Tyler Liddell

Sacramento, CA | tliddell13@gmail.com | GitHub | Website

Evaluating and improving the robustness of large language models through systematic linguistic analysis. Specifically focused on understanding how LLMs process syntactic structure versus surface-level patterns, with implications for model interpretability and benchmark design.

Education

MSci Computer Science | City University of London, UK | 2020 – 2024 *Second Class Honors (69%)*

Master's Dissertation: "Experiments With Word Order in Large Language Models"

- Supervised by Pranava Madhyastha
- Developed novel experimental framework for testing LLM performance across 7 benchmarks (MMLU, TruthfulQA, Hellaswag, Winogrande, ARC, GSM8K)
- Extended EleutherAI's lm-evaluation-harness with 500+ lines of Python code for systematic perturbation analysis
- Demonstrated that models retain comparable performance (within 10%) when evaluated on the shuffled versions and the standard version of these benchmarks
- Currently working on a publishable version of my findings

Relevant Coursework: Deep Learning, Natural Language Processing, Machine Learning, Computer Vision,

Technical Skills

Programming: Python, Java, C++, SQL, Haskell
ML/NLP Frameworks: PyTorch, Hugging Face Transformers, NLTK, spaCy, scikit-learn
Tools: Git, Linux, SLURM, Bash, Visual Studio, MATLAB
Specializations: Large Language Model Evaluation, Transfer Learning, Prompt Engineering, Model Interpretability

Projects

SMS Spam Detection

- Built spam classification models using NLP and CBOW embeddings; achieved 98% accuracy on unseen test data.
- Implemented full ML pipeline: data preprocessing, feature extraction (TF-IDF), model training, and evaluation and presented results in a formal paper.

Emotion Analysis

- Performed multi-label emotion classification using MATLAB, decision trees, and random forests.
- Designed end-to-end pipeline including preprocessing, TF-IDF vectorization, and model evaluation.
- Presented findings in a poster project.

Higgs Boson Detection

- Developed ML models for particle classification from CERN's ATLAS experiment data using scikitlearn.
- Improved model accuracy via hyperparameter tuning and cross-validation