CHAPTER 1
INTRODUCTION

1.1. Background

In recent years, the use of e-commerce or online shops has increased considerably. Various online stores have sprung up on the internet, both small and large-scale. This has a very important effect on the effective use of time and the level of sales figures. With so many e-commerce emerging and even with almost the same type or product, it will be quite difficult to reach potential buyers for them. The intentions of consumers will also be different for each online store they visit. So the problem with the number of e-commerce that has sprung up is that shop owners must have data that is effective enough to see the intentions of consumers in their stores whether to buy their products or not, regardless of the marketing strategy they apply.

Every e-commerce site has visitor history data especially nowadays that everything is online, it will make it easier for store owners to track consumers who visit their online store. This data includes all information regarding the traces of consumers who visit online stores, both new and repeated. With this data, processing will be carried out to make predictions about consumers' shopping intentions. The data contains attributes that will later be used for processing with the algorithm to be compared. Several algorithms will be used to process this data in order to find which one is the best in predicting consumers' intention to make a purchase or not. The comparison of algorithms will be carried out using the value of the Precision, Recall, F1-Score, Accuracy. Also some values that are considered to support the results of the comparison.

This prediction, online shop owners can know for sure whether their shop invites visitors to make a purchase or not, so that this will greatly save time in analyzing data on the intentions of visitors to shop on e-commerce. The results of this algorithm comparison also determine which algorithm is better in predicting consumer intentions on available e-commerce. With the results of this prediction, marketing strategies will be based on this data so that online shop owners can evaluate or develop their business even more against the ever-changing market situation.

1.2. Problem Formulation

There are several problem formulations in this study, especially on the issue of comparison and impact on current e-commerce.

1. Can predictions be made to find out the intentions of consumers when visiting online stores?

2. Can the comparison of the algorithm to this data be carried out with significant results?

3. Does the hyperparameter tuning with RandomizedSearchCV really impact each algorithm?

1.3. Scope

The scope of this research is based on currently available e-commerce. Some e-commerce may not have data on visitors who have visited their online store, and it also depends on the policies of the respective store owners. The problem limitation lies in the data attributes available in each e-commerce, some data contain different attributes or descriptions of visitor data per individual. This makes the data processing process must be evenly distributed and can be used universally in the comparison of algorithms to be carried out in order to support the concreteness of the comparison results.

1.4. Objective

The purpose of this study is to find out which algorithm results from the comparison is better in predicting the intentions of online shoppers. Meanwhile, the results of the comparison can help online store owners to find out more about the impact of their store on buyers' intentions to make a purchase or not. Moreover, with near-perfect predictions, this will change the marketing strategy so that they will adjust to the conditions of their online store efficiently with definite data results.