

A versatile data scientist with a comprehensive skill set, combining expertise in the implementation and fine-tuning of state-of-the-art models like GPT with an adept understanding of statistics and probability theory. Experienced in the development of machine learning techniques and proficient in handling vast data structures and optimizing their performance. Skilled in simplifying and conveying technical concepts to varied audiences via straightforward explanations, visual tools, and relatable examples. Poised to thrive in settings seeking a data-centric professional with a deep grounding in statistical theories, machine learning, deep learning methodologies, and a track record of applying them in diverse domains, from academia to AI-driven business applications.

## SKILLS

<b>PROGRAMMING LANGUAGES</b>	PYTHON   R   MATLAB
<b>WEB TECHNOLOGIES &amp; DATABASES</b>	JAVASCRIPT   HTML   CSS   POSTGRESQL   MYSQL   SQLITE   RSQLITE
<b>FRAMEWORKS &amp; LIBRARIES</b>	JUPYTER   BEAUTIFULSOUP   NUMPY   PANDAS   SCIKIT-LEARN   TENSORFLOW   PYTORCH   KERAS   XGBOOST   LIGHTGBM   SQLALCHEMY
<b>TOOLS, SOFTWARE &amp; FORMATTING</b>	ARCGIS   STATA   GIT   LATEX   MARKDOWN

## EXPERIENCE

### DATA SCIENCE EXPERT

APRIL 2023 - PRESENT

SCALE AI

SAINT LOUIS, MO

- Oversaw quality assurance and provided critical evaluation as a review specialist for RLHF projects, ensuring adherence to the highest standards
- Contributed to a reward model of reinforcement learning from human feedback (RLHF) to align with client-specific needs, emphasizing improvements in accuracy, conversational fluidity, and user engagement
- Incorporated sophisticated AI capabilities, including semantic comprehension and topic adherence to optimize model performance
- Formulated detailed training scripts for a range of data science applications, strengthening the chatbot's proficiency in code interpretation and generation
- Effectively synthesized information while maintaining strict writing style and structure protocols across various domains

### PYTHON INSTRUCTOR

JUNE 2020 — JUNE 2022

WASHINGTON UNIVERSITY IN SAINT LOUIS

SAINT LOUIS, MO

- Developed a curriculum teaching Python programming, data analysis techniques, Git, and JupyterLab for public health master's students
- Taught Python fundamentals, data manipulation, and visualization using pandas, numpy, seaborn and matplotlib libraries
- Supervised hands-on projects, highlighting the use of GitHub for version control and JupyterLab for interactive analysis
- Adapted teaching methods for diverse learners, preparing students to address public health problems with data-driven approaches

- Implemented advanced ensemble machine learning techniques, such as Bayesian model averaging and random forests, to develop an R package (EBMAforecast) for predicting heterogeneous treatment effects in political research
- Conducted extensive data pre-processing, exploratory analysis, handling of missing values, outlier detection, and feature engineering to ensure high-quality data input for model development
- Optimized model performance through rigorous hyper-parameter tuning, cross-validation, and appropriate feature set selection, leading to improved predictive accuracy and robustness across contexts
- Collaborated with interdisciplinary teams of political scientists, statisticians, and computer scientists to apply methodologies and communicate complex findings effectively

## PROJECTS

---

### OPENAI CHATBOT INTEGRATION

**APRIL 2023-AUGUST 2023**

WEB DEVELOPMENT, NLP, TEXT PARSING, OPENAI GPT-4 INTEGRATION

- Developed a web app allowing users to upload their resumes and get suggestions for improvement using OpenAI's GPT-4
- Integrated GPT-4 for a range of NLP tasks, enhancing the content of the submitted resumes
- Crafted various Python utility scripts vital for the app's functions, including resume parsing, summary extraction, and establishing connections to OpenAI
- Demonstrated the application's potential by refining resume summaries and work descriptions, leading to enhanced clarity and attractiveness

### HOSTING LOCAL LARGE LANGUAGE MODEL

**JULY 2023-OCTOBER 2023**

DEEP LEARNING, DATA MANAGEMENT, LOCAL HOSTING, USER INTERFACE

- Utilized PyTorch and NumPy for the implementation and training of the large language model, leveraging pre-trained weights and GPU acceleration for efficient computation
- Became proficient in the theories related to the general structure of LLMs, which includes neural network components like recurrent, feedforward, embedding, and attention layers
- Hosted the project on a public repository and provided basic guidance to allow other users to implement LLM
- Designed a local hosting environment that enables the deployment of the large language model within an on-site infrastructure
- Ensured efficient storage and retrieval processes, handling large-scale data through optimization techniques

### NEWS DISCOURSE ANALYSIS VIA ADVANCED NLP TECHNIQUES

**JUNE 2020-JUNE 2023**

BIG DATA, MACHINE LEARNING, DATA ENGINEERING, POLITICAL COMMUNICATION

- Constructed a large-scale dataset of Twitter data from 30 news outlets to examine communication patterns and styles, involving billions of tweets
- Designed and implemented a scalable ETL pipeline using Python packages such as nltk and numpy to develop efficient pre-processing in the management of big data
- Applied generalized linear mixed models and Bayesian statistics to analyze communication patterns, utilizing advanced NLP techniques such as sentiment analysis, topic modeling, and text classification using packages like BERT, gensim, and scikit-learn

- Demonstrated proficiency in handling API requests, pagination, and error management with Tweepy, ensuring a robust data collection process from Twitter
- Efficiently utilized SQLAlchemy ORM capabilities with SQLite for database storage and retrieval, employing features like declarative mapping, query composition, and connection pooling for optimal performance
- Presented complex quantitative analyses and findings to key stakeholders in a clear, concise, and actionable manner

#### **FORECASTING THE 2020 PRESIDENTIAL ELECTION**

**JULY 2019**

MACHINE LEARNING, ENSEMBLE METHODS, BAYESIAN INFERENCE, TIME SERIES ANALYSIS, ELECTION FORECASTING

- Conducted extensive data pre-processing, feature engineering, and model evaluation using cross-validation techniques to optimize model performance and ensure the validity of the forecasts
- Developed a Bayesian model averaging approach to combine predictions from multiple forecasting models, leveraging expertise in experimental design and multivariate statistics
- Applied the ensemble method to the 2020 presidential election, resulting in a highly accurate and precise forecast with a 95% credible interval using 11 historical out-of-sample forecasts
- Designed and implemented an MCMC sampling scheme in R for rigorous statistical inference and uncertainty quantification, enabling a more comprehensive understanding of model performance and potential biases

#### **POLITICAL ADVERTISEMENT PERCEPTION STUDY WITH ADVANCED ANALYTICS**

**DECEMBER 2021**

DATA ENGINEERING, DATA VISUALIZATION, CONJOINT EXPERIMENT, POLITICAL ADVERTISING

- Conducted a conjoint experiment to understand user preferences and perceptions of political ads, leveraging expertise in non-parametric analysis with tools such as `cjoint`, `coefplot`, and `gmodels` packages for advanced statistical modeling and analysis
- Examined political and social issues survey data with `descr`, `labelled`, and `haven` packages to identify key trends and insights in public opinion
- Ensured accurate interpretation of results using statistical concepts such as weighted means, confidence intervals, and hypothesis testing
- Designed and developed informative visualizations with `ggplot2` and `plotly` packages, effectively communicating complex findings
- Utilized strong communication competencies to present actionable insights to extended teams and small groups of key stakeholders

#### **ANALYZING MISINFORMATION EFFECTS**

**JULY 2022**

RESEARCH DESIGN, DATA ANALYSIS, MACHINE LEARNING, DATA VISUALIZATION, PUBLICATION, MISINFORMATION

- Developed a novel research framework integrating experimental and observational data to investigate misinformation effects on public opinion and behavior
- Employed machine learning techniques such as regression and generalized linear mixed models, classification, and clustering algorithms to evaluate causal factors and predict misinformation susceptibility
- Ensured validity and generalizability by using nationally representative samples and cross-validation techniques in the R
- Fostered transparency and reproducibility by submitting research data, code, and documentation to the Harvard Dataverse repository, enabling further collaboration and research in the field

- Devised a novel survey experiment to assess the efficacy of misinformation corrections using advanced statistical techniques, including non-linear mixed models and Bayesian statistics
- Implemented a pilot survey and deployed the final survey to a diverse, nationally-representative sample, ensuring generalizability of the results
- Analyzed survey data using R programming, employing packages such as `dp1yr`, `tidyverse`, and `ggplot2` for data manipulation, visualization, and application of multivariate statistics
- Developed data-driven insights on the effectiveness of various misinformation correction techniques and presented findings through clear and informative visualizations
- Contributed to grant writing efforts, securing funding for future research on misinformation and its impact on public opinion and behavior
- Effectively communicated complex quantitative analyses to extended teams and key stakeholders, utilizing strong communication competencies to present findings in a clear, concise, and actionable manner

## EDUCATION

---

PH.D. IN POLITICAL SCIENCE	WASHINGTON UNIVERSITY IN SAINT LOUIS	AUGUST 2017 — JULY 2024
M.A. IN POLITICAL SCIENCE	SAINT LOUIS UNIVERSITY	AUGUST 2016 — MAY 2017
B.A. POLITICAL SCIENCE (CUM LAUDE)	SAINT LOUIS UNIVERSITY	AUGUST 2013 — MAY 2016
A.A. COMMUNICATION	IVY TECH COMMUNITY COLLEGE	AUGUST 2009 — MAY 2012

## PUBLICATIONS

---

Guess, Andrew, Dominique Lockett, Benjamin Lyons, Brendan Nyhan, Jacob M. Montgomery, and Jason Reifler. 2020. "‘Fake news’ may have limited effects beyond increasing beliefs in false claims." *The Misinformation Review*.

Edelson, Laura, Dominique Lockett, Jacob Montgomery, Damon Mccoy, Tobias Lauinger, Celia Guillard "US Public Opinion Towards Platform Regulation of Political Advertisements: Discontent and Consensus for Reform" (Forthcoming)

Lockett, Dominique. Using Objectivity to Improve Argument Evaluations. (Forthcoming)

## AWARDS AND CERTIFICATES

---

RESEARCH SEED GRANT	WASHINGTON UNIVERSITY IN SAINT LOUIS	NOVEMBER 2020
DIVERSITY FELLOWSHIP	SAINT LOUIS UNIVERSITY	JUNE 2016
FUNDAMENTALS OF G.I.S.	UNIVERSITY OF CALIFORNIA, DAVIS	SEPTEMBER 2020