

# CHAPTER 1

## INTRODUCTION

### 1.1 Background

Weather is a condition of atmosphere in earth, the condition is significantly change every time. Lot of factors influence the change in weather in a place, such as temperature, air pressure, humidity, light intensity, and rainfall. The changes in weather attract scientists to conduct research on weather classification based on factors above. There are many methods used by scientists to do weather classification, one of this methods is data mining.

Data Mining is an act which involves data collection , the use of historical data to find an order, pattern or connection in big data. Data Mining is a combination of four disciplines namely statistic, visualization, database and machine learning. Machine learning is a part of artificial intelligence that related with developing programming techniques based on past data learning and intersect with statistic and also optimization.<sup>1</sup>

Nowadays, there are many projects uses data mining technique to classify the weather. There are some data mining methods that already being used to classify the weather such as Random Forest, C.45, k-NN, Neural Network, and Naive Bayes. Naive Bayes Classification is one of the classification algorithm with basic concept using probability to calculate the classification. This project use Naive Bayes Classification to analyze the accuracy of this algorithm to handle weather classification with parameter data temperature, air pressure, humidity, dew point, wind speed, visibility, heat index, and weather condition.

---

1 Subekti Mujiasih, “Pemanfaatan Data Mining Untuk Prakiraan Cuaca”. Meteorologi dan Geofisika. Vol. 12 No. 2, 2011,189-195.

## 1.2 Scope

Based the background , scope for this project are:

1. What factors affect weather classification?
2. How Naive Bayes Classification works in weather classification?
3. How accurate is Naive Bayes Classification to classify weather?
4. Which parameters are dominant on affecting the weather classification result?

Constraint for this project:

1. Parameter that used in this project are temperature, air pressure, humidity, dew point, wind speed, visibility, heat index, and weather condition.
2. Data is taken from API in Wunderground.com
3. Data focused on weather history data of Jakarta from 2015 to 2016.

## 1.3 Objective

The objective from the project is to find the accuracy of Naive Bayes Classification to classify weather using eight parameter as reference that are temperature, air pressure, humidity, dew point, wind speed, visibility, heat index, and weather condition. Also determine the most dominant parameter that affect weather classification.