

CHAPTER 1

INTRODUCTION

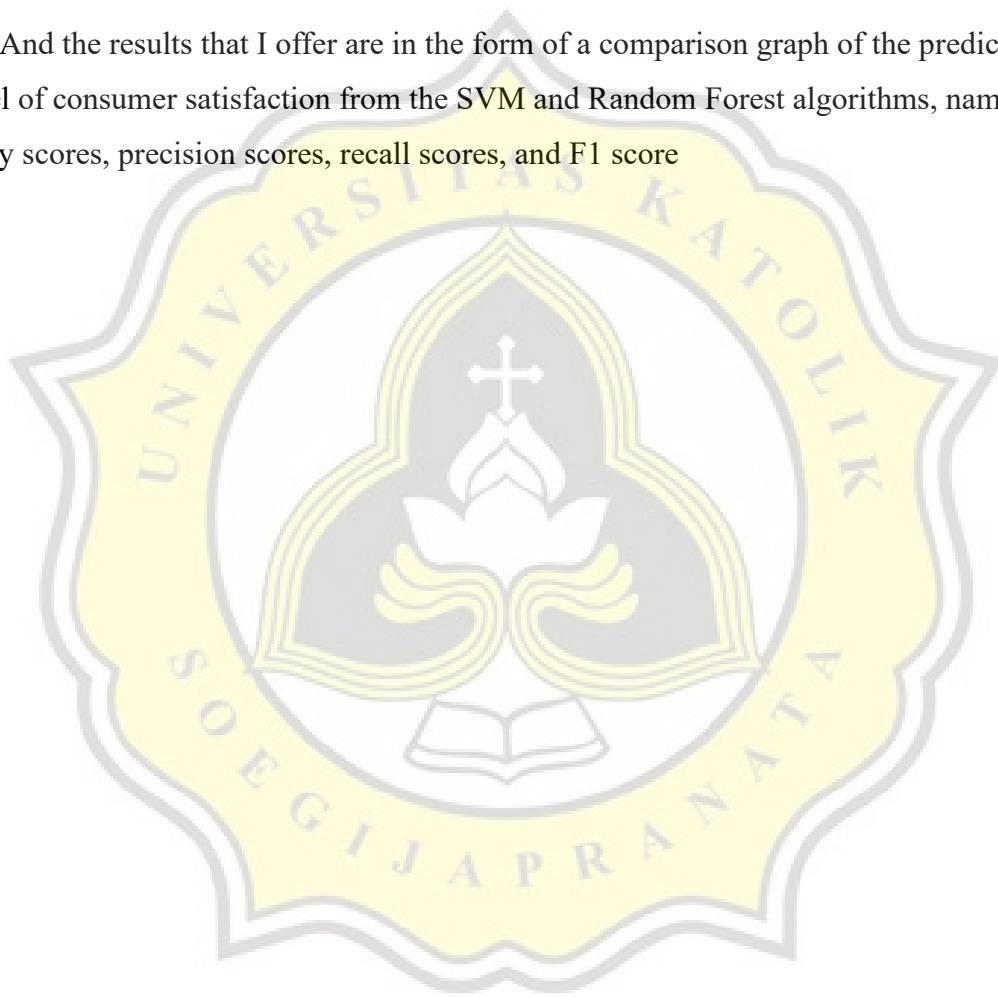
1.1. Background

Indonesia is one of the countries visited and favored by foreign tourists' tourism. One of them is Bali, which has beautifully both in terms of tourism and culture. Natural beauty in terms of tourism is an attraction for domestic and foreign tourists. The vehicles commonly used by foreign tourists are airplanes. In-flight facilities are one of the factors that can make passengers comfortable on a long flight schedule. During the pandemic, the government has aggressively opened various tourism sectors to restore the economy, especially in the tourism sector. Passenger comfort for foreign and local tourists is one of the factors that is the center of attention of the government which is addressed. Various facilities on the plane have been improved by the government to attract tourists to come to Indonesia. People sometimes are satisfied and dissatisfied with the facilities provided. This problem can occur due to less-than-optimal public services such as WIFI which is often problematic, toilet cleanliness, and many other things.

To overcome the existing problems, the SVM algorithm and random forest are used. The data structure of the entire data will be divided into training and test data as much as 70% training data and 30% test data for both algorithms. The data structure used in the form of data on passenger satisfaction with public servants at the airport in the form of satisfaction, age, flight distance, gate location, in-flight WIFI service, in-flight entertainment, online support, online booking ease, in-flight service, cleanliness, late departure flight. Minutes and late arrivals by minutes. The SVM algorithm is a classification method that works by finding the hyperplane with the largest margin. The margin in question is the distance between the hyperplane and the closest data in each class. SVM algorithm also performs the classification of linear and non-linear data. SVM maps training examples to points in space to maximize the width of the gap between the two categories. The SVM algorithm uses non-linear mapping to convert the initial training data to a higher dimension which is also called the kernel. While the Random Forest algorithm is an algorithm used to classify large amounts of good data.

Random forest classification is carried out by combining trees and conducting training on the sample data held. The use of more trees will affect the accuracy that will be obtained for the better. Determination of random forest classification is taken based on the voting results of the trees formed. The classification process in random forests begins by breaking down the existing sample data into a random decision tree. After the tree is formed, voting will be carried out on each class from the sample data. Then, combine the votes from each class and then take the most votes. By using the random forest in the data classification will produce the best votes.

And the results that I offer are in the form of a comparison graph of the prediction of the level of consumer satisfaction from the SVM and Random Forest algorithms, namely accuracy scores, precision scores, recall scores, and F1 score



1.2. Problem Formulation

1. How to classify data from the SVM Algorithm and Random Forest Algorithm?
2. How is the comparison between the SVM Algorithm and the Random Forest Algorithm to predict the level of passenger satisfaction with public services in a plane?
3. How do you calculate the accuracy result?
4. Is the SVM Algorithm better than the Random Forest algorithm based on the results of the Accuracy value?

1.3. Scope

In this final project the author uses Python version 3.0, for the data used in this study is the original data that I took from the kaggle website totaling 129,881 data. The author uses the process of dividing the data into training data and test data in data classification, after dividing into 2 parts, the next step is to calculate the accuracy, precision, recall, and f1 score of the two algorithms, namely the SVM algorithm and the Random Forest Algorithm. This project will also focus on the comparison of the SVM Algorithm and the random forest algorithm in classifying data on the level of buyer satisfaction in terms of accuracy.

1.4. Objective

The main objective to be achieved in this project is to classify the level of passenger satisfaction with public services at the airport from the SVM algorithm and the deep forest algorithm. The second objective of this project is to calculate the accuracy of the two algorithms. So, from the accuracy value obtained, it can be concluded that the comparison of the two algorithms.