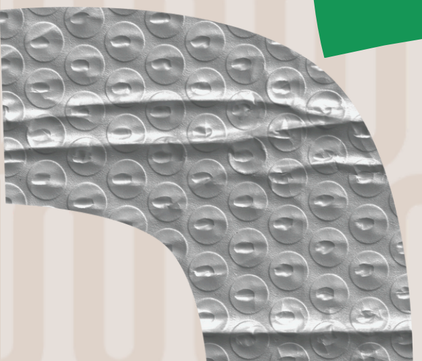
Track Sustainability

ESTIMATION AND SIMULATION OF THE CO_2 IMPACT OF LLM (LARGE LANGUAGE MODEL) QUERIES



Team#16

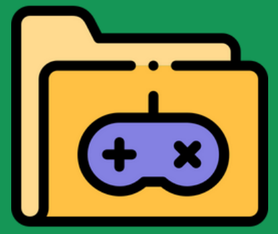
Presented by: Yasmine MAARBANI, Yannis BRIK, Sami CHELLIA & Aya BOUANANE

What is the impact of LLM on the environment ?

Every AI answer has a physical footprint —electricity consumed, carbon emitted, water evaporated, and hardware worn.

Environmental sustainability involves making responsible choices that ensure the long-term health of our planet.

Data Exploration Shape & Visualizations



Dataset:
llm_inference_energy_consumption_final.csv

Exploration Tools

Watsonx Studio (correlation plots, scatter plots, heatmaps)

Key Insights

Strong correlation between total_duration_s and energy_consumption_llm_total
Token lengths and durations show diverse patterns across models

GOAL

Build a complete solution (model + interface) to estimate and simulate the carbon footprint of LLM inference in real time.

Variable Statistics & Data Quality

DESCRIPTIVE STATS

Mean, Median, STD for durations, Tokens, Energy

MISSING VALUES

Numeric: filled with median
Categorical: handled via one-hot encoding with unknowns

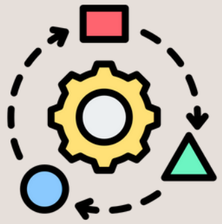
OUTLIERS

Handled via clamping in UI

FEATURE INTERACTIONS

Prompt vs response durations, token ratios .

Feature Engineering



Transformations

Durations normalized to seconds
total_tokens = prompt + response
prompt_response_ratio created

Standardization

StandardScaler(with_mean=False) for numeric features
OneHotEncoder for categorical variables

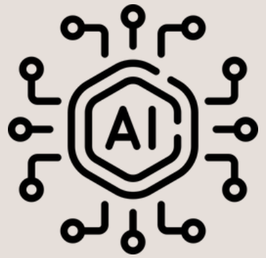
LEAKAGE PREVENTION

excluded direct energy/power columns

Why these models ?

Linear Regression

Simple
Interpretable
LightWeight



Random Forest

Captures non-linearities
Better Accuracy

TWO TARGETS

Per-request (kWh/request)
Per-token (kWh/token × total_tokens)

Results & Metrics

Per-request (kWh/Request)

Model	MAE (kWh)	RMSE (kWh)	MAPE	R2	Acc (terciles)	Recall macro	Recall weighted
LinearReg	0.00006012	0.00025289	109.53%	0.643	0.642	0.642	0.642
RandomForest	0.00000292	0.00004863	1.90%	0.987	0.993	0.993	0.993

Per-TOKEN (kWh/Token)

Model	MAE (kWh/token)	RMSE (kWh/token)	MAPE	R2	Acc (terciles)	Recall macro	Recall weighted
LinearReg	0.0000000513	0.000000228	12.45%	0.944	0.879	0.879	0.879
RandomForest	0.0000000159	0.0000001048	3.54%	0.988	0.962	0.962	0.962

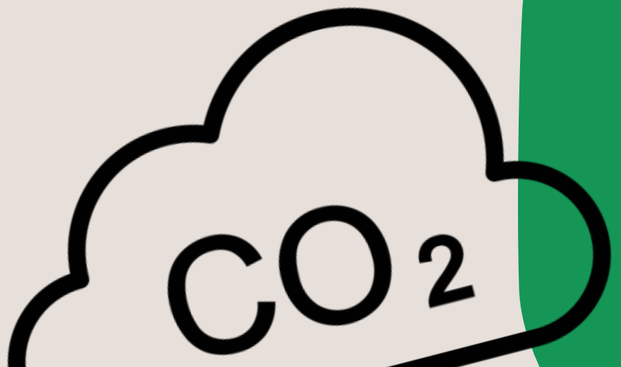
WHY WE USE THESE METRICS ?

MAE: absolute error (real-world impact) MAPE: relative error (scalability)

R^2: variance explained Tercile accuracy: ranking consistency

CO_2 Conversion & Simulation

Formula: CO_2e (kg) = kWh × grid intensity (kg/kWh)



Simulation capabilities: Vary token count
Change energy mix (EU, US, India, Renewable)
Optimize for CO_2 budget

Use case: scenario planning for sustainable AI usage

WEB INTERFACE

Using Streamlit and plotly

Features

Single request estimator
Multi-turn conversation simulation
Batch CSV upload
CO_2 heatmaps
Pipeline comparison
Explainability via feature importance

User-friendly

Sliders for quick adjustments
Presets for easy setup
Downloadable results for sharing and analysis

DEMONSTRATiON WEB INTERFACE

Impact & Next Steps

Impact

Raises awareness of AI's environmental cost
Empowers users to make greener decisions

Next steps

Integrate into real-world LLM workflows
Expand to more models and regions
Enhance UI and model robustness

THANK YOU

Presented by: Yasmine MAARBANI, Yannis BRIK, Sami CHELLIA
& Aya BOUANANE
Team #16