Estimation of the Alcoholic Degree in Beers through Near Infrared Spectrometry Using Machine Learning [†]

Brais Galdo *[ID]**, Daniel Rivero and Enrique Fernandez-Blanco**

Faculty of Computer Science, CITIC, University of A Coruna, 15071 Galicia, Spain; daniel.rivero@udc.es (D.R.); enrique.fernandez@udc.es (E.F.-B.)

* Correspondence: brais.cgaldo@udc.es
† Presented at the 2nd XoveTIC Conference, A Coruña, Spain, 5–6 September 2019.

Abstract: It is a fact that, non-destructive measurement technologies have gain a lot of attention over the years. Among those technologies, NIR technology is the one which allows the analysis of electromagnetic spectrum looking for carbon-link interactions. This technology analyzes the electromagnetic spectrum in the band between 700 nm and 2500 nm, a band very close to the visible spectrum. Traditionally, the devices used to measure are utterly expensive and enormously bulky. That is why this project was focused on a portable spectrophotometer to make measures. This device is smaller and cheaper than the common spectrophotometer, although at the cost of a lower resolution. In this work, that device in combination with the use of machine learning was used to detect if a beer contains alcohol or it can be labeled as non-alcoholic drink.

Keywords: NIR; Electromagnetic spectrum; Neural networks

1. Introduction

Spectroscopy is one of the main techniques when a non-destructive analysis on the composition of a sample is required, however those analysis usually has to be performed by expensive equipment host only on special laboratories. However recent advances allow the reduction of that expensive equipment into something portable. Those devices allow the analysis in field conditions of the samples, but what is now laking is the expertise to understand the collected spectra. There is where machine learning can help. The main aim of this work is to being able to classify "lagger" beers according to their alcohol level into "Non-alcohol" and "With alcohol". In order to accomplish this objective, a portable NIR scanner from Texas Instrument ([1]) was used. Because, the idea behind this development, was to perform the measurements with out destroying the sample. Therefore, the liquid has to be poured into an special container called cell to be scanned. The result of those scanner are a particular spectrogram of the beer which has to be processed in order to solve the problem. In this particular work, 36 beers of different brands were used, as well as different graduation levels. Following an strong methodology, the dataset was collected and subsequently analyzed by using ANNs. We must underline that the alcohol is situated by most of the works in the range of 1110 to 1300 nm [2] and there are some articles that use NIR technology with big spectrophotometer to treat problems like this [3–6].

2. Materials and Methods

The first task to be tackle in this study was to build the dataset, which requires:

- Spectrophotometry cells. These are the containers into which the sample will be poured.
- A glass container. Liquids would be poured into them to reduce the amount of gas as much as possible.
- Coffee filters. They were used to dramatically reduce the amount of gas in a short time period.

Once all the material was collected, the beers were leaved in a room for 3 days in order to ensure that the temperature of the beers was the same before performing the measurements. Those measurements were performed by a portable NIR scanner which operates in the range between 900 and 1700 nm and, therefore, contains the objective range of the alcohol. Attending to the spectrometer setup, an average of 30 measurements performed with the Hadamard's algorithm was chosen. This process was repeated 3 times for each beer.

Once the liquid was in the cells, the spectrometer was 2mm from the sample while the scan was performed. The resulting samples were labeled as "without alcohol" when the beer contains less than 1 degree, while the label "with alcohol" was reserved for the ones containing more than 1 degree.

3. Results

All ANNs develop in this work used TensorFlow as its development framework by following a K-fold Record-wise cross-validation schema. Therefore, all three measurements from the same beer were used only on train or on test. The developed ANNs are fully-connected forward networks with a single hidden layer. The optimizer was set to Adam and the loss function was the binary crossentropy. Table 1 shows the average of 50 repetitions of the training for each of the beers used in test. Additionally to the size of the hidden layer, the table shows the accuracies for train and test together with the corresponding standard deviation.

Table 1. Preliminary results.

Number of Neurons	Activation Function	Train		Train	
		Accuracy	STD	Accuracy	STD
32	Sigmoid	0.917	0.047	0.814	0.388
64	Relu	0.908	0.030	0.817	0.386
128	Relu	0.913	0.030	0.812	0.389

4. Discussion

First and foremost, we must emphasize that this are barely preliminary results although those have been encouraging. In this work a first dataset has been collected following a methodology pre-established by an expert in the field and it has proved that the machine learning technique can be of great utility in order to move the analysis from the laboratory to the field. As the measurements were collected, it has been possible to appreciate the difference between beers with alcohol of those without alcohol in the range between 1100 and 1300 nanometers, so it was to be expected that, making use of techniques of machine learning, it was possible to solve this problem. However, the high standard deviation observed on the test is something that requires additional study being the most probable reason the small size of the dataset and the use of a highly restricted technique as the K-fold Record-wise cross-validation. Consequently, the plan for the close future is to increase the size of the dataset and perform more test to solve it.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

NIR   Near Infrarred
ANN   Artificial Neural Network

References

1. Texas Instruments. DLPNIRNANOEVM. Available online: http://www.ti.com/tool/DLPNIRNANOEVM (accessed on 15 July 2019).
2. Wang, X.; Bao, Y.; Liu, G.; Li, G.; Lin, L. Study on the best analysis spectral section of NIR to detect alcohol concentration based on SiPLS. *Procedia Eng.* **2012**, *29*, 2285–2290.
3. Castritius, S.; Kron, A.; Schafer, T.; Radle, M.; Harms, D. Determination of alcohol and extract concentration in beer samples using a combined method of near-infrared (NIR) spectroscopy and refractometry. *J. Agric. food Chem.* **2010**, *58*, 12634–12641.
4. Iñón, F.A.; Garrigues, S.; de la Guardia, M. Combination of mid-and near-infrared spectroscopy for the determination of the quality properties of beers. *Anal. Chim. Acta* **2006**, *571*, 167–174.
5. Dorado, M.P.; Pinzi, S.; De Haro, A.; Font, R.; Garcia-Olmo, J. Visible and nir spec- troscopy to assess biodiesel quality: Determination of alcohol and glycerol traces. *Fuel* **2011**, *90*, 2321–2325.
6. Maudoux, M.; Yan, S.H.; Collin, S. Quantitative analysis of alcohol, real extract, original gravity, nitrogen and polyphenols in beers using nir spectroscopy. *J. Near Infrared Spectrosc.* **1998**, *6*, 363–366.