

MILI SHAH

248 Amherst Road, Apt. I-01
Sunderland, MA 01375
413-800-9382

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 MORGAN STANLEY | Technology Analyst Intern

May 2016 – July 2016

- Worked to automate level-1 support for java developers of Morgan Stanley
- Constructed a system in Python to cluster queries to analyse topics and to match new query e-mails with mined previous discussions and wiki pages

INFORMATION EXTRACTION AND SYNTHESIS LAB | Graduate Student Researcher

- Working with Prof. Andrew McCallum for multi-sentence relation extraction from biomedical text
- Previously worked on extending rowless universal schema LSTM model for automatic knowledge base completion using complex embeddings to capture asymmetry in relations
- Achieved a mean reciprocal rank of 33 with implementation in Python using Tensorflow

PROJECTS GOOGLE: LARGE-SCALE COMMONSENSE AS LEXICAL ENTAILMENT | Graduate Student Researcher

January 2018 – Present

- Working with Mr. Chris Welty and Prof. Lora Aroyo to construct a common-sense hypernym taxonomy from Microsoft Concept Graph, WordNet and crowdsourcing
- Use the constructed taxonomy to train partial-order structure preserving embeddings for better performance on downstream tasks like semantic web search and text understanding

CHARACTER IDENTIFICATION ON MULTI-PARTY DIALOGUES

September 2017 – December 2017

- SemEval 2017 task to build an efficient character identification system using supervised learning
- Achieved a mean precision of 71% in coreference resolution with agglomerative Convolutional Neural Nets implemented in Python using Tensorflow

STUDYING IMPACT OF INTERNATIONAL STOCK MARKETS ON INDIAN STOCK MARKETS

August 2016 – November 2016

- Built Support Vector Regression predictive models for stock markets in Python achieving a mean absolute error of 1.1%
- Performed a causality analysis study between different stock markets using these models

CLASSIFICATION OF MALWARE FILES

July 2015 – November 2015

- Built a system in Python using NLTK, scikit-learn modules to classify malware files
- Achieved an accuracy of 91.5% using weak Decision Tree learners with AdaBoost ensembling

SHUTTERING PLATES MANAGEMENT SYSTEM

October 2014 – Dec 2014

- A desktop application of stock management developed for a business
- Implementation in Java using Swing with MySQL database

COURSES Machine Learning, Neural Networks, Algorithms for Data Science, Natural Language Processing, Big Data Analysis, Advanced Data Structures, Database Management Systems, Operating Systems

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

Module Familiarity: Tensorflow, Keras, Scikit-learn, Numpy, Pandas, Spacy, CoreNLP, NLTK, Gensim, FastText, RDFLib

Databases and Big Data Frameworks: SQL, MongoDB, Hadoop