Stella Li

Baltimore, MD 21218 | sli136@jhu.edu | LinkedIn | Personal Website

RESEARCH INTERESTS

Natural Language Processing, Computational Socio/psycholinguistics, Multilingual Low-Resource Machine Translation, Human-Centered NLP, Ethical NLP

EDUCATION

Aug. 2019 – Current

Johns Hopkins University | Baltimore, MD
B.S. in Computer Science
Thesis: Learning from Gibberish: Code-Mixing Data Augmentation for Sentiment Analysis
Other Majors: Cognitive Science (linguistics focus), Applied Mathematics (statistics focus)
Minor: Mathematics
Cumulative GPA: 3.99/4.0, Major GPA: 4.0/4.0

Relevant Coursework

Natural Language Processing, Machine Learning: Deep Learning, Machine Translation, Human Language Technology, Data Science, Human-Computer Interaction, Math Models of Language

PUBLICATIONS

2022

[1] **Shuyue Stella Li** and Kenton Murray. 2022. Learning from Gibberish – Synthetic Language Agnostic Data Augmentation for Code-Mixing Sentiment Analysis. Under ACL Rolling Review.

2022

[2] **Shuyue Stella Li**, Hannah Peeler, Andrew N. Sloss, Kenneth N. Reid, and Wolfgang Banzhaf. 2022. Genetic improvement in the shackleton framework for optimizing LLVM pass sequences. In Proc. of the Genetic and Evolutionary Computation Conference Companion (GECCO'22).

2022

[3] Hannah Peeler, **Shuyue Stella Li**, Andrew N. Sloss, Kenneth N. Reid, Yuan Yuan, and Wolfgang Banzhaf. 2022. Optimizing LLVM pass sequences with shackleton: a linear genetic programming framework. In Proc. of the Genetic and Evolutionary Computation Conference Companion (GECCO'22).

2022

[4] **Shuyue Stella Li***, Xiangyu Zhang*, Shu Zhou, Hongchao Shu, Ruixing Liang, Hexin Liu, Leibny Paola Garcia. 2022. PQLM - Multilingual Decentralized Portable Quantum Language Model for Privacy Protection. Under review at ICASSP'23.

2022

[5] Xiangyu Zhang, **Shuyue Stella Li**, Zhanhong He, Roberto Togneri, and Leibny Paola Garcia. 2022. End-to-End Lyric Recognition with Self-Supervised Learning Models. Under review at ICASSP'23.

2022

[6] Yu Xuan, Xiangyu Zhang, **Shuyue Stella Li**, Zihan Shen, Leibny Paola Garcia, and Roberto Togneri. 2022. A new approach to extract fetal electrocardiogram using convex combination of adaptive filters. Under review at ICASSP'23.

RESEARCH EXPERIENCE

Language & Speech Processing

May 22 – Current

Multilingual Generalizability of Self-Supervised Models | Johns Hopkins CLSP
This study uses large self-supervised models trained on English as a feature extractor for ASR tasks in languages from diverse language families to investigate multilingual generalizability of these models. We use deepCCA to examine whether semantics or phonology is learned by SSL.

May 22 – Oct. 22

Language Agnostic Code-Mixing Data Augmentation | Johns Hopkins CLSP
This work analyzes the underlying linguistic phenomenon of Code-Mixing. I invented a language agnostic data augmentation algorithm for Code-Mixing languages by leveraging cross-lingual information. Results are under ACL Rolling Review.

Mar. 22 – Oct. 22 **Quantum Language Model** | Johns Hopkins CLSP
This project addresses the concern of data privacy in large pre-trained language models. We train language embeddings on quantum machines to preserve privacy without sacrificing performance. We also propose a portable method to transfer quantum information to classical machines and achieve competitive performance on downstream tasks. Results are under review at ICASSP'23.

Feb. 22 – Oct. 22 **Self-Supervised Learning for Lyric Recognition** | Johns Hopkins CLSP
This work establishes an end-to-end lyrics recognition baseline and uses self-supervised learning models to outperform state-of-the-art by 2.4%. Moreover, we study the generalization ability of the SSL features using attention weights. Results are under review at ICASSP'23.

General Machine Learning

Sep. 20 – Current **Delineo Covid-19 Modeling** | Johns Hopkins & University of Tsukuba & Microsoft Research
Designed expansion-based clustering algorithm for community detection using covid-19 mobility data. Built agent-based stochastic model for disease spread simulation and intervention modeling.

Jul. 22 – Current **Cryptographic Code Optimization** | Michigan State University & University of Adelaide
Optimizing cryptographic code generation with genetic improvement. This project aims to automatically reduce the computation time for encryption algorithms.

Mar. 22 – Oct. 22 **Fetal Electrocardiogram Extraction** | Johns Hopkins School of Medicine
This project uses a convex combination of adaptive filters to detect the fetal heart rate from the ECG signals from the mother's abdominal and thorax. Our algorithm out-performs both single adaptive filters and convex combination of RLS filters. Results are under review at ICASSP'23.

May 21 – Jul. 22 **Genetic Programming – LLVM Optimization** | Michigan State University
Designed and implemented novel GP algorithm for LLVM compiler flag optimization (20%)
Published work at GECCO; second author of GP paper; first author of GI paper

Jan. 20 – Jan. 22 **Language and Cognition Lab** | Johns Hopkins Language and Cognition Lab
Investigated developmental spatial cognition using Lego Block building
Created ML model for movement prediction and stability analysis using motion sensor data

TEACHING EXPERIENCE

Sep. 2022 – Current **Human-Computer Interaction (EN.601.490)** | Johns Hopkins University | Baltimore, MD
Course Assistant
Supervised design projects, held weekly office hours, graded homework & exams

May. 2022 – Aug. 2022 **Computer Ethics (EN.601.104)** | Johns Hopkins University | Baltimore, MD
Head Course Assistant
Finding supplemental materials, organizes in-class presentations, graded homework & exams

Sep. 2021 – May 2022 **Introduction to Statistics (EN.553.430)** | Johns Hopkins University | Baltimore, MD
Teaching Assistant
Taught recitation sections, held weekly office hours, graded homework & exams

Sep. 2021 – May 2022 **Intermediate Programming (EN.601.220)** | Johns Hopkins University | Baltimore, MD
Course Assistant
Oversaw in class exercises, held weekly office hours, graded homework & exams

INDUSTRY EXPERIENCE

May 22 – Aug. 22 **Yext** | Arlington, VA
Software Engineering Intern
Integrated client data to Yext platform for real-time site information updates using Go
Created a Figma Style Picker to improve developer workflow and scalability using ReactJS

May 20 – Aug. 20	**Bytedance (TikTok) AI Lab** | Beijing, China **Algorithms Engineering Intern** Trained neural networks for text normalization in TTS tasks Implemented algorithms for theme clustering and complexity ranking for TikTok videos
May 19 – Aug. 19	**IBM AI Doctor** | Beijing, China **Data Science Intern** Created ML models to predict diseases from symptoms using EHR records Improved classification accuracy from 74% to 99% with a hybrid algorithm of GA with SVM

HONORS

2022	Upsilon Pi Epsilon Computer Science Honor Society
2021 & 2022	Omicron Delta Kappa National Leadership Honors Society
2022	Omega Psi National Cognitive Science Honors Society
2019	Cum Laude Society

AWARDS

Jul. 2022	GECCO-GI Best Presentation Award
Jul. 2022	ACM Student Travel Grant
May 2022	PAJH Greek Scholars Award
Sep. 2021	Grace Hopper Scholarship Award
Every Semester	Dean's List (GPA > 3.50)

LEADERSHIP

Oct. 19 – Current	**HopHacks at Johns Hopkins**, Organizer & Head of Sponsorships (21-22)
Oct. 19 – Current	**Omega Psi Cognitive Science Honor Society (JHU Chapter)**, President (22)
Apr. 20 – Current	**Hopkins Undergraduate Society of Applied Math (HUSAM)**, President (22)
Oct. 19 – Current	**Alpha Phi Omega National Service Fraternity**, Executive Board Member (20)
Feb. 20 – Current	**Phi Mu Fraternity**, Academics Excellence Chairwoman (20-21), Honors Member (22)
Oct. 20 – Current	**Women in Computer Science (WiCS)**, Executive Board Member (21)
Nov. 21 – Current	**Theta Tau Engineering Fraternity,** Social Chair (22)

SKILLS

Programming Languages	Python, C/C++, Java, R, MATLAB, HTML/CSS, JavaScript, ReactJS, Go
Languages	English (native); Mandarin Chinese (fluent); Spanish (intermediate)