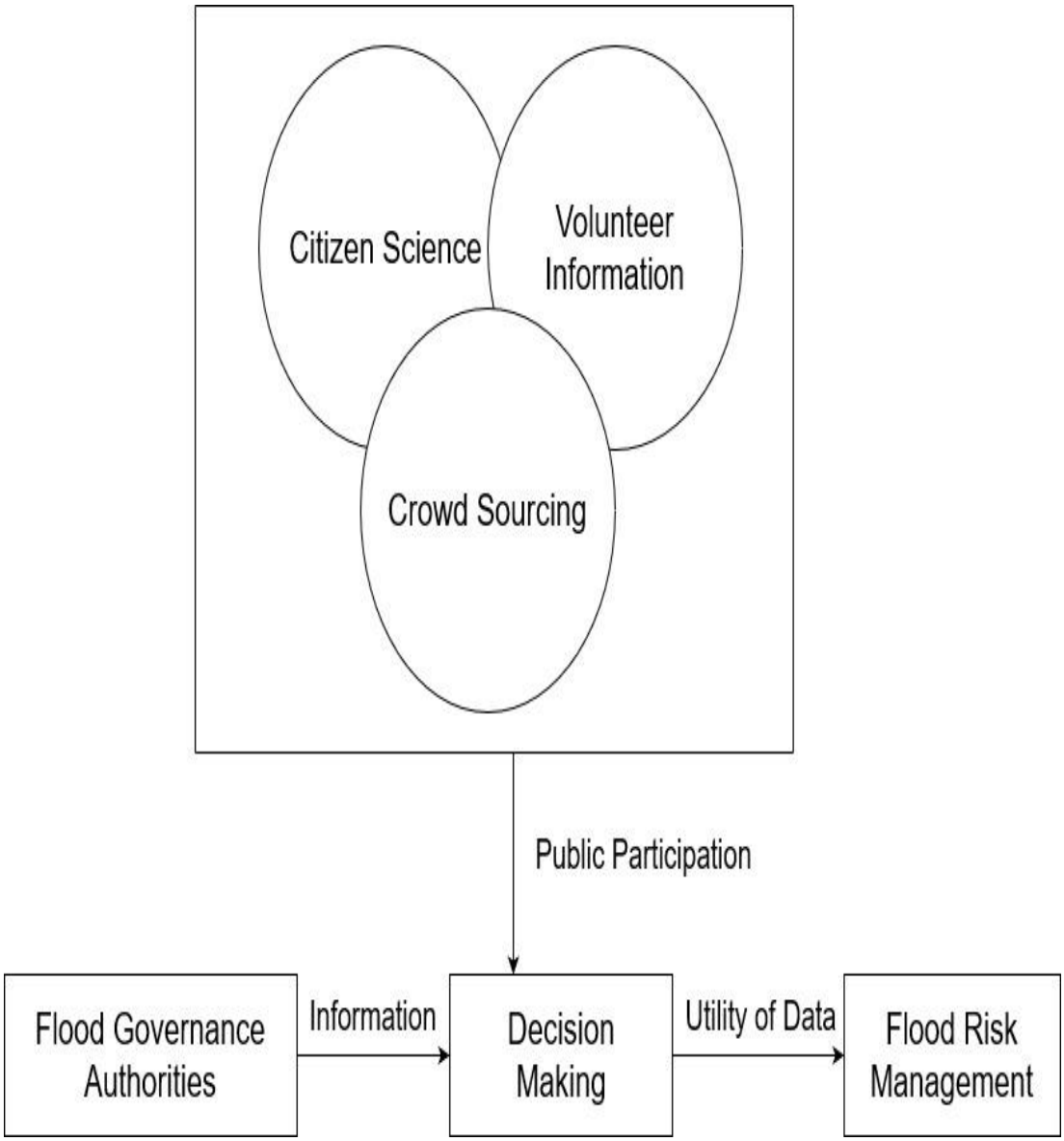

Figure 1 Crowdsourcing Framework

Crowd sourcing approach on disaster risk reduction focuses on collecting and analyzing data in a disastrous event. The similar studies have been done in Brazil to implement a crowd source-based approach to collect weather information from volunteers to improve and forecast flood risk management [11]. similarly, Taiwan developed a system named TSER (Taiwan scientific earthquake report system) for disaster monitoring which obtains information from trained volunteers, and it is used to identify earthquakes and disastrous events before and during disasters [12]. The crowdsourced data sets that are gathered from volunteers are unstructured and unorganized. These data sets should be organized and analyzed before using for further processes. To overcome this issue, research on crisis information management in the web 3.0 age implemented a machine learning based approach for transforming and classifying crowdsourcing data sets to a usable format to support making decisions in a disaster [13]. Authors on this research focus on classifying and analyzing crowdsourcing data sets in more technological way. Crowdsourcing establishes better and improved risk communication between affected people, volunteers and relief providing authorities [14].

Sahana is a volunteer coordination system which was implemented to manage and coordinate possible risks related to disasters. This crowdsourcing system mainly focused on the outbreak of 2004 Asian Tsunami to find missing victims, relief persons and identify damages to buildings, houses, and belongings [15]. Crowdsourcing has been an innovational approach for solving problems by participating public crowd and it plays a major role in pre disaster consciousness and readiness, post disaster rehabilitation phase where information related to disaster events are being captured to identify and help the affected people. Applying the practices of crowd sourcing-based information technology to analyze flood risk management and detecting floods will make it more productive for public and officials, therefore having an early warning structure consisting above mentioned crowd sourcing factor will be more helpful for public to overcome the flooding situations and to be prepared in any circumstances.

7.3 Research Gap

Similar studies and research have been conducted in worldwide by approaching crowdsource based model to capture disaster data from community. When it comes to volunteer-based systems, active participation of crowd plays a major role. Extracting the existing research, the research [] has successfully conducted crowd source-based approach to enable disaster response by implementing a mobile app feature. The research [] implemented a crowdsource based approach to gather weather information in Brazil. The research [] has focused on improving self-organizing and disaster resilience by participating volunteers in a disaster situation. These are the major highlights of crowdsourcing research papers.

In the gap analysis of crowdsourcing approach by going through available research papers the authors have directed future works on the data accuracy and validity of crowdsourcing data and those data need to be received in a precise and concise format to enforce accuracy of information in crowdsourcing data. The proposed crowdsourcing approach is consisting with following sections.

Existing Research	Gather Information Regarding Flood	Accuracy and Validity of Data	Data Comparison with Weather API and IOT device
[1]	✓	×	×
[2]	✓	×	×
[3]	✓	×	×
Proposed System	✓	✓	✓

Table 1 Comparison with Research

Almost every research has successfully collected crowdsourcing data to forecast weather information. Going through the available research papers, authors have mentioned accuracy and validity of crowdsourcing data should be improved further in order to make decisions correctly. Therefore, the proposed system is able to select the most correct and accurate data set based on majority decision function. Since the public and authorities makes decisions based on these disaster responses its highly important that most accurate and validated should be published through the system. Misleading crowdsourcing information can result in executing poor flood risk management strategies. Therefore, the data sets should be accurate in order to implement better flood risk management frameworks. Majority decision data should be organized and well-structured before publishing through the. To overcome that issue crowdsourcing data sets are being analyzed using data analysis techniques.

Apart from all these data manipulations final condensed crowdsourcing data set will be compared with third party weather API and flood and weather monitoring IoT device in order to make the volunteer information more accurate. Finally, the proposed solution will have a complete package of above-mentioned functions. Gathering data from public crowd, improving Accuracy and validity of data, and receiving data in precise and concise format and crowdsourcing data comparison with other data sources such as weather API's and IoT device will support the mentioned future works of crowdsourcing approach

7.4 Research Problem

When a flooding situation takes place at a specific location as the water rises quickly it will take some time to get fully prepared for public to evacuate from the area. When the water level reaches the peak, it will start to flood in their living area and cause immediate damage, therefore with the help information technology-based crowdsourcing we can implement a solution for this major matter.

At present, when a flooding situation takes place at a specific location as the water rises quickly, it will take some time to get fully prepared for public to evacuate from the area. When the water level reaches high, it will start to flood in their living area and cause immediate damage. When it comes to disaster management and flood tracking, the major problem identified is not merely the unavailability of technology or basic information, but its communication and unavailability of accurate set of information when a flooding situation occurs.

Most of the times, it might take some time for the authorities to give warnings about disasters and floods because of the late communication of information. Therefore, the research problem in this study is to develop an IT solution to resolve the information problem in the flood risk management system in Sri Lanka. By implementing crowdsourcing solution and active participation of public crowd will bridge the gap between government officials and public crowd

The crowdsourcing strategy is useful implement a solution for this major information problem. The purpose of implementing crowdsourcing solution is to gather inputs from public crowd to forecast severe weather conditions which may cause flooding to a specific area. In addition to the other whether prediction solution provided within the scope of “Early Warning for Pre and Post Flood Risk Management” project, crowdsourcing solution is mandatory to validate other source of information. Usually, the input source of data for a crowdsourcing solution gathered from the public crowd.

Hence, when designing the solution for this project, we may need to overcome following challenges specific to the crowdsourcing.

1. Sourcing the right crowd
2. Validate the accuracy of data (Data integrity)
3. Receive data in precise and concise format
4. Periodically receive live data

Above extracted factors shows that an early warning tool is the key to overcome above situations, therefore there is a need of having an early warning tool handle the flooding situation and reduce the economic and social losses in a country which is affected by floods.

During this research comprehensive analysis on above challengers required to be done based on the existing studies on the related topics and the data collection on the occurrences of flooding situations by evaluating the practicality of Crowdsourcing solution in real world. During the solution evaluation, inputs of subject matter experts will be considered to optimize and improve the proposed solution.

8 Research Objectives

8.1 Main Objective

Main objective is to collect weather information from public crowd, analyze and validate gathered information using statistical analysis techniques.

8.2 Specific Objectives

To full fill above mentioned main objective, the specific objectives that need to be covered as follows,

1. Implement a way to source the right crowd

In order to overcome the first challenge of choosing right crowd, first step is to identify the specific area and the volunteers in that identified area to gather inputs from them in a situation of a natural disaster therefore to source the right crowd we will provide a mobile application with registration page. Also, we will get data inputs from people through the mobile application. Therefore, it is mandatory for people to register through the mobile application to provide data inputs. This will partially solve the problem of sourcing right crowd. In addition to that, providing fault information will be resulted to block users.

2. validate the accuracy of crowd sourcing data

Data accuracy plays the major role of this crowdsourcing problem. Therefore, it is mandatory to validate data received from people. Therefore, data will be validated against data received through other sources such as third-party APIs, IoT devices before publishing them through the system. In addition to that, when user enters data inputs, mobile application will automatically capture his current location and that will be sent along with the input data. This will help to validate whether user is currently located into the specific area whether he send data inputs regarding whether.

3. Structure the data in precise and concise format

Also, our system should be intelligent enough to validate user input data. Since no admin user manually validate these data, it is mandatory for users to send these data in precise and concise format. In order to enforce that to the mobile app users, we will be providing a questionnaire with predefined set of questions as an application form. This will help system to easily identify and validate data

4. periodically receive live data

In order to overcome the active participation in public crowd in a disaster situation the systems need to receive live data periodically Therefore, during a situation of severe weather, mobile application will provide periodic based alerts for the user to send the current status of the weather.

9 METHODOLOGY

This section highlights approach for solving the research problem by applying suitable methodologies. This section consists with the overall system architecture, crowdsourcing component overview, logical design of crowdsourcing, development process, data collection and requirement gathering, feasibility study, design components and commercialization.

9.1 Overall System Architecture

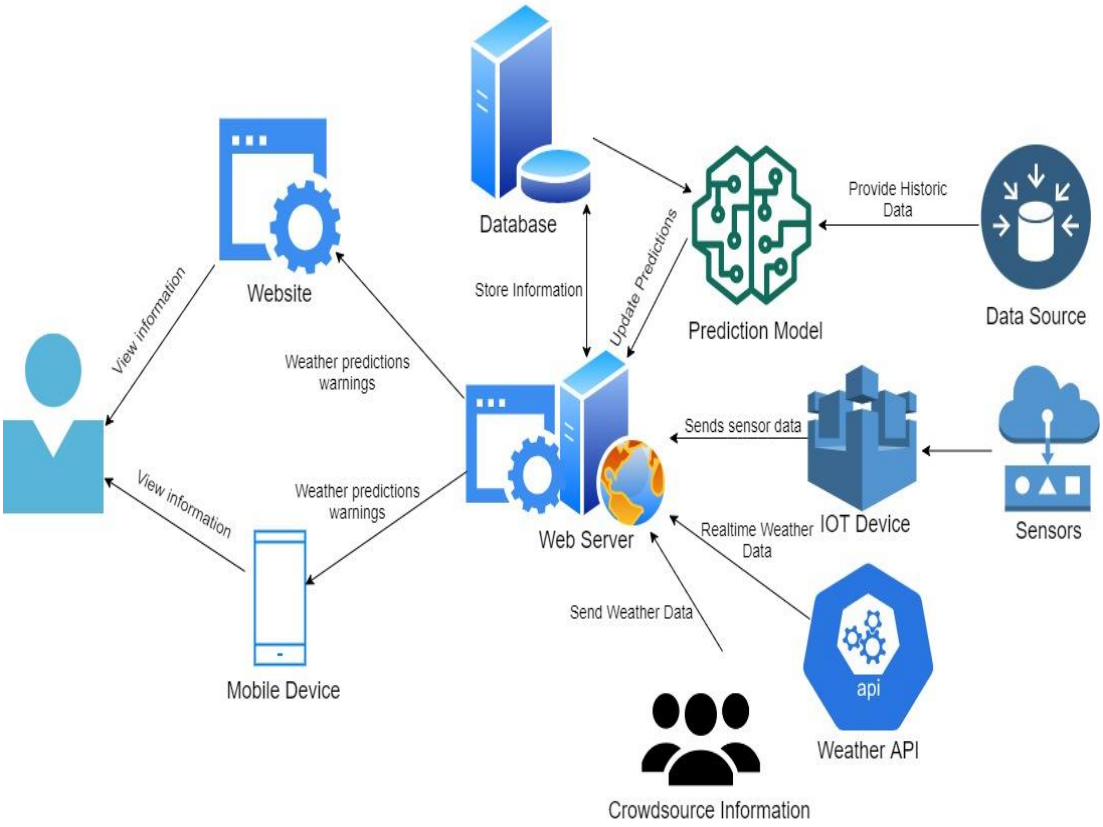

Figure 2 Overall System Architecture

9.2 High-level Component Diagram of Crowdsourcing

This section highlights the component diagram of crowdsourcing. Volunteer weather information is gathered through the mobile application and stored in the real-time firebase database. When visualizing the crowdsourced data, the most accurate crowdsource data is displayed to the users with IoT device data and weather API data

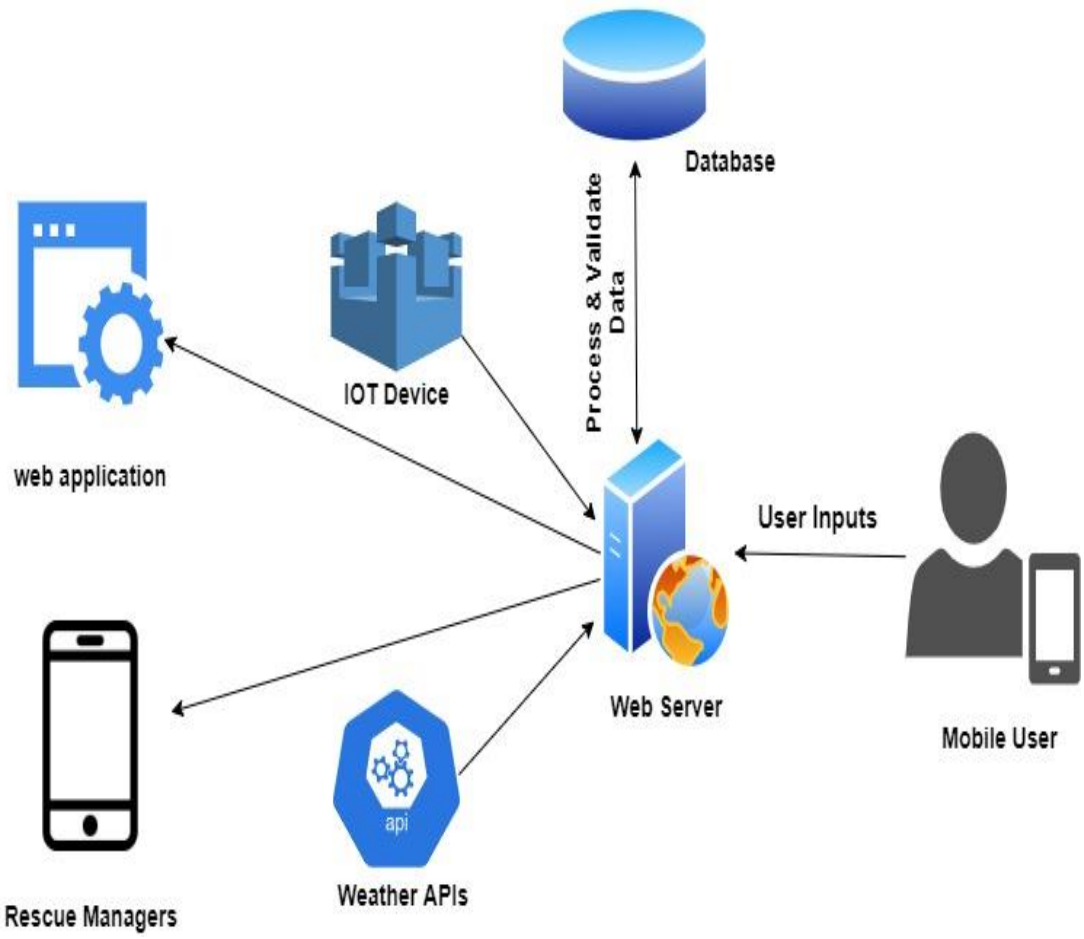

Figure 3 High-level Component Diagram

9.3 Logical View of Crowdsouricng

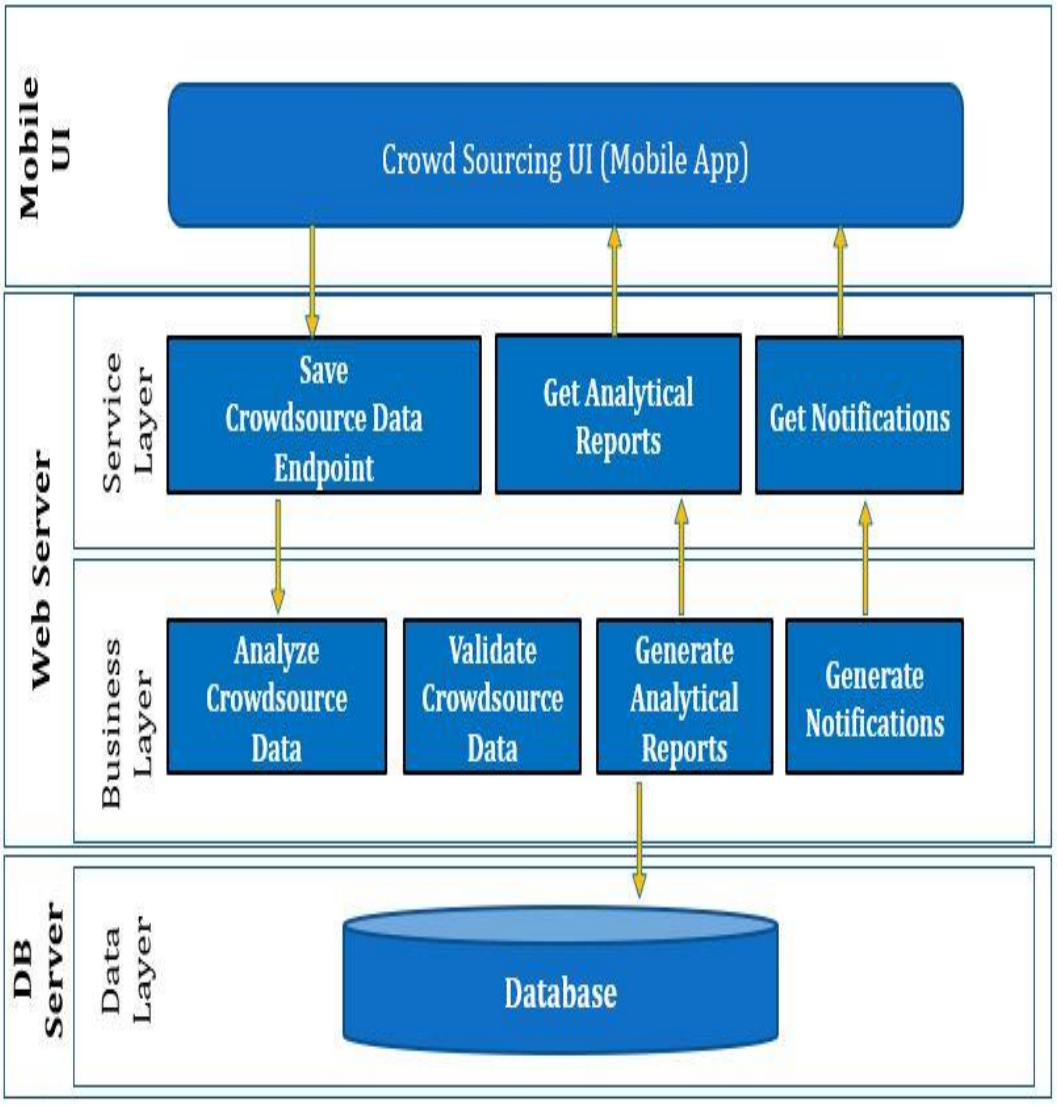

Figure 4 Logical View

9.4 Crowdsourcing Process Design

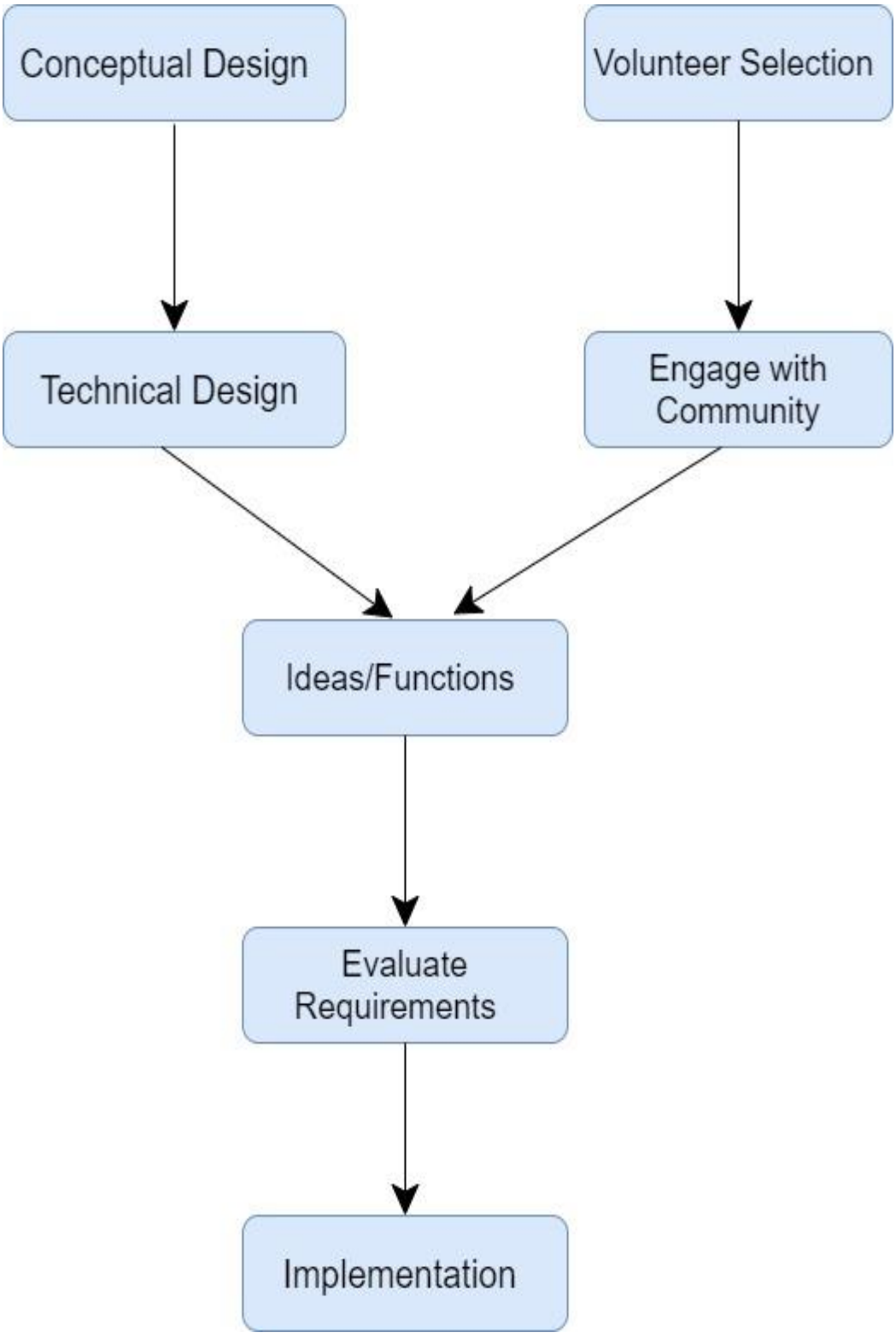

Figure 5 Crowdsource Process

9.5 Crowdsouricng Approach Development

The proposed early warning system is capable and have the below mentioned functionalities in terms of crowdsourcing,

- Gathering data from public crowd
- Validating and analyzing the gathered data
- Structuring data in precise and concise format
- Crowdsource data comparison with weather API and IoT device
- Receiving weather information periodically from public crowd
- Providing analytical reports and notifications

Crowd sourcing offers a technique of devolved and a low-cost approach for gathering and exchanging information. The overall intention of providing this feature is to approach the public crowd to collect their current weather information and analyze and validate those information before publishing them to users. The crowdsourcing section consists with a UI containing predefined questionnaire to gather information from public crowd and it enables to receive real time updates in a specific location and helps users to monitor floods and extreme weather conditions in nearby locations during the crisis.

Since the solution gather information from the public crowd, the gathered data should be validated and analyzed. The gathered data sets will be stored in firebase live database. The data analysis will be done using Statistical Data Analysis methodologies to identify the most common and matching crowd sourcing data set based on the location and comparing datasets to identify how they are statistically different each other.

K means algorithm is used to identify the similar groups of crowdsourcing data. Using this algorithm, reduces the number false entries and find the most accu The validity of analyzed data (outcome of data analysis) will be done by comparing with live data sources such as third-party weather APIs, inputs from the proposed flood tracking and weather monitoring IOT device. Apart from these implementations active participation of crowd plays a major role in crowd source-based systems, therefore, to enforce active

participation the system will provide periodic based alerts for the users to send status of weather in their living area which will help nearby people to be in touch with the current weather situation. The final dashboard does a major role in visualizing data, and it provides real time data management After processing these series of data manipulations, the final condensed crowd sourcing data set is published through the system based on the location. The bellow mentioned flow charts highlights the data manipulation process of the crowdsourcing data.

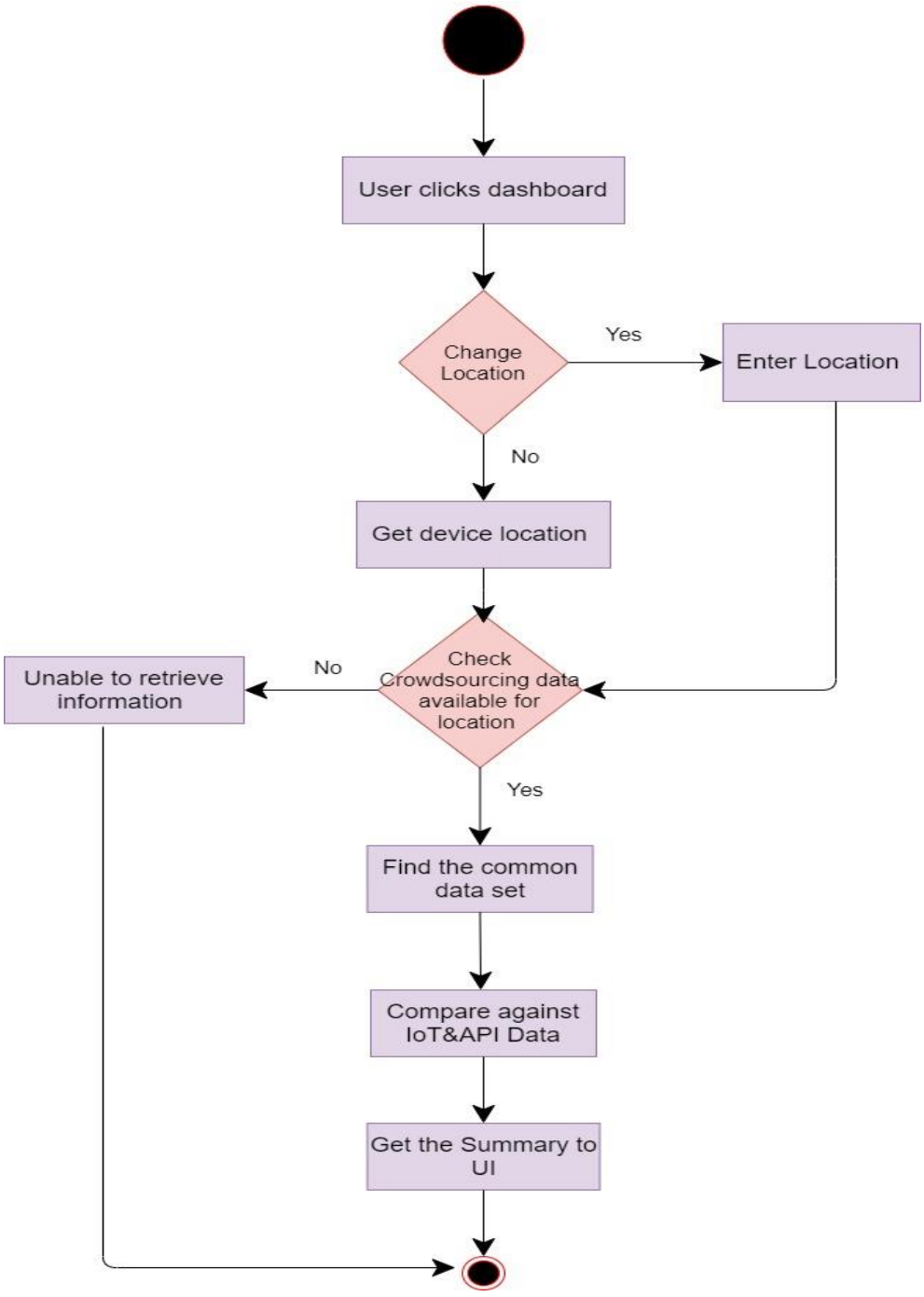

Figure 6 Crowdsourcing Data Flow

9.6 Crowdsource Data Analysis

Crowdsource data analysis is done by using K means algorithm to reduce number of false entries and get the most common data. The K means algorithm starts with first, picking the number of clusters which is known as K. After picking the K vale, next we randomly select the centroid of each cluster. As an example, if we want two clusters the K will be equal to two and then randomly select the centroid. Once centroids are initialized it assigns each data point in the cluster to the closest cluster centroid. calculating the distance from centroids to each data point is performed by Euclidean distance calculation. After assigning all the data points into two clusters the next step will be computing the centroids of newly formed clusters and it updates the new centroid. This iteration will continue till the centroid no longer changes.

In our mobile application we used K-means clustering algorithm to reduce invalid data error and get most common data from crowdsource data. When we consider current weather situation, there can be several values because of different users are experiencing different weather situations in their location. So, we must extract most common data from these different data set. To execute that using k-means algorithm, we divided our dataset into three clusters. Cluster is a set of data that contains similar data. Similarity is calculated by the distance of each data point to the centroid of the cluster. Centroid refers to the middle point of a cluster. From these 3 clusters we can find one cluster that contains more data values. It represents the cluster that contains higher number of similar kind of data. So, we can choose that cluster to achieve our goal. By choosing this cluster we can reduce error of invalid data, because invalid data are in remaining two clusters. even though we chose this cluster sometimes it will not contain exact same data. In that case again we must get most common data inside that cluster. To perform that operation, we can simply count the occurrences of each data values and get the most common value. After getting that common data value it will displayed in the text field in the relevant section of the dashboard.

9.7 Development Process

Software development life cycle (SDLC) is the approach for developing software applications. It is split up to six phases: Planning, requirement analysis, design, implementation, testing & integration, and maintenance. The phases are depending on the software scope. As per our requirements these are the main components approved for software development. Software development life cycle is the methodology for enhancing the software development process and it improves the efficiency of each phase of software development life cycle.

To develop the solution for the early warning for pre and post flood risk management system it is mandatory to adhere to proper software development methodologies. Planning stage is an important stage in the development process since that phase should be completed carefully with a proper idea of the development. Requirements gathering and analysis is done using various methodologies to gather available information and requirements for implementing flood risk management methodologies in terms of crowdsourcing. Volunteer communication, field observations are done at this phase for implementing design strategies for mobile application. Field observations are important to develop strategies for flood risk management.

Depending on the requirements, designing and sketches are done in the next phase of the software development life cycle. After selecting the technologies and tools the implementation is started adhering to latest implementation strategies and best practices. After completing implementation phase, software testing phase begins, and it will be followed by deployment and maintenance phase. The overview of the software development life cycle model which we selected to develop early warning for pre, and post flood risk management is outlined in bellow figure

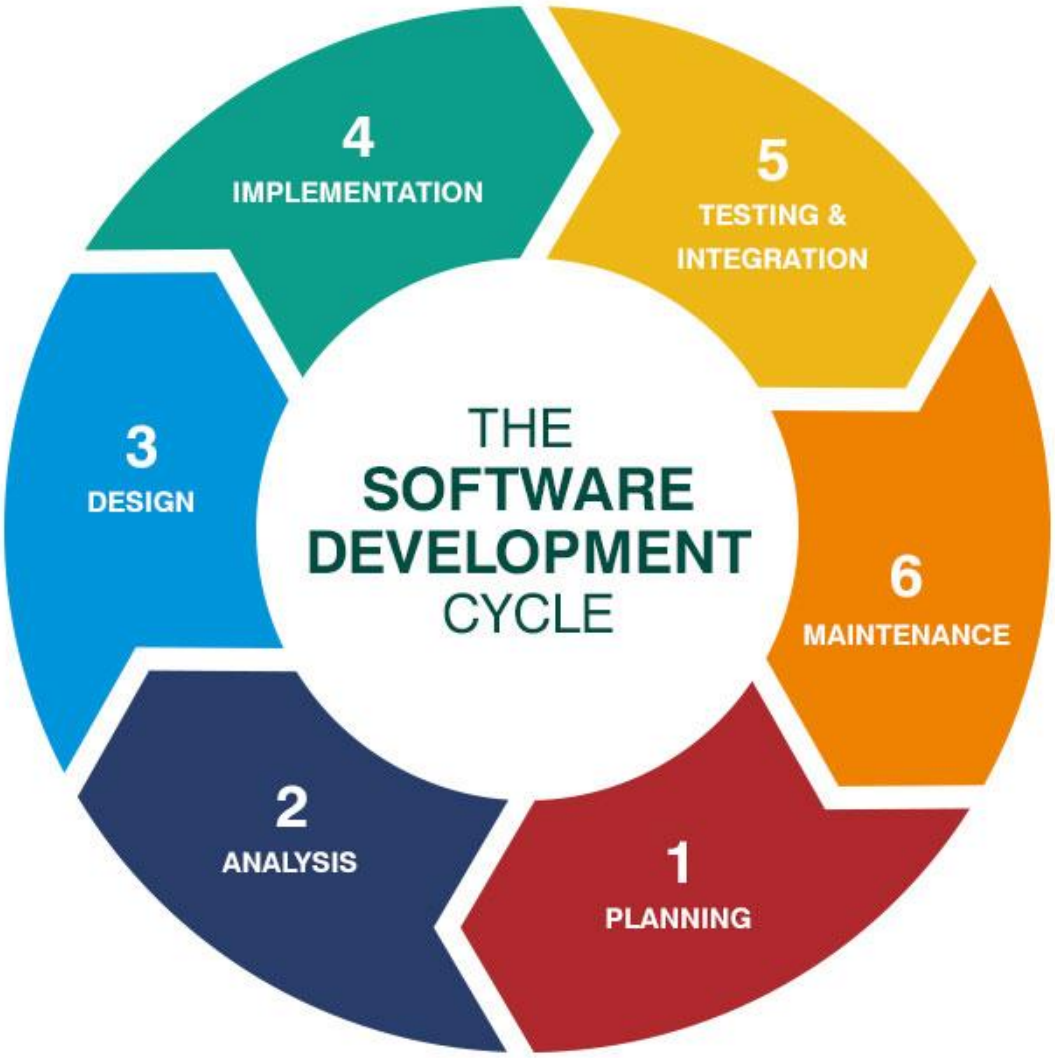

Figure 7 Software Development Lifecycle

9.8 Requirement Gathering and Data Collection

9.8.1 requirements

Requirements stands for set of features and functionalities of a target system. When defining the requirements that are needed for implementing the early warning system the research team identified what are the functional requirements and nonfunctional requirements for crowdsourcing approach in early warning system. Identified functional requirements and nonfunctional requirements are emphasized in bellow sections.

Functional requirements

Functional requirements are known as what the functions that a system should offer. Functional requirements express software system or its components. In this field of study, main functional requirements are identified as follows

- Authentication of the user
- Crowdsource data submission
- Validate crowdsourcing data
- Visualization of Crowdsource information in the dashboard

Nonfunctional requirements

Nonfunctional requirements describe the quality attribute of a system. Nonfunctional requirements determine the software system in terms of responsiveness, security, usability, and portability. Failing to implement nonfunctional requirements will lead to user dissatisfaction. Below mentioned factors are the identified nonfunctional requirements of crowdsourcing approach in the early warning system.

- Usability of the system
- Availability
- Reliability
- Performance and scalability
- Portability and compatibility

9.9 Feasibility Study

A feasibility study of a research project evaluates the feasibility of implementing the system. The main goal of conducting a feasibility study is to determine whether the proposed system is practical to implement or not. The feasibility study explores the success of the project. Feasibility study reveals whether the product is worth the investment. By conducting a feasibility study, it was identified that below mentioned factors should be considered when implementing the project

- **Schedule feasibility**

  The proposed early warning structure should be finalized and ready deploy within the time limits specified and by completing each stage of project maintaining a timeline and finally display the final outcome of the product on the arranged due date.

- **Economic feasibility**

  The proposed system should be limited to a specific budget and the final product must perform without any errors. The components need to be utilized correctly and having a low cost the quality of the final product and the benefits of the product should not be changed. The goal of the team is to provide a reliable product without much cost.

- **Technical feasibility**

  The knowledge about programming languages and statistical data analyzation methods will be needed under this section when implementing the solution

9.9.1 Software specifications

- Android Studio

  The proposed system includes an android mobile application. The mobile application plays a major role when it comes to crowd sourcing. The weather data is gathered from public crowd through the mobile application. Usability and user friendliness will be concerned more when developing the mobile application. Android studio IDE will be used as the development platform and java as the programming language.

- Database Handling

  The generated data will be stored in a database to retrieve the data when it is needed to display to end-users. Firebase will be used store these crowdsourcing data since interacting with public will generate lots of data. Firebase is used as the main data storage location

- GitLab

  GitLab is a web-based git repository management system that enables users to track the development lifecycle. GitLab is used as the main repository location of the project.

9.9.2 Hardware Specifications

Backend development process

- RAM : 8 GB RAM
- CPU : Any latest CPU
- Disk Space : 2.5 GB and another 1 GB for caches
- Monitor Resolution : 1024x768

Requirements to execute this mobile application in a desktop

- Operating System : Windows 07, Windows 10
- CPU : Any latest CPU
- Software : Android Studio

8.9.3 Communication Specifications

Weather app needs to identify users' current location for weather forecasting, Gmail, and Facebook authentication for login. Therefore, mobile application should be connected to Wi-Fi or mobile data to access these functions.

9.10 Data Collection

Requirement gathering and data collection is an important section in developing this system. Since the crowdsourcing approach directly interact with volunteers it is important to gather relevant information about flooding conditions and weather situations. Requirement gathering and data collection is done by using techniques such as conducting surveys, existing documents analyzation, prototyping. By investigating the survey results the research team gathered some important factors which cause floods and heavy rainfall and the details about flooding occurrences. As an example, which part of the year people who live in a certain area will get affected by flooding and heavy rainfall. Selecting a specific area to get crowdsourcing information about weather conditions will be mainly done through analyzing the survey results. After the implementation of the crowdsourcing solution volunteers will be trained to use the application

Gap analysis is done by reviewing the existing documents. Prototyping the solution is used to build an initial version of the solution before it gets approved and then adding additional requirements, changes will be done according to user friendliness of the system because the crowdsourcing UI will be directly interacted with the general public. The overall results that gathered from surveys and different data collection techniques is outlined below figures. Based on the gathered data, the need of implementing the early warning system was clearly highlighted from the participants responses. Bellow figures represents the participants responses

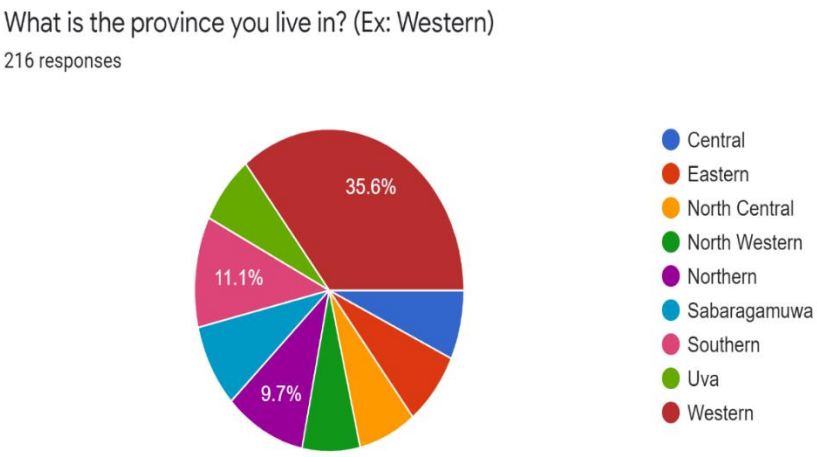

Figure 8 Provinces Selection

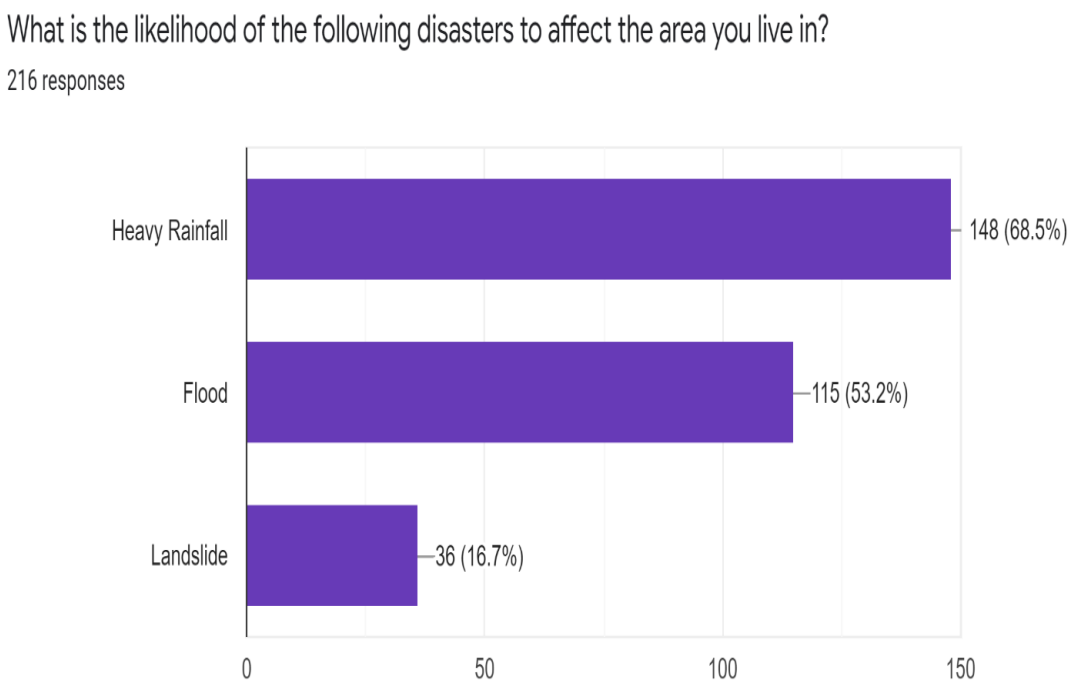

Figure 9 Likelihood of Disasters

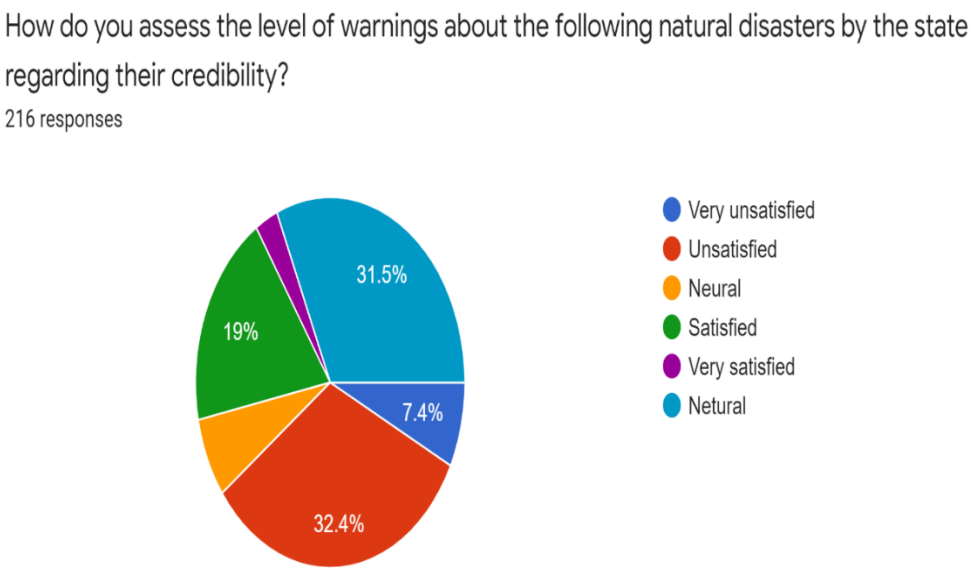

Figure 10 Warning Satisfaction

9.11 Designs

The mobile application plays a major in interacting with volunteers to obtain crowdsourcing data. This section consists with sample wireframe designs which are used in design process in the development of mobile application.

Figure 11 Login UI

Figure 12 Home UI

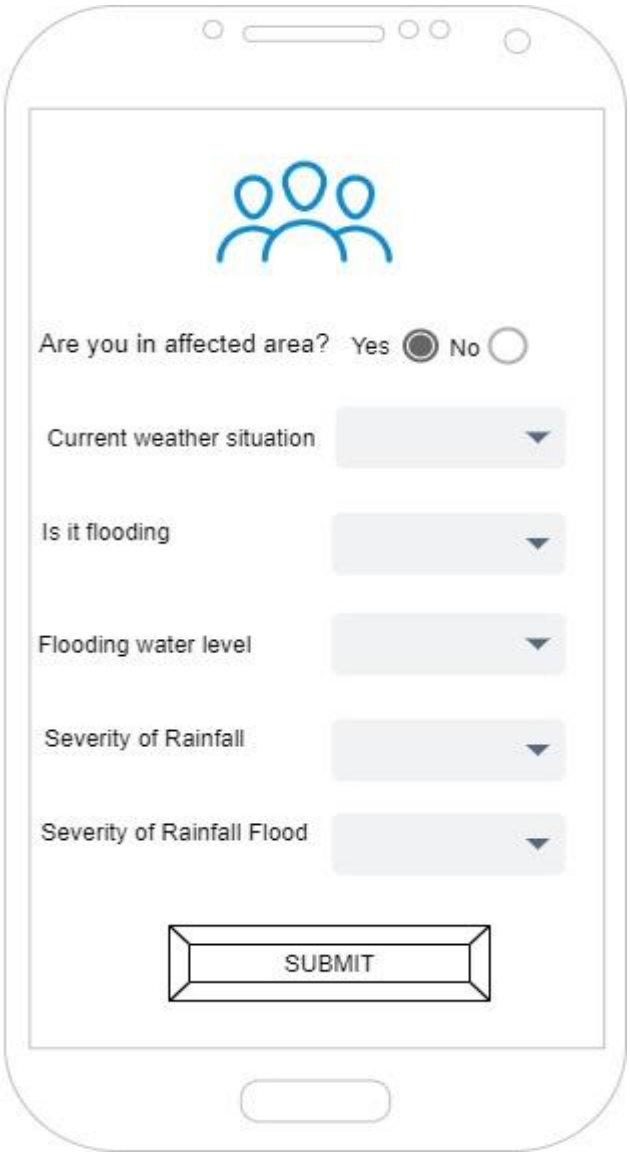

Figure 13 Crowdsourcing UI

9.12 Commercialization

This risk management software is developed as a private initiative at a cost to the developers. This cost is considered as capital employed in a business venture. Therefore, the return to the developers depends on the extent of commercial operations of the venture. Unlike in many commercial ventures that involve in production of goods and services for profit margins in markets, this software development is a social responsibility project that requires the engagement and management by the state authorities. Therefore, its commercialization involves in contracting with the relevant state authorities responsible for flood risk management to take ownership of the project or risk management system.

This can take two forms. First, the software is sold to the state authorities at a price with a service maintenance period at a fee as agreed. Accordingly, the rest of the risk management is the responsibility of the authorities. Second, the software ownership along with the patent rights is retained with the developers while an annual service price is charged for utilization. As such, the developers become software contractor to the risk management system. In either case, the developers should be able to recover the capital with a competitive rate of return on capital within a short period calculated on the basis of investment appraisals such as internal rate of return (IRR) and pay-back periods. The government will benefit by effectively managing flood risk and minimizing the damaged to people and properties that politically detrimental to the ruling parties. As such, commercialization requires skills in negotiation for selling the product to the state and the society.

10 Testing and Implementation

10.1 Testing

Software testing is a method for evaluating the functionality of software is matching its specified requirements and to check whether the software is error free. software testing is an essential part in a development process. Software testing involves using manual tools and automated tools for testing. The main goal of software testing is to identify errors, bugs, software gaps or missing software requirements. There are series of levels in software testing and those software testing strategies are outlined in bellow mentioned figure. Since this an early warning system for pre and post flood risk management, it is highly encouraged to develop the application with minimum number of errors. As a result of lack of software testing the entire crowdsourcing will fail to meet its specified requirements. Therefore, there is a need of executing software testing strategies to keep functioning the system without any interruptions.

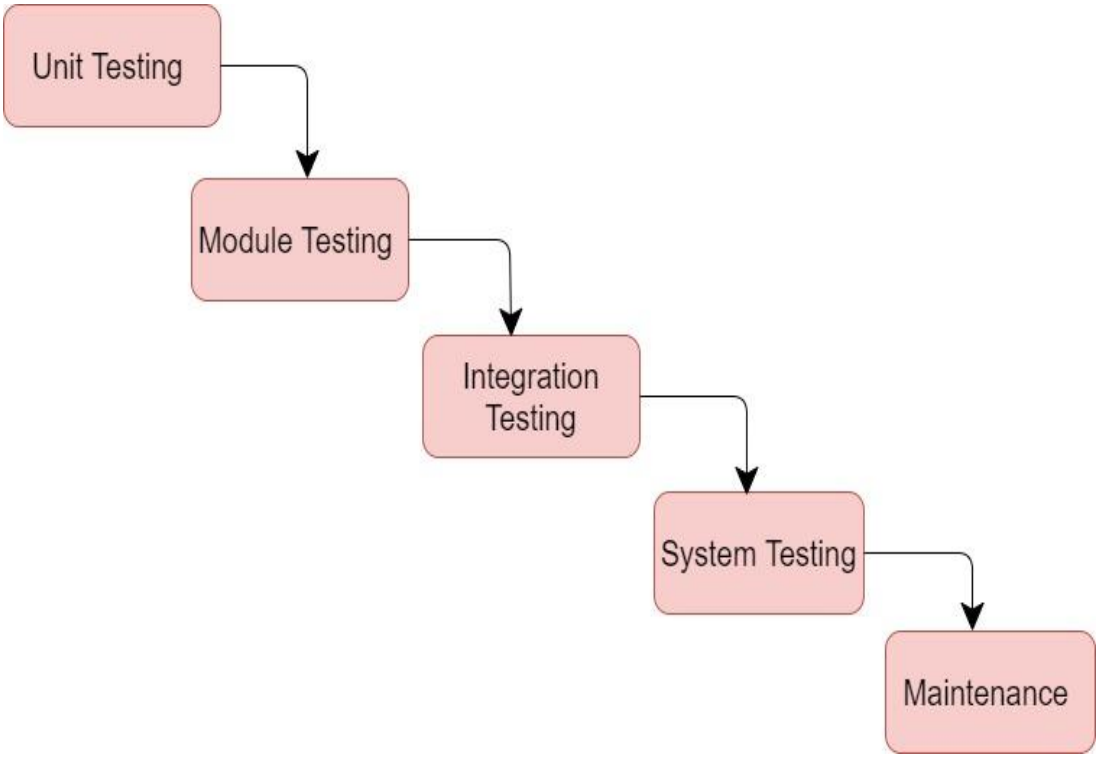

Figure 14 System Testing Flow

Unit Testing

Unit testing is a testing methodology that software is tested as individual units or components. Unit testing is the first phase of the software testing level. The main intention of unit testing is to check whether the code of each unit or the components is working as it is expected. Unit testing technique comes under white box testing. Unit testing is performed in mobile application development process to enforce better functionality of mobile application

Module Testing

Module testing refers to software testing type, which is focusing on individual sub programs, subroutines, classes, procedures in a software. The software is tested dividing into smaller modules and identifies the modules are functioning properly and error free. Module testing is highly encouraged before testing the systems as a whole

Integration Testing

The above-mentioned individual components and modules are tested as group integration testing. integration testing is essential for a mobile application since it is collection of a large group of subcomponents. The main intention of integration testing is to identify the errors when these subcomponents are executed and integrated together

System Testing

The system testing is done by integrating all the components mentioned in above testing strategies. The system is tested as whole in system testing. All the components will be integrated and tested as a whole to identify whether the system satisfy its requirements

User acceptance Testing

Since this mobile application is directly associated with users and volunteers it is mandatory to perform user acceptance testing. User acceptance testing is done at final phase after performing a successful integration testing and system testing. Performing

a user acceptance testing and getting volunteers feedback is done using this testing strategy.

Maintenance

When the mobile application deployed, volunteers and users will use the mobile application to retrieve weather information and to update weather information. Maintenance, upgrades, and improvements for the mobile application are done at this phase.

Sample Test Cases

Test case ID	01
Test case scenario	User Registration
Test steps	a. User navigate to the Registration UI b. User enters incorrect email c. Sign up
Test data	User data
Expected result	Please enter a valid email address
Actual result	As expected
Pass/Fail	Pass

Table 2 Test Case 01

Test case ID	02
Test case scenario	User Login
Test steps	a. User navigate to the Login UI b. User enters email and keep password empty c. Log in
Test data	User data
Expected result	Please enter a valid password
Actual result	As expected
Pass/Fail	Pass

Table 3 Test Case 02

Test case ID	03
Test case scenario	Weather API
Test steps	a. User navigate to the weather forecast UI b. User enters location as Malabe
Test data	Live weather data
Expected result	Successfully retrieved weather information in Malabe
Actual result	As expected
Pass/Fail	Pass

Table 4 Test Case 03

Test case ID	04
Test case scenario	Weather API
Test steps	a. User navigate to the weather forecast UI b. User retrieve current location weather data
Test data	Live weather data
Expected result	Successfully retrieved weather information in user current location
Actual result	As expected
Pass/Fail	Pass

Table 5 Test Case 04

Test case ID	05
Test case scenario	Weather API
Test steps	a. User navigate to the weather forecast UI b. User has disabled the location service
Test data	Live weather data
Expected result	Please enable the GPS location function
Actual result	As expected
Pass/Fail	Pass

Table 6 Test Case 05

Test case ID	06
Test case scenario	Crowdsource Data Submission
Test steps	a. User navigate to Crowdsource UI b. User enters whether information c. Submit information
Test data	Crowdsource data
Expected result	Submitted successfully and database has been updated
Actual result	As expected
Pass/Fail	Pass

Table 7 Test Case 06

Test case ID	07
Test case scenario	Crowdsource Data Retrieve
Test steps	a. User navigate to the dashboard UI b. User enters location as Malabe
Test data	Crowdsource data
Expected result	Successfully retrieved validated crowdsource weather information in Malabe
Actual result	As expected
Pass/Fail	Pass

Table 8 Test Case 07

10.2 Implementation

Figures below are the sample screenshots of the implemented mobile application using above mentioned methodologies.

10.2.1 UI implementations

Figure 16 Startup Screen UI

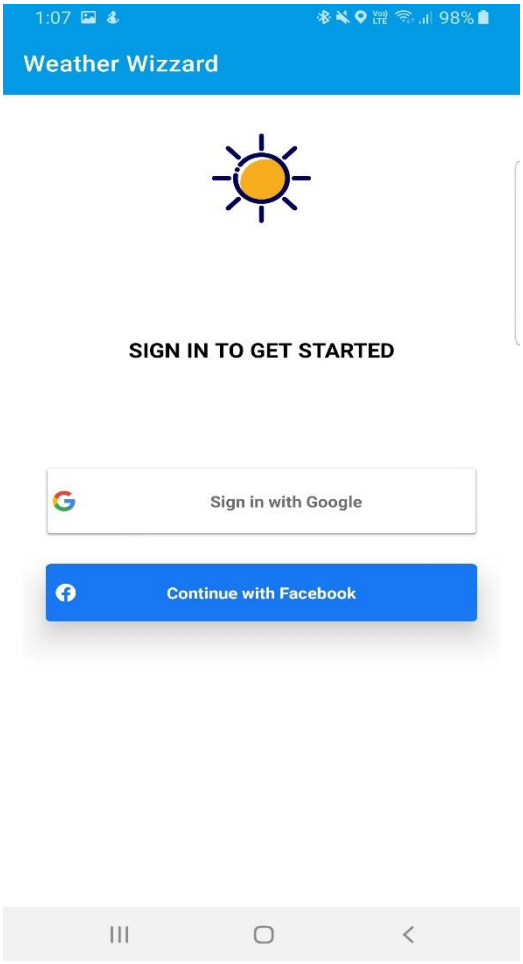

Figure 15 Login UI

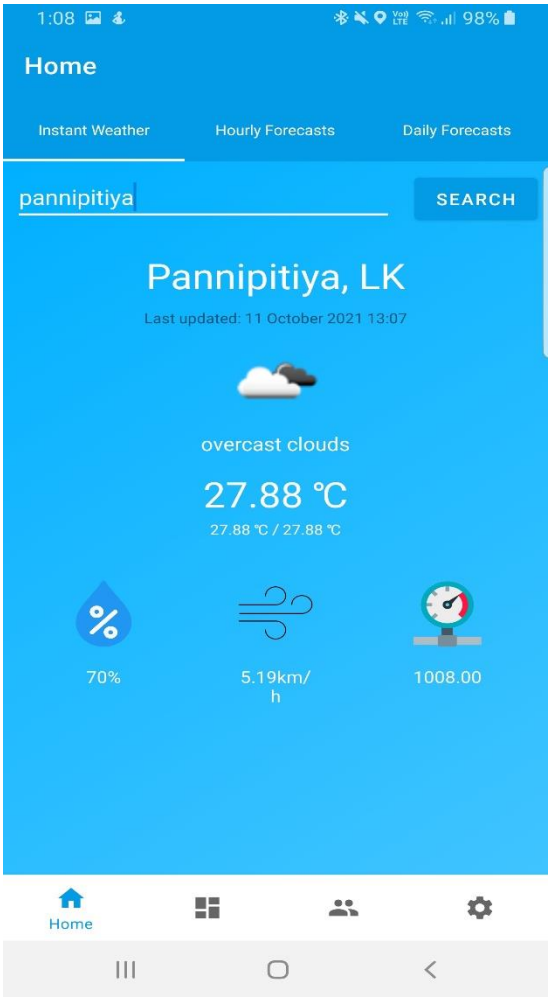

Figure 18 Instant Weather UI

Figure 17 Hourly Forecast UI

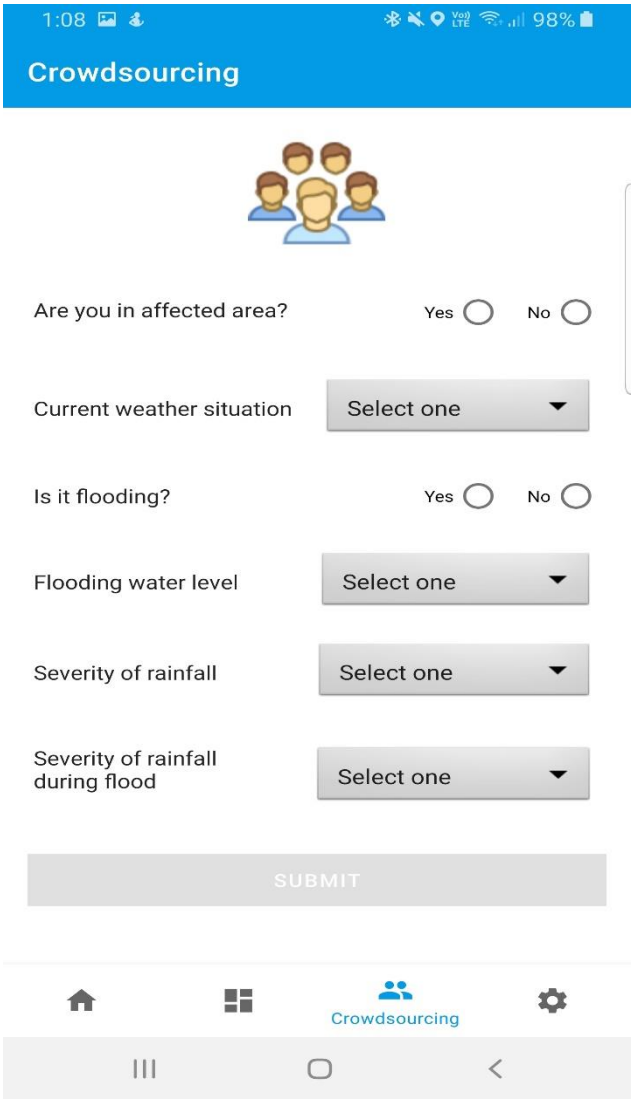

Figure 19 Crowdsourcing UI

11 Results and Discussion

11.1 Results

The crowdsourcing data is gathered through the implemented mobile application. bellow tables represent the sample data sets that we gathered in different weather conditions in three different locations by participating volunteers

Location	Weather condition	Flood	Flood water level	Severity of rainfall	Severity of rainfall with flooding
Pannipitiya	Overcast Clouds	No	0	None	None
	Light Rain	No	0	Low	None
	Heavy Rain	No	0	High	None

Table 9 Crowdsource Dataset 01

Location	Weather condition	Flood	Flood water level	Severity of rainfall	Severity of rainfall with flooding
Nugegoda	Clear Sky	No	0	None	None
	Thunderstorm	No	0	High	None
	Heavy Rain	No	0	High	None

Table 10 Crowdsource Dataset 02

Location	Weather condition	Flood	Flood water level	Severity of rainfall	Severity of rainfall with flooding
Maharagama	Overcast Clouds	No	0	None	None
	Light Rain	No	0	Low	None
	Heavy Rain	No	0	High	None

Table 11 Crowdsource Dataset 03

The above tables represent the gathered crowdsourcing datasets. The next step is to visualize gathered crowdsourcing results. The below UI shows the visualization dashboard of the gathered crowdsourcing results in a specific location

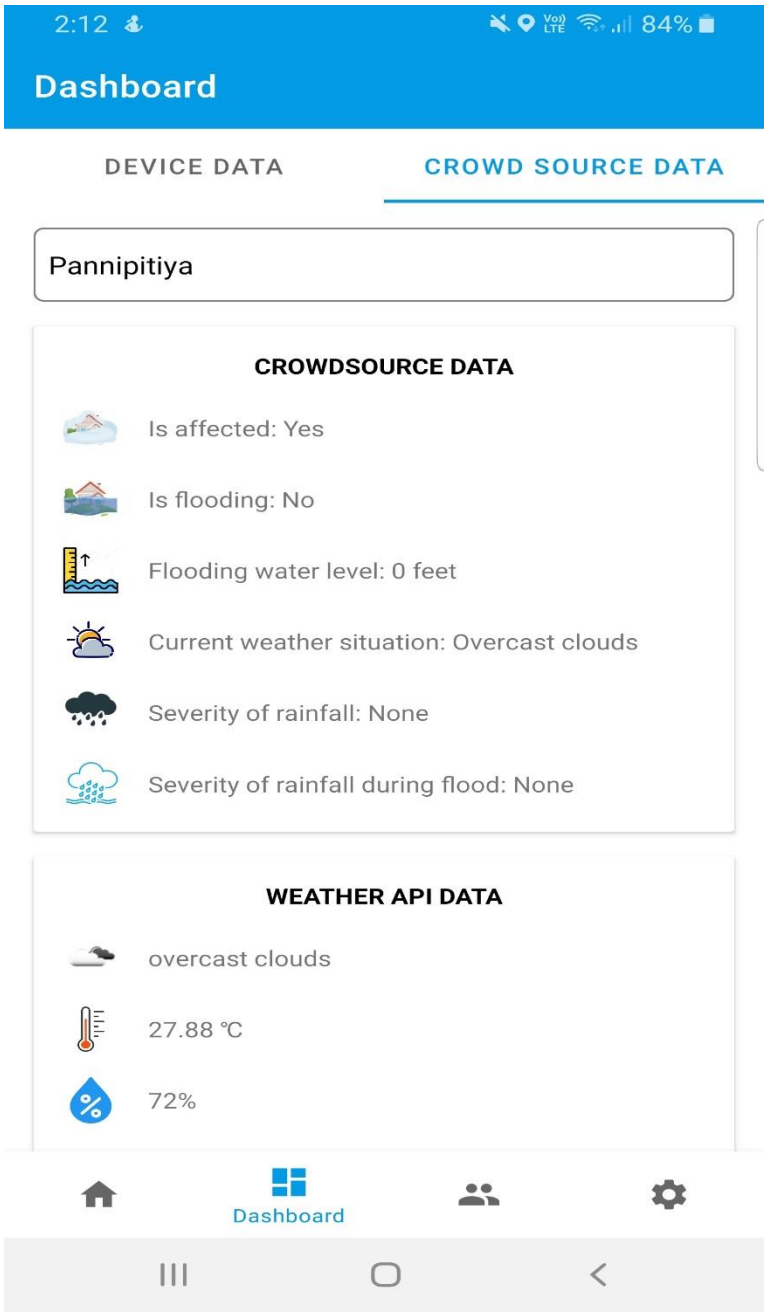

Figure 20 Crowdsourcing Dashboard

11.2 Research Findings

The above tables represent the crowdsourcing data sets that have been gathered in different weather conditions in cities. Integrating mobile technology with crowdsourcing has experimented a positive impact in public during a disastrous events. Authors identified crowdsourcing approach is a conventional way of collecting information in a natural disaster. The outcome of the crowdsourcing data is capable of saving millions of lives by promoting availability of disaster data. The mobile application helps users to view and track weather conditions and flood information in specific areas.

Based on the research results authors have come to a decision that crowdsourcing platform is effective in collecting accurate volunteer information with the help of majority decision function. Majority decision is used to reduce number of false entries. Since the mobile application gather crowdsource data from citizens it is important reduce number of false entries. Majority decision function is based on a K means clustering algorithm to eliminate number of false entries and analyze the most common crowdsource data set. When displaying the crowdsource data, majority decision function experimented positive results by displaying most accurate crowdsource data set in the mobile dashboard. These accurate crowdsource data sets will support the decision-making procedures in disaster risk management

11.3 Discussion

We tested the mobile application by participating volunteers in Colombo area and gathered responses in different weather conditions accordingly. Using the mobile application and the location tracking function enables users to submit correct and accurate information. Based on the gathered data sets it will enforce disaster risk limiting. Disaster risk limiting can directly benefit from crowdsourcing as an example by actively participating and communicating information will overcome unavailability of information before, during and after an extreme weather situation. Finally integrating overall crowdsourcing data sets will be a great aid for increasing the accuracy of other weather forecasts and early warning systems.

11.4 Summary of Student Contribution

Member	Component	Task
V.Y Samarasiri	Development of crowdsourcing component	Feasibility study to perceive the requirement of the aspect. Implement the mobile app user registration Establishing connection to the open weather map API, using API unique key and retrieve live weather data and display using a suitable way Finalize crowdsourcing data that has to be gathered from volunteers Implement the crowdsourcing UI Save crowdsource data and retrieve crowdsourcing weather data from DB and find the common data set and validate the data set against IOT data & weather API data Implement the dashboard for visualize validated crowdsource data

Table 12 Student Contribution

12 Conclusion

After conducting a successful research on crowdsourcing for flood risk management, the presented work fills the gap that has been mentioned in the research gap section in terms of crowdsourcing approach. In this crowdsourcing approach volunteers are taken as human sensors that will contribute information about metrological factors such as weather condition, flooding water level and flooded areas. Regarding the crowd sourcing-based approach further improvements and research can focus in these core areas. Weather forecast efficiency can be improved in collecting more factors from the crowd such as picture uploading function in a crisis with combining qualitative and quantitative inputs from the crowd and it will further benefit early warning systems. The success of the crowd sourcing solution depends on active participation of the crowd therefore there is a need for enabling and improving the contributing information in a disaster event. Further research can be done in implementing an offline method for involving crowd in a natural disaster and expanding the involvement in crowd for more cities.

References

[1] H. Thilakarathne and K. Premachandra, “Predicting Floods in North Central Province of Sri Lanka using Machine Learning and Data Mining Methods,” no. October, pp. 44–49, 2017.

[2] E. J. Plate, “Flood risk and flood management,” *J. Hydrol.*, vol. 267, no. 1–2, pp. 2–11, 2002, doi: 10.1016/S0022-1694(02)00135-X.

[3] E. M. Mendiondo, “Flood Risk Management of Urban Waters in Humid Tropics: Early Warning, Protection and Rehabilitation,” *Work. Integr. Urban Water Manag.*, no. April, pp. 2–3, 2005.

[4] A. García-olivares, “Self-Organization of Complex Systems,” no. July, 2014.

[5] K. B. Wells, B. F. Springgate, E. Lizaola, F. Jones, and A. Plough, “Community Engagement in Disaster Preparedness and Recovery: A Tale of Two Cities – Los Angeles and New Orleans,” *Psychiatr. Clin. North Am.*, vol. 36, no. 3, pp. 451–466, Sep. 2013, doi: 10.1016/J.PSC.2013.05.002.

[6] Z. Song, H. Zhang, and C. Dolan, “Promoting disaster resilience: operation mechanisms and self-organizing processes of crowdsourcing,” *Sustain.*, vol. 12, no. 5, pp. 1–14, 2020, doi: 10.3390/su12051862.

[7] S. Frigerio, L. Schenato, G. Bossi, M. Mantovani, G. Marcato, and A. Pasuto, “Hands-on experience of crowdsourcing for flood risks. An android mobile application tested in Frederikssund, Denmark,” *Int. J. Environ. Res. Public Health*, vol. 15, no. 9, 2018, doi: 10.3390/ijerph15091926.

[8] J. Pedersen *et al.*, “Conceptual foundations of crowdsourcing: A review of IS research,” *Proc. Annu. Hawaii Int. Conf. Syst. Sci.*, pp. 579–588, 2013, doi: 10.1109/HICSS.2013.143.

[9] M. N. Fienen and C. S. Lowry, “Social.Water—A crowdsourcing tool for environmental data acquisition,” *Comput. Geosci.*, vol. 49, pp. 164–169, Dec. 2012, doi: 10.1016/J.CAGEO.2012.06.015.

[10] L. Machado, J. Kroll, R. Prikladnicki, C. R. B. De Souza, and E. Carmel, “Software crowdsourcing challenges in the Brazilian IT industry,” *ICEIS 2016 - Proc. 18th Int. Conf. Enterp. Inf. Syst.*, vol. 1, no. April, pp. 482–489, 2016, doi: 10.5220/0005835004820489.

[11] L. C. Degrossi, J. P. De Albuquerque, M. C. Fava, and E. M. Mendiondo, “Flood citizen observatory: A crowdsourcing-based approach for flood risk management in Brazil,” Proc. Int. Conf. Softw. Eng. Knowl. Eng. SEKE, vol. 2014-January, no. January, pp. 570–575, 2014.

[12] W. T. Liang, J. C. Lee, and N. C. Hsiao, “Crowdsourcing platform toward seismic disaster reduction: The Taiwan scientific earthquake reporting (TSER) system,” Front. Earth Sci., vol. 7, no. April, pp. 1–12, 2019, doi: 10.3389/feart.2019.00079.

[13] A. Schulz, H. Paulheim, and F. Probst, “Crisis information management in the Web 3.0 age,” ISCRAM 2012 Conf. Proc. - 9th Int. Conf. Inf. Syst. Cris. Response Manag., no. April, pp. 2–6, 2012.

[14] O. Mejri, S. Menoni, K. Matias, and N. Aminoltaheri, “Crisis information to support spatial planning in post disaster recovery,” Int. J. Disaster Risk Reduct., vol. 22, pp. 46–61, Jun. 2017, doi: 10.1016/J.IJDRR.2017.02.007.

[15] M. Careem, C. De Silva, R. De Silva, L. Raschid, and S. Weerawarana, “Sahana: Overview of a disaster management system,” 2nd Int. Conf. Inf. Autom. ICIA 2006, no. January, pp. 361–366, 2006, doi: 10.1109/ICINFA.2006.374152.