Prepared by Deivyansh Singh

Parakeet Model

Emission Factor Recommendation for Carbon Footprinting using Generative AI

Key learnings from the study:

1. Quantifying greenhouse gas emissions from products and business activities is important to measure environmental impact and mitigating action.
2. LCA (Life Cycle Assessment): It is the estimation of environmental impact of the products from extraction to end-of-life.
3. EFs (Emission Factors): These factors are used to indirectly estimate GHGs per unit of activity.
4. Parakeet Model leverages LLMs (Large Language Models) to recommend EFs.
5. It parses business activity descriptions.
6. It has shown impressive results with an average precision of 86.9%.
7. The goal of the model is to support organisations in achieving net-zero carbon emissions.
8. Practitioners acquire EFs to map each activity with the appropriate EF and estimate the footprint by scaling the EF.
9. Total carbon footprints = sum of individual emission estimates.
10. Prior works (previous models) used to work using semantic text matching to recommend EFs. They do not provide human-interpretable justification for their recommendations.
11. The Parakeet Model combines semantic-text matching and text generation using LLMs for EF mapping.
12. Some Features of the Parakeet Model:
    - Carefully designed Prompts
    - Retrieval Augmented Generation
    - Human-in-loop Validation

13. It directly ingests real-world business data, whether it is structured or unstructured.
14. Provides human-readable justifications to enable expert verifications.
15. LCA (Life Cycle Assessment) Approaches:
    - EEIO-LCA (Environmentally extended Input-Output based LCA): It uses a top-down macroeconomic approach using supply-use tables and sector-level environmental data.
    - pLCA (Process-based LCA): It uses a bottom-up approach by tracking input-output across the supply chain. It requires specific EFs to operate.
16. Parakeet's Approach:
    - LLM generates a plain language description of input query.
    - The embedded model finds the top-K semantically similar EFs.
    - Using both LLMs recommends top EF.
    - The human confirms or overrides the recommendation.
17. Implementation Details:
    - Uses the general text-embedded (gte-large) model.
    - Uses Claude 3 Sonnet Model
    - Uses "one-shot" prompts
18. Results:
    - Based on six real-world datasets
    - Public:
        1. Government of Austin Invoices (EEIO-LCA)
        2. Katana Invoices (EEIO-LCA)
        3. Food.com Ingredients (pLCA)
    - Proprietary:
        1. Procurement (pLCA)
        2. Heavy Equipment (pLCA)
        3. Packaging (pLCA)
    - Metrics:
        1. Precision@1: Top recommendation Accuracy
        2. Precision@K: Accuracy within Top-K Recommendation
    - Results: High Prediction@K rates show that the model was able to stand on human expectations, as most of the time EF recommendations were confirmed by humans.

Note: Using LLMs for both paraphrasing and recommendation improves performance by >10%.

19. Challenges with pLCA:

- High Error Rate
- Required Detailed
- Requires more metadata and specifics.

20. LLM Limitations:

- Non-Deterministic Responses
- A small percentage of responses are not schema-compliant.

21. Other Challenges: Inconsistent human annotations hinder model fine-tuning.