MACHINE LEARNING CHEMOMETRIC MODEL FOR RAMAN SPECTROSCOPY BASED HONEY QUALITY ASSESSMENT

Stefan M. Kolašinac[1], Ilinka Pećinar[1], Zora P. Dajić Stevanović[1]

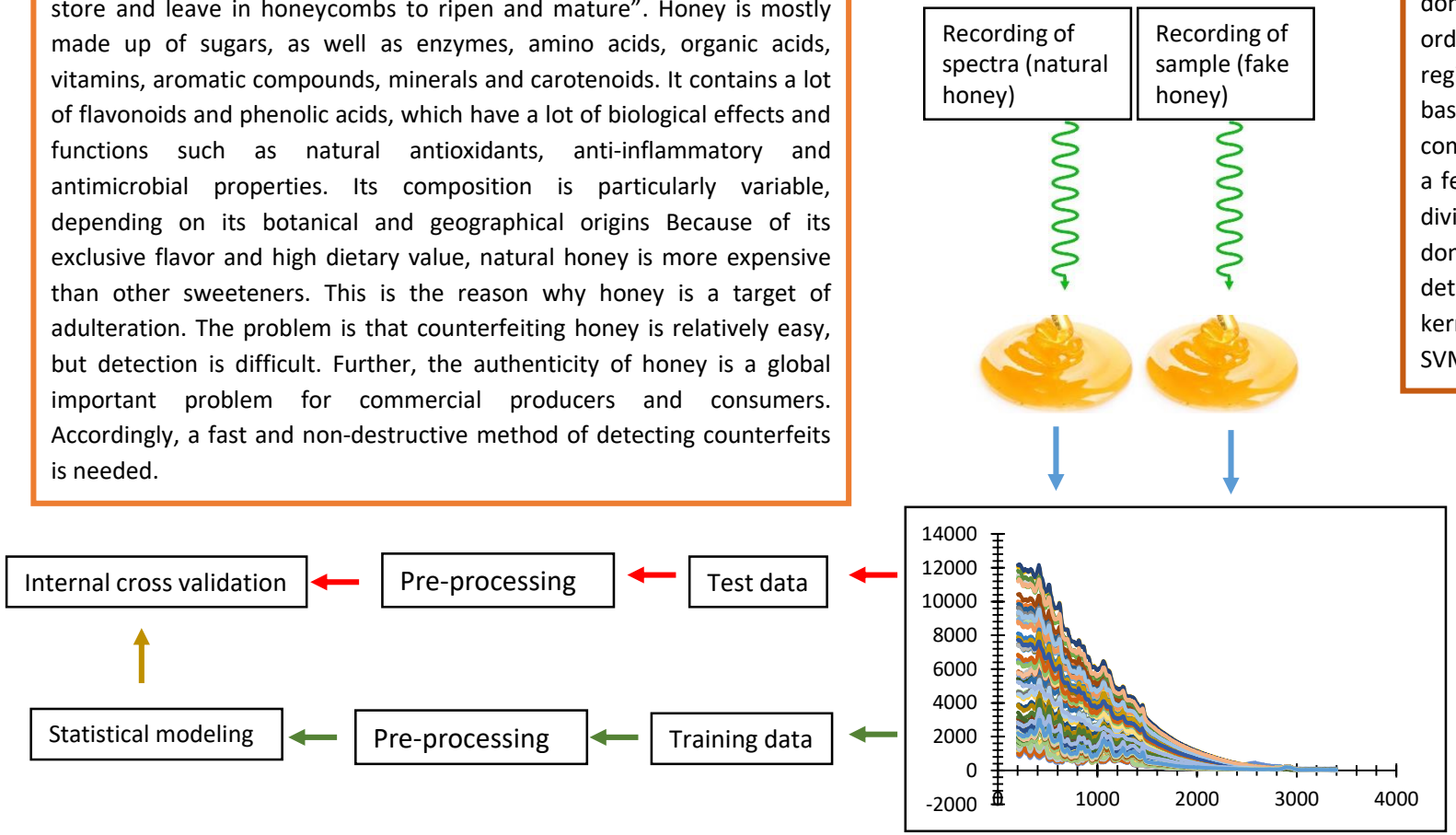
[1]*University of Belgrade, Faculty of Agriculture, Nemanjina 6, Belgrade, Serbia*

INTRODUCTION:

According to Codex Alimentarius (2001), "honey is the natural sweet substance, produced by honeybees from the nectar of plants or from secretions of living parts of plants, or excretions of plant-sucking insects on the living parts of plants, which the bees collect, transform by combining with specific substances of their own, deposit, dehydrate, store and leave in honeycombs to ripen and mature". Honey is mostly made up of sugars, as well as enzymes, amino acids, organic acids, vitamins, aromatic compounds, minerals and carotenoids. It contains a lot of flavonoids and phenolic acids, which have a lot of biological effects and functions such as natural antioxidants, anti-inflammatory and antimicrobial properties. Its composition is particularly variable, depending on its botanical and geographical origins Because of its exclusive flavor and high dietary value, natural honey is more expensive than other sweeteners. This is the reason why honey is a target of adulteration. The problem is that counterfeiting honey is relatively easy, but detection is difficult. Further, the authenticity of honey is a global important problem for commercial producers and consumers. Accordingly, a fast and non-destructive method of detecting counterfeits is needed.

OBJECTIVES:

The aim of this paper is to verify the possibility of Raman spectroscopy and Support Vector Machine (SVM) for classification of two different honeys and their fake duplicates. For this purpose, meadow and acacia honeys were selected.



METHOD / DESIGN:

Spectra of homemade and counterfeits honey were recorded using XploRA Raman spectrometer (Horiba Jobin Yvon). Raman scattering was excited by laser at a wavelength of 785 nm equipped with a 600 lines/mm grating; spectra were recorded by applying exposure time 10 s and accumulated from 10 times scans, using 100% filter. Spectral resolution was 3 cm^{-1} and autocalibration was done each time before recording of spectra by 520.47 cm^{-1} line of silicon. In order to assess a possible sample inhomogeneity, thirty Raman spectra in the region from 200-3400 cm^{-1} were recorded for each sample. All spectra were baseline-corrected, normalized and smoothed. After that PCA (Principal component analysis) was conducted and obtained PCs (first two PCs) served as a features for support vector machine (SVM) classification method. Data were divided into training model (70 %) and training data (30 %). Pre-processing was done by Unscrambler X 10.4 software (CAMO software, Norway). In order to determine the best shape of the hyperplane and decision boundary, several kernel function were used: linear, radial basis and polynomial function. The SVM was conducted by Python and Scikit-learn package.

RESULTS:

Support vector machine showed high accuracy in classification of different honey samples. Accordingly, the best discrimination power showed SVM with polynomial function (100%), followed by radial basis (96.67%) and linear (81.82 %).

CONCLUSIONS:

Results showed that SVM algorithm can be used as a tool for detection of fraudulent honey products.