Classification of Wood Powder Adulteration Level in Cilantro Seed Powder

Rishabh Goyal, Zirun Ye

1 Introduction

Ensuring food authenticity and safety of foods is a critical challenge in the food industry, especially with the increasing incidences of adulteration. Cilantro seed powder, a widely used spice, is often subjected to adulteration with wood powder. This not only compromises its quality, but also poses significant health risks to consumers. Detecting adulteration can be done by Fourier transform infrared (FTIR) spectroscopy. However, interpretation of FTIR spectra for the adulteration level is still challenging.

2 Objectives

The primary objective of this project is to develop models to classify the level of wood powder in cilantro seed powder through FTIR spectroscopy data collected by the Department of Food Process Engineering, National Institute of Technology, Rourkela, India. By achieving this, our aim is to protect the health and well-being of consumers of consumers.

3 Data

The dataset includes the FTIR spectroscopy data for different levels of wood powder in cilantro powder from 0 to 100 percent with increments of 10 percent (w/w) in different replication sizes. Each datapoint includes hundreds of light intensity against the different wavenumbers and an output adulteration level.

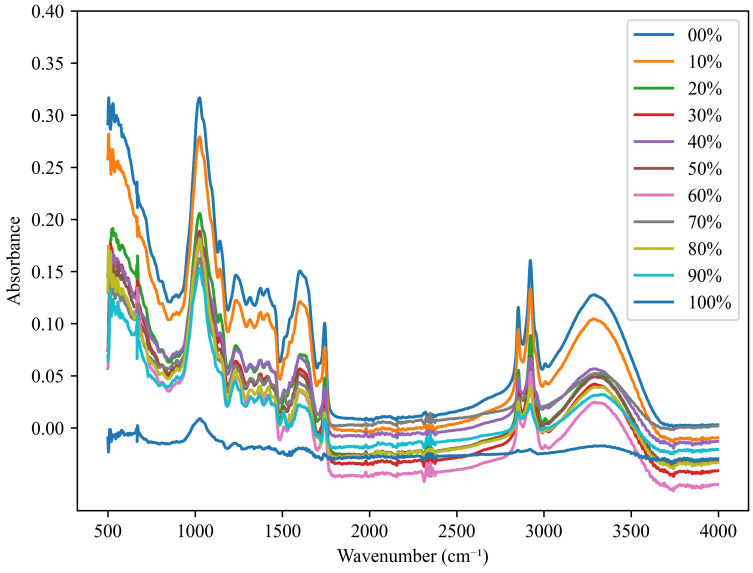

Figure 1: Wood powder level from 0 to 100% with 10% interval

4 Methodology

To address the problem, we plan to conduct data analysis, data preprocessing, feature engineering, and model development. The most challenging part in the project is to develop models in a manner to tackle the temporal and random noise in the dataset. The dataset also having lower datapoints as compared to number of features.

a) Data analysis: The data will be analyzed to understand the relationship between the light intensities (features) and against the adulteration level (output). The noisy data will be removed using Z-score outlier method.

b) Data preprocessing: The data consists the temporal and random noise which need to be reduced. For this project, we are planning to utilize the Savitzky–Golay (SG) smoothing, standard normal variate (SNV), and extended multiplicative signal correction (EMSC). These algorithms are generally used for smoothing the spectral data.

c) Feature engineering: The preposed dataset will be utilized to identify the key wavenumber (features) and feature extraction will be utilize to see the impact of feature transformation[1]. We are planning to utilize the recursive feature elimination for feature selection and principal component analysis for feature extraction.

d) Modelling: We are planning to utilize the common models such as logistic regression, decision tree, support vector machine, and artificial neural network[2] for detecting the wood powder level in the cilantro seed powder.

We will evaluate the performance of the model using accuracy, precision and confusion matrix.

5 Time Line

Task	**Date**
Research, data analysis, proposal	March 28th
Data preprocessing and Feature engineering	April 4th
Presentation for Checkpoint 3	April 10th
Compare and optimize models	April 14th
Final touch up and presentation for Project	April 20th

Table 1: Project Schedule

6 Expected Outcome

By the end of this project, a functional artificial intelligence model for estimating the percentage of wood powder in Cilantro powder will be implemented, tested, optimized, analyzed, and validated to ensure high accuracy and reliability. The findings will demonstrate the effectiveness of artificial intelligence in identifying food adulteration and improving food safety.

References

[1] Q. Li et al. "Mid-infrared spectra feature extraction and visualization by convolutional neural network for sugar adulteration identification of honey and real-world application". In: *LWT* 140 (2021), p. 110856. DOI: 10.1016/j.lwt.2021.110856.

[2] B. Yegnanarayana. *Artificial Neural Networks*. Google Books, 2009.