1 E University Pkwy, #1302
Baltimore MD 21218

Yash Kumar Lal

(443)-207-3261
yash@jhu.edu
https://www.cs.jhu.edu/∼ylal1/

Education

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

- M.S.E. in Computer Science. GPA: 4.0. Thesis: Low resource problems in NLP. Advisor: Philipp Koehn.
- 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. GPA: 3.34.

Employment

Head Teaching Assistant **Johns Hopkins University** **September 2018 – Present**

- Courses: Object Oriented Software Engineering, Gateway Computing: Python.

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

Summer Intern **Snapdeal India** **May 2016 – July 2016**

- Performed credit analysis and risk evaluations to help SMEs obtain loans for Snapdeal's online platform (R)
- Work was used as preliminary check to lend money to over 20,000 sellers to start their business

Technical Experience

Publications

- **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 Engine**. SIGIR, 2018.
- **SWDE: A Sub-Word And Document Embedding Based Engine for Clickbait Detection**. Computational Surprise Workshop, SIGIR, 2018.

Projects

- **Optimizing Job Scheduling in Datacenters** - Built a simluation of Mesos architecture and experimented with simple machine learning techniques to affect job scheduling in data centers. (Python, SocketIO)
- **Ping** - A smart platform for language independent communication between users, businesses and their customers. It uses artificial intelligence to get all your messages to you, in the language you prefer. (Swift 3, Python, Redis, Travis CI, AWS, Postman)
- **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)

Languages and Technologies

- Python; C++; C; Java; Objective-C; SQL; Swift; R; Shell
- Fairseq, Sockeye, PyTorch; SpaCy, FastText; NumPy, scikit-learn
- Heroku; Javalin; Travis CI; Version Control; Postman

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
- **Founder:** MUPy, Manipal's Python developers conference, in association with Python Software Foundation