

Youri Peskine Raphäel Troncy Paolo Papotti



EURECOM at FrugalAI Challenge

youri.peskine@eurecom.fr



Bert browsing internet memes

2025 Frugal Al Challenge

https://frugalaichallenge.org/



FEBRUARY 11th, 2024 PARIS, FRANCE

By: Theo Alves da Costa, Juliette Fropier, Sasha Luccioni



WHY FRUGALITY?





Resources are constrained globally (human pressure on raw materials / energy / water) and locally (deployment in specific environments). We need to make sure that **AI really brings net benefit** (what it can enable to solve/reduce vs what it costs)

Change the narrative -There is a current emphasis on "bigger is better" in AI - that bigger models, with more parameters, perform better

→ Apply digital service ecodesign practices
→ Set up a metrology for the consumption of development system resources

https://www.boutique.afnor.org/en-gb/standard/afnor-spec-2314//fa208976/421140#



DETECTING CLIMATE CHANGE MISINFORMATION CLASSIFYING REGIONS WITH WILDFIRE RISKS DETECTING ILLEGAL LOGGING

We want to make sure that our AI applications match real life needs in the fight against climate change and environmental destruction.

TASKS

EVALUATION APPROACH

- We evaluated both performance and efficiency on each of the tasks, based on the real-world constraints of each task
- We will benchmark the energy consumption of approaches using:
 - Code Carbon, an open-source Python package
 - Hugging Face Spaces, a backend for running code and demos



CARBO

THE FRUGAL AI CHALLENGE

- 4-weeks challenge
- 400 subscribers in more than 10 countries (academics, companies, individuals)
- 64 final submissions in total evaluated and reviewed manually

rugal Al	Challenge - S	ubmission Por	tal			
omit your mod	el results for any of the thr	ee tasks: Text, Image, or Au 🛎 Image Classification	dio classification.			
Space URL			,	NPI route (Advanced)		
username/y	our-space			/text		
	1. Evaluate model		2. Submit to public lead	lerboard (optional)	3. Submit to	final evaluation form
Accuracy		En	ergy Consumed (Wh)		Emissions (gCO2eq)	
0		C			0	



Factors

- Automatic fact-checking is "opaque" → veracity Binary labels
- Shift focus on adding context to understand documents
- Textual features play an important role

Factors:

- Persuasion Techniques
- Conspiracy theories
- Emotion
- Sentiment
- Political-leaning
- Tropes

Misinformation-related factors

General factors

Misinformation-related factors

• Persuasion techniques

- Push agendas and narratives
- Leveraging emotions
- Using fallacies



- Conspiracy theories
 - Malevolent schemes
 - Disproved or unproven accusations



SemEval 2024 - Results

Our approach: Ensembling of BERT-based models

"EURECOM at SemEval-2024 Task 4: Hierarchical Loss and Model Ensembling in Detecting Persuasion Techniques", Peskine Y., Troncy R., Papotti P., Proceedings of the 18th International Workshop on Semantic Evaluation

	Team	F1H
1	914isthebest	0.752
2	BCAmirs	0.699
3	OtterlyObsessedWithSemantics	0.697
4	TUMnlp	0.674
5	GreyBox	0.670
6	NLPNCHU	0.663
7	Puer	0.660
8	EURECOM	0.655
9	SuteAlbastre	0.652
10/33	UMUTeam	0.648

Code shared on github: https://github.com/D2KLab/semeval-2024-task-4

MediaEval 2021 - Results

Our winning approach: Ensembling of Covid-Twitter-BERT models



"Detecting COVID-19-Related Conspiracy Theories in Tweets", Peskine Y., Alfarano G., Harrando I., Papotti P., Troncy R., Multimedia Benchmark Workshop (MediaEval 2021)

	Team	MCC	
1	D2KLab	0.733	
2	SELAB_HCMUS	0.648	-
3	Deltamap	0.632	
4	SELAB-HMUS-Junior	0.599	
5	Upsilon Labs	0.454	
6	MG-UCB	0.450	
7	FakeINA	0.446	
8	OTS-UEC	0.413	
9/17	Delta_IITKGP	0.347	

Code shared on github: https://github.com/D2KLab/mediaeval-fakenews

MediaEval 2022 - Results

"Detection of COVID-19-Related Conspiracy Theories in Tweets using Transformer-Based Models and Node Embedding Techniques", Peskine Y., Papotti P., Troncy R., MediaEval 2022, Multimedia Evaluation Workshop

	Team	MCC		Team	MCC		Team	MCC	
1	Anticon	0.738	1	AIDA_UPM	0.459	1	D2KLab	0.719	
2	D2KLab	0.710	2	D2KLab	0.355	2	Anticon	0.699	
3	NLP-SRL	0.705	3	Anticon	0.283	3	NLP-SRL	0.698	
4	Zephyros	0.596	4	CS@QAU	0.241	4	Zephyros	0.558	
5	MMSys	0.488	5	Zephyros	0.111	5	MMSys	0.452	
6	CS@QAU	0.056	6	NLP-SRL	0.041				
			7	HPC-SIMULA	-0.008				

Code shared on github: https://github.com/D2KLab/mediaeval-fakenews

Frugal AI Challenge



https://frugalaichallenge.org/

Detecting Climate Change related Disinformation:

- Textual data
- Identify misleading or false information related to climate change
- 8 classes: "climate change is not happening", "It is not human-caused", "science is unreliable", etc.

Evaluation:

- Accuracy
- Energy consumption

Frugal AI Challenge

Proposed approaches:

- Fine-tuned BERT
- Embeddings + Classifier
- Fine-tuned LLMs



We train each model on the same 80% of the data and test them on the same 20%. Energy and emissions might vary depending on hardware used.

Model	Size (M)	Accuracy	F1 macro	MCC average	energy_c onsumed _wh	emission s_gco2eq	0_not_r elevant	1_not_ha ppening	2_not_ human	3_not _bad	4_solutions _harmful_u nnecessary	5_scienc e_is_unr eliable	6_propo nents_b iased	7_fossil _fuels_n eeded
Modern- BERT- large	395	0.758	0.744	0.716	2.78716	0.15619	0.782	0.812	0.803	0.732	0.744	0.725	0.734	0.631
CT- BERT	336	0.753	0.741	0.711	1.9166	0.1074	0.746	0.805	0.832	0.742	0.744	0.731	0.741	0.615
gte- large	434	0.747	0.737	0.704	2.39456	0.13419	0.743	0.818	0.796	0.784	0.719	0.669	0.763	0.662
gte- base	136	0.726	0.714	0.68	0.74074	0.04151	0.723	0.812	0.745	0.732	0.75	0.719	0.64	0.631
Modern- BERT- base	149	0.716	0.702	0.667	1.493	0.08367	0.765	0.792	0.715	0.639	0.738	0.669	0.676	0.569
Sbert + SVM	х	0.713	0.699	0.661	0.18236	0.01022	0.788	0.792	0.701	0.629	0.662	0.65	0.748	0.523
Sbert + MLP	0.065	0.705	0.689	0.655	0.01268	0.00071	0.72	0.818	0.686	0.649	0.656	0.706	0.698	0.615

Frugal AI Challenge

Proposed approaches:

- Fine-tuned BERT
- Embeddings + Classifier
- Fine-tuned LLMs

Our winning submission:

• Sentence-BERT + MLP



Code shared on github: <u>https://github.com/D2KLab/frugal-ai-challenge-eurecom</u>

TASK COMPARISON



*Using average kWh per task for text generation and image captioning from "Power Hungry Processing: Watts Driving the Cost of AI Deployment?" Luccioni & al

What's the problem?

- "The internet currently produces approximately 3.8% of global carbon emissions, which are rising in line with our hunger to consume more data." (<u>https://sustainablewebdesign.org/</u>)
- What can we do?
 - > Explore technologies that will reduce energy usage
 - Explore methodologies that reliably measure energy usage, and tools that communicate it



- 19

s12y

W3C Sustainability Community / Working Group: <u>https://www.w3.org/community/sustainability/</u>

- Establish and help evaluate environmental sustainability of new web technologies (e.g. energy use, device obsolescence impact)
- Introduction: <u>https://w3c.github.io/sustyweb/intro.html</u>
- Guidelines at a glance: <u>https://w3c.github.io/sustyweb/glance.html</u>
- STAR (Sustainability Tooling And Reporting): <u>https://w3c.github.io/sustyweb/drafts/star.html</u>
- W3C Sustainability Web Design Community Group: <u>https://www.w3.org/community/sustyweb/</u>

Focus on creating sustainable web sites

References: <u>https://www.w3.org/community/sustyweb/wiki/References</u>



Framework

Sustainable principles for the web, game and interactive world

Nathan Shredoff. Design is the Problem: The Future of Design Must Be Sustainable

General Sustainability Principle	Sustainable Web Design Goals
Make meaningful products	Make websites that are have real value, not fashion or tech-tricks
Easy design rollback	Iterative or Agile design workflow
Source Renewable Materials	Switch to a "Green" webhost
Design products to work in the future	Implement classic design strategies
Design with the user in mind	Create effective User Experience (UX)
Ensure democratic access	Build accessible, responsive websites
Interchangable Parts	Apply standards-based design
Minimize energy and resource consumption	Web Performance Optimization (WPO)
Don't corrupt the virtual system	Search Engine Optimization (SEO)

Permaculture Sustainability Principle	Sustainable Web Design Goals
Observe and Interact	Build sites as part of an interdependent community
Catch and store energy	Cache information, update sites rather the build completely new ones
Obtain a yield	The site should provide positive value to the client, and larger web community, not a time or money sink
Apply self-regulation and accept feedback	The site should have "reporter" technology for use, efficiency and ultimately carbon footprint can be tracked and used for revisions
Use and value renewable resources and services	Use efficient virtual services (e.g. green webhosts)
Produce no waste	Sites should be steady-state, not causing an accumulation of e-junk (files, stored data) in themselves or on the Internet
Design from patterns to details	Start with group-design techniques like Progressive Enhancement
Integrate rather than segregate	Connect your site to others, and create value by interconnection of websites instead of portal-style content provisions
Use small & slow solutions	Design for the low end first, instead of starting with the bleeding edge
Use & value diversity	Use local designers, developers, webhosts of big "cloud" services with low green scores
Use edges & value the marginal	Support communication at the edges – old browsers, platforms, slow networks
Creatively use & respond to change	Use, don't avoid new technologies that promise sustainability, e.g. imageless design with CSS

Inspiration from the permaculture movement



Calculating digital emissions

Websites have a carbon footprint



https://www.websitecarbon.com/

- Estimate: Internet consumes 416.2 TWh / year
 - Is kWh/GB a suitable metric?
 - Calculate the carbon intensity for software applications:

https://github.com/Green-Software-Foundation/sci

 PageSpeed Insights <u>https://pagespeed.web.dev/report?</u> <u>url=http%3A%2F%2Flemonde.fr</u>

 Mobile Efficiency Index <u>https://mobile-efficiency-</u> <u>index.com/en/results/df82124d-</u> 06db-422d-afe8-8c4ce15f3175

CO2.js:

https://github.com/thegreenwebfo undation/co2.js/



Methodology and Challenge

- Data transfer over the wire
 - Energy used roughly proportional to the amount of data transferred
- Energy intensity of web data
 - Energy used at data center, telecom networks and by the end user's computer or mobile device
- Energy source used by the data center
 - Green Web Foundation (TGWF), <u>https://www.thegreenwebfoundation.org/</u> database to see if a data center uses green energy
- Carbon intensity of (grid) electricity, <u>https://ember-climate.org/data/data-explorer/</u>
- Website traffic: carbon per page view x number of annual page views



- 23

Methodology and Challenge

- Why do estimates for internet energy consumption vary so drastically? (<u>source</u>)
- Orders of magnitude: 0.004 kWh/GB vs 136 kWh/GB?
- Meta analysis (2015): 0.06 kWh/GB

24/04/2025 -





System Segments (SWD model)

- Consumer device use: end users interacting with a product or service. This accounts for an estimated 52% of the system. Returning visitors are assumed to be 25%, loading 2% of data.
- Network use: data transferred across the network. This accounts for an estimated 14% of the system.
- Data center use: energy required to house and serve data. This accounts for an estimated 15% of the system.
- Hardware production: embodied energy used in the creation of embedded chips, use of data centers, use of networks, and the use of consumer communication devices. This accounts for an estimated 19% of the system.



Improving the accuracy of website carbon emissions

- The SWD model uses the following percentages for each:
 - >75% first time visitors
 - ≥ 25% return visitors
 - > 2% of data downloaded by return visitors

Test: <u>https://www.webpagetest.org/</u>

- How much data is downloaded for the first view?
- > How much data is re-downloaded for a repeat view?

Reference: <u>https://fershad.com/writing/improving-the-accuracy-of-website-carbon-emissions-estimates/</u>



Is this effective?

- Team UX (French):
 - Approach: <u>https://www.team-ux.com/demarche-eco-conception/</u>
 - Study: <u>https://lowwwcarbon.com/case-study/team-ux/</u>
- Craft Code:

https://www.cantankerouscoder.com/p/craft-code-is-up



- 27



Thank you!



<u>/D2KLab/frugal-ai-challenge-eurecom</u>



Bert correctly identifies persuasion technique for his friend ernie