

Simmi Mourya

linkedin.com/in/simmi-mourya-34406886

simmimourya.github.io | simmim@seas.upenn.edu
Philadelphia, PA 19104 | Mobile : +1-267-968-8186

EDUCATION

- **University of Pennsylvania** Philadelphia, PA
Master of Science in Computer and Information Science, GPA: 3.74/4.0 **Graduating May 2021**
- **Coursework:** Analysis of Algorithms, Internet & Web Systems, Software Systems (Audit), Advanced Machine Perception, Comp. Linguistics, Computer Vision, Machine Learning. **Teaching Assistant:** CIS 581- Computer Vision, C++ Programming
- **Cluster Innovation Center, University of Delhi** Delhi, India
Bachelor of Technology in Information Technology, GPA: 8.2/10 Aug. 2013 – July 2017

SKILLS

- **Software:** Java, Apache: (Spark, Storm, Bench), Oracle BDB, PHP, HTML/CSS, Javascript, SQL, MongoDB, Jenkins
- **Research:** Python, PyTorch, FastAI, Keras, Scikit-Learn, Numpy, Pandas, Caffe, Cython, Python/C API, MATLAB
- **Accepted Talks:** "Scientific Computing using Cython: Best of both worlds!" - EuroScipy '18, Europython '17, Pycon India '17

SOFTWARE PROJECTS

- **Search Engine: Goal:** To have a functioning, reasonable search engine which retrieved relevant pages. Create meaningful indexes and page rank scores for all the webpages crawled.
Built a scalable web crawler hosted on Amazon AWS complete with a crawler, indexer, pagerank, and a front end. Worked majorly on developing a **TF-IDF** and **Map-Reduce** Based Indexer based scalable Indexer. Also worked on DevOps for Gradle, EMR, Hadoop, EMRFS and minor Hadoop DevOps for PageRank. (Team size: 4)
- **Multi-threaded web server and Service framework:** A Java based web **HTTP 1.1 compliant web server** developed from scratch. Later merged it with a custom-built web service framework which emulates the behaviour of **Java Spark**. Services implemented: **Route registration, Session/Cookie management**, Filter handler, Request and Response handlers.
- **Multithreaded Web crawler:** Developed a **multithreaded web crawler** with a custom XPath Parser and to query and store matched HTML, XML documents into a persistent data store.

EXPERIENCE

- **University of Pennsylvania** Philadelphia, PA
Graduate Research Assistant May 2020 - Present
 - **Multimodal Question Answering framework:** Working on intersection of NLP and Computer Vision. Developing a novel task framework for Goal-Step inference and Step membership inference using multimodal Wikihow data. Also contributing towards extending the tasks for 1 million HowToVideos dataset.
- **ESRI** Delhi, India
Software Developer May 2019 - July 2019
 - **ArcGIS Python API:** Developed framework for **Multispectral support for Pixel classification** in **ArcGIS** Python API. Developed Pyramid scene parsing backbone support of **object segmentation** for the API.
 - **Spatial Dataframes:** Optimized validation checks in `arcgis.geometry` package using pre-compiled Cython binaries. This processes **0.1 million entries in less than 2 ms**, which earlier took **45-55 ms**.
- **IIIT Delhi** New Delhi, India
Research Associate Feb 2018 - March 2019
 - **Article:** Mourya, S., Kant, S., Kumar, P., Gupta, A. and Gupta, R., 2018. LeukoNet: DCT-based CNN architecture for the classification of normal versus Leukemic blasts in B-ALL Cancer.
 - **Accepted Challenge:** Classification of Normal versus Malignant Cells in B-ALL White Blood Cancer Microscopic Images, challenge selected at IEEE ISBI '19, Venice, Italy.
- **Predible Health** Bangalore, India
Software Developer August 2017 - December 2017
 - **Development:** Developed **U-Net** based framework for Lung nodule segmentation from 3D CT scans (LIDC-IDRI dataset) Also developed classifiers to analyze nodule level malignancy and emphysema. Built POC for identifying cancerous lung nodules from Radiomics data. Streamlined prototyping and testing via **parallelization** of the data pre-processing pipeline.
- **Google Summer of Code** Portland State University
Software Developer Intern May 2016 - August 2016
 - **Cyvlfeat:** Designed and developed 12 new features for a **high-performance Python/Cython wrapper** of computer vision library, VLFeat. Emulated the wrapper from **MATLAB MEX** scripts. (Added algorithms specializing in image understanding and local features extraction and matching such as LBP, SIFT, hierarchical k-means, SLIC).
 - **Continuous Integration and Tests:** Extensively used **Continuous Integration platforms** such as Jenkins, Travis. **Built unit and integration tests** using Python's Nose test suite.

RESEARCH PROJECTS

- **Computer Vision:** Mask-RCNN from scratch: Built an **attention mechanism** in form of **Region Proposal network (RPN)** for Object detection task. Implemented **vectorized ROIAlign** for FPN-ROI Mapping.
- **Computational Linguistics:** Developed **Bilingual Named Entity Recognition** using Bi-LSTM CRF and Self Attention.