CHAPTER I
INTRODUCTION

1.1. Background

Red chili peppers are a strategic horticultural commodity in Indonesia, playing a vital role in people's daily lives. Not only are they a staple ingredient in various dishes, but they also directly impact economic stability, particularly in the food expenditure group. High chili consumption makes price fluctuations highly sensitive and have a broad impact on public welfare.

The phenomenon of price fluctuations in red chili peppers has long been a national issue, particularly in production centers like East Java. As one of the largest chili-producing provinces in Indonesia, East Java frequently faces unstable price dynamics. Spikes in red chili pepper prices can drive food inflation and suppress purchasing power, while drastic price drops harm farmers because selling prices are not commensurate with production costs [1].

According to data from the Central Statistics Agency (BPS), food inflation in Indonesia is often triggered by rising prices of red chili peppers, particularly during the rainy season and when distribution disruptions occur. This situation demonstrates the close relationship between chili prices and national inflation stability. In fact, in 2020–2022, the price of cayenne pepper spiked by more than 100% in a short period of time due to extreme weather and inefficient distribution [2].

The price of red cayenne pepper in East Java exhibits a fairly sharp seasonal fluctuation pattern. When production is abundant, prices plummet, resulting in losses for farmers. Conversely, when supply is low, prices can soar to over IDR 100,000 per kilogram at the consumer level. This situation creates economic uncertainty for both consumers and producers, necessitating an accurate price prediction system to anticipate market changes [3].

The XGBoost, LSTM, and Neural Prophet methods are widely used approaches in time series data forecasting due to their ability to capture nonlinear and complex patterns. XGBoost, as an implementation of extreme gradient boosting, offers advantages in computational efficiency, regularization capabilities, and stable prediction performance across various types of structured data. Research by B. W. Sari and D. Prabowo in Intellect: Indonesian Journal of Learning and Technological Innovation shows that XGBoost produces better accuracy than Random Forest and Gradient Boosting in house price prediction, confirming its strength in handling both numerical and categorical variables simultaneously [4]. This advantage makes XGBoost relevant for research based on tabular and time series data that has undergone feature engineering.

However, XGBoost has limitations in directly modeling long-term temporal dependencies because it is not inherently a sequential model. Its reliance on feature engineering means that its performance is heavily influenced by the quality of the lag features and derived variables created by the researchers. Furthermore, model interpretation can become complex as the number of trees and model depth increase. This differs from recurrent neural network-based approaches such as LSTM, which are inherently designed to capture temporal dynamics.

LSTM, as an extension of Recurrent Neural Networks, has a memory mechanism through cell states and gates that can retain long-term information. A study by A. F. Alkayes and T. Sugihartono in the Indonesian Journal of Education and Technology showed that LSTM can provide competitive, even superior, performance in predicting Tesla stock prices compared to XGBoost on highly fluctuating data [5]. The main advantage of LSTM lies in its ability to capture seasonal patterns and complex trends without the need for extensive feature engineering. However, this model requires large amounts of data, takes longer to train, and is susceptible to overfitting if not accompanied by adequate regularization techniques.

Neural Prophet is an extension of the Prophet model, based on a decomposable time series model combined with neural network components to increase nonlinear flexibility. Research by A. Primawati, F. A. Mustika, and Y. Wibawanti in Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer (Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer) shows that the Prophet and LSTM approaches have different performance characteristics in predicting gold prices, with the decomposition-based model excelling in capturing clear trends and seasonality [6]. Neural Prophet extends these capabilities by incorporating autoregression and neural components, making it more adaptive to nonlinear patterns. Its advantage lies in the interpretability of the trend and seasonality components, but its performance may decrease on data with extreme volatility or high noise.

Based on these characteristics, the research hypothesis can be formulated that there are significant differences in predictive performance between XGBoost, LSTM, and Neural Prophet on specific time series data. The first hypothesis states that LSTM will provide higher accuracy on data with long-term dependencies and complex fluctuations. The second hypothesis states that XGBoost will excel on data with structured features and nonlinear patterns that can be represented through feature engineering. The third hypothesis states that Neural Prophet will demonstrate stable performance on data with strong trends and seasonality and relatively structured data.

The selection of these three methods was based on theoretical and empirical considerations. Theoretically, XGBoost represents a boosting-based ensemble approach, LSTM represents a sequential deep learning architecture, and Neural Prophet combines statistical decomposition with neural learning. Empirically, all three have competitive performance in the context of price and economic time series prediction. By comparing these three approaches within a systematic experimental framework, the research can contribute comprehensively to identifying the most optimal method based on specific data characteristics and prediction objectives.

Unfortunately, there is still very little empirical research in Indonesia that directly compares these three models in the context of red cayenne pepper price prediction, particularly in East Java. The majority of previous research has focused on conventional methods or on commodities other than red cayenne pepper. This gap indicates a research gap that needs to be filled to make chili price modeling in Indonesia more comprehensive and accurate [7].

In addition to methodological factors, another challenge is the availability of sufficiently long and clean historical data. Daily price fluctuations available from the East Java Department of Industry and Trade (DISPERINDAG) from September 2024 to August 2025 provide an opportunity to build a predictive model with a high level of accuracy. With this data, the study can test the performance of various algorithms and determine the best model for red cayenne pepper in East Java.

The urgency of this research is also supported by the fact that spikes in red cayenne pepper prices have a significant impact on household consumption and government policy. Fauzi and Andriani (2023) emphasized that rising chili prices can drastically affect supply and demand, thus triggering market imbalances [8]. Therefore, accurate price prediction is not only academic but also practical in supporting decision-making.

Based on the description above, this study aims to compare the performance of three modern models (XGBoost, LSTM, and Neural Prophet) in predicting the price of red chili peppers in East Java. This research is expected to provide new contributions to the field of food commodity price forecasting, addressing the limitations of previous studies that predominantly used conventional methods, and providing practical benefits for farmers, traders, consumers, and the government in anticipating price dynamics.

1.2. Problem Formulation

Based on the background, the problems to be addressed in this research are as follows:

1. How do the XGBoost, LSTM, and Neural Prophet models predict the price of red cayenne pepper in East Java based on data available from the East Java Provincial Department of Industry and Trade (DISPERINDAG)?
2. How do you evaluate the performance of the XGBoost, LSTM, and Neural Prophet models that produce the best accuracy?
3. How do you predict the price of red cayenne pepper in East Java using the best method among XGBoost, LSTM, and Neural Prophet?

1.3. Research Objectives

Based on the problem formulation outlined, the objectives of this research consist of several points, which will be explained below:

1. Applying the XGBoost, LSTM, and Neural Prophet models to the price of red cayenne pepper in East Java based on data available from the East Java Provincial Department of Industry and Trade.
2. Evaluate the performance of the XGBoost, LSTM, and Neural Prophet models to determine the best model.
3. Determine the price prediction of red cayenne pepper in East Java using the best model among XGBoost, LSTM, and Neural Prophet.

1.4 Research Benefits

The benefits of this research cover two aspects: theoretical and practical, which are outlined as follows:

1. Theoretical Benefits
    a. This research can make a significant contribution to the development of science, particularly in the field of machine learning-based agricultural commodity price prediction and neural forecasting.
    b. This research can enrich the literature related to the different models of XGBoost, LSTM, and Neural Prophet.
    c. This research can fill the gap in research in Indonesia, which is still limited in the use of the Neural Prophet model, thus providing a new perspective on commodity price prediction methods.

2. Practical Benefits
    a. For Local Governments
        - This research can be used to design an early warning system for potential price spikes that impact inflation and public welfare.
    b. For Farmers
        - This research can provide a reference for planning planting times, distribution, and marketing strategies to maximize profits while minimizing the risk of losses due to price fluctuations.
    c. For Market Actors
        - This research can assist in making stock, distribution, and pricing decisions due to more accurate price predictions.
    d. For Authors
        - Provides an opportunity for authors to apply the knowledge learned in this research and make a real contribution.
        - Gain practical experience that can be useful in career development.

1.5 Research Limitations

Delimitations are an important component of any research, serving to prevent overly broad assumptions. The description of the problem limitations will be presented in bullet point format as follows:

1. This research will focus on comparing models for predicting the price of red cayenne pepper in East Java.
2. The models being compared use XGBOOST, LSTM, and Neural Prophet.
3. Data comes from the price of red cayenne pepper in East Java which is sourced from the Department of Industry and Trade (DISPERINDAG) of East Java Province on the website siskaperbapo.jatimprov.go.id.