

MILI SHAH

206-591-6095

E-mail: milishah224@gmail.com

LinkedIn: linkedin.com/in/milishah224

GitHub: github.com/itsmilishah

EDUCATION UNIVERSITY OF MASSACHUSETTS AMHERST

Master's in Computer Science, September 2017 – May 2019

NIRMA UNIVERSITY

Bachelor of Technology, Computer Engineering, August 2013 – May 2017

EXPERIENCE AMAZON | Software Development Engineer

August 2019 - Present

- As part of the ProductGraph team, developing workflows that build a knowledge graph and use it to clean, correct and extract new information for supporting several use cases of Amazon Music.
- Developing and maintaining pipelines in Spark that run on AWS Elastic MapReduce and scale out to ingest, process and disambiguate billions of facts and millions of entities on a daily basis.

VIASAT | Software Engineering Intern

May 2018 – August 2018

- Developed a PyTorch based pipeline in AWS that provides real-time predictions about when multiple airplanes will fall under the coverage area of the same satellite, for data multicasting optimization.
- Resulted in projected savings of 500 GB in data transfer per month for in-flight WiFi services.

MORGAN STANLEY | Technology Analyst Intern

May 2016 – July 2016

- Developed a system to automate level-1 support for Java developers of Morgan Stanley.
- Achieved good qualitative performance, with implementation of the system in Python using NLTK, that clusters noisy e-mail queries to analyse topics and mines old discussions and wiki pages to answer new e-mails.

PROJECTS MACHINE READING COMPREHENSION QUESTION ANSWERING

February 2018 – May 2018

- Built models for Question Answering on SQuAD based on BiDAF and transformers in PyTorch. Fused linguistic information, obtained using spaCy, into the models.
- Achieved an F1 score of 72.14% by adding a dependency parse layer, implemented using a transformer, to BiDAF - an improvement over AllenAI's BiDAF model's score of 71.49%.

IMAGE TAG PREDICTION USING KNOWLEDGE GRAPHS

September 2018 – December 2018

- Developed models to incorporate semantic information into image tagging models by exploiting tag relations from Knowledge Graphs like ConceptNet.
- Implemented techniques that incorporate information from GloVe embeddings using transfer learning and Knowledge Graphs using multi-task learning.

COMPLEX EMBEDDINGS FOR UNIVERSAL SCHEMA

November 2017 – December 2017

- Developed a Knowledge Graph construction model that uses complex embeddings to capture asymmetry in rowless Universal Schema.
- Achieved an improved mean reciprocal rank of 0.33 in relation extraction, while joint learning from unstructured text and KBs, with implementation in Tensorflow.

GOOGLE: LARGE-SCALE COMMONSENSE AS LEXICAL ENTAILMENT

January 2018 – April 2018

- Performed crowdsourcing experiments in collaboration with Google using CrowdFlower aimed towards constructing a common-sense hypernym taxonomy.
- Constructed an Elasticsearch database of 205 million sentences to use in the experiments as questions.

COURSES

Natural Language Processing, Advanced Machine Learning, Deep Learning, Algorithms for Data Science, Reinforcement Learning, Computer Vision, Advanced Data Structures, Databases, Operating Systems

TECHNICAL SKILLS Languages: Python, Java, R, C, C++, Javascript

Modules/Frameworks: Tensorflow, PyTorch, Keras, spaCy, CoreNLP, Spark, Gensim, FastText, RDFLib, Elasticsearch, AllenNLP, Scikit-learn

Other: AWS, Hadoop, SQL