CHAPTER I
INTRODUCTION

This chapter aims to provide a general overview of the importance of developing a local data-based rainfall prediction system in the agricultural sector. Weather uncertainty caused by global climate change makes it difficult for farmers to determine the appropriate time for irrigation, fertilization, and pesticide spraying, thereby potentially reducing productivity and resource efficiency. To address this issue, this study proposes a rainfall prediction system based on the Internet of Things (IoT), LoRa, and the Extreme Gradient Boosting (XGBoost) algorithm as a solution capable of providing daily rainfall information to support farmers' decision-making.

1.1 Background

Weather is an important factor that affects various aspects of human life, ranging from the transportation, trade, and tourism sectors. Accurate and timely weather information is highly needed to support community activities, reduce disaster risks, and support more effective development planning [1], [2]. In tropical climate regions such as Indonesia, weather conditions often change suddenly, making them difficult to predict accurately. Discrepancies between forecasts and actual conditions can cause various impacts, such as transportation delays, disruption of economic activities, and public health problems [3]. Among the various affected sectors, agriculture is one of the fields most vulnerable to weather uncertainty. This is because agriculture depends heavily on weather and environmental conditions [4].

The agricultural sector is the backbone of national food security and also plays an important role in supporting the economy and public welfare. Despite its strategic role, this sector still faces various challenges that hinder its sustainability and productivity. One of the main challenges is the uncertainty of weather conditions, which are becoming increasingly difficult to predict due to climate change. Unpredictable weather often makes it difficult for farmers to properly plan agricultural activities. In addition, the inefficient use of resources such as water, fertilizers, and pesticides also contributes to high production costs and reduced crop yields [4].

To address these challenges, the concept of smart agriculture or Smart Farming 4.0 has emerged as a new approach to managing the agricultural sector by utilizing digital technologies. These technologies include the Internet of Things (IoT), Artificial Intelligence (AI), and Machine Learning (ML), which are integrated to support decision-making processes in the field [5]. The purpose of this approach is to increase productivity, reduce operational costs, and optimize the use of resources such as water, fertilizers, and pesticides. With an accurate and adaptive monitoring system, agricultural activities can be carried out more effectively and precisely. This ultimately contributes to increased crop yields while minimizing resource waste [4].

However, the implementation of this approach in the field still faces various challenges, particularly related to limitations in technological infrastructure. The available internet network generally cannot be directly utilized in agricultural areas due to the long distances between devices and the absence of communication channels such as cables or WiFi networks that can reach sensor locations. In addition to these technological infrastructure limitations, farmers are still required to make quick decisions in response to rapidly changing weather conditions. In such situations, the timing of irrigation and pesticide spraying becomes highly crucial. If these activities are carried out shortly before or during rainfall, water and pesticides may be washed away by runoff, making them ineffective, causing environmental pollution, and resulting in wasted labor, time, and costs [6]. On the other hand, global climate change has led to increasing uncertainty in rainfall patterns, particularly in tropical regions, making them more difficult to predict. This condition poses a major challenge to the agricultural sector because it affects daily farming practices such as planting, irrigation, and fertilization processes [7].

Based on these problems, this study proposes a rainfall prediction system designed to provide weather forecast information for the following day by utilizing sensor data collected hourly over a continuous 24-hour period. This system is expected to assist farmers in determining the optimal time to carry out agricultural activities, thereby reducing the risk of losses and improving farming efficiency.

Internet of Things (IoT) technology enables the collection of weather data, including temperature, humidity, and wind speed parameters, directly from agricultural land through SHT31 and anemometer sensors. The data are transmitted to an AWS EC2-based cloud server using LoRa (Long Range). LoRa was selected as a communication solution because it is capable of transmitting data over long distances with low power consumption, making it highly suitable for implementation in large agricultural areas with limited internet infrastructure [8]. In this study, AWS EC2 functions as both a cloud server and a gateway that connects the sensor devices to the prediction system. It is used to receive data from LoRa, store the data, and run the XGBoost prediction model.

The measurement data are processed using the Extreme Gradient Boosting (XGBoost) algorithm to generate rainfall predictions. XGBoost is an ensemble learning algorithm developed to overcome the limitations of single decision tree models such as J48. This algorithm utilizes the gradient boosting method with multi-threading support to accelerate computation, sparse-aware processing to handle missing data, and regularization to improve accuracy and prevent overfitting. These advantages make XGBoost effective for modeling complex and diverse weather data [9]. Based on previous research [10], the XGBoost algorithm has been proven to outperform other methods in rainfall forecasting because it is capable of modeling nonlinear and complex data patterns. In addition, XGBoost demonstrates stable and accurate performance even when the available training data are relatively limited. Through this approach, the proposed rainfall prediction system is expected to support the implementation of digital technology in the agricultural sector.

1.2 Research Problems

Based on the background that has been described, the research problems in this study are formulated as follows:

1. How can a rainfall prediction system architecture based on the Internet of Things (IoT) and LoRa data transmission be designed to predict rainfall in agricultural areas?

2. How can the Extreme Gradient Boosting (XGBoost) algorithm be implemented to predict the probability of rainfall based on weather data obtained from IoT sensors?
3. How does the rainfall prediction system process weather data to generate prediction outputs that support agricultural activities?

1.3 Research Objectives

The objectives of this study are as follows:

1. To design and develop an IoT-based sensor system to collect weather data from agricultural areas and transmit the data using LoRa.
2. To implement the Extreme Gradient Boosting (XGBoost) algorithm to predict the probability of rainfall based on data obtained from IoT sensors.
3. To design a rainfall prediction system mechanism that processes weather data into prediction outputs to support farmers' decision-making.

1.4 Research Benefits

This study is expected to provide benefits to several parties, including:

1.4.1 For the Author

1. To provide practical experience in developing an integrated system based on IoT, LoRa transmission, and machine learning.
2. To enhance understanding of the implementation of the Extreme Gradient Boosting (XGBoost) algorithm for weather prediction based on environmental data.
3. To serve as a medium for applying informatics knowledge to real-world solutions in the agricultural sector.

1.4.2 For Readers

1. To provide insight into the application of IoT and machine learning technologies in precision agriculture.
2. To serve as a reference for researchers or developers of weather data-based systems to support agricultural activities.
3. To provide an understanding of the integration between environmental sensors, LoRa transmission, and weather prediction algorithms.

1.4.3 For Farmers and Agricultural Practitioners

1. To support decision-making based on weather predictions so that agricultural activities can be carried out more efficiently.
2. To reduce the potential waste of resources caused by activities performed under unfavorable weather conditions.
3. To improve productivity through the support of affordable and easy-to-use weather prediction technology.

1.5 Research Limitations

This study has the following limitations:

1. The developed system is a prototype intended for testing purposes, focusing on the functionality of data collection, transmission, and processing from sensors.
2. The prediction system is developed using the Extreme Gradient Boosting (XGBoost) algorithm as the main focus of the study without discussing comparisons with other methods.
3. The research data used for training and testing the model are obtained from BMKG historical datasets, while the primary sensor data are used for end-to-end system testing and are not utilized in the model training process.