

Lynn Samson

📞 978-876-6196 📍 Boston, MA ✉ lnsamson@gmail.com [/in/lynnsamson](https://www.linkedin.com/in/lynnsamson) [lasamson](https://github.com/lasamson)

Education

- University of Massachusetts Amherst** Sep 2018 - May 2020
Master of Science in Computer Science Amherst, MA
- **Coursework:** Neural Networks, Deep Learning for Natural Language Processing (NLP), Software Engineering, Advanced Algorithms
- University of Massachusetts Amherst** Sep 2014 - May 2018
Bachelor of Science in Computer Science and Mathematics Amherst, MA
- **Coursework:** Machine Learning, Probability Theory, Statistical Inference, Mathematical Modeling, Numerical Methods, Databases

Experience

- Sensor Tower** Aug 2021 – May 2023
Data Scientist San Francisco, CA
- Utilized Ruby and MongoDB to develop statistical demographics models and productionize estimates as part of the Usage Intelligence team.
 - Delivered improved estimates to clients via high-impact projects, including accurate bias adjustments from a population growth model and a restructured cross-platform blending algorithm.
 - Released several key product enhancements that drove client retention and new sales, such as demographics estimates by country and region.
 - Performed custom data analysis of active users estimates and communicated results to clients, sales team during quarterly reviews, as well as in response to ad-hoc client tickets.
- Amazon** Jan 2020 – May 2020
Applied Scientist Intern Cambridge, MA
- Developed machine learning baselines in Python and PyTorch for a semi-supervised natural language classification task in a streaming data setting.
 - Implemented self-training baseline using deep learning architectures (feedforward, LSTM) to prove SSL as a viable solution.
 - Researched cutting-edge algorithms such as consistency regularization and developed prototypes in an online deep learning setting via Hedge Backpropagation.
- WW International, Inc.** May 2019 – Aug 2019
Data Scientist Intern New York, NY
- Created word embedding representations for food-items from food journal data using spaCy and FastText.
 - Validated embedding performance using qualitative analysis on downstream tasks such as substitute food extraction, and refined results using in-house food ontology data.
 - Integrated SQL and Python code as an end-to-end data pipeline within the internal data science library; work was featured in *HealthRecSys 2019 Workshop*.

Technical Skills

Programming Languages: Python, Ruby, R, SQL, Java, C#, JavaScript
Data and Machine Learning: Scikit-Learn, PyTorch, NumPy, Pandas, Matplotlib, Google BigQuery, FastText, spaCy, MongoDB
Data Science & Miscellaneous Technologies: Predictive Modeling, Exploratory Data Analysis (EDA), Statistics, Databases, Git, APIs, CI/CD, LaTeX

Projects and Publications

- German-to-English Machine Translation | Python, PyTorch**
- Implemented sequence-to-sequence models such as LSTM with attention and transformers using PyTorch.
 - Achieved BLEU score of 34.2 (highest in class) using a 6-layer transformer model with 8 attention heads.
- Modeling Affect Intensity in Tweets – SemEval 2018 Task | Python, Scikit-Learn, Keras**
- Experimented with machine learning models (e.g. random forest, neural networks) using Scikit-Learn and Keras.
 - Achieved accuracy of 0.68 using a deep neural network trained on GloVe embedding features.