CHAPTER I
INTRODUCTION

1.1 Background

Pasuruan Regency is one of the regencies in East Java with high vulnerability to hydrometeorological disasters, especially in the rainy season [1] . Erratic rainfall often causes hydrometeorological disasters, such as floods, flash floods, and landslides in some areas, resulting in disruption of economic and social activities of the community. This condition is exacerbated by the diverse and complex topography of Pasuruan Regency, which affects the spatial and temporal distribution of rainfall [2]. In addition, global climate change that has had an impact on weather patterns in East Java, including Pasuruan [3], has resulted in an increase in rainfall intensity by an average of 15-20 percent in the 2011-2022 period compared to the 1981-2010 period, with a trend that is increasingly non-linear and difficult to predict [4]. This condition confirms that rainfall predictions are increasingly complicated and full of uncertainty. Therefore, rainfall forecasts are not only a technical meteorological instrument, but also a basis for decision-making because they have a significant impact on various aspects of life [5].

In the 2020-2025 period, Pasuruan Regency experienced a significant increase in the frequency of hydrometeorological disasters [6]. Based on an official report from the Regional Disaster Management Agency (BPBD) of Pasuruan Regency, throughout 2025 there have been a total of 448 disaster events [7]. This figure places Pasuruan as one of the disaster red zones in East Java [1]. The number of disasters reflects a failure to anticipate the dynamics of rainfall. A total of 112 cases of flooding, inundation of floods paralyzed residential areas and road infrastructure, while 101 landslides destroyed access in mountainside areas [7]. Behind the number of disasters, the community bears the burden of material and non-material losses due to inaccurate weather predictions. The affected groups are food crop farmers, salt farmers and communities in disaster-prone areas.

This condition shows that the high frequency of hydrometeorological disasters in Pasuruan Regency is inseparable from the limitations of the rainfall prediction system. An effective disaster mitigation strategy requires accurate weather forecast information as the basis for community and related agency preparedness [8]. The Meteorology, Climatology, and Geophysics Agency (BMKG) through the online data website BMKG Pasuruan Geophysics Station has provided standardized daily meteorological data according to international standards [9]. However, the availability of existing historical data has not been optimally utilized to build a prediction system that is able to anticipate rainfall movements. Accurate and reliable rainfall information can help reduce the impact of disasters while providing guidance for stakeholders in decision-making.

Research related to rainfall prediction using traditional statistical methods such as ARIMA (Autoregressive Integrated Moving Average) is a conventional method that is widely used in forecasting rainfall time series [10]. However, ARIMA has limitations in capturing non-linear patterns and temporal complexity found in fluctuating rainfall data [10]. Other approaches such as Triple Exponential Smoothing (TES) have also been carried out for rainfall prediction, but the results have not been sufficient to handle dynamic patterns [11]. The limitations of these conventional models encourage research with a more adaptive approach and are able to model long-term temporal dependence.

The deep learning approach then became a more promising solution due to its ability to capture long-term temporal dependencies on complex and non-linear time series. One of the most widely used algorithms for time series data analysis is Long Short-Term Memory (LSTM) [12]. A study by Badriyah et al. (2022) used a multi-attribute LSTM to predict rainfall in the city of Surabaya and obtained the evaluation results of MSE 0.489 and MAE 0.537 [13], confirming that LSTM is able to model rainfall time series data although accuracy can be improved. Another study by Patro and Bartakke (2024) applied an AWS data-driven LSTM for daily precipitation prediction and found that the model was able to capture temporal patterns well using the Adam optimizer [14]. In a study conducted by Arassah et al. (2025), applying a combination of SVC and RNN reinforced by the Random Forest (missForest) imputation method to predict rainfall in Bogor, it succeeded in lowering the RMSE to 16.19 and achieving a classification accuracy of 98.5% [15]. The three studies show great potential in deep learning while revealing one weakness, namely, that hyperparameter optimization has not been implemented systematically and globally.

Based on the review of these studies, there are gaps that need to be overcome, the main limitation in conventional [13] LSTM research lies in the hyperparameter tuning process which is still manual or based on simple grid search. This approach tends to be less efficient because it only explores a small portion of the available parameter space, with no guarantee that the combination found is the optimal solution globally. In addition, grid search has increasing computational complexity as the number of parameters optimized increases, making it less practical for complex architectures [16]. The complexity of rainfall patterns in areas with diverse topography such as Pasuruan Regency demands an optimization approach that is able to explore parameter space in an adaptive, efficient and comprehensive manner to produce a model that is representative of local climatic characteristics.

In this case, Particle Swarm Optimization (PSO) provides an alternative optimal parameter search mechanism based on the collective behavior of particles that move iteratively towards the best solution [16]. Compared to the gradient-based optimization method, PSO does not require objective function derivative information so that it is more flexible and robust for the search space. The effectiveness of PSO in LSTM optimization has been proven by previous research. Umamaheswari (2025) showed that the integration of LSTM-PSO was able to significantly improve the performance of rainfall prediction, characterized by the results of the MAE evaluation of 16.43 and RMSE of 20.78 as well as an increase in the correlation coefficient compared to conventional LSTM, with PSO that succeeded in optimizing hyperparameters such as learning rate, number of neurons, and epoch more convergently and accurately [17]. Jangid P et al. (2023) developed a hybrid model of LSTM-Modified PSO and noted a decrease compared to LSTM without optimization, with an MSE evaluation value of 8,146 and an RMSE of 2,009, with a more stable model in handling non-linear data [18]. Wu et al. (2023) developed the EEMD-LSTM-PSO model for long-term hydrometeorological time series prediction and produced better RMSE, MAPE and MSE than ARIMA, single

LSTM, and other combination models [19]. These three studies confirm that PSO is able to improve the performance of LSTM measurably in the context of rainfall prediction and hydrometeorology.

Based on the literature study that has been conducted, the LSTM-PSO integration provides increased accuracy as well as computational efficiency in the modeling process that is relevant for daily rainfall prediction, so that research was carried out related to the Optimization of the Long Short-Term Memory Algorithm using Particle Swarm Optimization for Rainfall Prediction in Pasuruan Regency. The data that will be used in this study comes from the online data website of the BMKG Geophysics Station of Pasuruan Regency, covering the variables of daily average temperature, daily average humidity, and daily rainfall for the period January 1, 2021 to December 31, 2025 [20]. The data will go through the pre-processing stage including handling missing values using Random Forest Imputation, transformation and normalization before being used as an input model for the LSTM-PSO model. The results of the model's prediction will be evaluated using RMSE, MAE, and NSE. This research is expected to produce a rainfall prediction model that is adaptive to climate change and weather patterns, as well as a decision-making support tool in mitigating the risk of hydrometeorological disasters in Pasuruan Regency.

1.2 Problem Formulation

Based on the background that has been stated earlier, the formulation of the problem in this study is as follows:

1. What is the process of developing the Long Short-Term Memory (LSTM) model to predict daily rainfall in Pasuruan Regency based on BMKG meteorological data?
2. How can the Particle Swarm Optimization (PSO) method be effectively applied to LSTM model optimization?
3. How accurate is the LSTM-PSO model in predicting daily rainfall in Pasuruan Regency compared to the LSTM model without optimization?
4. How to implement the LSTM-PSO model into the prototype website to display the results of daily rainfall prediction in Pasuruan Regency?

1.3 Research Objectives

In line with the formulation of the problem that has been explained earlier, the purpose of this study is to:

1. Build and implement a Long Short-Term Memory model to predict daily rainfall in Pasuruan Regency based on BMKG data.
2. Apply the Particle Swarm Optimization (PSO) algorithm to optimize parameters on the LSTM model.
3. Analyze and evaluate the performance of the LSTM model that has been optimized using PSO in predicting daily rainfall in Pasuruan Regency.
4. Developing a prototype website as a medium for the implementation of the LSTM-PSO model to display the results of rainfall predictions.

1.4 Research Benefits

This research is expected to provide benefits in the form of increasing scientific insight and understanding related to the application of Long Short-Term Memory (LSTM) algorithms and Particle Swarm Optimization (PSO) optimization for rainfall time series data and hydrometeorology, as well as contributing to research related to rainfall prediction. In addition, this research also produces a systematic framework starting from data collection, pre-processing, handling missing values, model training and optimization, to evaluation, as well as providing the results of model performance evaluation based on MAE, RMSE, NSE, and as a basis for assessing the accuracy of daily rainfall predictions in Pasuruan Regency. In addition, the results and findings of this study are expected to be a reference for further research in the development of rainfall time series predictions, for example through the addition of other meteorological features or variables, expansion of regional coverage, and comparison with other methods such as GRU, ARIMA or optimization methods other than PSO.R^2

1.5 Research Limitations

This research has several limitations that need to be clarified so that the scope of research is more focused, namely:

1. The data source in this study only comes from the official website of BMKG online data, UPT Pasuruan Geophysics Station, without other data sources.
2. The location of this study is limited to the scope of Pasuruan Regency.
3. The meteorological variables used in this study were limited to the daily average temperature, daily average air humidity, and daily rainfall.
4. The data observation period is limited starting from January 1, 2021 to December 31, 2025 with a total of 1815 days of data.