Andrew Fogarty

apfogarty@berkeley.edu
For applied ML/stats demos, visit: seekinginference.com
434-806-2891
Active TS//SCI Clearance

Education

University of California, Berkeley
Master of Information and Data Science — *All But Capstone: December 2020*

University of Virginia
Master of Political Science — *2015-2017*

Methods

Statistical Models: Linear, Unordered Categorical, Ordered Categorical, Binary, Count, Survival, Time Series, Time Series Cross-Section, Multiple Imputation, Numerical Optimization

Machine Learning Models: K-Nearest Neighbors, Decision Trees, Naive Bayes, Support Vector Machines, Neural Networks, Dimensionality Reduction, Cluster Analysis

Natural Language Processing: Classification: CNN, Capsules, T5, BERT, Generation: GPT-2, Summarization: T5

Qualitative: Research Design and Causal Inference, Comparative Methods, Essential Empirical Methods

Technologies: Python, R, AWS, Spark, DataBricks, SQL

Key Projects

Research and Analysis: Translated research puzzles into insight and opportunity for U.S. Government and military action through 44 written assessments that include case studies, comparative analysis, executive assessments, visual intelligence products, and statistical graphs.
Taliban Mobility: Conducted causal inference research on the determinants of a military group's mobility across geographic units and time using a time series cross-sectional linear model, treatment and control case studies, and custom maps. This research influenced the President's Daily Brief, was cited in defense of the U.S. National Intelligence Estimate's key judgments, and was selected for presentation at a data science conference.
Cities Under Attack: Proposed causal explanations for why cities in Afghanistan are attacked using a time series logit model, treatment and control case studies, and dashboards to animate time series data. This research was selected for presentation at a data science conference.
Enemy at the Gates: Delivered novel cross-sectional research through interactive graphics and visual executive summaries that estimated the most important factors associated with the deaths of tens of thousands of allied foreign military forces in Afghanistan.
Natural Language Propaganda: Created novel NLP data set from scratch on Taliban propaganda messages from 2014-2020 for use in classification, text generation, and summarization tasks on an array of transformers and baseline models including: BERT, T5, and GPT-2 in PyTorch.

Work Experience

National Ground Intelligence Center — Charlottesville, VA
Data Scientist — *2010 - Present*

Technical: Provided technical consultation to my colleagues on methodological matters such as overcoming non-random assignment to treatment, data summation, data visualization, data gathering strategies, hypothesis testing, and research design.

Customer Service: Tailored customer service support to a wide variety of military and government leaders at tactical, senior, and executive levels for over a decade via in-person training and video conferences.

Innovative: First analyst to publish interactive intelligence assessments via state-of-the-art data science methods at NGIC. To view my programming, replicate my research, review my coding criteria and coding justifications, and view ongoing live analysis, please see my JWICS GitHub: github.devops.cia.ic.gov/frfogap-USA/

Invited Data Science Conference Presentations

5th Annual Data Science Technical Exchange — Dulles, VA
Central Intelligence Agency — *September 2019*

ConMap Meet: Conflict Mapping — Springfield, VA
National Geospatial-Intelligence Agency — *May 2019*

4th Annual Data Science Technical Exchange — Bethesda, MD
Central Intelligence Agency — *September 2018*