Lynn Samson

978-876-6196 | lnsamsn@gmail.com | linkedin.com/in/lynnsamson | github.com/lasamson

Education

University of Massachusetts Amherst — Amherst, MA
Master of Science in Computer Science — *Sep 2018 – May 2020*

University of Massachusetts Amherst — Amherst, MA
Bachelor of Science in Computer Science and Mathematics (Double Major) — *Sep 2014 – May 2018*

Experience

Data Scientist — Aug 2021 – May 2023
Sensor Tower — *San Francisco, CA*

- Developed statistical models in Ruby to analyze user demographics for mobile applications, storing results in a MongoDB database.
- Improved bias adjustment accuracy, cross-platform consistency, and productionized estimates at finer granularities to boost client retention and acquire new sales.
- Analyzed active user estimates and presented findings, including insights and trends, to address client queries and supported quarterly reviews.

Applied Scientist Intern — Jan 2020 – May 2020
Amazon — *Cambridge, MA*

- Developed a semi-supervised machine learning model for natural language classification in a real-time streaming environment using Python, PyTorch, and data pipelines.
- Optimized self-training baselines with feed-forward and LSTM architectures to showcase the benefits of semi-supervised learning for intent recognition.
- Prototyped and validated an online semi-supervised LSTM model, leveraging adaptive learning methods like hedge backpropagation and consistency regularization.

Data Scientist Intern — May 2019 – Aug 2019
Weight Watchers — *New York, NY*

- Developed word embeddings for food items from food journal data, testing multiple approaches and implementing FastText for subword embeddings.
- Preprocessed and structured data using SQL (BigQuery), Pandas, and spaCy. Evaluated embeddings via substitute food extraction and recipe recommendations.
- Built and deployed an end-to-end model pipeline within the company's internal data science library, streamlining food recommendation tasks. Project showcased at *HealthRecSys 2019 Workshop*.

Projects

Wildlife Species Image Classification | *Keras, Flask, Docker, Kubernetes, AWS, Git* — Feb 2025

- Developed a custom CNN with 5 convolutional layers and 3 dense layers, optimizing parameters such as learning rate, dropout, and batch normalization, achieving 85% accuracy on test data.
- Containerized model using Docker, deployed on AWS Lambda, and exposed as a web service via API Gateway.
- Packaged model for inference using TensorFlow Serving, implemented a Flask-based gRPC gateway, and deployed to an AWS EKS-managed Kubernetes cluster using Docker Compose.

Predicting Vaccination Status | *Pandas, Scikit-Learn, Flask, Docker, AWS, Git* — Dec 2024

- Developed and optimized multi-label classification models using XGBoost, Random Forest, and Logistic Regression, achieving a best mean AUC of 0.85 (within 1% of the top score).
- Deployed the XGBoost model as a Flask web service, containerized it with Docker, and deployed it to AWS Elastic Beanstalk, demonstrating end-to-end machine learning workflow.

Technical Skills

Languages: Python, Ruby, Java, SQL, JavaScript, HTML/CSS, R
Libraries and Frameworks: Pandas, NumPy, Matplotlib, Scikit-Learn, PyTorch, FastText, spaCy, NLTK, Flask
Developer Tools and Cloud: Git, Docker, Kubernetes, CI/CD, Google BigQuery, Amazon Web Services (AWS Elastic Beanstalk, AWS Lambda, AWS Elastic Kubernetes Service)