Yash Kumar Lal

ylal@cs.stonybrook.edu
(443)-207-3261

Education

Stony Brook, NY **Stony Brook University** **August 2020 –**

- PhD in Computer Science. Advisor: Niranjan Balasubramanian. Research Area: Question Answering

Baltimore, MD **Johns Hopkins University** **August 2018 – May 2020**

- M.S.E. in Computer Science. Thesis: Low resource problems in NLP. Advisor: Philipp Koehn. CGPA: 3.94
- Grad Courses: Cloud Computing; Comp. Psycholinguistics; Machine Translation; Semantics.

Manipal, India **Manipal Institute of Technology** **Aug 2014 – June 2018**

- B.Tech. in Computer Science Engineering. CGPA: 3.34

Employment

Research Intern **Salesforce Inc.** **May 2022 – Aug 2022**

- Mentors: Semih Yavuz, Ye Liu, Yingbo Zhou
- Develop global question decomposition model to improve multi-hop question answering

NLP Engineer Intern **ThreatLandscape** **Jan 2018 – June 2018**

- Created auto-annotated data for relation extraction of cyber entities using weak supervision. (Python, Prodigy)
- Improved precision of the MVP of the company - an NLP engine for threat actor analysis - by 3 points (PyTorch)

Research Intern **IIIT-Hyderabad** **May 2017 – July 2017**

- Extended previous work to enhance neural embeddings for polysemous words (Python, Word2Vec)
- Created more fine-grained representations than Word2Vec

Technical Experience

Publications

- **TellMeWhy: A Dataset for Answering Why-Questions in Narratives**. Findings of ACL-IJCNLP, 2021.
- **IrEne-viz: Visualizing Energy Consumption of Transformer Models**. EMNLP Demo, 2021.
- **IrEne: Interpretable Energy Prediction for Transformers**. ACL-IJCNLP, 2021.
- **Temporal Reasoning in Natural Language Inference**. Findings of EMNLP, 2020.
- **Sentence-Level Adaptation for Low-Resource Languages**. LoResMT workshop, MT Summit, 2019.
- **De-Mixing Sentiment from Code-Mixed Text**. ACL Student Research Workshop, 2019.
- **Johns Hopkins University Submission for WMT News Translation Task**. WMT, 2019.
- **Identifying Clickbait: A Multi-Strategy Approach Using Neural Networks**. SIGIR, 2018.
- **SWDE: A Sub-Word And Document Embedding Based Engine for Clickbait Detection**. Computational Surprise Workshop, SIGIR, 2018.

Projects

- **Ping** - A smart platform for language independent communication between businesses and their customers. Uses AI to present messages to users, in their preferred language. (Swift 3, Python, Redis, Travis CI, AWS)
- **hello friend** - SMS service utilising natural language processing to bring essential smartphone features to simple feature phones. (Python, Heroku)
- **simmrr** - New metric for Microsoft MSMARCO tasks using Microsoft Gen Encoder (Python)

Additional Experience and Awards

- **Winner, Acceleprise Award** at AngelHack Global Demo Day ’17, San Francisco; AngelHack Hyderabad ’17
- **Top 10 in India:** Microsoft code.fun.do National Showcase, 2017
- **Chair**: NAACL 2022 Reproducibility Track
- **Organizer:** Commonsense Representation and Reasoning Workshop, ACL 2021
- **Reviewer:** ECIR ’19, ’21; WMT ’19; TALLIP (’20); EMNLP ’21-’22; ACL Rolling Review Nov ’21 - present
- **Program Committee:** ACL SRW ’20-’21; AACL SRW ’20; NAACL ’21; NAACL SRW ’21