Food Research International 99 (2017) 206–215

Contents lists available at ScienceDirect

Food Research International

journal homepage: www.elsevier.com/locate/foodres

An automated ranking platform for machine learning regression models for meat spoilage prediction using multi-spectral imaging and metabolic profiling

Lucia Estelles-Lopez[a], Athina Ropodi[b], Dimitris Pavlidis[b], Jenny Fotopoulou[b], Christina Gkousari[b], Audrey Peyrodie[a], Efstathios Panagou[b], George-John Nychas[b], Fady Mohareb[a,*]

[a] Bioinformatics Group, Department of Agrifood, School of Water, Energy and Environment Cranfield University, College Road, Cranfield, Bedfordshire MK43 0AL, UK
[b] Laboratory of Microbiology and Biotechnology of Foods, Department of Food Science and Technology, Agricultural University of Athens, Iera Odos 75, Athens, GR 11855, Greece

ARTICLE INFO

Keywords:
Machine learning
Pattern recognition
Meat spoilage
Metabolic profiling
Data science
Food quality

ABSTRACT

Over the past decade, analytical approaches based on vibrational spectroscopy, hyperspectral/multispectral imagining and biomimetic sensors started gaining popularity as rapid and efficient methods for assessing food quality, safety and authentication; as a sensible alternative to the expensive and time-consuming conventional microbiological techniques. Due to the multi-dimensional nature of the data generated from such analyses, the output needs to be coupled with a suitable statistical approach or machine-learning algorithms before the results can be interpreted. Choosing the optimum pattern recognition or machine learning approach for a given analytical platform is often challenging and involves a comparative analysis between various algorithms in order to achieve the best possible prediction accuracy.

In this work, "MeatReg", a web-based application is presented, able to automate the procedure of identifying the best machine learning method for comparing data from several analytical techniques, to predict the counts of microorganisms responsible of meat spoilage regardless of the packaging system applied. In particularly up to 7 regression methods were applied and these are ordinary least squares regression, stepwise linear regression, partial least square regression, principal component regression, support vector regression, random forest and *k*-nearest neighbours.

MeatReg" was tested with minced beef samples stored under aerobic and modified atmosphere packaging and analysed with electronic nose, HPLC, FT-IR, GC–MS and Multispectral imaging instrument. Population of total viable count, lactic acid bacteria, pseudomonads, Enterobacteriaceae and *B. thermosphacta*, were predicted. As a result, recommendations of which analytical platforms are suitable to predict each type of bacteria and which machine learning methods to use in each case were obtained. The developed system is accessible via the link: www.sorfml.com.

1. Introduction

Consumers demand food products, which should not be only perfectly safe for human consumption, but visually attractive as well. Additionally, to be more likely to meet the customers' expectations, foodstuff should include fewer additives, be minimally processed and high in quality. (Van Wezemael, Verbeke, de Barcellos, Scholderer, & Perez-Cueto, 2010). In general, food products are considered spoiled if it is unacceptable by the customer, though not necessarily unsafe (Koutsoumanis, 2009). This is also the case with fresh meat, where spoilage is quite a subjective judgement among consumers, often based in the presence of gross discoloration, strong off-odours and development of slime due to intrinsic or extrinsic factors (Nychas, Skandamis, Tassou, & Koutsoumanis, 2008).

Meat spoilage is a very complex phenomenon, which involves significant changes and activities of very different microbial groups

Abbreviations: SVM-R, Support Vector machines regression; RF-R, Random forests regressions; kNN, k-nearest neighbour; PCA, Principal Component Analysis; OLS-R, Ordinary least squares regression; SL-R, Stepwise Linear regression; ML, Machine Learning; RMSE, Root mean square of error; PC-R, Principal Component regression

* Corresponding author at: Bioinformatics Group, Department of Agrifood, School of Water, Energy and Environment, Cranfield University, College Road, Cranfield, Bedfordshire MK43 0AL, UK.
E-mail address: f.mohareb@cranfield.ac.uk (F. Mohareb).

http://dx.doi.org/10.1016/j.foodres.2017.05.013
Received 6 March 2017; Received in revised form 16 May 2017; Accepted 18 May 2017
Available online 20 May 2017
0963-9969/ © 2017 Elsevier Ltd. All rights reserved.

depending on the storage conditions e.g. packaging, temperature (Doulgeraki, Ercolini, Villani, & Nychas, 2012). However, according to EU authorities (Commission regulation, EC, 2005), the quality of fresh meat is evaluated only by viable counts of bacteria able to grow on very generic medium (Total viable count) or on counts of the Enterobacteriaceae family. On the other hand, it is well established that pseudomonads are the major cause of spoilage in aerobic conditions as once they have used all glucose and lactate available, they start to metabolise the nitrogen sources, producing slime and off-odours (Mohareb et al., 2015). Under modified atmosphere packaging, other microorganisms like *B. thermosphacta*, Enterobacteriaceae and lactic acid bacteria are dominating, whilst under vacuum packaging the dominant species are *Pseudomonas* spp., *B. thermosphacta* and *S. putrefaciens* (Nychas, Marshall, & Sofos, 2007). Counting colonies is certainly time-consuming, costly and provide retrospective information (Nychas et al., 2008). Moreover, both the analysis of limited samples and/or their low counts, can significantly underestimate the microbial contribution to meat quality. Recently, rapid, non-invasive methods relying on processing large datasets using computational analysis are gaining popularity (Nychas, Panagou, & Mohareb, 2016). Although such instruments represent an efficient alternative to conventional microbiological analysis, the experimental output is far more complex and usually needs processing before the results can be interpreted. In the food sector, a plethora of machine learning approaches has been followed by different authors in order to predict spoilage in meat samples using metabolomics data (Comprehensive examples are highlighted in (Ropodi, Panagou, & Nychas, 2016). internet technologies and more specifically open sources platforms will enhance food safety management system (Nychas et al., 2016) while will allow supply chains to use virtualizations dynamically in operational management processes. This will improve support for food companies dealing with perishable products, unpredictable supply variations and stringent food safety and sustainability requirements.

While research involving metabolomics data in tandem with machine learning techniques is extensive, guidelines to choose the machine learning method that provides the best results for a specific type of data are still needed. Furthermore, the actual procedure of optimising and validating the spoilage prediction models is computationally extensive, and often requires the availability of suitable resources and statistical knowledge.

Therefore, the aim of this work is (i) to develop spoilage prediction models from data derived from different analytical instruments, and (ii) to implement an accuracy ranking system through a platform (MeatReg), which assesses the suitability of machine learning methods to specific type of metabolic data provided by a certain analytical process. For this study, metabolomics data from minced beef samples stored under aerobic and modified atmosphere packaging were collected using 5 different analytical and imaging instruments: electronic nose (eNose), High Performance Liquid Chromatography (HPLC), Fourier Transformed Infrared Spectroscopy (FT-IR), Gas Chromatography coupled to Mass Spectrometry (GC–MS) and multispectral imaging (MSI).

In particular in this study, seven machine learning methods; namely Ordinary Least Squares regression (OLS-R), Stepwise Linear regression (SL-R), Principal Component regression (PC-R), Partial Least Squares Regression (PLS-R), Support Vector Regression (SVM-R), Random Forests Regression (RF-R) and k-Nearest Neighbours' Regression (kNN-R) were used to predict bacterial counts for Pseudomonads, Lactobacilli, *B. thermosphacta* and Enterobacteriaceae, as well as for the bacterial total viable count. This way, the most suitable analytical platforms to predict bacterial counts for each type of bacteria present in meat stored under aerobic or modified atmosphere conditions were identified and machine-learning methods were ranked for each scenario according to their performance. Additionally, “MeatReg” was made available online as a web-based application in order to provide a flexible and setup-free mean to automate the whole analysis process since internet technologies and more specifically open sources platforms will enhance food safety management system (George-John E. Nychas et al., 2016) while will allow supply chains to use virtualizations dynamically in operational management processes. This will improve support for food companies dealing with perishable products, unpredictable supply variations and stringent food safety and sustainability requirements.

2. Material and methods

2.1. Sample preparation & microbiological analyses

Fresh minced meat was obtained from a central butcher shop in Athens and transported under refrigeration to the laboratory within 30 min. Portions of approximately 75–80 g were placed on styrofoam trays, were stored in air or under modified air packaging (MAP) conditions (20% CO_2/80% O_2) at 4 and 10 °C. For aerobic storage, samples were covered with plastic food membrane for domestic use and for MAP storage, samples were packed into plastic pouches of gas permeability at 20 °C and 50% relative humidity of ca. 25 and 90 cm^3/m^2 per day/10^5 Pa for CO_2 and O_2 respectively, using a HenkoVac 1900 Machine. At appropriate time intervals (approximately every 24 and 12 h for the case of 4 °C and 10 °C respectively), multispectral images of duplicate samples were captured and samples were analysed microbiologically until spoilage was pronounced and sub-samples were stored (− 20 °C) for FTIR, HPLC, GC–MS and e-nose measurements. Additionally, three more samples at 0 h (control samples) were analysed. In total, 105 samples (three control samples and 11–14 duplicate samplings per packaging condition per storage temperature) were analysed. Generally speaking, meat spoilage and sample discoloration was not evident to the naked eye, except between extreme storage times (See Supplementary Fig. S4). Twenty-five gram-portions from each meat samples were weighted aseptically in 400 ml sterile stomacher bags (Seward Medical, London, United Kingdom), containing 225 ml of sterile quarter Ringer's solution (LabM Limited, Lancashire, United Kingdom) and were homogenized for 60 s (Lab Blender 400, Seward Medical). Appropriate serial dilutions were prepared with the same Ringer's solution and duplicate 0.1 or 1 ml samples of the appropriate dilutionswere spread or mixed on the following media: plate count agar (PCA, Biolife 4,021,452, Milano, Italy) for total viable counts (TVC), incubated at 30 °C for 48–72 h; Pseudomonas agar base (PAB, Biolife 401,961,Milano, Italy) for Pseudomonas spp., incubated at 25 °C for 48–72 h; streptomycin thallous acetate-actidione agar (STAA, Biolife 402,079,Milano, Italy) for *B. thermosphacta*, incubated at 25 °C for 72 h; and de Man–Rogosa–Sharpe medium (MRS, Biolife, 4,017,282, Milano, Italy) with pH adjusted to 5.7 with 10 N HCl, for lactic acid bacteria overlaid with the same medium and incubated at 30 °C for 48–72 h. All plates were examined visually for typical colony types and morphological characteristics that were associated with each growth medium. Moreover, the selectivity of each medium was routinely checked by Gram staining and microscopic examination of smears prepared from randomly selected colonies obtained from the media.

2.2. Fourier transform infrared (FTIR) spectroscopy

FTIR spectra were acquired using an FT/IR 6200 JASCO spectrometer (Jasco Corp., Tokyo, Japan). Small portions (~3 g) were placed on the surface of a ZnSe 45° HATR (Horizontal Attenuated Total Reflectance) crystal (PIKE Technologies, Madison, Wisconsin, United States) and, using the Spectra Manager software version 2 (Jasco Corp.), spectra were collected from 4000 to 400 cm^{-1} (100 scans, resolution of 4 cm^{-1}) within a period of 2 min. Prior to the measurements, reference spectra were acquired using the crystal with no added meat. After each measurement, the crystal's surface was cleaned using firstly detergent and distilled water and secondly with analytical grade

Download English Version:

https://daneshyari.com/en/article/5768093

Download Persian Version:

https://daneshyari.com/article/5768093

Daneshyari.com