Kushagra Mahajan

kmahajan@cs.cmu.edu +1-412-214-2036 mahajan-kushagra Kushagra Mahajan kushagramahajan.me

Education

Carnegie Mellon University, School of Computer Science **Aug. 2021 – Present**
Master of Computational Data Science — GPA: 4.06/4.00 *Pittsburgh, PA*

- *Coursework*: Machine Learning, Machine Learning for Large Datasets, NLP, Cloud Computing, Visual Learning and Recognition, Multilingual NLP, Multimodal ML
- ***Capstone Project*** *(Advisor: Prof. Yonatan Bisk)*: **Amazon Alexa Prize Simbot Challenge** – Building state-of-the-art systems and pipelines for navigation, localization, mapping and interaction of a virtual assistant bot with the Arena environment based on user-assigned real-world tasks in collaboration with **Amazon Alexa AI**.

Indraprastha Institute of Information Technology Delhi (IIIT Delhi) **Aug. 2014 – Dec. 2018**
Bachelor of Technology in Computer Science — GPA: 8.78/10.00 *Delhi, India*

- Coursework: Computer Vision, Probabilistic Graphical Models, Data Mining, Collaborative Filtering, Machine Learning (Teaching Assistant - Fall '18), Data Structures and Algorithms, Compiler Design, Database Systems, Operating Systems.

Experience

Amazon **May 2022 – Aug 2022**
Software Engineering Intern — Installments Team *Seattle, WA*

- Improved web page notification system for payment failures to be more descriptive and provide rectification steps to customers. Reduced customer tickets regarding payment failures by 72%.

Tata Research and Innovation Labs **Feb. 2019 – April 2021**
Machine Learning Engineer — Computer Vision Team | Advisors: Dr. Lovekesh Vig & Dr. Gautam Shroff *Delhi, India*

- Designed a meta-learning based framework for skin lesion and chest x-ray classification, and segmentation of medical and natural scene images. Published **3 papers** and built **2 products** for chest x-ray analysis, and skin lesion detection.
- Explored disentangling biological signals from noise in cellular images and achieved top-5 percentile in NeurIPS 2019 challenge with test accuracy 96.06%. Also worked on distributed training, abnormality detection in X-rays using GANs, visual attention.
- Built an end-to-end alignment and information extraction system for document images using a novel keypoint extraction algorithm. Product sold to the **Landmark Group**. Published **1 paper** and filed **1 US patent**.

Intel Corporation **Aug. 2017 – Dec. 2017**
Machine Learning Intern | Advisor: Mr. Tigi Thomas *Bangalore, India*

- Created a highly optimized sensor-based gesture detection and recognition model for on-device deployment surpassing benchmarks for memory constraints and output latency. Tested by deploying model on laptop hardware.

CVML Lab, IIIT Delhi **Aug. 2016 – Dec. 2018**
Machine Learning Intern | Advisor: Prof. Chetan Arora *Delhi, India*

- Used texture descriptors to improve clothing segmentation by 3% for **visual fashion image and attribute search** systems.
- Exploited the pose structure to enhance SoA fine-grained classification performance by 2-3% across standard FGVC datasets. Curated an Amazon pose-aware apparel dataset. Published **2 papers**. Work was in collaboration with **Staqu Technologies**.

Projects

Twitter Cloud Native Web Service with Microservices **Spring 2022**
Course Project: Cloud Computing | Advisor: Prof. Majd Sakr

- Created an ETL pipeline using Spark for processing 1.2TB Twitter data, and a microservice based architecture using Kubernetes for data retrieval and running analytic jobs in a cost constrained setting.

Multimodal Image to Recipe Generation **Fall 2022**
Course Project: Multimodal ML | Advisor: Prof. Louis-Philippe Morency

- Recipe generation from food images using a novel approach comprising multimodal alignment before fusion (ALBEF), CLIP for improved image encoding and transformer for recipe generation.

Natural Language Inference for Code-Switched Hinglish **Spring 2022**
Course Project: Multilingual NLP | Advisor: Prof. Alan Black

- Improved state-of-the-art NLI performance on GLUECoS benchmark by 6% for code-switched Hindi-English by translating to matrix or embedded language, and adaptation of language models to the code-switched domain using PyTorch. [Report]

Skills & Achievements

Programming Languages, FrameWorks and Tools: Python, C, C++, Java, SQL, Matlab, R, Tensorflow, Pytorch, Keras, PySpark, Caffe, OpenCV, Scikit, NumPy, Pandas, SciPy, Kafka, Samza, HBase, MongoDB, AWS, Azure, Kubernetes
Achievements: *Travel Grant*: AICTE-INAE for ICIP 2018, *Dean's List*: 2017-2018, Teaching Fellow at Vivekananda Kendra, Delhi.