Elise Jing

+1 (812) 360-8298 — jingyz42@gmail.com — yzjing.github.io

Education

Ph.D. Complex Systems, Indiana University — January 2021

- Minor: Computational Linguistics
- Main areas: Data Science, Natural Language Processing, Network Science

B.A. Sun Yat-sen University — June 2014

- Dual major in Information Science and Anthropology

Work Experience

Scientist, Sirius XM Pandora — November 2020 – Current

- Developed machine learning models to automatically create topic labels for more than 182,000 podcast episodes based on transcripts.
- Created embeddings for over 22,000 podcasts and tags to use in podcast recommendation.
- Deployed a topic pipeline collaborating with content experts that supplies topic pages, pills, and topic-based recommendations.

Content Science Intern, Sirius XM Pandora — June 2019 – August 2019

- Created a labeled dataset of over 500 podcast episode transcripts.
- Developed a BERT-based deep learning model that identifies the introduction in podcasts.

Skills

Languages: Python (primary), R, SQL, Apache Hive
Libraries:

- Data science: NumPy, SciPy, Pandas, Scikit-Learn, Jupyter, Matplotlib
- Deep learning: PyTorch

Cloud computing: Google Cloud Platform, Apache Airflow
Other: Git, LaTeX

Research Experience

Characterizing Partisan Political Narratives about COVID-19 on Twitter (arXiv:2103.06960)

- Constructed a dataset of over 440,000 Tweets by U.S. politicians during the COVID-19 pandemic.
- Identified and quantified framings in political tweets automatically using an embedding-based machine learning model.
- Discovered and analyzed major semantic roles about COVID-19 in political tweets.

Sameness Attracts, Novelty Disturbs, but Outliers Flourish in Fanfiction Online (arXiv:1904.07741)

- Scraped websites to collect a large dataset of more than 4 million pieces of fanfiction.
- Quantified the novelty of fanfiction writing using TF-IDF and LDA models.
- Presented findings that challenge the classical theory of novelty and liking.

Global labor flow network reveals the hierarchical organization and dynamics of geo-industrial clusters (Nature communications 10 (1), 1-10, Patent US10592535B2)

- Analyzed LinkedIn's employment history data from more than 500 million users over 25 years to construct a labor flow network of over 4 million firms across the world.
- Characterized the labor force flows between 147 industries and more than 3,000 regions.

Honors & Awards

- NSF Research Trainee scholarship in Complex Networks and Systems — 2018
- Santa Fe Institute alumna of the Complex Systems Summer School — 2016
- LinkedIn Economic Graph Challenge (One of the 11 winning teams) — 2015