JENNIFER J. RUSSELL

Contact: (585) 645-3468 | jjrussell10@gmail.com | https://www.linkedin.com/in/jennifer-j-russell/
Portfolio: https://github.com/jenna-russell/Russell_Jennifer_Portfolio

Research Interests

Natural Language Generation, Text Summarization, Truthful Generation

Education

Bachelor of Science in Information Science, Statistics *Aug 2017 - Dec 2020*
Cornell University
GPA:3.8, Magna Cum Laude

Work Experience

Data Scientist, Bank of America, Erica Conversational AI Research & Development
Plano, TX, July 2021 - Present

- Created Semantic Role Labeling system specific to Conversational AI to improve the contextual understanding of chatbots
- Improved generative dialogue summarization model for call centers by creating a summarization dataset focused specifically on task-driven dialogue summaries for customer service and invented a hybrid extractive-abstractive modeling technique for real-time summarization
- Improved machine-translation system using weakly supervised methods of data generation for Spanish for Erica solution
- TODO make this a one-liner

Global Technology Summer Analyst, Bank of America
Remote, May 2020 - Aug 2020

- Team lead creating a forecasting model for ATM utilization during the pandemic, reduced MAE from 8.6% to 6.3%

Data Analyst Intent, Corning Incorporated
Corning, NY, May 2019 - Aug 2019

- Improved emerging trend identification by analyzing news data using topic modeling to track the rise and fall of industry trends

Current Research Projects

Abstractive Dialogue Summarization

- Creating an Issue-Resolution Summarization system for call center calls with the issue a customer is calling about and how the agent resolved a call

- Built a Bart-based summarizer fine-tuned on DialogSum & XSum datasets
- Employed methods to generate more faithful/truthful summaries such as training with a contrastive loss function and re-ranking beams by automatic faithfulness metrics

Extreme Summarization for Topic Extraction

- Research methods of extreme summarization to provide short descriptions for the purpose of extracting the main topic of a text
- Created topic summarization system using the following methodology
    - Used fine-tuned issue-resolution Bart to generate 3 summary beams for about 15k call transcripts.
    - Used Few-shot label generation with MPT-7B based off of Unlabeled Data Generation methodology to generate extreme summaries from the issue-resolution summaries
    - Used transfer learning to train Bart to learn extreme summaries from original dialogue text.

Semantic Role Labeling

- Researching semantic role systems for dialogue systems to improve contextual understanding of low-resource systems
- Proposed new semantic role schema specifically for chatbots
- Demonstrated improved contextual understanding of chatbot when using the semantic role system to gain a better underlying understanding of language

Teaching

Teaching Assistant, Introduction to Data Science (INFO/CS 2950), Spring 2020 & Fall 2020
Teaching Assistant, Introduction to Computing Using Python (CS 1100), Spring 2019 & Fall 2019

Membership

Women in Computing at Cornell (2017-2020)
Information Science Student Association (2018-2020)
Women in Data Science at Bank of America (WiDS) (2021 - present)

Leadership/Service

Executive Board Member, Women in Data Science at Bank of America (2022 - present)
Program Lead, Girls Who Code of North Texas Summer Immersion Program (2023 - present)
Mentor, The Coding School (2021 - 2022)

Patent Applications

1. "***Selection System for contextual prediction processing versus classical prediction processing***". US Patent Application No. 17/993,048, filed November 23, 2022.
2. *"**Action-topic Ontology**"*. US Patent Application No. 17/993,038, filed November 23, 2022.
3. *"**Semantic frame builder**"*. US Patent Application No. 17/993,029, filed November 23, 2022.
4. ***"Dynamic semantic role classification"***. US Patent Application No. 17/993,019, filed November 23, 2022.
5. *"**Dual-pipeline utterance output construct**"*. US Patent Application No. 17/993,013, filed November 23, 2022.
6. Russell, J., Noorzidaeh, E., Dibia, E., Jhaveri, R. *"**Iterative Processing System for Small Amounts of Training Data**"*. US Patent Application No. 18/199,073, filed May 18, 2023.
7. Russell, J., Noorzidaeh, E., Yannam, R., Jhaveri, R. *"**Multilingual Chatbot**"*. US Patent Application No. 17/993,063, filed November 23, 2022.
8. "***Performance Optimization for Real-time Large Language Speech-to-text Systems***". US Patent Application No. 18/204,981, filed June 2, 2023.
9. *"**Call center voice system for use with a real-time complaint identification system**"*. US Patent Application No. 18/144,925, filed May 9, 2023.

Relevant Coursework

Introduction to Data Science, Natural Language Processing, Machine Learning for Intelligent Systems, Machine Learning for Data Science, Statistical Computing, Data-Driven Web Applications, Interactive Information Visualization

References

David Mimno, Cornell University
Emad Noorzidaeh, Bank of America