Tim Schott

Charlotte, NC
timschott@berkeley.edu | 571-420-7308
timschott.com | github.com/timschott | linkedin.com/in/timothy-schott

Education

University of California, Berkeley — Berkeley, CA
Master of Information Management and Systems — May 2023

University of California, Berkeley — Charlottesville, VA
Bachelor of Arts, Computer Science | Bachelor of Arts, English — May 2019

Professional Experience

Data Scientist — Charlotte, NC
Red Ventures — Intern Summer 2022 | Full-time 2023-present
- Leveraged word embeddings from Large Language Models to score the compatibility of ad copy on web pages.
- Increased ad clicks by 40% by using compatibility scores as contextual features in Multi-Armed Bandit tests.
- Productionized enhancements with a data pipeline and machine learning workflow using PySpark and Databricks.
- Authored a technical summary of internship experience for company engineering blog that garnered 500+ views.

Software Engineer, Client Services — Reston, VA
Brightspot — Intern Summer 2018 | Full-time 2019 - 2021
- Performed object-oriented data modeling, data engineering and CMS customizations in a cross-functional setting. Collaborated with creative staff to tailor features to business needs. Trained client developers to use Brightspot APIs.
- Produced a data ingestion framework with Java for mattressfirm.com and sleep.com e-commerce products that automated the maintenance of a 5000-product inventory. Eliminated the need for editors to manually curate inventory.
- Delivered a cloud computing ETL Pipeline to Janes UK that transforms creative content into usable web data. Simplified creative workflows and enhanced visibility to editorial management. Increased janes.com publishing activity by 60%.
- Integrated the Apple News API into bcg.com's CMS and wrote documentation explaining how editors can syndicate content to Apple News. Differentiated BCG's presence in the competitive landscape.

Projects

Polyglot or Not? https://arxiv.org/abs/2305.13675 — Spring 2023
- Created a novel factual recall benchmark for large language models. Discovered significant performance lapses when models are tasked with retrieving non-English information.
- Contributed a new, multilingual dataset containing 303k factual associations in 20 different languages.
- Wrote modular, extensible model assessment code using Python and automated evaluations with Google Cloud.
- First-author on technical manuscript currently under review for EMNLP 2023, a leading conference in AI and NLP.

Detail in the Novel — Fall 2022
- Annotated 500 passages from canonical fiction novels and evaluated their usage of details using heuristics developed after extensive research into narrative structure, formalism, and corpus linguistics.
- Concluded that the use of prepositions is highly associated with detailed modes of writing via chi-squared tests.
- Visualized attention weights from large language models using PyTorch to reinforce findings.

Academic Experience

Teaching Assistant, Natural Language Processing — Berkeley, CA
Supervisor: Professor David Bamman — Spring 2022 | Spring 2023
- Supported 250+ students in technically rigorous NLP course. Wrote weekly quizzes and created Python programming assignments which were automatically graded using unit tests and shell scripts.
- Reviewed student code and reinforced class concepts during highly-attended weekly office hours. Fielded a wide range of queries related to class administration, personal projects, internship searches, etc.

Research Assistant, UVA Humanities Institute — Charlottesville, VA
Supervisor: Professor Jennifer Petersen — 2016 – 2019
- Mined and cleaned unstructured text data as part of an interdisciplinary research group studying the history of Free Speech and the U.S. Supreme Court.
- Ran regressions, supervised classifications, and hypothesis tests. Improved statistics skills and learned R programming.
- Gained experience explaining technical findings from data analysis to non-technical stakeholders.