

Curt Tigges

ct@curttigges.com | [in/curttigges](https://in.curttigges) | github.com/curt-tigges

PAPERS

First-Name Papers

Language Models Linearly Represent Sentiment (accepted at Blackbox NLP) | 2023 [\[arxiv\]](#)

Investigation of how LLMs build representations and world models of sentiment.

LLM Circuit Analyses Are Consistent Across Training and Scale (accepted at NeurIPS) | 2024 [\[arxiv\]](#)

Investigation of algorithmic, component, and size stability of circuits across training and over scale.

Other Papers

Transformer-Based Models Are Not Yet Perfect At Learning to Emulate Structural Recursion (TMLR) |

[\[arxiv\]](#)

RELEVANT EXPERIENCE

Science Lead | Decode Research | Jul 2024 - Present

- Lead and execute on mechanistic interpretability research projects, including external collaborations
- Collaborate with engineering to design & review features for our mech interp research platform (Neuronpedia)
- Maintain and develop features for open-source SAE libraries like SAE Lens and other parts of our stack

Research Scientist | EleutherAI Institute | Jan 2023 - Jul 2024

- Conduct and assist in LLM mechanistic interpretability & other empirical research projects
- Train LLMs on GPU cluster for various projects as needed, and maintain and improve the GPT-NeoX library

Data Scientist (part time) | NCSU Ops Research & Education Lab | Raleigh, NC | Jun 2022 - Nov 2022

- Built models to predict student population changes and to predict optimal locations for new schools

Senior Data Analyst | Taroko.io | Taipei, TW & Raleigh, NC | Jan 2020 - Dec 2022

- Built & deployed ML & statistical solutions for churn & revenue prediction, key conversion path analysis, etc.
- Planned & built out data warehouse in BigQuery, integrating PostgreSQL database, Heap Analytics & more

Business & Search Engine Marketing Analyst | Taroko.io | Taipei, TW | Jun 2016 - Dec 2019

- Analyzed customer data & produced SQL queries and reports to produce various metrics & ad-hoc analyses
- Developed bidding algorithm that rescued failing products, decreasing CPA by 21% and increasing volume

Software Engineer | KPIT Extended PLM | Raleigh, NC | Mar 2014 - Aug 2015

- Deployed enterprise software (PTC Windchill) customizations to product lifecycle management systems
- Planned & conducted detailed QA tests for software products

PROJECTS

Lingwave Language Learning App

Led project to design an app using innovative pedagogical & algorithmic methods to teach Chinese

EDUCATION

SERI MATS (Neel Nanda Interpretability Stream) | 2023

Alignment Research Engineer Accelerator (ARENA) | 2022

Master of Computer Science (Data Science Track) | University of Illinois Urbana-Champaign | 2021

Bachelor of Science in Science, Technology and Society | NC State University | 2012

SKILLS

Programming Languages - Proficient In: Python | SQL - **Familiar With:** R | C++ | PHP

Packages: PyTorch | PyTorch Lightning | Transformers | Scikit-Learn | Matplotlib | Pandas | Numpy

Higher-Level Skills: Mechanistic Interpretability | Deep Learning | Fine Tuning | Prompt Engineering | Dataset Generation | Software Engineering