Cuong (Johnny) Nguyen

Georgia Institute of Technology, Atlanta, GA 30308
Email: johnny.nguyen@gatech.edu | Homepage: linkedin.com/in/johnny-nguyen-ckmjx/

Education

Georgia Institute of Technology — Atlanta, GA, USA
Joint B.S/M.S in Computer Science | Overall GPA: 3.95; Major GPA: 4.0 — *September 2020 – May 2023*

Georgia Institute of Technology — Atlanta, GA, USA
Bachelor of Science in Mathematics | Overall GPA: 3.95; Major GPA: 3.89 — *September 2020 – May 2022*

Purdue University — West Lafayette, IN, USA
Bachelor of Science in Computer Science | Overall GPA: 3.94; Major GPA: 4.0 — *August 2019 – August 2020*

Publications

Refereed Journal Publications

1. **Nguyen, C.**, Lu, N., Kane, J. M., Birnbaum, M. L., De Choudhury, M (2022). Cross-Platform Detection of Psychiatric Hospitalization via Social Media Data: A Comparison Study. JMIR Mental Health [**Impact Factor 6.33**]

Refereed Workshop Publications

1. **Nguyen, C.**, Nkemelu, D., Mehta, A., and Best, M. (2022). Why So Inflammatory? Explainability in Automatic Detection of Inflammatory Social Media Users. In ICLR 2022 Workshop on Practical Machine Learning for Developing Countries (Oral Presentation, Top 20% of accepted submissions)

In Preparation

1. **Nguyen, C.**, Nkemelu, D., Mehta, A., and Best, M. Detection of Inflammatory Users on Twitter in the Global South

Presentations

1. *Understanding Users Spreading Inflammatory Content on Twitter in the Global South.* GVU Fall Research Showcase 2022

2. *Cross-Platform Multimedia Detection of Schizophrenia Relapse Through Patient-Contributed Social Media Data.* Georgia Tech Annual Undergraduate Reearch Symposium 2022

Research Experience

GT Social Dynamics and Wellbeing Lab, Graduate Research Assistant — January 2021 - Present
Research Advisor: Munmun De Choudhury — *Atlanta, GA*

- Wrote a first-author paper that seeks to measure and compare the effectiveness of data from multiple social media platforms for training models to distinguish between patients with near-term hospitalization due to Schizophrenia Spectrum Disorder (SSD) and healthy controls.
- Collaborated with clinical psychiatrists at Northwell Health in order to access, clean and normalize their raw data using Python and SQL on social media usage and hospitalization records among psychiatric patients and healthy controls
- Feature-engineered a set of user-level features highly indicative of psychiatric hospitalizations due to SSD can be uniformly used to build models across social media platforms
- Built and tuned RNN and BERT models to differentiate multi-modal, multi-platform social media data (Facebook, Twitter, Instagram) belonging to diagnosed patients and healthy controls with comparable state-of-the-art F1 of 75%
- Implemented a React-based prototype for a clinician-facing tool that integrates social media-based machine learning insights for early intervention of relapse hospitalizations due to SSD

GT Technology and International Development Lab, Research Assistant — January 2021 - Present
Research Advisor: Michael Best — *Atlanta, GA*

- Worked on the back-end for Aggie, an open-source platform for election monitoring on social media deployed in 10+ countries and provided real-time assistance to Aggie's deployment in Ethiopia by the Center for Advancement of Rights and Democracy (CARD).
- Inspired by the needs of non-profit social media trackers in Ethiopia and previous deployments, formulated research questions on characterizing and identifying users who spread inflammatory (i.e. violence-provoking) content in the Global South context
- Created a novel dataset of tweets and Twitter users through hand-crafted search terms related to the discourse surrounding the recent Ethiopian civil conflict
- Identified key differences between users who did and did not spread inflammatory content regarding the Ethiopian civil conflict and adapted Graphical Neural Networks (GNNs) to identify inflammatory users

- Wrote a first-author paper presenting the aforementioned results and its implications to civil society-led social media tracking initiatives
- Collaborated with another graduate student in the lab to verify the effectiveness of a novel target-aware data augmentation technique for low-resource hate speech detection on two Vietnamese hate speech datasets.

Purdue AKRA-NLU Lab, Research Assistant March - December 2020
Research Advisor: Julia Rayz *West Lafayette, IN*

- Designed a PL/SQL script to calculated similarity between jokes based on the General Theory of Verbal Humor (GTVH)
- Worked on the development of a full-stack website using FuelPHP and OracleSQL to facilitate joke annotation
- Maintained an OracleSQL database of 20000 jokes and associated annotations

Notable Course Projects

Detection and Generation of Tiktok Euphemisms September - December 2022
Professor: Srijan Kumar; Course: CSE 8803 Data Science for Social Networks *Atlanta, GA*

- Implemented a novel Transformer-based euphemism generation algorithm and finetuned it on a corpus of Gab posts.
- Quantified the prevalence and usage of euphemistic hashtags (i.e. violating Terms of Service) on TikTok generated by the aforementioned algorithm through NLP techniques (sentiment analysis, LIWC, etc.) and qualitative analysis

The Nerdfighteria Discord Server: A Gathering Place for Awesomeness September - December 2022
Professor: Amy Bruckman; Course: CS 6470 Online Communities *Atlanta, GA*

- Performed interviews with users of the Nerdfighteria Discord and subsequent qualitative analysis of interview transcripts
- Wrote a 40-page research paper alongside 2 other graduate students on how Nerdfighteria Discord acts as an effective third place and safe space for its members

Dissecting Large Language Models' Rationale for Implicit Hate Classification September - December 2021
Professor: Diyi Yang; Course: CS 4650 Natural Language Processing *Atlanta, GA*

- Trained multiple Large Language Models (LLMs) to classify implicit hate speech based on a recently-introduced hate speech taxonomy
- Dissected the rationale of the aforementioned models using SHAP values, with a focus on lexicon related to protected characteristics and punctuations

Nutribuzz: A Dietary Assistant App for Georgia Tech Students September - December 2021
Professor: Chris Le Dantec; Course: CS 3873 Human Computer Interaction *Atlanta, GA*

- Designed a novel wearable interface to assist Georgia Tech students with dietary monitoring
- Conducted qualitative user study to measure the design's usability

Industry Experience

Software Engineering Intern May 2022 – August 2022
Intuit *Mountain View, CA*

- Created a Python library for the creation of customizable Kubeflow configuration files for finetuning Large Language Models (LLMs) on various natural language tasks across multiple Intuit products (QuickBooks, TurboTax, Mint, etc.).
- Designed templates for AWS Sagemaker inference endpoints and API contracts for Multilabel Text Classification LLMs
- Trained and fine-tuned a novel BERT-based model using the aforementioned library on the task of mapping natural language queries to QuickBooks use cases, achieving a Hamming Loss score of 0.02 and a inference time 3 times less than previous models.

Award

President's Undergraduate Research Award (PURA) Salary Award Summer 2021, Spring 2022

- 1500$ stipend to undergraduate students who conduct research with a Georgia Tech faculty member (30% acceptance rate)
- Awarded in Summer 2021 for work on investigating novel methods to prevent the spread of hate speech and misinformation on social media in low-resource conditions
- Awarded in Spring 2022 for work on investigating the effectiveness on distinguishing between patients suffering from Schizophrenia Spectrum Disorder (SSD) from healthy controls using multimodal data from multiple social media streams

Georgia Tech Faculty Honors Fall 2020 - Spring 2022
Georgia Tech Highest Honors Spring 2022

References

Dr. Michael Best

Title: Professor at the Georgia Institute of Technology, **Email:** mikeb@gatech.edu

Dr. Munmun De Choudhury

Title: Associate Professor at the Georgia Institute of Technology, **Email:** munmun.choudhury@cc.gatech.edu

Dr. Julia Rayz

Title: Professor at the Purdue Univerity, **Email:** jtaylor1@purdue.edu