AI MODEL DEVELOPMENT

1st February 2026

Group 4 : Dhiksha Rathis, Shreya Verma

PHASE 1 - FRAMING BRIEF

Domain: Green AI, sustainable computing

Topic: Carbon Footprint Measurement of Large Language Models

Main Research Question: How do we accurately measure and compare the carbon footprint of different LLMs across their lifecycle?

Sub-questions:

1. What are the major sources of emissions in LLM training vs. inference?
2. How do different studies measure and report carbon metrics?
3. What factors (model size, hardware, location) most impact carbon footprint?
4. How do carbon estimates vary across different LLM families?
5. What data is missing or inconsistent in current carbon reporting?

Scope:

- Include: Published research on LLM carbon emissions, energy consumption measurement methods, lifecycle analysis
- Exclude: General data center emissions not specific to LLMs, theoretical efficiency improvements without empirical data, carbon offset programs

Chosen Tasks:

1. Claim-evidence extraction: Extract specific claims with verbatim supporting quotes and citations
2. Cross-source synthesis: Compare multiple sources to identify agreements and disagreements

Models: Claude Opus 4.5, Claude Sonnet 4.5, GPT-5, Gemini 3

Sources Chosen: (Papers)

[1] Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and Policy Considerations for Deep Learning in NLP. ACL 2019.

[2] Luccioni, A.S., Viguier, S., & Ligozat, A.L. (2022). Estimating the Carbon Footprint of BLOOM. arXiv:2211.02001.

[3] Patterson, D., et al. (2021). Carbon Emissions and Large Neural Network Training. arXiv:2104.10350.

[4] Schwartz, R., Dodge, J., Smith, N.A., & Etzioni, O. (2020). Green AI. Communications of the ACM.

[5] Henderson, P., et al. (2020). Towards the Systematic Reporting of the Energy and Carbon Footprints of ML. arXiv:2002.05651.