



**UNIVERSITY COLLEGE OF  
ENGINEERING, DINDIGUL**

**ECO-FRIENDLY**

**ALTERNATIVES TO**

**EVERYDAY PRODUCTS**

- Gayathri S
- Siva D
- Niranjana C
- Hariharan R

**Date:30/04/2025**

The Sustainable Lifestyle Rating Dataset, curated by Naveen Nas, provides a comprehensive overview of individuals' lifestyle choices and their alignment with sustainability principles. Here's an in-depth analysis covering various aspects:

---

## Dataset Overview

This dataset encompasses a range of variables that capture personal habits, consumption patterns, and environmental awareness. It includes features such as:

- **Transportation Modes:** Indicates whether individuals prefer public transport, cycling, or personal vehicles.
- **Energy Consumption:** Details on household energy usage and renewable energy adoption.
- **Dietary Choices:** Information on vegetarianism, veganism, or meat consumption.
- **Waste Management:** Practices related to recycling, composting, and waste reduction.
- **Water Usage:** Insights into water conservation efforts.
- **Sustainability Rating:** A composite score reflecting the overall sustainability of an individual's lifestyle. ([Sustainable Power Enterprise Dataset - Kaggle](#))

The dataset is structured to facilitate both classification and regression analyses, making it suitable for various machine learning applications.

---

## Problem Statement

The primary objective is to develop predictive models that can assess and classify individuals based on their sustainability ratings. By analyzing lifestyle attributes, the goal is to:

- **Predict Sustainability Scores:** Estimate the sustainability rating of individuals based on their lifestyle choices.
- **Identify Key Influencers:** Determine which factors most significantly impact sustainable living.
- **Promote Sustainable Practices:** Use insights to encourage behaviors that lead to higher sustainability scores.

---

## Data Processing

Effective data preprocessing is crucial for accurate modeling. The steps include:

- **Data Cleaning:** Handling missing values, correcting inconsistencies, and removing duplicates.
- **Normalization:** Scaling numerical features to ensure uniformity.
- **Encoding Categorical Variables:** Transforming categorical data into numerical formats using techniques like one-hot encoding.
- **Feature Selection:** Identifying and retaining features that contribute most to the predictive power of the model.

---

## Methodology

Several machine learning algorithms can be employed to model the data:

- **Logistic Regression:** For binary classification tasks, such as determining whether an individual has a sustainable lifestyle.
- **Decision Trees and Random Forests:** To capture nonlinear relationships and interactions between variables.
- **Support Vector Machines (SVM):** Effective for classification with clear margins of separation.

- **Neural Networks:** For capturing complex patterns in the data.  
([Sustainable Power Enterprise Dataset - Kaggle](#))

Model evaluation metrics include accuracy, precision, recall, F1-score, and ROC-AUC.

---

## Results

Upon training and evaluating the models:

- **Random Forests:** Achieved high accuracy, indicating strong performance in classification tasks.
- **Feature Importance Analysis:** Revealed that transportation choices and energy consumption are among the top predictors of sustainability ratings.
- **Model Interpretability:** Decision trees provided clear insights into decision-making paths, aiding in understanding the influence of various features.

---

## Discussion

The analysis underscores the significant impact of daily choices on sustainability. Transportation and energy use emerge as critical areas

where behavioral changes can lead to substantial improvements in sustainability scores. However, the dataset's limitations, such as potential self-reporting biases and lack of demographic diversity, should be considered when generalizing findings.

---

## Solution Impact

By identifying key lifestyle factors that influence sustainability, targeted interventions can be designed:

- **Policy Formulation:** Informing policies that promote public transportation and renewable energy adoption.
  - **Educational Campaigns:** Raising awareness about the environmental impact of dietary choices and waste management.
  - **Personalized Recommendations:** Developing tools that provide individuals with actionable steps to enhance their sustainability scores.
-

# Conclusion

The Sustainable Lifestyle Rating Dataset serves as a valuable resource for understanding the interplay between personal habits and environmental impact. Through meticulous data analysis and modeling, it is possible to identify actionable insights that can drive both individual and collective efforts toward a more sustainable future.

---

For further exploration and hands-on experience with the dataset, you can access it on Kaggle: ([Sustainable Lifestyle Rating Dataset - Kaggle](https://www.kaggle.com/datasets/naveennas/sustainable-lifestyle-rating-dataset))

---

# REFERENCES

Dataset:

<https://www.kaggle.com/datasets/naveennas/sustainable-lifestyle-rating-dataset>

Project Link:

<https://github.com/GayathiriSundarakani/ai-green-skills.git>

