KSHITIZ SHARMA

E-mail: kshitiz.sharma38@gmail.com
Mobile: (+91) 9717013056
LinkedIn: https://in.linkedin.com/in/sharmakshitiz
GRE : 324 + 4, TOEFL : 111 (2019)

RESEARCH INTERESTS

- Deep Learning
- Computer Vision
- Machine Reading Comprehension
- Neural Machine Translation
- DevOps and MLOps
- Reinforcement Learning

EDUCATION

INDIAN INSTITUTE OF TECHNOLOGY DELHI (IITD) **| B.Tech. in Mathematics and Computing, Delhi, India**, 2018

D.A.V. SCHOOL **| All India Senior School Certificate Examination, Kota, India**, 2014

MAYOOR SCHOOL **| All India Secondary School Examination, Ajmer, India**, 2012

PROFESSIONAL EXPERIENCE

VMOCK INDIA PVT. LTD., Gurgaon, India **2018 - Present**

Principal Data Scientist (July 2021 – Present)

- Led all Data Science teams at VMock across all business verticals contributing to Technical Management and R&D processes. Transformed the organization structure of Data Science teams, implemented Monorepo and MLOps practices using MLFLow and DVC.
- Managed research team working on standardizing SOTA research to train and make large scale transformer models productionable using onnxruntime, deepspeed integrations, mixed precision training, quantization and gradient accumulation. Subsequently adopted these models across various NLP problems within the company.

Senior Data Scientist (December 2019 – July 2021)

- Led a team of 3 people overseeing all reviews, development, improvements, deployment and research projects for the Resume Parser Product.
- Tasked with developing a multilingual parser capable of parsing resumes in French Language. Solved the Reverse Translation and Token Alignment problem using unsupervised IBM Model 1 and HMM models like Berkeley Aligner and string-matching heuristics using Levenshtein distance. Worked on developing an in-house Joint Machine Translation and Token Alignment system using MBart Transformer based encoder-decoder model.
- Deployed production ready prediction models with TensorFlow Serving and implemented Async. Task Queue pipeline using Celery and AWS SQS.

Data Scientist (July 2018 – December 2019)

- Tasked with improving and maintaining a robust resume parser that required converting unstructured data into structured data and extracting useful information using techniques of NLP.
- Trained in-house BiLSTM CRF based NERs with Character Embeddings and Attention Mechanism delivering high recall and precision for recognizing various entities as positions, organizations, locations, universities, degrees, section titles etc.

RESEARCH EXPERIENCE

AUTONOMOUS OBJECT TRACKING WITH UAVS, New Delhi, India **2018 - 2019**

B.Tech Project supervised by Prof. Subhashis Banerjee at IIT Delhi

- Developed an autonomous agent capable of tracking an object using techniques of Computer Vision and Deep Reinforcement Learning. Implemented Q-Learning networks like DQN/Double DQN with Experience Replay and

simulated the task with a drone tracking a car on the ground moving in simple motions, governed by real world like physics. Deployed a modified version of Microsoft AirSim Simulator for training and testing.

OBJECT DETECTION AND TRACKING FOR DRIVERLESS CARS, New Delhi, India **2016**

Research Project supervised by Prof. Subhashis Banerjee at IIT Delhi

- Tackled the problem of object detection and tracking using machine learning algorithms. We explored slower region proposal based classification models (R-CNN and its variants - Fast, Faster R-CNN), and faster end-to-end regression based models (such as YOLO) for detection. Subsequently, leveraged detection models for feature extraction and generating detection heatmaps along with LSTMs (Long Short Term Memory) to capture temporal information in sequences/videos for robust tracking and occlusion handling.

INTERNSHIPS

ARTIFACIA INC., Bangalore, India **2017**

Research Intern

Media Engagement Prediction

- Devised a framework for predicting the engagement (number of likes) over time, given an Instagram user handle and a query image. Using image features from a ConvNet and hand-crafted features from User Profile (Follower's count, Average Engagement, tagged users, day posted), created the entire pipeline systematically; from data collection, data preprocessing to training the model, evaluation, and deployment for a real-time production system.

Visual Recognition (Multi-Label Classification)

- Implemented a Machine Learning algorithm to identify all the tags/labels relevant to a given image. Implemented an LSTM-based recurrent model for modelling label dependencies and label prediction and a VGG16 ConvNet for extracting image features (Trained on NUS-WIDE dataset). The model was capable of identifying both the co-occurrence of labels, and the attention regions in images. Used Beam-Search algorithm for decoding at test time.

KEY ACADEMIC PROJECTS

PARALLEL PROGRAMMING ASSIGNMENTS, New Delhi, India **2017**

Course taught by Prof. Subodh Kumar at IIT Delhi

- Implemented LU decomposition with OpenMP and OpenMPI(shared and distributed memory model) [Github].
- Implemented parallel Merge Sort with Cuda Programming. [Github].
- Worked on Parallel DBSCAN algorithm for density-based clustering (OpenMP/MPI implementation).

OPERATING SYSTEMS ASSIGNMENTS, New Delhi, India **2017**

Course taught by Prof. B.S. Panda at IIT Delhi

- Designed an IO library supporting read/write/append modes with buffering using UNIX system calls. [Github] Worked on a client/server-based chatting system that uses UNIX message queue for message passing.
- Performed copy using n processes to share the load of copying a file using semaphores for synchronization.

VQA - VISUAL QUESTION ANSWERING [GITHUB], New Delhi, India **2016**

Course project supervised by Prof. Subhashis Banerjee at IIT Delhi

- Worked on the task of open-ended VQA, that is given an image and a natural language question about it, the task is to generate an accurate natural language response. Used CNN(VGG16/InceptionV3) for extracting image features and encoded questions and answers using a 1-layer LSTM based recurrent network over the word embeddings.

COMPUTER VISION ASSIGNMENTS, New Delhi, India **2016**

Course taught by Prof. Subhashis Banerjee at IIT Delhi

- Executed Lucas-Kanade for optical flow estimation and deployed it for stabilizing video.
- Implemented single-view metrology-based algorithm for comparing heights of two objects in the same vertical plane in an image using cross ratio invariance of projective transformations. [Github].
- Worked on scalable image retrieval system using hierarchical vocabulary tree with SIFT-Descriptors.

TECHNICAL SKILLS

Programming Languages:	Python, C, C++, CUDA, Java, OpenMP/MPI, Golang, OCaml
Machine Learning:	Tensorflow & Keras, PyTorch, Hugging Face, OONX, Deepspeed, Caffe, Darknet
Development Tools:	MLFlow, DVC, Celery, Redis, SQS, Mongo, Apache, SQLAlchemy, AWS Console, Git
Web Programming:	PHP, HTML, CSS, JavaScript, Laravel, Jekyll

KEY ACCOMPLISHMENTS

2014 - All India Rank 485 in IIT - Joint Entrance Examination (amongst 1.4 million candidates)
2012 - Secured 11th position in Rajasthan's State Science Talent Search Examination (SSTSE)
2010 - Awarded the prestigious National Talent Search Examination (NTSE) Scholarship

POSITIONS OF LEADERSHIP AND RESPONSIBILITY

INDIAN INSTITUTE OF TECHNOLOGY (IIT) DELHI, New Delhi, India **2016 - 2017**

Web Management Coordinator

- Web Management Coordinator, Alumni Affairs & International Programmes (AAIP).
- Developed and maintained IIT Delhi's AAIP's website and portal, which handled more than 8000 alumni, students, and faculty registrations by the end of tenure. Also contributed in organizing AAIP events as a Coordinator.

INDIAN INSTITUTE OF TECHNOLOGY (IIT) DELHI, New Delhi, India **2015 - 2016**

House Fine Arts Club (FAC) Representative

- Managed and Conceptualized various events (House Day, BHM Night), and various competitions (inter/intra house). Held workshops for juniors, managed events for the FAC during cultural festivals, at the university level as FAC Representative. Our house finished in the top-3 positions at two events.

EXTRA-CURRICULAR ACHIEVEMENTS

2015 - Our 5-member team stood 1st in Split Painting competition at IIT Delhi; Theme: Paradigm Shift.
2015 - Our 5-member team stood 4th in Wall Painting competition at IIT Delhi.
2012 - 1st position - Documentary making Competition on theme Life in Mayoor' at Mayoor School.
2011 - 1st position - Product Advertisement Making Competition at Mayoor School Cultural Festival.
2010 - Received Best Artist in Secondary School award at Mayoor School Cultural Festival.

COURSEWORK DONE RELEVANT TO RESEARCH INTERESTS

- MTL100: Calculus
- MTL101: Linear Algebra & Differential Equations
- MTL102: Differential Equations
- MTL103: Optimization Methods & Applications
- MTL104: Linear Algebra & Applications
- MTL105: Algebra
- MTL106: Probability & Stochastic Processes
- MTL107: Numerical Methods & Computation
- MTL180: Discrete Mathematical Structures
- MTL342: Analysis & Design on Algorithms
- MTL390: Statistical Methods
- MTL458: Operating System
- MTL783: Theory of Computation
- MTL782: Data Mining
- ELL305: Computer Architecture
- COL106: Data Structures and Algorithms
- COP290: Design Practices in Computer Science
- COL334: Computer Networks
- COL730: Parallel Programming
- COL780: Computer Vision

LINKS

86 Stars : 35 Forks | github.com/ktzsh/object tracking
36 Stars : 16 Forks | github.com/ktzsh/autonomous drone for tracking
06 Stars : 03 Forks | github.com/ktzsh/AirSim (Forked from Microsoft/AirSim)
03 Stars : 04 Forks | github.com/ktzsh/merge_sort_cuda
01 Stars : 01 Forks | github.com/ktzsh/DeepRLTensorflow (Inspired from tensorpack/tensorpack)