

COURSE PRESENTATION & SYLLABUS

Introduction to Machine Learning Applications for Sustainability Research

Prof. Lucas Zanon & Prof. Tiago Sigahi

University of São Paulo (USP), Brazil · Graduate course · 30h · English

Remote lectures · July 13 – 22, 2026

Who are we?

Lucas Zanon – Professor at University of São Paulo, São Carlos School of Engineering, Department of Production Engineering

Tiago Sigahi – Professor at University of São Paulo, Polytechnic School, Department of Production Engineering

Lídia Silva – Post-Graduate Secretary at University of São Paulo, Polytechnic School, Department of Production Engineering

Renato Rodrigues – Monitor, PhD student at University of São Paulo, Polytechnic School, Department of Production Engineering



What we do?

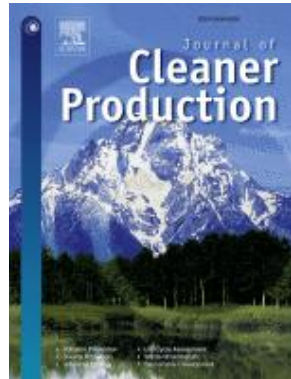
**Data and Decision
Modelling**



Sustainability



What we do?



From isolated actions to systemic transformations: Exploring innovative initiatives on engineering education for sustainable development in Brazil

Tiago F.A.C. Sigahi ^{a,*}, Laerte Idal Sznclwar ^b

Thematic evolution and trends linking sustainability and project management: Scientific mapping using SciMAT

Randal Victor Gibbin ^a, Tiago F.A.C. Sigahi ^{b,*}, Jefferson de Souza Pinto ^{a,c}, Izabela Simon Rampasso ^d, Rosley Anholon ^a



Exploring ESG Indicators and the Sustainable Development Goals: A Fuzzy Multi-Criteria Approach to Cause-and-Effect Analysis

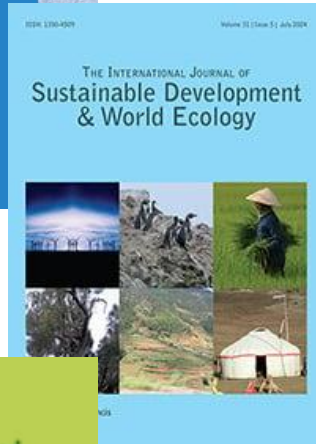
[Ícaro Guilherme Félix da Cunha](#) ✉, [Renata Veloso Santos Policarpo](#), [Paula Cristina Senra de Oliveira](#), [Lucas Gabriel Zanon](#), [Daisy Aparecida do Nascimento Rebelatto](#)



Machine Learning in Reverse Logistics: A Systematic Literature Review

Abner Fernandes Souza da Silva ¹ , Virginia Aparecida da Silva Moris ¹, João Eduardo Azevedo Ramos da Silva ¹, Murilo Aparecido Voltarelli ² and Tiago F. A. C. Sigahi ^{1,3,*} 

What we do?



Relations between supply chain performance and circular economy implementation: A fuzzy cognitive map-based analysis for sustainable development

[Lucas Gabriel Zanon](#) ✉ [Ana Carolina Bertassini](#), [Tiago Fonseca Albuquerque Cavalcanti Sigahi](#), [Rosley Anholon](#), [Luiz Cesar Ribeiro Carpinetti](#)



A multicriteria approach for assessing the maturity of supply chains regarding the implementation of circular economy practices in Brazil

Gustavo Tietz Cazeri, [Tiago F. A. C. Sigahi](#) ✉, Izabela Simon Rampasso, Gustavo Hermínio Salati Marcondes de Moraes, Lucas Gabriel Zanon, Muriel de Oliveira Gavira, João Henrique Paulino Pires Eustachio, Walter Leal Filho & Rosley Anholon ...show less



Exploring the Barriers to Circular Economy Implementation in Agricultural Sector: A Way Forward to Achieve Sustainable Development Goal 2 (Zero Hunger)

Júlia Santos Humberto¹ · Christian Eric Barrantes Briceño¹ · Lucas Gabriel Zanon¹ · Marcelo Seido Nagano¹



What we do?



Beyond the hype: A sociotechnical perspective on AI's potentials and constraints in bridging the gap to basic-needs sustainability in higher education institutions

Walter Leal Filho ^{a,b,c}, Tiago F.A.C. Sigahi ^{d,e,*}, Antonis Skouloudis ^f,
Izabela Simon Rampasso ^g, Rosley Anholon ^h



Classical paradigms versus complexity thinking in engineering education: an essential discussion in the education for sustainable development ✓

Tiago F.A.C. Sigahi ✉; Izabela Simon Rampasso; Rosley Anholon; Laerte Idal Szelwar

Teaching the “10 Rs” of circular economy in technological-level undergraduate courses offered in Brazil: an exploratory analysis via fuzzy TOPSIS class ✓

Ricardo Favaro; Gustavo Hermínio Salati Marcondes de Moraes; Tiago F.A.C. Sigahi ✉; Jefferson de Souza Pinto; Izabela Simon Rampasso; Suzana Regina Moro; Rosley Anholon



Promoting sustainable development via stakeholder engagement in higher education

Walter Leal Filho ^{1,2,3}, Tiago F. A. C. Sigahi ^{4,5*}, Rosley Anholon ⁶, Bianca Gasparetto Rebelatto ⁷, Inga Schmidt-Ross ⁸, Susanne Hensel-Börner ⁸, Dirk Franco ^{9,10}, Thomas Treacy ¹¹ and Luciana Londero



ABOUT THE TEAM

What we do?



Difficulties experienced by Latin American oil and gas companies in the integrated adoption of practices aligned with the UN SDGs ✓

Fabiola M.M.G. Borges ✉; Rosley Anholon; Izabela Simon Rampasso; Tiago F.A.C. Sigahi; Gustavo Herminio Salati Marcondes de Moraes; Walter Leal Filho

Critical points regarding verifiability in Brazilian companies' sustainability reports: a fuzzy TOPSIS class perspective ✓

Isabela Caroline de Sousa; Tiago F.A.C. Sigahi ; Lucas Gabriel Zanon; Izabela Simon Rampasso; Gustavo Herminio Salati Marcondes de Moraes; Walter Leal Filho; Rosley Anholon

Analysis of the quality of sustainability reports published by Brazilian companies: An analytic hierarchy process-grey clustering approach

[Isabela Caroline de Sousa](#), [Tiago F. A. C. Sigahi](#) ✉, [Izabela Simon Rampasso](#), [Jefferson de Souza Pinto](#), [Lucas Gabriel Zanon](#), [Walter Leal Filho](#), [Rosley Anholon](#)



What we do?



Challenges in adopting energy management systems by Brazilian companies: A fuzzy DEMATEL approach

Rodrigo César Gomes, Tiago F.A. C. Sigahi, Jefferson de Souza Pinto, Izabela Simon Rampasso, Lucas Gabriel Zanon, Milena Pavan Serafim, Walter Leal Filho & Rosley Anholon [✉](#) ...show less

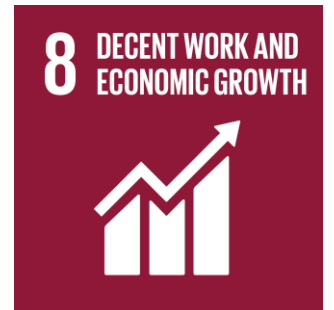
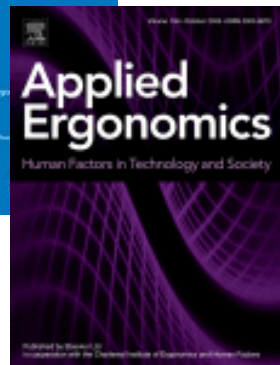
Selection of photovoltaic panels for floating systems: an analysis based on Entropy, CRITIC and TOPSIS

Alessandra Brito Leal, Hélio Nunes de Souza Filho, Lucas Gabriel Zanon, Tiago F. A. C. Sigahi, Izabela Simon Rampasso & Rosley Anholon [✉](#)



Actionable insights for human factors and ergonomics to achieve the Sustainable Development Goals: mapping past advancements and proposing future strategic paths

Tiago F. A. C. Sigahi [✉](#) [ID](#), Andrew Thatcher [ID](#), Ivan Bolis [ID](#), Izabela Simon Rampasso [ID](#), Patrícia Saltorato [ID](#) & Rosley Anholon [ID](#)

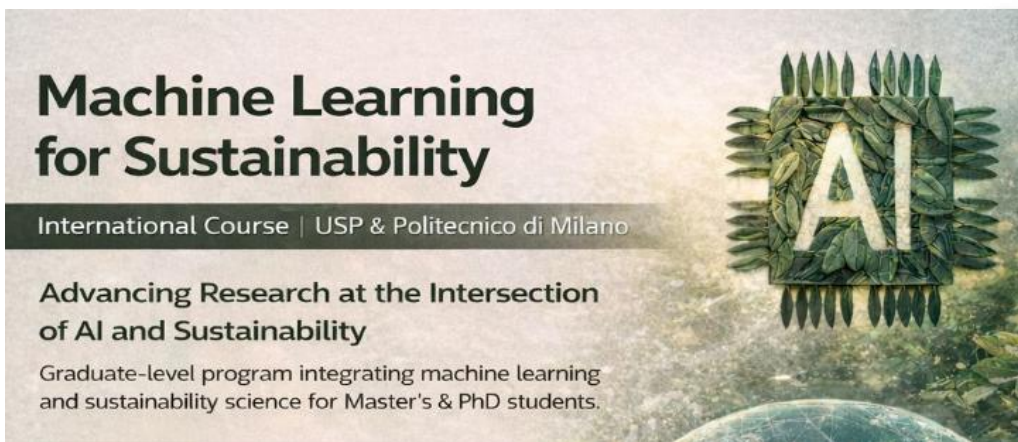


Ergonomics and human factors for a sustainable, resilient, human-centered future

Ivan Bolis ^a [✉](#) [ID](#), Tiago F.A.C. Sigahi ^b, Andrew Thatcher ^c, Rosley Anholon ^d



What this course is about



An introduction to supervised machine learning, with a focus on applications in sustainability research. It gives a solid foundation in core ML concepts and workflows, with hands-on implementation in Python.

Students work through **hands-on exercises, case studies and a final project** on real sustainability challenges — bridging technical methods and sustainability-oriented research questions.

30h

workload

Graduate

Master's & PhD

English

instruction

Remote

lectures

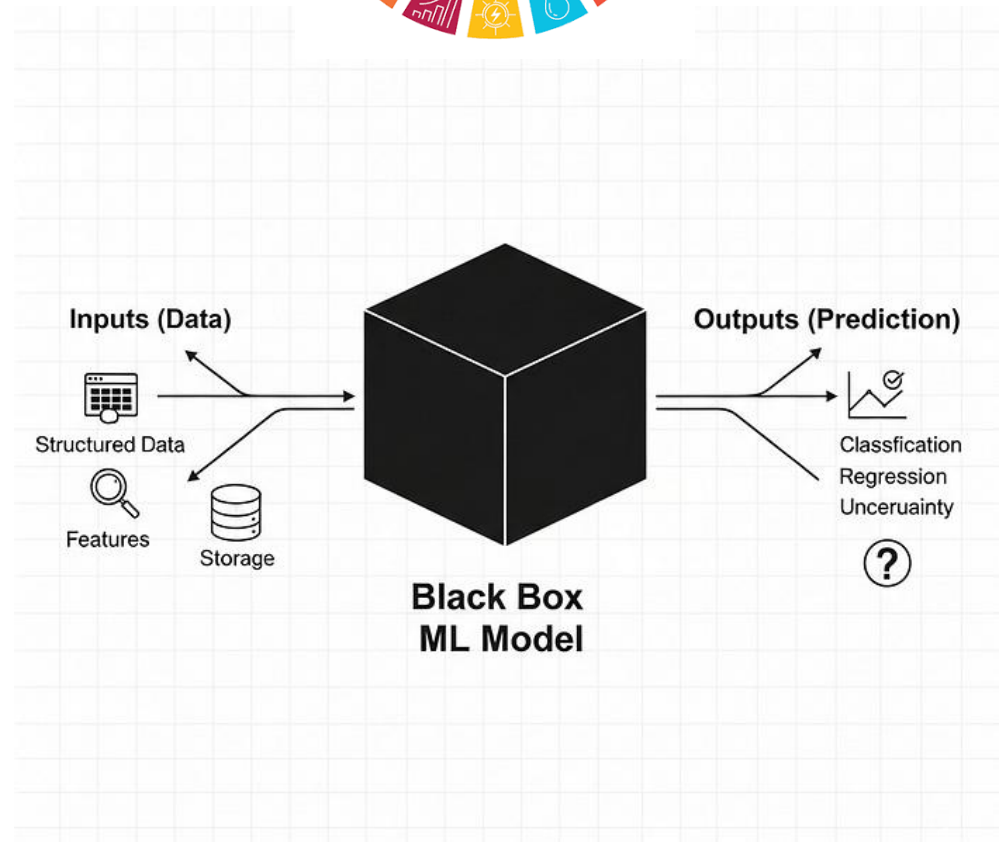
Jul 13–22

2026

LEARNING OBJECTIVES

By the end, students will be able to

- Understand the fundamentals of supervised machine learning.
- Differentiate the main problem types — classification and regression — and their applications.
- Apply data preprocessing, exploratory analysis and visualization in Python.
- Implement and interpret core models: linear & polynomial regression, KNN, Naive Bayes, decision trees and ensembles.
- Evaluate and compare models with appropriate metrics, and tune simple hyperparameters (e.g., K, tree depth).
- Develop an applied project applying ML to a real sustainability problem.



PLANNED SCHEDULE

Ten days, from data to final project

Mon 13/07	Lecture 1 - Course intro & group formation	Syllabus, datasets and course settings
Tue 14/07	Lecture 2 - Introduction to ML and Sustainability <i>Sustainability context: CO2 emissions</i>	Foundations of supervised & unsupervised ML
Wed 15/07	Lecture 3 - Data exploration & preprocessing <i>Sustainability context: Energy transition</i>	Explore & preprocess the SDG data
Thu 16/07	Lecture 4 - Regression models (linear, polynomial, KNN) <i>Sustainability context: Wedding Cake model e SDG 8</i>	Predict SDG 8 Define the group's proposition
Fri 17/07	Lecture 5 - KNN, Naive Bayes & model evaluation <i>Sustainability context: SDGs & sustainability frameworks – Part 1</i>	Classify the SDG Index
Mon 20/07	Lecture 6 - Decision trees & ensembles <i>Sustainability context: SDGs & sustainability frameworks – Part 2</i>	Classify the SDG Index
Tue 21/07	Lecture 7 - Practical day — final project	Final project development
Wed 22/07	Lecture 8 - Final presentation	Group presentations

*Application context may change.

How you are graded

20%

Participation &
in-class exercises

40%

Assignments
(Python notebooks)

40%

Final project
(research-oriented)

Key references

James, G., Witten, D., Hastie, T. & Tibshirani, R. (2023). An Introduction to Statistical Learning: With Applications in Python. Springer.

Agudelo, M. A. L., Jóhannsdóttir, L. & Davídsdóttir, B. (2019). A literature review of the history and evolution of corporate social responsibility. Int. J. of CSR, 4(1).

Busch, T., Barnett, M. L., Burritt, R. L. et al. (2023). Moving beyond “the” business case: how to make corporate sustainability work. Business Strategy and the Environment.

Wickert, C. (2021). CSR research in the Journal of Management Studies: a shift from a business-centric to a society-centric focus. JMS, 58(8).

Who are you?

- ✓ **Name and country**
- ✓ **Background and university**
- ✓ **What you are researching now**
- ✓ **Motivation to take this course**



THE COURSE DATASETS

Sustainable Development Goals (SDG) data

Where it comes from, how it is built, and how we use it in the course

The Sustainable Development Report 2025

The data come from the **Sustainable Development Report 2025 (SDR2025)**, published annually by the UN Sustainable Development Solutions Network (SDSN).

Three source sheets of SDR2025 are combined:

- All Raw Data — a country-year indicator panel, 2000–2024
- Backdated SDG Index — historical aggregate scores
- SDR2025 Data — 2025 aggregate scores

Raw indicators are merged with the aggregate SDG Index Score (**sdgi_s**) on country code (ISO3) and year.

5,382

country-year rows

208

countries

2000–
2024

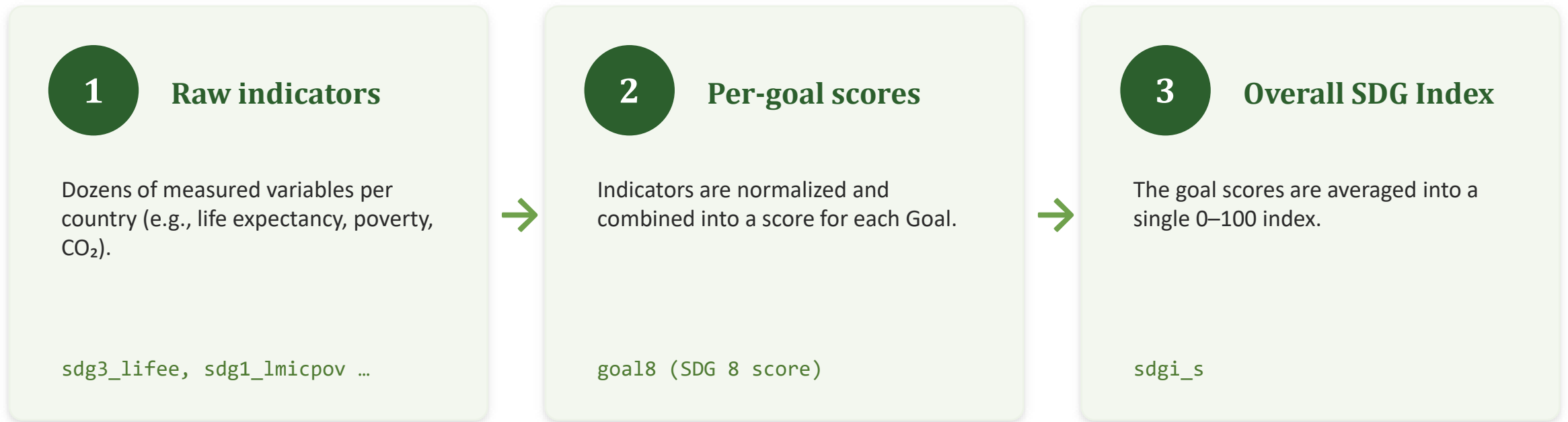
time span

17

SDGs covered

From indicators to the SDG Index

The SDG Index summarizes how far each country is from achieving the 17 Goals, on a 0–100 scale. It is built in three steps:



HOW THE TABLES ARE ORGANIZED

One row = one country-year

Both files share the same shape. Every row is a single country in a single year, described by many SDG indicators.

id	ISO3 country code (e.g., AFG)
Country	country name
year	2000 – 2025
indexreg	world region (e.g., OECD)
sdg1_* ... sdg17_*	indicators, grouped by SDG

Identifiers (id, Country) and the region are not predictors — the SDG indicators are.

The `sdg{goal}_{indicator}` convention

Each indicator column is named `sdg<goal number>_<short name>`. A few examples:

`sdg3_lifec`

Life expectancy at birth

SDG 3 · Health

`sdg1_lmipov`

Poverty headcount

SDG 1 · No poverty

`sdg6_water`

Access to drinking water

SDG 6 · Water

`sdg7_elecac`

Access to electricity

SDG 7 · Energy

`sdg9_intuse`

Internet use

SDG 9 · Infrastructure

`sdg13_co2gcp`

CO₂ emissions

SDG 13 · Climate

Note: these short codes (e.g., lifec, elecac, intuse, lmipov) are the original SDR2025 abbreviations — the odd spellings are correct, not typos. Use the column names exactly as they appear in the file.

Regression — predicting the SDG 8 score

`sdg8_dataset_final.xlsx`

Target: goal8 — the SDG 8 score (“Decent Work & Economic Growth”), a continuous number (~40–91).

You predict goal8 from indicators of the OTHER SDGs. There are no `sdg8_*` predictor columns, so the target cannot be trivially reconstructed.

Two sheets, split by time:

- Treino (2000–2018) — 3,743 rows
- Teste (2019–2024) — 1,182 rows

4,925 rows

197 countries

60 columns

DATASET B

Classification — high vs. low SDG performers

`sdgi_classification_dataset.xlsx`

Target: `sdgi_class` — 1 if the overall SDG Index (`sdgi_s`) is above the median (a high performer), 0 otherwise.

Training is balanced (1,976 vs 1,976); the test set is imbalanced (949 vs 481) — a good reason to look beyond accuracy.

- Treino (2000–2018) — 3,952 rows
- Teste (2019–2025) — 1,430 rows

⚠ **Data leakage:** `sdgi_s` is the index the label is built from — it must be dropped from the predictors.

5,382 rows

208 countries

59 columns

Split by time, not at random

Each file already comes divided into two sheets. We train on the past and test on the future.

TRAINING

2000 – 2018

the model learns here

TEST

2019 +

judged here, unseen

Why by time? For data ordered in time, predicting future years from past ones is the honest test. A random split would let the model peek at the future — a subtle form of data leakage.

How each lecture uses the data

Lecture 2

Introduction to ML and Sustainability

Foundations

Lecture 3

Data exploration & preprocessing

Both files — explore, clean, encode, scale, split

Lecture 4

Regression models

Dataset A — predict goal8 (linear, polynomial, KNN)

Lecture 5

KNN, Naive Bayes & evaluation

Dataset B — classify sdgi_class

Lecture 6

Trees, forests & boosting

Dataset B — classify + feature importance

OPENING THE FILES

Two sheets per file — Treino & Teste

- In Excel: click the sheet tabs at the bottom (Treino = train, Teste = test).
- In Colab: `pd.read_excel(FILE, sheet_name=0)` for train, `sheet_name=1` for test.

Sources

- Sachs, J. et al. (2025). Sustainable Development Report 2025. SDSN / SDG Transformation Center.
- Rockström, J. & Sukhdev, P. (2016). The SDG “Wedding Cake”. Stockholm Resilience Centre.