LAURA HAMBRIDGE

Data Analyst | Computational Linguist

@ Ihambrid@buffalo.edu

in linkedin.com/in/laura-hambridge

github.com/lhambrid

(716) 200 8134

Buffalo, NY

EDUCATION

University at Buffalo

M.S. Computational Linguistics

❷ Buffalo, NY

CGPA: 3.7/4.0

EXPERIENCE

Capstone, University at Buffalo

"Comparing Per-Language Performance in a Many-to-Many Machine Translation System"

Mov 2021 - Jan 2022

Rui P. Chaves, Associate Professor, Linguistics

- Sought to find if NMTs have more difficulty translating between different pairs of languages and whether NMTs struggle more with lexical or syntactic choices.
- In Python, fine-tuned a many-to-many NMT on translating to and from 7 different languages. Scored translations using sacrebleu.
- Wrote tests to assess the syntactic and semantic similarity between the model's translations and human translations.

Causality Across Languages, SUNY Research Assistant

₩ Aug 2021 - Apr 2022

Jürgen Bohnemeyer, Professor, Linguistics

- Sought to find if the way people assign responsibility for events and outcomes is dependent on their culture. So CAL blog
- In R, used linear regression and ANOVA for confirmatory analysis and conditional inference trees and random forests for exploratory analysis on ordinal and categorical data from sociolinguistic experiments.

The Baldy Center for Law & Social Policy, SUNY Research Assistant

🛗 Sept 2020 - Jan 2022

- Amy Semet, Associate Professor, Law
- Sought to find if a legal case opinion's author could be predicted based on word choice alone.
- Extracted information from approx. 7k legal documents using regular expressions and normalized to optimize for analysis.
- Built both a simple neural network and a statistical model driven by tf-idf to predict the author (out of about 27 authors).

Natural Language Understanding Laboratory, Jacobs School of Medicine and Biomedical Sciences, SUNY

May 2021 - Aug 2021

Peter Elkin, Professor, Biomedical Informatics

• Developed a web-based user interface for biomedical researchers to extract a csv of medical terms, codes for such terms, and relationships between such terms from any input text. So NLU Lab

Tesserae, University of Notre Dame & SUNY Volunteer Backend Engineer

Mar 2020 - Nov 2020

Neil Coffee, Professor, Classics

• Implemented sound-matching feature and accompanying tf-idf based algorithm in Python for version 5 of Tesserae, a project developed to aid scholars in comparing classical Greek and Latin texts.

PROGRAMMING

Python R SQL JavaScript Bash Markdown LaTeX



TOOLS

PyTorch huggingface scikit-learn numpy pandas regex spaCy
WordNet NLTK matplotlib
ggplot2 git Excel

SKILLS

tf-idf Text Mining NLP NLU
Distributional Semantics
Language Modeling Machine Learning
Clustering Statistical Analysis
Linear & Logistic Regression
Analysis of Variance

INTERESTS

natural language understanding
language modeling machine translation
historical linguistics linguistic typology
under-served languages
automated speech recognition
cognitive science data analysis

COURSEWORK

- Computational Linguistics
- Machine Learning
- Information Retrieval
- Quantitative Methods
- NLP & Text Mining

LANGUAGES

English Spanish German French

