CHAPTER I
INTRODUCTION

1.1.Background

The limitations of weather forecasting accuracy in Indonesia have become a public interest, particularly when the Indonesian Meteorological, Climatological, and Geophysical Agency (BMKG) forecasts deviate from reality. BMKG acknowledges the inherent challenges associated with predicting weather in Indonesia. These challenges stem from the country's unique geographical location: its equatorial position leading to unpredictable and rapidly fluctuating weather patterns, and its archipelagic nature surrounded by the vast Indian and Pacific Oceans [1].

Prior research by Jumianti et al. [2] employed three brightness temperature (BT) bands (B11, B13, B15) to evaluate the performance of three different Brightness Temperature Difference (BTD) combinations (BTD1, BTD2, BTD3). Their findings revealed a significant influence of temperature and rainfall intensity on BT accuracy. Notably, BT exhibited superior performance for light rain at a temperature of 250 K. BTD1 demonstrated improved accuracy, enabling extreme rainfall prediction with a 10-20 minute lead time. However, BTD2 showed potential for extreme rainfall anticipation but resulted in a higher False Alarm Rate (FAR) for moderate to heavy rainfall events. Conversely, BTD3 generally yielded improved accuracy and Critical Success Index (CSI), offering a 10-20 minute lead time for heavy rainfall predictions. These results highlight the importance of selecting appropriate BTD combinations for optimal prediction accuracy.

Putranto et al. [3] investigated a deep learning approach utilizing a Convolutional Long Short-Term Memory (ConvLSTM) model with an encoder-forecaster architecture. This model leveraged Himawari-8 and Rapidly Developing Cumulus Area (RDCA) data for heavy rain prediction. Their findings demonstrated a positive correlation between the number of training iterations and model performance. Additionally, the model generated probability values during testing that closely resembled the RDCA index values. This high degree of similarity, reflected in a

Structural Similarity Index Measure (SSIM) value approaching one, suggests the effectiveness of the ConvLSTM model in capturing the underlying patterns within the data.

Kawasaki and Wang's research [4] explored the application of a Convolutional Long Short-Term Memory (ConvLSTM) model for insolation forecasting using Himawari-8 cloud image data. They compared the forecast results obtained with their proposed method to those generated without satellite image data and those incorporating only the latest satellite image data. Their findings revealed that integrating satellite image data and cloud image forecasts improved the understanding of insolation characteristics, leading to a reduction in erroneous predictions and a subsequent increase in prediction accuracy.

Leveraging insights from previous research, this study seeks to develop an accurate heavy rain prediction model for Bandung City, Indonesia. The model will utilize cloud data from the Himawari-8 satellite and employ deep learning algorithms. The performance of the LSTM forecasting architecture will be compared for heavy rain prediction for 1-hour and 2-hour predictions, using 13 key heavy rain prediction parameters derived from satellite data (channels 3, 8, 10, 11, 13, 15, and 16). In addition, it also provides recommended parameter combinations for heavy rain prediction with the aim being to optimize these parameters to improve prediction accuracy.

1.2.Problem Identification

Based on the background that has been described previously, the following problems can be identified:

1. How to minimize the lack of accurate heavy rain prediction models or existing methods for predicting heavy rain in Bandung City?
2. How to find the best model with LSTM prediction for heavy rain in the next 1 hour or 2 hours?
3. How to determine the appropriate combination of heavy rain detection parameters for heavy rain prediction in Bandung City?

1.3.Objective

The objectives obtained from the thesis are as follows:

1. Developing an accurate heavy rain prediction model with cloud data obtained from Himawari-8 satellite data and deep learning algorithms in the city of Bandung.
2. Comparing LSTM algorithm for heavy rain prediction between 1-hour prediction model and 2-hour prediction model in Bandung city.
3. Provide recommendations for the combination of 13 parameters as features in the heavy rain prediction model for the Himawari-8 satellite 1-hour and 2-hour models.

1.4.Scope Limitations

The scope of this research is limited by several factors, as follows:

1. The study's focus on Bandung City limits the generalizability of the findings to other regions with potentially different weather patterns and precipitation characteristics.
2. The analysis is restricted to 13 key rain detection parameters derived from Himawari-8 satellite data.
3. The study employs solely the LSTM algorithm for heavy rain prediction.
4. This research only uses bands 3, 8, 10, 11, 13, 15, and 16 of the Himawari-8 satellite data.
5. This research uses the Himawari 8 satellite dataset in CSV format, which has 8,738 rows and 14 columns (13 feature columns and 1 label column) of data.
6. The training architecture with LSTM uses 50 hidden layers and 100 hidden layers with 20% test data and 80% training data.
7. This research also uses early stopping with a minimum delta value of 0.001 and monitors the validation loss.
8. The test data uses new data, which has 669 rows and 14 columns (13 feature columns and 1 label column).

1.5.Hypothesis

The LSTM model can predict the intensity of heavy rain with higher accuracy for longer periods than shorter ones. Since the LSTM is designed to capture long-term dependencies in time data, the model will likely better predict heavy rain for longer periods compared to shorter periods. At longer periods, external factors and higher variability may affect the prediction accuracy. In this research, 3 time periods will be used, namely 1-hour, 2-hours, and 3-hours.

The use of additional weather features (such as temperature and cloud data) will improve the performance of the LSTM model in predicting heavy rain for all three time periods (1 hour, 2 hours, and 3 hours). By adding additional relevant features, the LSTM model can utilize more information to make more accurate predictions. These features can provide additional context that assists the model in identifying heavy rainfall patterns.

In addition, LSTM models trained with longer historical data (e.g., several years) will provide better predictions for heavy rainfall in longer periods compared to models trained with shorter historical data. Models trained with longer historical data have access to more seasonal and long-term patterns, which can help improve prediction accuracy for longer periods.

1.6.Contribution

Some of the contributions made in this thesis research are described in the following points:

1. Developed an accurate heavy rain prediction model, the model can aid in improving preparedness and risk mitigation associated with heavy rain events in Bandung.
2. Compared LSTM Forecasting Algorithm for 1 hour prediction and 2-hour prediction, the comparison provides insights into the effectiveness of different deep learning algorithms in the context of heavy rain prediction, which can inform the selection of optimal algorithms for similar applications in the future.

3. Providing a recommended combination of 13 parameters to feature in Himawari-8 satellite heavy rain prediction, this optimization identifies the most significant subset of parameters to improve the accuracy of heavy rain prediction, thereby increasing the usefulness of Himawari-8 satellite data for weather prediction.

1.7.Research Methodology

This thesis is divided into four work packages to produce high-quality results.

1. Literature study: Searching for information related to this thesis sourced from books, journals, and discussions to facilitate the completion of this thesis.
2. System design: This simulation system design uses Python 3.9 and additional software such as FileZilla Client version 3.63.2.1.
3. System analysis: Observing the results of system testing according to scenarios and parameters, concluding the problems in this thesis.
4. Conclusion: From all the steps carried out above with input from the supervisor, conclusions can be drawn from the results that have been carried out.