JIAMAN BETTY **WU**

 [/JiamanBettyWu](#) | [/jiamanbettywu](#) | jiaman.wu@duke.edu | (530) 220-7908 | [jiamanbettywu.github.io](#)
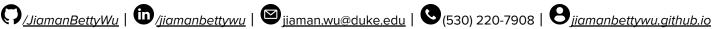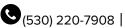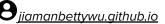
EDUCATION

Duke University | Durham, NC   Aug 2020 - May 2022
- M.S. in Data Science
- Coursework: Bayesian Statistics · Causal Inference · Cloud Computing · Machine Learning · Deep Learning · NLP · Data Visualization

University of California, Davis | Davis, CA   Sep 2015 - Jun 2019
- B.A. in Statistics and Economics — *highest honors*
- Coursework: Multivariate Data Analysis · Time Series · Non-parametric Statistics · Econometrics

SKILLS

Frameworks: Scikit-Learn · Pandas · PyTorch · Tableau · Git · AWS [*[ML Specialist Certified]*]
Programming: Python · R · SQL

EXPERIENCE

Capital One | McLean, VA   Jun 2022 - Present
Marketing Data Scientist
- Develop and maintain marketing models that score customers' conversion propensity on a bi-monthly basis. Manage model risks through active model monitoring. Address potential model errors.
- Contribute to building a culture of experimentation by advising business teams on A/B testing best practices. Collaborate with business partners to refine experimental designs. Build tools to enable appropriate statistical tests and power analysis for stakeholders.

Duke University Marine Lab | Durham, NC   Aug 2021 - Apr 2022
Data Scientist
- Partnered across multiple institutions to deliver a machine learning pipeline for estimating the abundance of endangered North Pacific whales with acoustic recordings from the sub-Arctic.
- Improved and automated the existing abundance estimation process by developing *CNN* models that extract whales' "gunshot" calls.

Capital One | McLean, VA   Jun 2021 - Aug 2021
Data Science Intern
- Spearheaded gaming behaviors analysis in incentive marketing campaigns. Collaborated with cross-functional teams to gather information and ensure data integrity in the collection process.
- Prototyped models to identify key drivers in gaming behaviors using Databricks, Snowflake, and Scikit-Learn.
- Communicated insights to business partners and stakeholders by incorporating principles of data storytelling and interpretable machine learning.

Duke Graduate School | Durham, NC   Aug 2021 - Nov 2021
Teaching Assistant
- Mentored graduate students' projects in applied statistical inference and statistical learning.
- Held weekly office hours. Responded to student questions on course forums. Graded assignments and projects.

PROJECTS

Self-supervised Learning Pretext Tasks for Language Modeling *[[Github](#)]*   Nov 2021 - Dec 2021
- Constructed a 3-layer Transformer Encoder. Pre-trained the model with various pretext tasks, including Masked Language Modeling (MLM), Next Sentence Prediction (NSP), Sentence Order Prediction (SOP), and Sentence Permutation (SP).
- Fine-tuned each pre-trained model on the IMDB sentiment classification dataset in order to compare model accuracies with different pretext tasks. Achieved the highest accuracy of 0.86.

2020 Healthcare Analytics Competition *[[Github](#)]*   Sep 2020 - Nov 2020
- Trained and validated classifiers to predict patients with transportation challenges.
- Introduced the KNN missing value imputation technique and PCA feature extraction framework to the team. Advanced as 2nd round finalists out of 279 competing teams.