

CHAPTER 1

INTRODUCTION

1.1. Background

Microplastics have been known since 2004 after a lot of sedimentation appeared in European waters. Research on microplastics began to develop from year to year to define what microplastics are. In 2019, Frias and Nash, defined microplastics as any synthetic solid particle or polymeric matrix which are insoluble in water. The size is from 1 micrometre to 5 millimetre and its shape can be regular or irregular of either primary or secondary manufacturing origin [1]. The consequences of microplastics include, affecting pregnant women and their babies, imbalances the hormones, carcinogenicity, and many more [2].

Fourier Transform Infrared Spectroscopy (FTIR) is a tool that shoots infrared radiation through an existing sample. The radiation will be fully absorbed, partially, or not even absorbed by part of the existing sample. From these data, a special spectrum will emerge that represents the characteristics of the existing sample [3]. The result of FTIR is a spectrum wave that changes from the one that was fired. Each spectrum will be a marker of the chemical group characteristics of the microplastics [4].

Spectrum data processing as a characteristic of chemical groups can use gaussian mixture and naive bayes. Gaussian mixture is a statistical distribution model used to measure the distribution of a category to its members. The heterogeneity in the population can cause a parametric family fails to model the data properly [5]. In this case, the gaussian mixture is used to determine the chemical groups of the spectrum present.

Naive Bayes is one of the probabilistic models to determine the class of a thing with Bayes theory [6]. In this paper, the Naive Bayes method used is Gaussian because it is used to calculate the microplastics probability of many plastic classes.

In this study, the author suspects that the use of gaussian mixture and naive bayes to determine the microplastics content can improve accuracy compared to manual matching (one by one) against existing references. This is because, the current reference is the single absorption bands number, not the absorption bands range.

1.2. Problem Formulation

1. How to answer the ambiguity of which polymer to classify at manual matching?
2. How to solve the difference in component data size of a classification?
3. How precise is the use of Gaussian Mixture and Gaussian Naïve Bayes?

1.3. Scope

The data used is the data that I got from the Faculty of Agricultural Technology, Soegijapranata Catholic University, Semarang. All the data that I use are assumed as ground truth for each microplastics. The data obtained was tested in the lab by deliberately creating microplastics contamination. The experiment was carried out in sterile conditions using 96% ethanol, not using plastic equipment, and covered with aluminum foil so as to minimize contamination.

1.4. Objective

The main objectives of this research are to get the scope range of absorption bands and create a model that can identify the spectrum of absorption bands and classify the group of absorption bands into what kind of microplastics pollution.