

Zhengyang (Jason) Qi

+1 (404) 797-9976 | zqi2@andrew.cmu.edu | <https://github.com/Jasonqi146>

EDUCATION

Carnegie Mellon University (CMU) <i>Master of Science in Intelligent Information Systems</i>	Pittsburgh, PA August 2022 - December 2023
Georgia Institute of Technology (Georgia Tech) <i>Bachelor of Science in Computer Science</i>	Atlanta, GA January 2020 - May 2022
Emory University <i>Bachelor of Arts in Mathematics & Economics</i>	Atlanta, GA August 2017 - December 2019

WORK EXPERIENCES

Intuit Inc. <i>Software Engineer Intern</i>	Mountain View, CA July - August 2022
<ul style="list-style-type: none">• Designed and constructed the complete lifecycle of an image captioning model to generate alt-texts for accessibility.• Researched and trained transformer-based mapping networks with state-of-the-art models (CLIP, GPT3, GPT2) as image encoders and text decoders on the MS COCO dataset with AWS Sage Maker.• Deployed the model on cloud services to auto-generate alt-texts during image uploads on Adobe DAM.	
EPI-USE America, Inc. <i>Software Engineer Intern</i>	Atlanta, GA June - July 2021
<ul style="list-style-type: none">• Implemented graph visualization interface of ADP's information networks through Neo4j graph database and Node.js.• Built a BERT classifier for privacy leakage detection (F1=.96) and graph algorithms for clustering and flow analysis.• Assembled interfaces and algorithms into a toolkit for optimizing network efficiency and information security.	
Kiwi Smart Shopping Assistant <i>CTO & Machine Learning Engineer</i>	Atlanta, GA December 2020 - December 2021
<ul style="list-style-type: none">• Led a team of six to develop Kiwi, a browser extension for cross-website price comparison on Google Chrome.• Trained ensemble models (Random Forests, AdaBoost, CatBoost) through 3 iterations of data collection, feature engineering, and model tuning (F1 = .89).• Delivered the application's beta-version backend that auto-updated database and ML algorithms.	

PROJECTS

CMU – Dialog Analysis and Persuasion Generation for Softbank Pepper	August 2022 - Present
<ul style="list-style-type: none">• Construct a dialog system to generate differentiated persuasions to a group of shoppers with different personas and stances.• Research and incorporate multi-agent dialog understanding and strategic language generation to Softbank Pepper robot.	
Georgia Tech - Visualizing Climate Change Discourses	March 2022 - Present
<ul style="list-style-type: none">• Provide analytic algorithms back boning visualization tools for political trends in climate-change-related Reddit channels.• Generate measurements of supportiveness, political stance, and demographics by text analysis and web data extraction.	
Kaggle - American Express Default Prediction (top 3%)	June - August 2022
<ul style="list-style-type: none">• Cleaned and downsized the large dataset and conducted feature engineering with aggregation and lag features.• Trained boosting models with 10-fold cross-validations including LightGBM, CatBoost, and XGBoost, and improved model performances by hyper-parameter tuning and constructing ensembles on out-of-fold predictions.	
Georgia Tech - Financial Question Generation for Earnings Conference Calls	August 2020 - December 2021
<ul style="list-style-type: none">• Formulated a dataset for question generation through matching questions to contexts by similarity metrics on name entities, LDA topic modeling, and BERT embeddings tuned on earning call texts.• Researched and trained classifiers for question intentions (F1 = .87) to enable conditional question generation.• Cleaned question corpus leveraging GPT-3 to improve transformer models' performance on the dataset.	

SKILLS

Programming Languages: Python (expert), Java (fluent), JavaScript (proficient), C++ (proficient), C (proficient)
Machine Learning: PyTorch, Scikit-Learn, HuggingFace, Keras, Gensim, SpaCy, NLTK, OpenCV, Pandas, NumPy, SciPy
Web Data Extraction: Selenium, BeautifulSoup, Scrapy
Cloud Services: AWS Sage Maker, AWS Lambda, AWS S3, Neo4j Graph Database, Kubeflow, Splunk
Software Development: Docker, Node.js, shell, Flair, Ubuntu
Analytics & Visualization: R, Stata, MATLAB, PowerBI