

# CHAPTER I

## INTRODUCTION

### 1.1. Background of the Study

Commercial chicken eggs constitute a vital component of Indonesia's national food security system. As an affordable and readily available source of animal protein, these eggs play a significant role in household food consumption patterns across various social strata. Data from the Badan Pangan Nasional (Bapanas) indicate that in 2023, the per capita consumption of commercial chicken eggs in Indonesia averaged 6.69 kilograms annually, with the total national demand reaching 1.86 million tons per year. This substantial consumption level establishes commercial chicken eggs as the primary source of animal protein consistently consumed by all demographic segments of Indonesian society.

East Java Province has consistently been one of the main centers of commercial chicken egg production in Indonesia. According to data from the Ministry of Agriculture of the Republic of Indonesia, East Java makes a significant contribution to national egg production, and in 2023 it is projected to see a surplus of 107.40 thousand tons above consumption needs [1]. Sidoarjo Regency, one of the regencies with high economic growth and population density in the Surabaya metropolitan area, has a complex egg distribution chain due to its strategic location along interprovincial logistics routes. Market dynamics in Sidoarjo Regency are influenced by patterns of urban consumer demand, food processing industry activities, and intensive interregional distribution movements.

Although the supply of commercial chicken eggs in East Java is relatively abundant, the price of this commodity continues to experience significant fluctuations and is difficult to predict. Based on historical daily data from the Sistem Informasi Pemantauan Pasar dan Kebutuhan Pokok (SIMPONI Ternak) of Sidoarjo Regency during the period from 1 January 2020 to 31 December 2025, the price of commercial chicken eggs fluctuated between Rp17,000 and Rp31,000 per kilogram. The average recorded price was Rp24,489 per kilogram, with a standard deviation of Rp3,709 per kilogram. The fluctuation range, which reaches more than

82% between the lowest and highest prices, indicates a very high level of price uncertainty for market participants

Fluctuations in commercial chicken eggs prices are caused by a variety of complex factors. Calendar factors, especially during major religious holidays such as Eid al-Fitr, Eid al-Adha, Christmas, and New Year, consistently lead to surges in demand that result in sharp price increases in a short period. On the other hand, changes in the prices of livestock feed ingredients, particularly corn and soybean meal, which are the main components in layer chicken feed, also play a crucial role in determining production costs and retail market prices [2]. Disruptions in the supply chain due to extreme weather conditions, poultry disease outbreaks, and the impact of the COVID-19 pandemic in 2020–2021 add layers of complexity that make conventional price prediction even more challenging.

Uncertainty in commercial chicken egg prices significantly impacts various stakeholders. For small and medium scale farmers, who dominate production in Sidoarjo Regency, the inability to anticipate changes in selling prices amid rising feed costs poses a risk of financial losses that can threaten the continuity of their businesses. For traders and distributors, price volatility increases the risk of inefficient inventory management. From a public policy perspective, fluctuations in the prices of strategic food commodities can exacerbate inflation in the food sector and disrupt the regional economic stability. Therefore, the ability to accurately predict commercial chicken egg prices holds strategic value for production planning, distribution management, and price stabilization policies.

Over the past ten years, advances in artificial intelligence techniques and deep learning have opened new opportunities for modeling commodity price time series. Traditional statistical models, such as the autoregressive integrated moving average (ARIMA), have fundamental limitations in capturing non-linear patterns and long-term dependencies that are often found in commodity price data, which are influenced by seasonal factors and external events [3].

Recurrent neural networks (RNN) and their variants, especially long short-term memory (LSTM) introduced by Hochreiter and Schmidhuber [4], have proven highly effective in modeling non-linear time series. However, the gated recurrent unit (GRU) introduced by Cho et al.[5], offers a more computationally efficient

architecture. The GRU unifies the forget and input gates of LSTM into a single update and reset gate, thereby reducing the number of parameters without significantly compromising its ability to model temporal dependencies. On medium-scale datasets, such as six years of daily price data (2,192 observations), GRU often demonstrates performance equal to or better than LSTM, with shorter training times [6].

One of the main challenges in using deep learning models such as GRU is determining the optimal hyperparameters. Hyperparameters, such as the number of hidden units, number of layers, learning rate, dropout rate, and batch size, have a major impact on model performance but cannot be learned directly from data during training. Conventional methods, such as grid and random searches, face significant efficiency limitations, particularly when the hyperparameter search space is high-dimensional.

The Tree-structured Parzen Estimator (TPE), introduced by Bergstra et al [7], is a component of the Bayesian optimization framework that provides a more efficient way to find optimal hyperparameters. TPE develops a probabilistic model that maps the relationship between hyperparameter configurations and model performance and then uses this model to focus the search on the most promising areas of the hyperparameter space. In the Optuna framework [8], the implementation of TPE has empirically been shown to be more effective in finding superior configurations compared to grid search and random search, with far fewer evaluations.

Various relevant previous studies have been conducted in this field, [2] which conducted a comparative study between LSTM and GRU to predict daily prices of three strategic food commodities in Indonesia, namely rice, broiler chicken meat, and chicken eggs, utilizing data from the National Strategic Food Price Information Center (PIHPS). The results of this study indicate that GRU provides higher accuracy based on the MAE and MAPE metrics and completes the computation process approximately 20 seconds faster than LSTM. However, this study did not include calendar-based exogenous features that could enrich the prediction context, nor did it systematically optimize the hyperparameters. Similar

results were obtained by Oni et al.[9] who used GRU to predict the prices of five foodstuffs, including chicken eggs, in traditional markets in Bandung.

In a broader scope, Gu et al.[10] predict agricultural commodity prices by utilizing multivariate data, including meteorological data and trading volume. They reported a significant increase in prediction accuracy compared to the standard LSTM model. Zhang et al.[11] also highlighted the advantages of deep learning models that integrate exogenous features to predict commodity futures prices. However, these studies have not specifically examined TPE-based hyperparameter optimization in the context of local livestock commodities in Indonesia.

Nevertheless, merely developing accurate prediction models is not sufficient to provide a tangible impact for stakeholders in the field. Small and medium farmers, traders, and government officials who are the primary users usually do not possess programming knowledge or the ability to operate machine learning models independently. Therefore, a user-friendly, interactive, and easily accessible interface that does not require specific technical expertise is needed so that model prediction results can be directly used in daily operational decision-making.

Based on a review of previous studies, three main gaps need to be addressed. First, no study has specifically examined commercial chicken egg price prediction in Sidoarjo Regency using GRU optimized with TPE. Second, the use of exogenous features related to proximity to national holidays, both in binary (*is\_holiday*) and quantitative (*days-to-holiday*) forms as input variables in local livestock commodity price prediction models, has rarely been discussed in the existing literature. Third, previous studies generally stop at the model evaluation stage without including an interface implementation that can be directly used by non-technical users. This study is designed to address these three gaps, as well as to implement the best model into a Streamlit-based web application as a prototype early warning system that can be directly operated by stakeholders without requiring programming expertise to support more effective decision-making.

## **1.2. Research Questions**

Based on the background described above, the research problems in this study are as follows :

1. How is the performance of the Gated Recurrent Unit (GRU) model optimized with Tree-structured Parzen Estimator (TPE) in predicting commercial chicken egg prices in Sidoarjo Regency?
2. To what extent does the addition of calendar-based exogenous features contribute to improving prediction accuracy?
3. How can the best GRU-TPE model be implemented into a streamlit-based web application as a price prediction interface?

## **1.3. Research Objectives**

The objectives of this study are :

1. To develop and evaluate a Gated Recurrent Unit (GRU) model optimized with Tree-structured Parzen Estimator (TPE) for predicting commercial chicken eggs prices in Sidoarjo Regency.
2. To evaluate whether the addition of calendar-based exogenous features contributes to commercial chicken eggs price prediction in Sidoarjo Regency.
3. To implement the best GRU-TPE model into a streamlit-based web application as a prototype for a commercial eggs price prediction system.

## **1.4. Research Benefits**

The benefits of this research are as follows :

1. Academic Benefits

This research makes a significant contribution to the development of knowledge in the fields of computational intelligence and data science through three main aspects. First, this study expands the empirical literature related to the use of the GRU architecture for predicting the prices of livestock commodities in Indonesia, which is currently still lacking compared to other agricultural commodities. The findings of this research can serve as a reference for other researchers who wish to develop similar models for other livestock commodities. Second, this research provides empirical evidence regarding the

effectiveness of the Bayesian optimization method, especially TPE as implemented in Optuna, in improving the performance of GRU models on non-stationary time series data. These findings can reinforce the argument that TPE is a more efficient alternative than grid search in time series modeling research for local commodity prices. Third, this study offers a methodological contribution in the experimental design of comparisons for calendar-based exogenous features, which can be adopted as a standard approach in commodity price prediction research affected by national holiday patterns.

## 2. Practical Benefits

For layer chicken farmers, the streamlit-based web application provides a simple price prediction interface that can be operated without programming expertise. By entering the data to be predicted, farmers can obtain estimates for commercial chicken egg prices for the following day along with their confidence intervals, thus enabling them to plan production quantities and marketing strategies to minimize the risk of losses due to price uncertainty.

For commercial chicken egg traders and distributors, the price prediction information can support decision-making related to inventory management and more data-driven selling price determination, potentially increasing the efficiency of the distribution chain.

For the Sidoarjo Regency Government, particularly the Livestock and Fisheries Agency, the developed streamlit application prototype can be integrated into the SIMPONI Ternak platform as an early warning feature to anticipate price spikes that could disrupt regional food stability.

### 1.5. Scope Limitations

To maintain the focus and scope of the research, the limitations in this study are as follows :

1. This study focuses on predicting the price of commercial chicken eggs at the retail market level in Sidoarjo Regency, East Java Province.
2. The data used were sourced from the Sidoarjo Regency Market and Basic Needs Monitoring Information System (SIMPONI Ternak) in the form of daily data.

3. The web application implemented in this study is a functional prototype and does not include direct integration with the SIMPONI Ternak server, user management, or a production-level system security.
4. The predictive model used is limited to the Gated Recurrent Unit (GRU) optimized with the Tree-structured Parzen Estimator (TPE).