Jacob Choi

(+1) (408) 828-3935
jacobjch@usc.edu
https://jacobxchoi.github.io/jacob/

RESEARCH INTERESTS

Exploring the applications and ethical implications of large language models (LLMs), with a focus on mitigating vulnerabilities such as bias, adversarial attacks, and backdoor exploits. Investigating novel techniques for improving the robustness and fairness of AI systems, particularly in the context of natural language processing and knowledge representation.

EDUCATION

University of Southern California, *Los Angeles, CA, US* *Expected May 2026*

Master of Science in Computer Science

Emory University, *Atlanta, GA, US* *May 2024*

Bachelor of Science in Computer Science with Highest Honors, Advisor: Dr. Jinho Choi Major GPA: 3.700

- Finetuned a large language model to map biblical entities and their relationships to develop a chatbot capable of answering context-specific questions from religious texts.

RESEARCH EXPERIENCE

Emory NLP Lab *Aug 2023 – Aug 2024*

Undergraduate Researcher, Advisor: Dr. Jinho Choi *Atlanta, GA*

- Fine-tuned Llama models using Low-Rank Adaptation (LoRA) to create a verse extraction tool for religious texts, employing large language models (LLMs) to answer user queries with relevant Bible verses
- Developed five custom NLP tasks to assess chatbot effectiveness in religious dialogues, including semantic similarity, cross-referencing, named entity recognition, and theological question answering.
- Implemented web scrapers to gather and structure large datasets from Bible versions and religious sources, leading to enhanced chatbot performance across multiple tasks compared to ChatGPT
- Evaluated model performance using various metrics such as exact match accuracy, cosine similarity, and clustering analysis to provide insights into theological NLP model capabilities and limitations

Language Information and Computation Lab (LINC) - Research Experience for Undergraduates (REU)

University of Colorado at Colorado Springs *May 2023 – Aug 2023*

Visiting Researcher, Advisor: Dr. Jugal Kalita *Colorado Springs, CO*

- Explored innovative approaches to creating latently inseparable backdoor attacks in neural networks, particularly for text-based models, addressing the limitations of latent separability assumptions in existing defense methods
- Designed and implemented backdoor attack mechanisms using partial and asymmetric triggers, enhancing attack stealth and robustness against latent-space-based defense methods
- Analyzed and documented experimental results using attack success rates (ASR), perplexity, grammar error rate, and clustering metrics, contributing insights for future research on backdoor attacks in NLP
- Prepared and Presented research findings, contributing to a broader understanding of vulnerabilities in large language models and informing future work on adversarial attacks in NLP

Emory Graph Mining Group, Emory University *Feb 2021 – Apr 2022*

Undergraduate Research Assistant, Advisor: Dr. Carl Yang *Atlanta, GA*

- Investigated utilization of large language models to automatically build structured knowledge (taxonomies and knowledge graphs) in the healthcare domain
- Performed evaluation and verification of structured knowledge (concept maps) derived from graph generative models

PUBLICATIONS

Conference Papers

[1] What is Your Favorite Gender, MLM? Gender Bias Evaluation in Multilingual Masked Language Models

MDPI 2024. Jeongrok Yu, Seong Ug Kim, **Jacob Choi**, Jinho Choi.

Jacob Choi

(+1) (408) 828-3935
jacobjch@usc.edu

WORK EXPERIENCE

Zybooks, *John Wiley & Sons* *Oct 2024 – Present*
Editor *Los Angeles, CA*

- Conducted final review of computer science textbooks from a student perspective, ensuring accuracy in spelling, grammar, and conceptual clarity.

Emory Computing Center, Emory University *Aug 2023 – May 2024*
Technology Consultant *Atlanta, GA*

- Provided technological assistance to students, faculty, and staff by troubleshooting hardware and software problems, setting up devices, and offering guidance on academic projects
- Oversaw usage and maintenance of computing resources, such as computers, printers, and other peripherals

RELEVANT COURSEWORK

Machine Learning, AI Research Practicum, Natural Language Processing, Deep Learning with Computer Vision, Data Structures and Algorithms, Database Systems, Linear Algebra, Multivariable Calculus, Analysis of Algorithms

TECHNICAL SKILLS

Programming Languages: Python, Java, C, R, x86, JavaScript, SQL, NoSQL, MapReduce, PostgreSQL, MySQL
Tools: Pytorch, Docker, Git, Jira

PROJECTS

AI Research Practicum | Gender Bias in Machine Translation

- Resolved existing biases in previous methodologies and enhanced accuracy by removing confounding biases
- Produced gendered lexicon to detect gendered sentences of different languages without parallel corpora
- Compiled research report, delivered oral and poster presentation under guidance of Dr. Jinho Choi

Personalized E-Sports Conversational Chatbot Agent

- Developed AI-powered conversational chatbot for League of Legends esports, using Emora STDM for structured dialogue management and OpenAI's GPT-3.5 Turbo API for dynamic, context-aware response generation

I Want Good Food – Finding Good Restaurants and Reviewers on Yelp Dataset

- Utilized advanced statistical machine learning models to identify key attributes correlating with positive and negative Yelp restaurant reviews to find customer preferences.

Breast Cancer Detection

- Implemented and fine-tuned a convolutional neural network-based object detection model to achieve a mean Average Precision (mAP) exceeding 90% for breast cancer detection in medical imaging

Course Calendar | Product Owner

- Developed online calendar (CRUD application) on SQLite and Python, implemented user authentication with Auth0m, implemented front-end features such as search and auto-layout using Twitter Bootstrap
- Led 7-person scrum team using agile methodology – organized standups, sprints, and retrospectives

AWARDS

Deans Honors List x2 *Fall 2022, 2023*

HackEmory Annual Hackathon *2nd Place* *Sep 2020*

- Developed self-updating Spotify playlist with YouTube trending music using Spotify API

Congress Bundestag Youth Exchange Scholar (CBYX) *Jul 2019 – Mar 2020*

- Year-long language immersion program in Aachen, Germany funded by US Department of State

LEADERSHIP AND COMMUNITY ENGAGEMENT

Undergraduate Student Conduct Peer Review Board Aug 2023 – May 2024
Member/Panelist *Atlanta, GA*

- Engaged in comprehensive training to evaluate potential violations of the Undergraduate Code of Conduct, honing skills in assessing incident reports, considering multiple perspectives, and contributing to fair and informed decisions on responsibilities and sanctions.

Journey Church of Atlanta *Aug 2022 – May 2024*
Small Group Leader *Atlanta, GA*

- Facilitated weekly discussions exploring scripture as it applies to daily life and manifestation of faith
- Coordinated and direct engaging musical worship sessions to foster spiritual engagement among college students
- Engaged in local outreach initiatives, providing tutoring and mentorship for children's educational development

Emory In Via *Aug 2022 – May 2024*
Treasurer *Atlanta, GA*

- Facilitated interactive faculty-student meetings, fostering dialogue on the intersection of faith with academia.
- Efficiently managed $2,000+ annual budget, strategically allocating funds to support publication of student journals
- Led fundraising initiatives and authored successful grant proposals to secure funding from interdisciplinary offices

ADDITIONAL SKILLS AND INTERESTS

Languages: Fluent in English, Proficient Mandarin Speaking, CEFR-Level B1 German (intermediate)
Fine Arts: Guitar, Violin, Piano, Chinese Landscape Painting