# Introduction

Indonesia has five types of vegetables which have the highest production levels, namely, cabbage, potatoes, tomatoes, shallots, and red chili. Where among the five types of plants, red chili has the most significant yield [1]. Red chili is an essential type of spice. According to the Ministry of Agriculture (MoA) stated that red chili has contributed to the Indonesian economy because the export value in 2011 reached 11 million US $ and created employment opportunities for many farmers. Exports of chili value increased in 2012 to 26 million US dollars and remained second as a vegetable commodity that has the highest export value [1]. Red chili is an vital plant that contributes significantly to the local and national economy. However, red chili is still one of the sources of food that experiencing price increases continuously [2].

The weather is very influential in the agricultural sector, especially with extreme weather that can affect the land, such as droughts or floods, making the planting process hampered and affecting crop production [3]. Other factors that influence the production process are climate change, supply, demand, price, policy, and technology [4].

Data Mining is a process to obtain knowledge or information from large and complex data [5]. Data Mining has various methods. Data Mining was used in multiple studies such as Vegetable Price Prediction Using Data Mining Classification Technique with the ANN algorithm produces monthly accuracy of 90% [6]. KNN is used to predict power plant power output, based on the Root Mean Squared Error (RMSE) to obtain a prediction model that occupies the second position in the Global Energy Forecasting Competition in 2012 [7]. Besides, KNN is also used in forecasting short-term traffic volumes with trials of values k = 10 to 40 having a value of Mean Absolute Percentage Error (MAPE) which decreased by 1% [8]. This study uses the K-Nearest Neighbors (KNN) classification algorithm in the process of predicting chili commodity prices.

The prediction of the price of agricultural investment in Bandung regency has previously been carried out in several studies, such as in the study of Recommended Planting Date for Chili and Tomatoes Based on Economic Value of Agricultural Commodity Prediction with ANN-NGA algorithm which obtained 81% and 100% results [9], Chili commodities in Bandung Regency using the Bayesian Network algorithm managed to obtain precision values of 0.92, recall 1, and accuracy of 83.5%, with unbalanced dataset [10], and other studies related to this are Predictions of Chili Prices in Regency of Bandung Using Support Vector Machine (SVM) Optimized with Adaptive Neuro-Fuzzy Inference System (ANFIS)[11].

Against the background of previous research [9,10,11], this study aims to prior development research with previous studies using data datasets that are not appropriate. Adaptive Synthetic (ADASYN) with oversampling techniques on unbalanced datasets resulting in better accuracy and F1-Score. ADASYN algorithm was invited in 2008 [12], it has been used in the research on Handling Unbalanced Data in Churn Prediction using ADASYN and Backpropagation Algorithm by testing 96.31% and F1-Score 0.4607 [13], and in that study Rainfall Forecasting uses Tree (CART) classification and regression Algorithms and Adaptive Synthetic Sampling (ADASYN) with the results of accuracy without ADASYN equal to 93.94%, while ADASYN speed is 98.18% [14].