H2O Open Ecosystem for State-of-the-art Large Language Models

Arno Candel, Jon McKinney, Philipp Singer, Pascal Pfeiffer,
Maximilian Jeblick, Chun Ming Lee, Marcos V. Conde

H2O.ai, Inc.
Mountain View, CA
{firstname.surname}@h2o.ai
https://gpt.h2o.ai/

Abstract

Large Language Models (LLMs) represent a revolution in AI. However, they also pose many significant risks, such as the presence of biased, private, copyrighted or harmful text. For this reason we need open, transparent and safe solutions. We introduce a complete open-source ecosystem for developing and testing LLMs. The goal of this project is to boost open alternatives to closed-source approaches. We release h2oGPT, a family of fine-tuned LLMs from 7 to 70 Billion parameters. We also introduce H2O LLM Studio, a framework and no-code GUI designed for efficient fine-tuning, evaluation, and deployment of LLMs using the most recent state-of-the-art techniques. Our code and models are licensed under fully permissive Apache 2.0 licenses. We believe open-source language models help to boost AI development and make it more accessible and trustworthy.

1 Introduction

Since the Transformer (Vaswani et al., 2017) was introduced in the Natural Language Processing (NLP) community, the advances in this field have increased exponentially (Wolf et al., 2020).

Starting from popular models such as BERT (Devlin et al., 2018a) or Generative Pre-trained Transformers (GPT) (Radford et al., 2018) -both introduced in 2018-, researchers have been pushing the limits of scaling and learned representations in language models (Liu et al., 2019; Radford et al., 2019; Brown et al., 2020; Chowdhery et al., 2022).

Recent advances in Large Language Models (LLMs) are all over the news; these models represent a revolution in Artificial Intelligence (AI) due to their real-world applications through natural language processing (NLP), from internet chatbots to virtual assistants and programmers. However, these also pose significant risks and challenges. The most popular models (*e.g.,* chatGPT (OpenAI, 2023)) are proprietary and not truly open-source, either transparent regarding their training data.

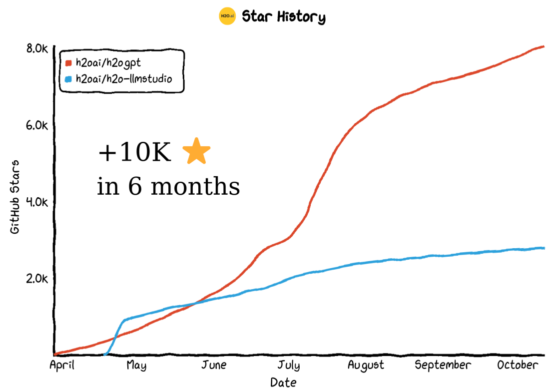


Figure 1: Evolution of our project in GitHub. Our tools have been widely adopted by the NLP community. See https://github.com/h2oai/h2ogpt.

This fast advance leads to a wide range of practical challenges that must be addressed in order for these models to be widely utilized and explored. The popularity and demand of LLMs call for systems to train, fine-tune, evaluate, scale, and deploy the models on a variety of platforms. Given the training costs (millions of dollars), practitioners increasingly rely on pre-trained general-purpose LLMs and fine-tune them for specific downstream tasks and datasets. This requires a wide catalogue of open-source pre-trained LLMs, and sophisticated procedures and tools for efficient fine-tuning. Moreover, considering the massive size of these models (usually from 7 to 100 Billion parameters), we also need compression techniques to deploy them successfully on different platforms.

We believe open-source language models help to boost AI development and make it more accessible and trustworthy. They lower entry hurdles, allowing people to tailor these models to their needs. This openness increases innovation, transparency, and fairness. As part of this effort, we **introduce two open-source libraries**: *h2oGPT* and *H2O LLM Studio*, for LLMs development, including Multi LLM deployment and evaluation — widely adopted in the NLP community (see Fig. 1).

Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, pages 82–89
December 6-10, 2023 ©2023 Association for Computational Linguistics

h2oGPT (`https://github.com/h2oai/h2ogpt`) is a library dedicated to supporting open-source LLMs research, and facilitating their integration while ensuring privacy and transparency. Most integrated models are designed for both research and production. The main use-case of this library is to deploy and test efficiently a wide variety of LLMs on private databases and documents. This tool allows users to compare different models on several tasks and datasets concurrently. An example of this application is `https://gpt.h2o.ai/`.

H2O LLM Studio (`https://github.com/h2oai/h2o-llmstudio`) complements the previous library, and allows users to efficiently fine-tune any LLM using the most recent *state-of-the-art* techniques such as LoRA adapters (Hu et al., 2021), reinforcement learning (RLHF), and 4-bit training. After fine-tuning (or training), the models can be easily exported and deployed at the Hugging Face Hub [1]. Moreover, the library includes a graphic user interface (GUI) specially designed for large language models.

h2oGPT and *H2O LLM Studio* are an ongoing effort maintained frequently by the team of engineers and researchers at H2O.ai with exciting support from the open-source NLP community and external contributors. Both are released under the Apache 2.0 license [2]. Tutorials and detailed documentation are available at the corresponding websites and the technical report (Candel et al., 2023).

2 Related Work

Large language models (LLMs) are designed to process and understand vast amounts of natural language data *e.g.,* internet questions, text in documents, financial data, textbook material, etc. As foundation models (Bommasani et al., 2021), these are trained from broad data at scale (Howard and Ruder, 2018), and can be adapted (*ie.* fine-tuned) to a wide range of down-stream tasks (Wang et al., 2018; Lewis et al., 2019).

They are built on the *Transformer* neural network architecture (Vaswani et al., 2017), which allows them to capture complex language patterns and relationships. Derived from the Transformer, we find BERT-like models (Devlin et al., 2018b; Le et al., 2020; Liu et al., 2019) focused on pre-training with bidirectional encoders. We also find the popular Generative Pre-trained Transformers (GPTs) (Radford et al., 2018, 2019; Brown et al., 2020; OpenAI, 2023), focused on generative pre-training. These serve as the engine of chatGPT.

Since 2022, we experience a new revolution in NLP with the rise of LLMs (over billion parameters models). These models usually follow a multi-stage training strategy, starting with a task-agnostic pre-training on large and diverse datasets. Some related LLMs are LLaMA (Touvron et al., 2023a), GPT-NeoX (Black et al., 2022), BLOOM (Scao et al., 2022), Palm (Chowdhery et al., 2022), OPT (Zhang et al., 2022), and GPT-4 (OpenAI, 2023). We also explore community models such as Falcon (Penedo et al.), Alpaca (Taori et al., 2023), and OpenAssistant (Köpf et al., 2023).

2.1 Why Open-Source LLMs?

While commercially hosted and centralized LLMs like ChatGPT -based on GPT-4 (OpenAI, 2023)-, Microsoft's Bing AI Chat, and Google's Bard are powerful and effective, they have certain risks and limitations compared to open-source LLMs:

- **Data Privacy and Security**: Many require sending data to external servers. This can raise concerns about data privacy, security, and compliance, especially for sensitive information or industries with strict regulations.
- **Dependency and Customization**: We want to allow users to train LLMs on private data safely, and customize the models to their specific needs and applications. Moreover the users can deploy them on their own infrastructure, and even modify the underlying code.
- **Traceability and Transparency**: To understand the risky behaviours of LLMs (*e.g.,* hallucinations, biases, private information etc.), and ensure their safe and trustworthy use, it is fundamental to analyze the dataset and training strategies used to produce such model.
- **Carbon footprint**: Users tend to adopt our open *state-of-the-art* models, instead of running expensive and complicated experiments (in most cases to replicate results). Therefore, we aim to reduce the overall carbon footprint (*ie.* GPU hours consumption) by providing high-quality models and tools.

Overall, open-source LLMs offer greater flexibility, control, and cost-effectiveness, while addressing data privacy and security concerns.

[1] `https://huggingface.co/models`

[2] `https://www.apache.org/licenses/LICENSE-2.0`

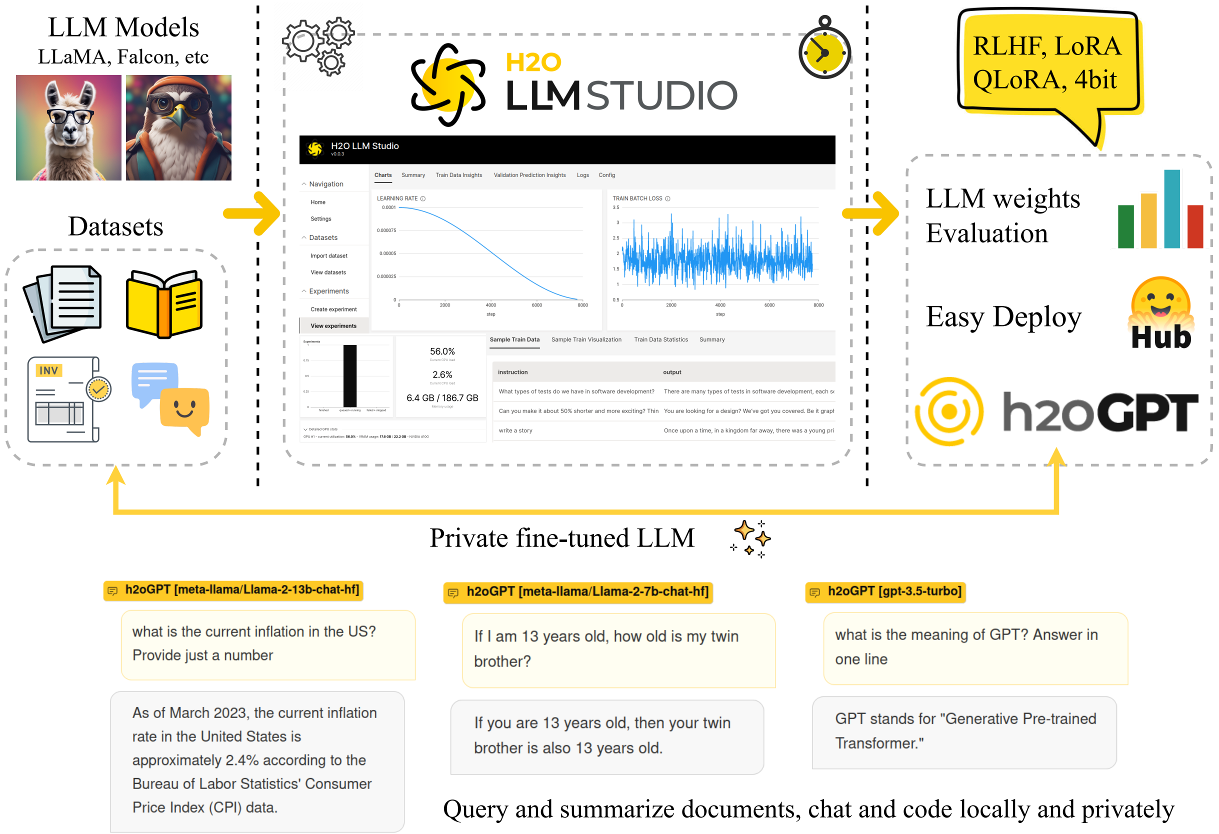


Figure 2: **Open LLM Ecosystem**. (left) The user does not need to transfer private data to 3rd parties, and can select any popular LLM *e.g.,* LLaMA, Falcon. (mid) H2O LLM Studio allows to train and fine-tune any language model using *state-of-the-art* techniques and a GUI without coding. (right) The models can be easily evaluated, exported and deployed. More information at `https://github.com/h2oai/h2o-llmstudio`. Apache 2 License.

3 H2O LLM Studio

An open-source framework for efficient fine-tuning LLMs without coding, using a graphic user interface (GUI) specially designed for large language models [3]. This is illustrated in Figures 2 and 4.

We use the most popular **adapters** for fast fine-tuning such as Low-Rank Adaptation (LoRA) (Hu et al., 2021) and QLoRA (Dettmers et al., 2023), as well as 8-bit (up to 4-bit) model training with a low memory footprint, and the corresponding **quantization**. This allows to fine-tune small LLMs in regular GPUs, even using Google Colab or Kaggle. For example $<$ 10B models (*e.g.,* LlaMa-2 7B) can be fine-tuned in a single NVIDIA-T4 (16GB).

We also integrate *Reinforcement Learning from Human Feedback (**RLHF**)* (Ouyang et al., 2022; Stiennon et al., 2020). This feature is inspired in TRL [4] (von Werra et al., 2020), with the Proximal Policy Optimisation (PPO) by (Ziegler et al., 2019).

[3] `https://github.com/h2oai/h2o-llmstudio`
[4] `https://github.com/lvwerra/trl`

LLM Studio allows complete **customization** of the experimental setup: dataset, *state-of-the-art* model selection, optimizer, learning rate schedule, tokenizer, sequence length (number of tokens), low-rank adapter, validation set and metrics, etc.

The users can **track** several simultaneous experiments, and easily **export** the logs and results. Moreover, the models can be easily exported to the Hugging Face Hub, to be shared with the community or deploy locally and privately.

The framework supports **any open-source language model**, we here highlight the most popular *state-of-the-art* large models: GPT-NeoX (Black et al., 2022), Falcon (Penedo et al.), LLaMa and Llama 2 (Touvron et al., 2023b), Vicuna (Chiang et al., 2023), WizardLM (Xu et al., 2023; Luo et al., 2023), h2oGPT (Candel et al., 2023), and MPT (MosaicML, 2023). We summarize these models in Table 1. Most models are trained on a large amount of data (over 1T tokens), they can handle extremely long inputs (large context length), and are licensed for commercial use.

Model	Size (B)
Llama 2 (Touvron et al., 2023b)	7 / 13 / 70
CodeLlama (Touvron et al., 2023b)	34
Falcon (Penedo et al.)	7 / 40 / 180
Mistral AI (Mistral AI, 2023)	7
GPT-NeoX (Black et al., 2022)	20
WizardLM (Xu et al., 2023)	7 / 13 / 70
Vicuna (Chiang et al., 2023)	13
MPT (MosaicML, 2023)	7 / 30
h2oGPT (Candel et al., 2023)	7 to 70
GPT-3.5 (*by OpenAI*)	?

Table 1: Most popular pre-trained LLMs for fine-tuning. We report the size in Billions (B) of parameters.

We acknowledge **other existing tools** such as LLMTune (Kuleshov, 2023) and EasyLM (Geng, 2023). However, these do not include as many features as LLM Studio (*e.g.,* GUI, supported models and techniques, etc), their licenses can be less permissive. Our tools are amongst the most adopted LLM-related software in GitHub (considering stars and forks by July 2023) — see Fig. 1.

4 Multi LLM Deployment and Evaluation

Any model produced from LLM Studio can be easily integrated into HuggingFace's space & models. We refer to our own space for more information and access to our models [5].

In Fig. 3 (top) we show a snapshot of our demo h2oGPT `https://gpt.h2o.ai/`. We deploy multiple *state-of-the-art* LLM models including Falcon (7/40B), Llama 2 (7/13/70B), and GPT-3.5. This allows us to compare different models and setups.

The user's prompt is evaluated by the different LLMs **concurrently**. We can see the answer generation progress for each model, at the same time. Using this software we can identify clear differences between LLMs easily, for example fast/low inference, hallucinations, common response patterns, bias, memorized data etc. Also, we can analyze the effect of **prompt engineering** on the different models and expose vulnerabilities. The users can deploy the models on a wide variety of inference servers (HF TGI server, vLLM, Gradio, OpenAI), and evaluate performance using reward models.

Document Analysis *h2oGPT* also allows to query and summarize documents in many formats (*e.g.,* PDFs, Word, Code, Text, MarkDown, etc). We implement an efficient use of context using instruct-tuned LLMs (no need for LangChain).

Note that this ecosystem can be reproduced locally, to analyze the models in a private and safe manner.We also provide a OpenAI-compliant Python client API for client-server control.

Guides & Material We provide a short **Video tutorial (2 mins)**, and a complete video overview of the ecosystem (16 min, 340K views) on YouTube.

Also a step-by-step tutorial **Make Your Own GPT With h2oGPT & H2O LLM Studio** (1hr).

We also host all of our models in HF: `https://huggingface.co/h2oai`. We refer the reader to our GitHubs for more demos, and documentation.

5 Future Work

Our open-source LLM Ecosystem is in constant development, *h2oGPT* and *LLM Studio* are updated based on the most recent research advances and demands. We plan to integrate new model quantization techniques, distillation and long-context training (context length over 100K tokens).

We also plan to support more multi-lingual models, and multi-modal models.

6 Limitations

Datasets Fine-tuning requires data text pairs of `instruction` and expected `result/answer`.

Biases and Offensiveness LLMs are trained on a diverse range of unfiltered internet text data, which may contain biased, racist, offensive, or otherwise inappropriate content. Therefore, the generated content by these models may sometimes exhibit biases or produce content that is offensive or inappropriate. We do not endorse, support, or promote any such content or viewpoints.

Usage The large language model is an AI-based tool and not a human. It may produce incorrect, offensive, nonsensical, or irrelevant responses. It is the user's responsibility to critically evaluate the generated content and use it at their discretion.

Carbon footprint Training LLMs is expensive and their use is associated to tons of CO_2 emissions (Touvron et al., 2023a).

Hallucinations LLMs are probabilistic, therefore, certain "random" behaviour is natural and expected, especially on complex prompts (*e.g.,* logical paradoxes, reasoning problems, etc) and "unknown content" not present in the training corpus.

[5] `https://huggingface.co/h2oai`

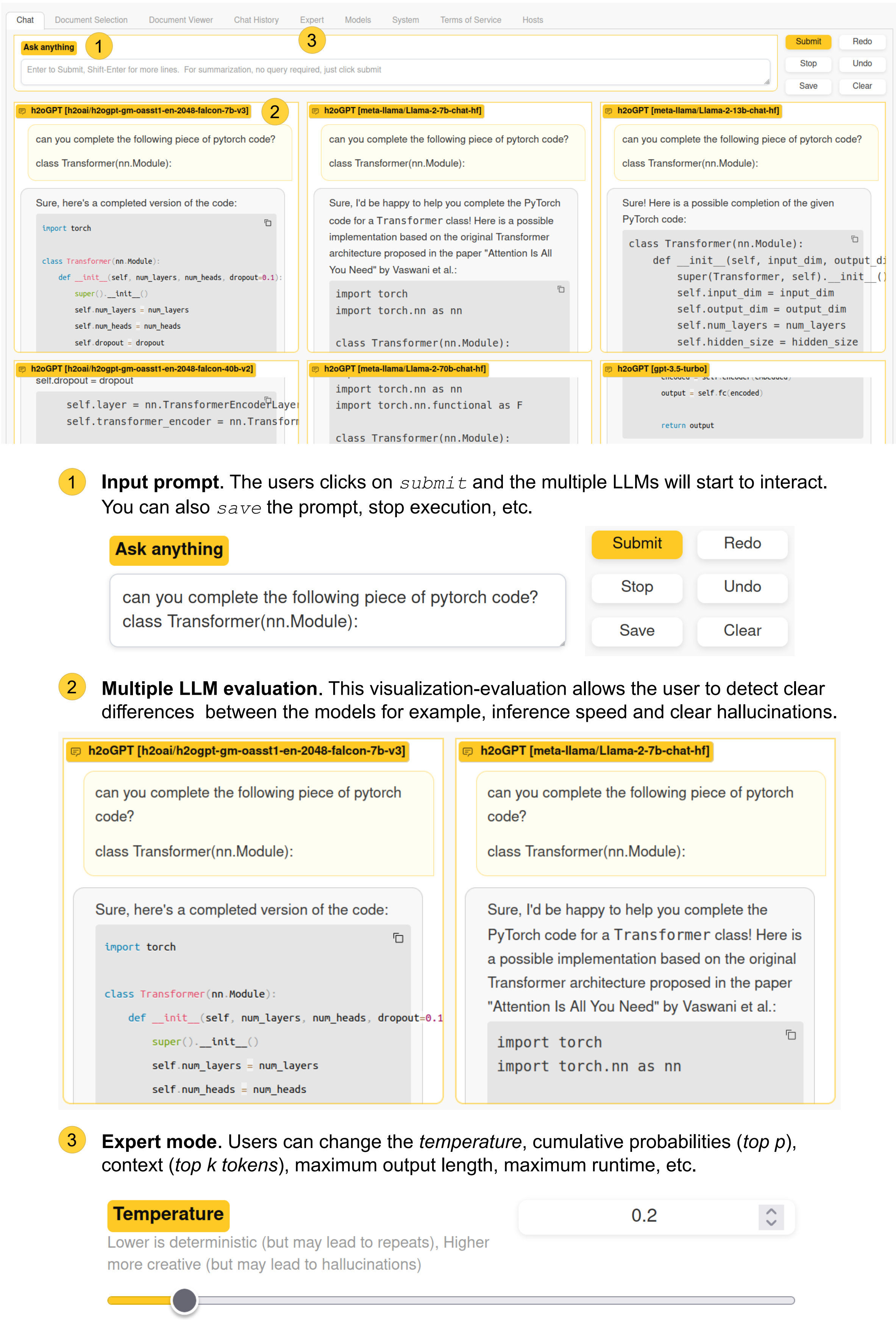
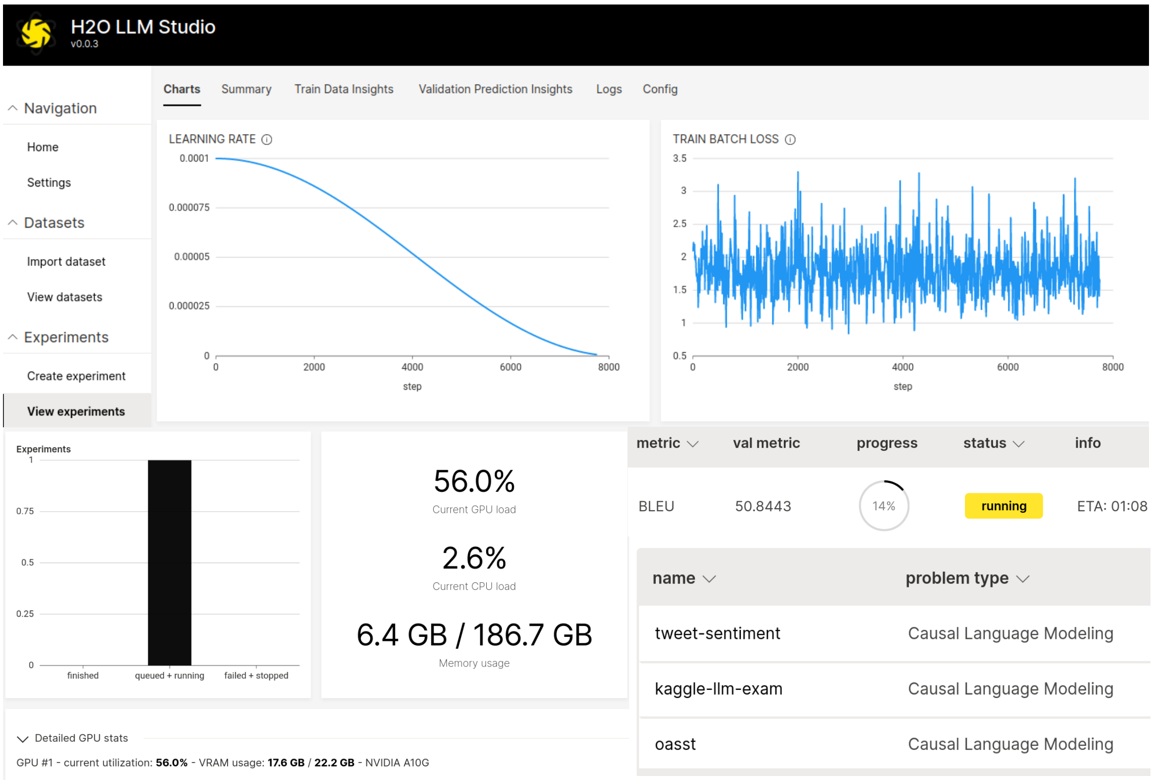
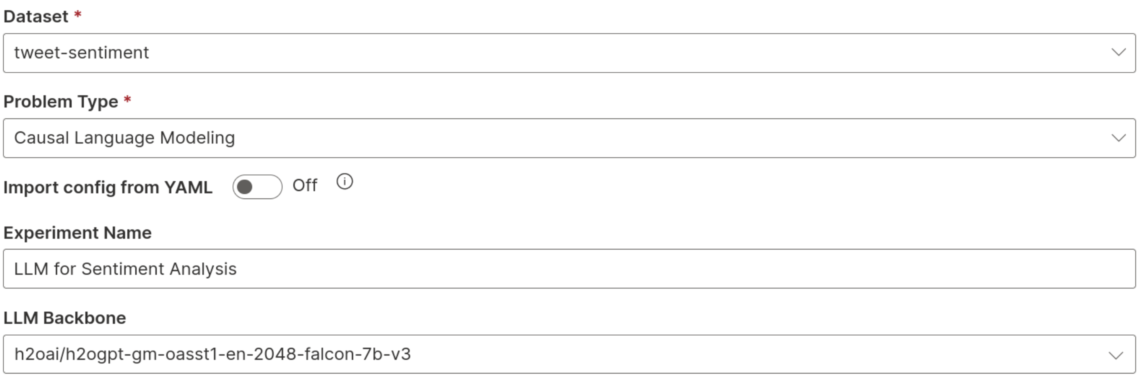


Figure 3: **h2oGPT**. Evaluation of multiple *state-of-the-art* LLM models using the same prompt. This visualization and evaluation allows the user to detect clear differences between the models *e.g.* faster or slower inference, clear hallucinations, common memorized patterns. Demo available at `https://gpt.h2o.ai/` completely free.



Complete LLM Framework. Users can track all the experiments and the system's status. The software allows complete *customization* of the experimental setup: dataset and model selection, validation and metrics, optimizer, adapters, RLHF, bit precision, etc.



Advanced Settings. Users can use *state-of-the-art* techniques to speed up training and obtain real-time performance metrics. Also we allow Tokenizer and context customization.

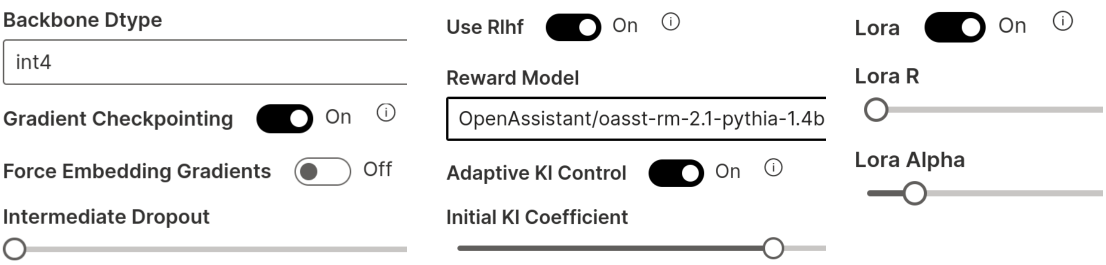


Figure 4: **LLM Studio** allows efficient training and fine-tuning of LLMs using *state-of-the-art* techniques (*e.g.*, advanced models, LoRA, int4, RLHF), and an intuitive GUI with complete experiment's customization. More information in `https://github.com/h2oai/h2o-llmstudio`. Apache 2 License.

Broad Impact

We advocate for the use of open-source LLMs to accelerate AI development and enhance its transparency, accessibility, security, and reliability. Our open framework for training, fine-tuning, deployment and analysis of LLMs enables this to any user, in a private and safe manner. We provide a detailed **Disclaimer** for users of our software.

References

Sid Black, Stella Biderman, Eric Hallahan, Quentin Anthony, Leo Gao, Laurence Golding, Horace He, Connor Leahy, Kyle McDonell, Jason Phang, et al. 2022. Gpt-neox-20b: An open-source autoregressive language model. *arXiv preprint arXiv:2204.06745*.

Rishi Bommasani, Drew A Hudson, Ehsan Adeli, Russ Altman, Simran Arora, Sydney von Arx, Michael S Bernstein, Jeannette Bohg, Antoine Bosselut, Emma Brunskill, et al. 2021. On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*.

Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language models are few-shot learners.

Arno Candel, Jon McKinney, Philipp Singer, Pascal Pfeiffer, Maximilian Jeblick, Prithvi Prabhu, Jeff Gambera, Mark Landry, Shivam Bansal, Ryan Chesler, et al. 2023. h2ogpt: Democratizing large language models. *arXiv preprint arXiv:2306.08161*.

Wei-Lin Chiang, Zhuohan Li, Zi Lin, Ying Sheng, Zhanghao Wu, Hao Zhang, Lianmin Zheng, Siyuan Zhuang, Yonghao Zhuang, Joseph E. Gonzalez, Ion Stoica, and Eric P. Xing. 2023. Vicuna: An open-source chatbot impressing gpt-4 with 90%* chatgpt quality.

Aakanksha Chowdhery, Sharan Narang, Jacob Devlin, Maarten Bosma, Gaurav Mishra, Adam Roberts, Paul Barham, Hyung Won Chung, Charles Sutton, Sebastian Gehrmann, Parker Schuh, Kensen Shi, Sasha Tsvyashchenko, Joshua Maynez, Abhishek Rao, Parker Barnes, Yi Tay, Noam Shazeer, Vinodkumar Prabhakaran, Emily Reif, Nan Du, Ben Hutchinson, Reiner Pope, James Bradbury, Jacob Austin, Michael Isard, Guy Gur-Ari, Pengcheng Yin, Toju Duke, Anselm Levskaya, Sanjay Ghemawat, Sunipa Dev, Henryk Michalewski, Xavier Garcia, Vedant Misra, Kevin Robinson, Liam Fedus, Denny Zhou, Daphne Ippolito, David Luan, Hyeontaek Lim, Barret Zoph, Alexander Spiridonov, Ryan Sepassi, David Dohan, Shivani Agrawal, Mark Omernick, Andrew M. Dai, Thanumalayan Sankaranarayana Pillai, Marie Pellat, Aitor Lewkowycz, Erica Moreira, Rewon Child, Oleksandr Polozov, Katherine Lee, Zongwei Zhou, Xuezhi Wang, Brennan Saeta, Mark Diaz, Orhan Firat, Michele Catasta, Jason Wei, Kathy Meier-Hellstern, Douglas Eck, Jeff Dean, Slav Petrov, and Noah Fiedel. 2022. Palm: Scaling language modeling with pathways.

Tim Dettmers, Artidoro Pagnoni, Ari Holtzman, and Luke Zettlemoyer. 2023. Qlora: Efficient finetuning of quantized llms. *arXiv preprint arXiv:2305.14314*.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018a. Bert: Pre-training of deep bidirectional transformers for language understanding. In *NAACL-HLT*.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018b. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Xinyang Geng. 2023. Easylm: A simple and scalable training framework for large language models.

Jeremy Howard and Sebastian Ruder. 2018. Universal language model fine-tuning for text classification. In *ACL*.

Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2021. Lora: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685*.

Andreas Köpf, Yannic Kilcher, Dimitri von Rütte, Sotiris Anagnostidis, Zhi-Rui Tam, Keith Stevens, Abdullah Barhoum, Nguyen Minh Duc, Oliver Stanley, Richárd Nagyfi, et al. 2023. Openassistant conversations–democratizing large language model alignment. *arXiv preprint arXiv:2304.07327*.

Volodymyr Kuleshov. 2023. Llmtune: Fine-tuning large language models on one consumer gpu. `https://github.com/kuleshov-group/llmtune`.

Hang Le, Loïc Vial, Jibril Frej, Vincent Segonne, Maximin Coavoux, Benjamin Lecouteux, Alexandre Allauzen, Benoît Crabbé, Laurent Besacier, and Didier Schwab. 2020. Flaubert: Unsupervised language model pre-training for french. In *Proceedings of The 12th Language Resources and Evaluation Conference*, pages 2479–2490, Marseille, France. European Language Resources Association.

Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Ves Stoyanov, and Luke Zettlemoyer. 2019. BART: Denoising Sequence-to-Sequence pre-training for natural language generation, translation, and comprehension.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar S. Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke S. Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *ArXiv*, abs/1907.11692.

Ziyang Luo, Can Xu, Pu Zhao, Qingfeng Sun, Xiubo Geng, Wenxiang Hu, Chongyang Tao, Jing Ma, Qingwei Lin, and Daxin Jiang. 2023. Wizardcoder: Empowering code large language models with evol-instruct.

Mistral AI. 2023. Mistral 7b introduction. https://mistral.ai/news/announcing-mistral-7b/.

MosaicML. 2023. Mpt-30b: Raising the bar for open-source foundation models.

OpenAI. 2023. Gpt-4 technical report.

Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. 2022. Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 35:27730–27744.

Guilherme Penedo, Quentin Malartic, Daniel Hesslow, Ruxandra Cojocaru, Alessandro Cappelli, Baptiste Pannier, Ebtesam Almazrouei, and Julien Launay. The refinedweb dataset for falcon llm: Outperforming curated corpora with web data, and web data only.

Alec Radford, Karthik Narasimhan, Tim Salimans, Ilya Sutskever, et al. 2018. Improving language understanding by generative pre-training.

Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, et al. 2019. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9.

Teven Le Scao, Angela Fan, Christopher Akiki, Ellie Pavlick, Suzana Ilić, Daniel Hesslow, Roman Castagné, Alexandra Sasha Luccioni, François Yvon, Matthias Gallé, et al. 2022. Bloom: A 176b-parameter open-access multilingual language model. *arXiv preprint arXiv:2211.05100*.

Nisan Stiennon, Long Ouyang, Jeffrey Wu, Daniel Ziegler, Ryan Lowe, Chelsea Voss, Alec Radford, Dario Amodei, and Paul F Christiano. 2020. Learning to summarize with human feedback. *Advances in Neural Information Processing Systems*, 33:3008–3021.

Rohan Taori, Ishaan Gulrajani, Tianyi Zhang, Yann Dubois, Xuechen Li, Carlos Guestrin, Percy Liang, and Tatsunori B. Hashimoto. 2023. Stanford alpaca: An instruction-following llama model. https://github.com/tatsu-lab/stanford_alpaca.

Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. 2023a. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*.

Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, et al. 2023b. Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288*.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Ł Ukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In I Guyon, U V Luxburg, S Bengio, H Wallach, R Fergus, S Vishwanathan, and R Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 5998–6008. Curran Associates, Inc.

Leandro von Werra, Younes Belkada, Lewis Tunstall, Edward Beeching, Tristan Thrush, and Nathan Lambert. 2020. Trl: Transformer reinforcement learning. https://github.com/lvwerra/trl.

Alex Wang, Amanpreet Singh, Julian Michael, Felix Hill, Omer Levy, and Samuel R. Bowman. 2018. Glue: A multi-task benchmark and analysis platform for natural language understanding. In *ICLR*.

Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, et al. 2020. Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations*, pages 38–45.

Can Xu, Qingfeng Sun, Kai Zheng, Xiubo Geng, Pu Zhao, Jiazhan Feng, Chongyang Tao, and Daxin Jiang. 2023. Wizardlm: Empowering large language models to follow complex instructions. *arXiv preprint arXiv:2304.12244*.

Susan Zhang, Stephen Roller, Naman Goyal, Mikel Artetxe, Moya Chen, Shuohui Chen, Christopher Dewan, Mona Diab, Xian Li, Xi Victoria Lin, et al. 2022. Opt: Open pre-trained transformer language models. *arXiv preprint arXiv:2205.01068*.

Daniel M Ziegler, Nisan Stiennon, Jeffrey Wu, Tom B Brown, Alec Radford, Dario Amodei, Paul Christiano, and Geoffrey Irving. 2019. Fine-tuning language models from human preferences. *arXiv preprint arXiv:1909.08593*.