

Sam Sims

📍 Stoneham, MA | ☎ 860-480-7317 | ✉ samhsims@gmail.com | 🏠 sims-s.github.io

Professional Experience

NLP Data Scientist

September 2021 - Present

Workhuman

Framingham, MA

- Built recommendation system that suggests the value of a recognition moment. Expected to drive \$18M annual revenue by improving user experience and reducing funnel dropoff.
- Fine-tuned and deployed BERT in AWS Sagemaker to classify and detect spans of unconscious bias in messages. Model helped boost user engagement by 14% for 3.3M users.
- Pre-trained large and small language models with distributed training algorithms on hundreds of millions of tokens in AWS Sagemaker.
- Enabled retention of eight \$75M+ clients by conducting program usage and attrition analytics.
- Managed modeling datasets and routinely collaborated with annotators to label data & optimize annotation process.
- Created a generative model evaluation aid tool, which reduced LLM manual evaluation time by 36%.
- Trained tf-idf text classification model to detect recognition moments in Slack. Projected to capture 50K new events annually and generate \$2.75M revenue.
- Created a configurable experiment tracking pipeline integrated with AWS Sagemaker and S3. Managed 400+ models from six team members.

Data Scientist

June 2018 - September 2021

Travelers

Hartford, CT

- Reduced chat handle time by 16% by automating service reps' after chat work with BERT intent classification & T5 abstractive summarization models.
- Decreased summary correction time for chat reps by 22% by inventing method for discovering numeric hallucinations in generative model outputs.
- Automated 100+ hours of manual review each month by enabling Customer Experience team to discover issues systematically using clustering, semantic search, and sentiment models.
- Provided business partners data driven insights into customer callback and policy cancellation reasons by inventing model interpretability method.
- Trained and deployed transformer classification model for lowering call volume by encouraging customers likely to call to go online instead. Reduced call volume by 12%.
- Mentored four data science interns and three NLP data scientists.

Projects

Spice Rack AI Search Assistant

May 2024

- Designed & 3D printed spice rack integrated with real time voice controlled spice search.
- Developed automatic spice scanning pipeline using CV2 and custom PyTorch object detection model for spice detection and GPT-4 for OCR.
- Used RAG pipeline with FAISS to retrieve spices and GPT-4 tools to generate rack API calls (e.g. locating spices). Achieved 100% retrieval accuracy and 96% acceptable response generation rate.
- Implemented RAG based feedback system to adjust LLM behavior to match user's desired output.

Education

Master's of Science in Applied Mathematics

May 2018

Rensselaer Polytechnic Institute

Troy, NY

Skills

Languages & Tools

Python, Git, AWS, Docker, LaTeX

Concepts & Algorithms

Natural Language Processing, Deep Learning, Distributed Training

Machine Learning Packages

PyTorch, Tensorflow, Numpy, Huggingface Transformers, FAISS, Pandas