DAVID BENJAMIN GOMEZ

dbgomez94@gmail.com (478) 747-6519 Cambridge, MA, USA Website

Data scientist with 4+ years of graduate research experience, across two master's degrees, applying machine learning, deep learning, and natural language processing to solve engineering, public health, and social computing problems

EDUCATION

Georgia Institute of Technology **Atlanta, Georgia**

- Master of Science in Computer Science, Highest Honors, GPA: 3.8 *Aug 2022 – Dec 2024*
- Master of Science in Aerospace Engineering, Highest Honoors, GPA: 3.7 *Aug 2018 – May 2020*
- Bachelor of Science in Aerospace Engineering, Highest Honors, GPA: 3.6 *Aug 2015 – May 2018*

PROJECTS

Disclosure Patterns of Suicidality on Social Media

Independent Research for Social Dynamics and Well-Being Laboratory *May 2023 – Dec 2023*

- Applied natural language processing and multivariate timeseries k-means clustering in Python to suicide related tweets collected from X (formerly Twitter) APIs to identify temporal response patterns to suicide disclosures
- Performed imbalanced classification via logistics regression, support vector machines, and multilayer perceptrons to predict which response pattern a user would exhibit using pre-disclosure data with the best model achieving F1=0.64

Identifying Linguistic Characteristics that Differentiate Mental Health Forums on Reddit

Team Leader for Course Project in Deep Learning *Jan 2023 – May 2023*

- Trained/fined-tuned text classifiers (LSTM, DistilBERT, and RoBERTa) in Python to perform multiclass classification assigning Reddit posts to one of three mental health-related subreddits, with the best model achieving F1=0.83
- Conducted model interpretability analyses for each model in Python using Shapley values to identify the key linguistic features that differentiate the three subreddits: r/suicidewatch, r/stopselfharm, and r/depression

Affective Desensitization to Public Mass Shootings on Social Media

Team Leader for Course Project in Social Computing *Jan 2022 – May 2022*

- Applied multiple linear regression in Python to data collected from X APIs with features derived from the Valence-Dominance-Arousal lexicon to find evidence that Americans have become desensitized to mass shootings
- Controlled for confounding variables using backward feature selection and variable inflation factor analyses

Improving the Gun Ownership Proxy for Firearm Research

Personal Project (turned Publication) *Jan 2020 – Dec 2020*

- Applied multiple nonlinear regression and a deep neural network with cross-validation in Python to develop two new proxies of state-level gun ownership for gun research improving state-of-the-art performance from $R^2 = 0.61$ to 0.94
- Conducted model diagnostics via residual symmetries, prediction intervals, and Cook's distance for outlier influence

Challenging the Inevitability of Suicide

Team Leader for Course Project in Accident Causation and System Safety *Aug 2019 – Dec 2019*

- Applied simple linear regression and hypothesis testing in STATA to suicide data collected from CDC APIs to argue that measures to reduce firearm suicide rates (i.e., gun regulations) do not exacerbate suicide rates by other means

EXPERIENCE

Social Dynamics and Well-Being Laboratory (SocWeB) **Atlanta, Georgia**

Graduate Research Assistant @ Georgia Tech *Aug 2022 – May 2024*

- Applied machine and deep learning to address research questions at the intersection of social media and mental health

Space Systems Design Laboratory (SSDL) **Atlanta, Georgia**

Graduate Research Assistant @ Georgia Tech *Aug 2021 – May 2022*

- Conducted reliability analyses of CubeSats in Python using Kaplan-Meier curves and time-to-failure distributions

Busek Space Propulsion and Systems **Natick, Massachusetts**

Research and Development Engineer *Aug 2020 – Aug 2021*

- Supported laboratory testing and data acquisition of electric propulsion (EP) devices for spaceflight qualification

NASA Jet Propulsion Laboratory (JPL) **Pasadena, California**

Electric Propulsion Intern (x3) *May – Aug 2018, 2019, 2020*

- Developed an uncertainty analysis procedure for EP thrust stands using linear regression and prediction intervals
- Built a custom LabVIEW-based Hall-effect thruster control interface for JPL's primary testing facility

SKILLS / FRAMEWORKS

Programming: Python, SQL, R, STATA, LabVIEW, MATLAB, LaTeX, Git/GitHub
Frameworks: NumPy, Pandas, Seaborn, Matplotlib, Plotly, Scikit-Learn, Keras, TensorFlow, PyTorch, Transformers