

William Braga

wmbraga7@gmail.com • (229) 712-1599 • www.wilabraga.com • www.linkedin.com/in/williambraga

EDUCATION

Georgia Tech | College of Computing

Master of Science in Computer Science

- TA for Natural Language Processing

December, 2021

Atlanta, GA

Georgia Tech | College of Computing

Bachelor of Science in Computer Science

- Concentration in Information Systems and AI
- Highest Honors

December, 2020

Atlanta, GA

WORK EXPERIENCE

Centuria, Inc. - National Data Buoy Center

Application Programmer

November 2022 – Present

Remote (Stennis Space Center, MS)

- Pioneered a new project to ingest hundreds of drifting buoys' realtime JSON data from a partner organization's AWS S3 bucket, transforming it into BUFR files with the ecCodes' Python library
- Lead the code review for their website redesign, correcting dozens of PHP, JS (ArcGIS), and HTML bugs
- Revamped their legacy realtime software in C to calculate nine new types of buoy weather measurements
- Wrote the developer standards document for safe user input sanitization and validation

College of Computing - Georgia Tech

Study Abroad Program and Teaching Assistant

May 2022 - August 2022

Barcelona, Spain

- TA'd the Intro to Databases and Discrete Math courses, designing the course project for Databases

CHAI Lab - Georgia Tech Research Institute

Graduate Research Assistant (NLP)

August 2021 - December 2021

Remote (Atlanta, GA)

- Datamined hundreds of online articles to compile an Excel sheet of US companies' Covid return-to-work policies, identifying key information using entity recognition and part-of-speech tagging with SpaCy
- Tracked the spread of Covid-19 by identifying potential Covid symptoms with comprehensive Regex matching across thousands of doctors' notes provided by the CDC

Assurant, Inc

Data Science Intern

May 2021 - August 2021

Remote (Marietta, GA)

- Lead the research initiative to find cutting-edge ML optimizations to speed up their claim decision times
- Tested caching techniques on hundreds of thousands of insurance claims data, containing natural language comments from the user and Assurant's metadata using PySpark on Databricks
- Achieved subsecond SLA time on cached claim decisions, a speed optimization of over 50%
- Presented methodology and accuracy/time results to multiple departments and VPs
- Conducted research on reducing model bias with the IBM Fairness 360 toolkit

PROJECTS

- Culinary Recipe Generation: Compared recipes made with GPT-2 and custom Keras Char/Word-RNNs
- Video Stabilizer: Implemented a Linear Program algorithm from a research paper to smooth video motion
- Artwork Cluster Visualization: Visualized K-Means clustered data from Cleveland Museum of Art's API

SKILLS

Programming: Python (Gensim, PyTorch, PySpark, SciKit, TensorFlow, NLTK), Java, SQL, C, Javascript, Bash, Matlab, R
Technical: ML, NLP, RestAPI, Databricks, Git, Docker, Spark, Jupyter, Tableau, CV, AWS S3, MySQL, MongoDB
World Languages: English (Native), Portuguese (Native, spoken only), Spanish (Advanced)
Hobbies/Volunteering: Reading (SciFi, History), Cooking, 3D Printing, Support Group Facilitator, Volunteer Webmaster