Jake Vasilakes

jvasilakes@gmail.com (763) 339-9216 jvasilakes.github.io github.com/jvasilakes

Experience

Oct 2017 - Present | **Natural Language Processing Researcher and Programmer**
Institute for Health Informatics: University of Minnesota - Minneapolis, MN

- Implementing a knowledge graph of dietary supplements using data integrated from multiple sources.
- Building and evaluating survival models on EHR data from over 30,000 patients.
- Researching methods to reduce the amount of labeled data required to build effective machine learning models.

Feb - Nov 2016 | **Research Assistant in Speech Processing**
IARPA Babel Project: University of Cambridge - Cambridge, UK

- Trained and evaluated state of the art machine learning systems for multilingual speech recognition on multiple datasets each containing over 80 hours of audio data.
- Developed a statistical model to predict system performance on unseen data to within 5%.
- Designed and implemented pipelines for building statistical language models using Python and shell.
- Supervised an undergraduate student's research project on optimizing a search graph, which was published in IEEE ICASSP 2017.

Education

Aug 2015 | **University of Edinburgh**
M.S. Speech and Language Processing, *Distinction*
Thesis: *Automatic Generation of Wide-scale Semantic Representations in NLTK*
Advised by Dr. Ewan Klein

June 2013 | **Loyola University of Chicago**
B.A. Philosophy with Honors, *Magna Cum Laude*
Thesis: *The World of Speech*
Advised by Dr. Hanne Jacobs

Skills

Data Science: Natural language processing, machine learning, deep learning, statistical analysis
Health Informatics: Biomedical concept detection, survival models, pharmacovigilance
Programming Languages: Python, R, C, *nix shell (`bash` & `tcsh`), SQL
Software and Libraries: NumPy, SciPy, Pandas, NLTK, Neo4j, Jupyter, Git, LaTeX

Publications

Vasilakes, J., Fan, Y., Rizvi, R., Bompelli, A., Bodenreider, O., Zhang, R. (2019). *Normalizing Dietary Supplement Product Names using the RxNorm Model.* MedInfo, Lyon, France. *Forthcoming*

Vasilakes, J., Rizvi, R., Zhang, J., Adam, T.J., Zhang R. (2019). *Detecting Signals of Dietary Supplement Adverse Events from the CFSAN Adverse Event Reporting System (CAERS).* American Medical Informatics Association (AMIA) Informatics Summit, San Francisco, CA

Vasilakes, J., Rizvi, R., Melton, G.B., Pakhomov, S., Zhang, R. (2018). *Evaluating Active Learning Methods for Annotating Semantic Predications.* Journal of the American Medical Informatics Association (JAMIA) Open.

Vasilakes, J., Wang, H., Ragni, A., Gales, M.J.F. & Knill, K.M. (2016). *Speech Recognition and Keyword Spotting Performance Analysis Across Languages.* Poster presented at UK Speech Conference, Sheffield, UK

Rizvi, R., Wang, Y., Nguyen, T., **Vasilakes, J.**, Bian, J., He, Z., Zhang, R. (2019). *Analyzing Social Media Data to Understand Consumers Information Needs on Dietary Supplements.* MedInfo, Lyon, France. *Forthcoming*

Xing, H., Zhang, R., Rizvi, R., **Vasilakes, J.**, Yang, X., Guo, Y., He, Z., Prosperi, M., Bian, J. (2018). *Prototyping an Interactive Visualization of Dietary Supplement Knowledge Graph.* IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Madrid, Spain

Rizvi, R., Adam, T.J., Lindemann, E., **Vasilakes, J.**, Pakhomov, S., Bishop, J., Meltion, G.B., Zhang, R. (2018). *Comparing Existing Resources to Represent Dietary Supplements.* American Medical Informatics Association (AMIA) Summits on Translational Science, San Francisco, CA

Ragni, A., Wu, C., Gales, M.J.F., **Vasilakes, J.**, Knill, K.M. (2017). *Stimulated training for automatic speech recognition and keyword search in limited resource conditions.* IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA

Ragni, A., Saunders, D., Zahemszky, P., **Vasilakes, J.**, Gales, M.J.F., Knill, K.M. (2017). *Morph-to-word transduction for accurate and efficient automatic speech recognition and keyword search.* IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA

Chen, X., Ragni, A., **Vasilakes, J.**, Liu, X., Knill, K.M., Gales, M.J.F. (2017). *Recurrent neural network language models for keyword search.* IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA