Simmi Mourya

simmimourya.github.io | **simmim@seas.upenn.edu**
linkedin.com/in/simmi-mourya-34406886 Philadelphia, PA 19104

Education

- **University of Pennsylvania** Philadelphia, PA
  Master of Science in Computer and Information Science, GPA: ***3.74/4.0*** ***Graduating May 2021***
- **Coursework**: Internet & Web Systems, Advanced Machine Perception, Comp. Linguistics, Computer Vision, Machine Learning
  Teaching Assistant: Computational Linguusitics, 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, Nose, Jenkins, ArcGIS
  Research: Python, PyTorch, FastAI, Keras, Scikit-Learn, Numpy, Pandas, Caffe, Cython, Python/C API

Software Projects

- **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.
- **Web crawler and Search Engine**: Developed a **multithreaded web crawler** with a custom XPath Parser and to query and store matched HTML, XML documents into a persistent data store. Also developed a **Map-Reduced Based Indexer** Developed and scaled Hadoop based Indexer. DevOps for Gradle, EMR, Hadoop, EMRFS. Minor Hadoop DevOps for PageRank.

Experience

- **University of Pennsylvania** Philadelphia, PA
  Graduate Research Assistant *May 2020 - Present*
  - **Multimodal Question Answering framework**: Advised by Prof. Chris Callison Burch. Developing a novel task framework for Goal-Step inference and Step membership inference using multimodal Wikihow data.
- **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.
- **Spark Infosystems** New Delhi, India
  Full Stack Developer Intern *May 2015 - August 2015*
  - **Job-seeker-Employer Platform**: Extended functionality of an **'Indeed' like platform** by adding the Job-seeker and custom Test-series module. Helped automate hiring by redesigning a **customized self-evaluation** framework for job-seekers.

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.

  Developed **YOLO** pipeline (end-to-end) for **object detection**, with a **Non Maximum Suppression** post-processing module.
  Computational Linguistics: Developed **Bilingual Named Entity Recognition** using Bi-LSTM CRF and Self Attention.