Life Cycle Assessment using Machine Learning

Sofia Gomes[1*], Brígida Mónica Faria[1,2], Alexandra Alves Oliveira[1,2], Edgar Pinto[1,4], Matilde Rodrigues[1,3], Manuela Vieira[1,4]

1. ESS, Polytechnic of Porto, Portugal; **2**. Artificial Intelligence and Computer Science Laboratory (LIACC member of LASI), Porto, Portugal; **3.** Center for Translational Health and Medical Biotechnology Research (TBIO), Porto, Portugal; **4**. REQUIMTE/LAQV, ESS, Polytechnic of Porto, Porto, Portugal.

* Corresponding author email: 10221255@ess.ipp.pt

Introduction: Life Cycle Assessment (LCA) is a scientific tool that allows calculating the impact of a product or service on the environment, considering the different phases from planting to transportation, commercialization, consumption, and disposal. (1) LCA requires comprehensive data collection of the inputs and outputs such as raw materials, energy, water, used chemicals and pollutants emissions at each stage of the life cycle. Data is usually obtained from different sources like producers or farmers (primary data), literature reviews, government reports and scientific publications (secondary data) or from associations, non-governmental organizations (NGOs) and international organizations. (2) Data processing and analysis are conducted with the aim of uncovering the resultant environmental impacts. This dissertation, integrated into the project REtail using Technology based on Artificial InteLLigence (RETAILL) (3) aims to apply Machine Learning (ML) techniques to develop surrogate LCA models that can be used to estimate LCA results for new products or services. Both public data and data from the Terras de Felgueiras Cooperative (4) will be used to develop the intended model. By preprocessing and modelling this data, the study aims to provide valuable insights for enhancing sustainability in the production of fresh fruits and vegetables. These insights can guide decision-making and drive continuous improvement in the supply chain. **Objectives:** The objective of this study is to develop a ML model that estimates LCA results for new products or services and that translates environmental indicators into measurable impacts on both the environment and human health, specifically focusing on the production process. Another objective is to establish clusters that represent similar environmental performance of producers or products. **Methods:** The first step was to review the existing literature on the subject. To accurately determine emissions from agricultural activities, validated equations from the Agri-footprint 6 methodology (5) were employed. Preliminary analyses and descriptive statistics of variables such as fertilizers, pesticides and fungicides applied on agriculture from public data assessments were conducted using tools like SPSS and RapidMiner. This latter was used to carry out the construction of decision trees and clusters. To evaluate clustering models, certain indices were considered namely the Davies-Bouldin Index and the Calinski-Harabasz Index. Meanwhile, for assessing decision trees, measures such as accuracy rate, F- measure, and confusion matrix are well-known evaluation criteria. Subsequently, the ML model was developed using Python programming language and libraries such as Scikit- learn, Pandas, and SciPy. **Results:** The analysis of public data reveals results from the cultivation of kiwi, watermelon, citrus, tea, and hazelnut across a total of 865 orchards (6). The results include the development of a tailored ML model for LCA phase that allows. the identification and translation of key environmental indicators into environmental and human health impacts. Furthermore, the clustering results of products and producers enables the observation of patterns in the environmental impact of the production process. **Conclusions:** Overall, this study contributes to the field of sustainability by providing a framework for integrating ML techniques with life cycle assessment, ultimately leading to more efficient and effective practices in agricultural production. The utilization of validated equations from Agri-footprint 6 enhances the reliability of emissions determination from agriculture, contributing to more accurate assessments of environmental impacts. Ultimately, the goal of an LCA is to support informed decision-making and promote more sustainable practices across industries.

Keywords: Life Cycle Assessment, Sustainability, Machine Learning, Data analysis

References:

1. The data you need to make a Life Cycle Assessment (LCA) [Internet]. Ecochain. [cited 2024 Apr 18]. Available from: https://ecochain.com/blog/input-data-for-lca/
2. Pergola M, Persiani A, D'Ammaro D, Pastore V, D'Adamo C, Palese AM, et al. Environmental and Energy Analysis of Two Orchard Systems: A Case Study in Mediterranean Environment. Agronomy. 2022 Oct;12(10):2556.
3. RETAILL's Consortium. REtail using Technology based on Artificial InteLLigence. 2022.
4. Cooperativa Terras de Felgueiras. Apresentação [Internet]. 2020 [cited 2024 Jan 17]. Available from: https://coopfelgueiras.pt/instituicao/apresentacao/
5. Sustainability B. Agri-footprint 6 Methodology Report.
6. Mostashari-Rad F, Ghasemi-Mobtaker H, Taki M, Ghahderijani M, Kaab A, Chau K wing, et al. Exergoenvironmental damages assessment of horticultural crops using ReCiPe2016 and cumulative exergy demand frameworks. J Clean Prod. 2021 Jan 1;278.