A Supervised Machine Learning Classification Framework for Clothing Products' Sustainability

Chloe Satinet and François Fouss

UCLouvain, Louvain Research Institute in Management and Organizations (LouRIM), Mons, Belgium
e-mail: chloe.satinet@uclouvain.be, francois.fouss@uclouvain.be

Keywords. Classification, data mining, environmental impact assessment, life-cycle assessment, supervised machine learning, sustainability

Extended abstract.

I. Introduction

These days, many sustainability-minded consumers face a major problem when trying to identify sustainable products. Indeed, as highlighted by Brad et al. (2018) and Gaasbeek et al. (2015), there are a variety of confusing certifications of sustainability, and there are few labels capturing the overall environmental impact of products as the existing procedures for assessing the environmental impact of products over their life-cycle are long, costly and require a lot of data and input from domain experts.

As suggested by previous research (e.g., Meinrenken et al. (2012) and Sousa et al. (2000)), we have decided to further explore the use of supervised machine learning tools to develop a model that could easily and quickly assess products' sustainability over their life-cycle. In particular, could we provide clothing distributors with a tool to assess the sustainability of their product catalogue with the limited data at their disposal, to further help consumers identify sustainable clothing products?

II. Methodology

Supervised machine learning tools train and learn a model from patterns that can be identified in a dataset with training examples that are pairs consisting of an input object and an output variable, so that the learned model can predict the output variable to new input data (see, e.g., Brownlee (2016)).

To correctly construct and assess machine learning models, knowledge from computer/data science is required. However, to develop a precise model to assess clothing products' sustainability over their life-cycle, we need a dataset of clothing products (observations) with their life-cycle characteristics (independent input variables) and known total environmental impact calculated with the life-cycle assessment (LCA) method (dependent output variable). And so far, few products have been subjected to a full LCA or LCA results are seldom shared. Consequently, to assemble such a dataset, a plurality of actors – researchers and companies - and knowledge from different other disciplines are required; knowledge from economics to understand the phases of a product's life-cycle, engineering knowledge to understand how the products are made (and disassembled), and knowledge from ecology/environmental sciences to understand the environmental impacts of the actions carried out to make, distribute, use, and dispose of the products.

We plan to join a research group composed of researchers and organisations to carry out the project. Thus, there is already a multidisciplinary research group within the Nottingham Trent University, i.e., the clothing sustainability research group, that focuses on creating knowledge to support a more sustainable future for clothing design and consumption. There is also Quantis, a consulting group that works with organisations to define and implement sustainability assessment metrics and tools.

III. Results

To get an idea of the results to expect, we assembled a first small dataset of clothing products with their life-cycle characteristics and corresponding known total environmental impact and tested, on a cross-validation basis, nine supervised machine learning algorithms for modelling the relationship between the products' sustainability level and the components and processes entering their life-cycle. Figure 1 shows the accuracy, Mean Squared Error (MSE), Kendall's Tau-b and the Ordinal Classification Index (OCI) calculated based on the overall confusion matrix obtained for each algorithm. According to these first performance metrics, the bagged decision trees and random forest perform best. The resulting models seem to provide rapid correct environmental feedback for a variety of clothing products.

Model	Accuracy	MSE	Tau-b	OCI
k-NN	0,89	0,13	0,91	0,16
LR	0,84	0,19	0,87	0,22
SVM	0,89	0,17	0,89	0,17
ANN	0,90	0,12	0,91	0,15
DT	0,84	0,31	0,83	0,24
Bag. DT	0,90	0,11	0,92	0,14
RF	0,91	0,11	0,92	0,14
Boost. DT	0,88	0,16	0,89	0,18
GBoost. DT	0,90	0,14	0,91	0,15
Dummy1	0,25	2,67	-0,03	0,84
Dummy2	0,41	1,64	-	0,63

Figure 1: Accuracy, MSE, Kendall's Tau-b and OCI of the 9 algorithms and 2 dummy classifiers.

IV. References

Brad, A., Delemare, A., Hurley, N., Lenikus, V., Mulrenan, R., Nemes, N., Trunk, U., & Urbancic, N. (2018). *The false promise of certification*. Changing Markets Foundation.

Brownlee, J. (2016). *Machine Learning Mastery With Python: Understand Your Data, Create Accurate Models, and Work Projects End-to-End*. Machine Learning Mastery.

Gaasbeek, A., Golsteijn, L., & Vieira, M. (2015). *LCA in the labelling industry—Understand the LCA landscape of the production of self- adhesive labels*. PRé Consultants bv.

Meinrenken, C. J., Kaufman, S. M., Ramesh, S., & Lackner, K. S. (2012). Fast Carbon Footprinting for Large Product Portfolios. *Journal of Industrial Ecology*, *16*(5), 669–679. https://doi.org/10.1111/j.1530-9290.2012.00463.x

Sousa, I., Wallace, D., & Eisenhard, J. L. (2000). Approximate Life-Cycle Assessment of Product Concepts Using Learning Systems. *Journal of Industrial Ecology*, *4*(4), 61–81. https://doi.org/10.1162/10881980052541954