# CHAPTER 3
# RESEARCH METHODOLOGY

The research techniques used to address the issues raised in the problem formulation are covered in detail in this chapter. How to carry out this research is explained in this chapter. This study must be finished in various phases.

## 3.1 Research Process

The first process is how to process some drink data that will be classified using the KNN and SVM algorithms to get accurate results for the comparison of the two algorithms by programming processing using the orange data mining application and how the application works to produce which accuracy has higher accuracy.

## 3.2 Collecting Datasets

The data set for this research was taken from several types of bottled beverage samples in minimarkets. These types of drinks include coffee, mineral water, tea and energy drinks. This dataset has 610 types of data, which contain the name of the drink, volume quantity, calories, caffeine quantity and drink type. The data is taken from the site https://www.caffeineinformer.com/the-caffeine-database. The data is compiled using nutrition labels with lab tests

## 3.3 Convolutional Neural Network

This research focuses on the usage of KNN and SVM in the classification process at this time. The classification method, as well as how the KNN and SVM algorithms are used in the "orange data mining" application programming, are explained in the sections that follow.

In data mining, classification is a category based on function. A procedure for defining and separating data ideas is classification. The objective is to utilize the model to forecast an object's class when the class label is unknown. analysis of models using training data sets or data objects with known class labels. C4.5, RainForest, Naive Bayesian, Neural Network, Genetic Algorithm Fuzzy, case-based reasoning, Support Vector Machine, and K-Nearest Neighbor are a few classification techniques.

In classification, there are a number of records called training data, which consist of several attributes that can be either continuous or categorical, one of the attributes indicates the class for the record, where the mapping can be seen in the figure 3.1.

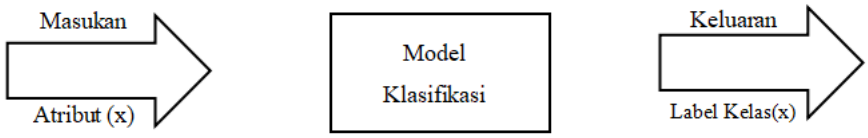


**Figure 3. 1 Mapping attributes (x) into class labels (y)**

**Orange Data Mining.**

Orange is an open source program that uses no code to analyze analytical and data mining data. The user does data mining for input and output utilizing a widget system by using the Orange program. When a dataset is entered into Orange, a number of algorithmic approaches may be employed to analyze it and display results that are appropriate for the situation.

Orange offers a variety of widgets for visualization that are drawn or put on a canvas and then connected to other widgets. Orange can easily do several analyses using a variety of algorithmic techniques. This medium makes it simpler to use data and carry out data analytics procedures in a natural way. Additionally, Orange offers a number of modules in fields including text mining, bioinformatics, data networks and social networks, model mapping, and prototyping processes.

The benefits of this orange application include the ability to perform machine learning and data mining without the need to acquire a challenging programming language. Orange's drawbacks include the fact that, although being simple for beginners to use, several features of the application still need computer code. When developing a work integration system, for instance, simply extracting data is insufficient; you must write specific computer code to ensure that you receive the desired results.