SUDHANSHU MISHRA

sudhanshu.ktj@gmail.com ghostktjMactavish

EDUCATION

Indian Institute of Technology, Kanpur 2017 - present
Master of Technology, Mechanical Engineering CGPA : 9.7/10
Bachelor of Technology, Mechanical Engineering CGPA : 8.4/10
ISC 2016, BHS Allahabad, Class XII 94.8%
ICSE 2014, BHS Allahabad, Class X 93.2%
Minor in English Literature

ACADEMIC HONORS

Selected for poster presentation at EEML Summer School. 2021
Won 2nd prize and prize money of $750 in NYU CSAW Hack3D competition 2021
Selected for student grant to attend FIRE 2019 and present paper. 2019
Awarded the Summer Research Fellowship by INSA, NASI and IAS. 2019
Secured All India Rank 3680 in IIT-JEE Advanced among 170,000 students. 2017
Secured All India Rank 7783 in IIT-JEE Mains among 1,200,000 students. 2017

EXPERIENCE

Machine Learning Engineer Aug 2022 - present
Vmock India Pvt. Ltd

- Led complete refactoring of Samples Service to OpenSearch, implementing transformer-based similarity matching and complex hybrid retrieval, resulting in 10x faster response times and 80% improved relevance in A/B testing
- Singlehandedly maintained and optimized entire Skills codebase for 12 months, reducing startup time from 250s to 90s and runtime latency by 2x while decreasing memory usage from 1.4GB to 90MB, leading to significant cost savings
- Architected and implemented new skills detection service utilizing contextual splitting, FAISS-based retrieval, and ColBert re-ranking, achieving superior recall and precision compared to existing pipeline while maintaining production latency requirements
- Developed and deployed enterprise-wide GenAI infrastructure, including LLM guardrailing framework, monitoring systems, and fine-tuned models that outperformed GPT-3.5/4 (78% vs 32% accuracy) for domain-specific tasks
- Created an LSH-based large-scale data deduplication pipeline, reducing data cleanup time from days to minutes, which was later adapted into a plagiarism detection product Optimized multiple critical services: reduced candidate generation pipeline from 120s to 4s, cover letter analysis from 10s to 2s, and skills rating service from 10s to 2s through batched inference and algorithmic improvements
- Designed and implemented automated testing framework covering 12 different endpoints, significantly reducing manual validation time while increasing test coverage and reliability
- Led development of agentic framework for large-scale data analysis, outperforming single/few-shot approaches from leading models, enabling rapid core data release for new client
- Created efficient smart compose framework progressing from n-gram models to fine-tuned small LLMs using SFT and DPO, implementing speculative decoding for real-time suggestions
- Mentored team members on ML model optimization, reducing training times by 3-6x, and led monthly learning sessions on advanced topics like Flash Attention, Speculative Decoding, and Agentic Workflows

INTERNSHIPS

Summer Research Fellowship Programme - Indian Statistical Institute, Kolkata May 2019- July 2019
Undergraduate Summer Research Program - NYU Tandon School of Engineering June 2021- Aug 2021

PUBLICATIONS

Goswamy, T., Parmar, N., Gupta, A., Shah, R., Tandon, V., Goyal, V., Gupta, S., Laud, K., Gupta, S., **Mishra, S.**, Modi,A. AI-based Monitoring and Response System for Hospital Preparedness towards COVID-19 in Southeast Asia at Healthcare Systems, Population Health, and the Role of Health-Tech (HSYS 2020), ICML 2020 Workshop. [link]

Mishra, S., Prasad, S., & Mishra, S. (2020). Multilingual Joint Fine-tuning of Transformer models for identifying Trolling, Aggression and Cyberbullying at TRAC 2020. In Proceedings of the Second Workshop on Trolling, Aggression and Cyberbullying (TRAC-2020). [link]

Mishra, S., & **Mishra, S.** (2019). 3Idiots at HASOC 2019: Fine-tuning Transformer Neural Networks for Hate Speech Identification in Indo-European Languages. In Proceedings of the 11th annual meeting of the Forum for Information Retrieval Evaluation. [link]

Mishra, Sudhanshu, Shivangi Prasad, and Shubhanshu Mishra. 2021. Exploring Multi-Task Multi-Lingual Learning of Transformer Models for Hate Speech and Offensive Speech Identification in Social Media. SN COMPUT. SCI. [link]

Mishra, S. and **Mishra, S.** (2020). Scubed at 3C task a - a simple baselinefor citation context purpose classification. InProceedings of the 8th Inter-national Workshop on Mining Scientific Publications, pages 5964, Wuhan,China. Association for Computational Linguistics[link]

PROJECTS

Intelligent Railway Platform Scheduling June 2021 - June 2022
M.Tech Thesis - Under Prof. N. S. Vyas

- Developed a real-time train scheduling system using Reinforcement Learning and Mixed Integer Linear Programming to optimize platform allocation at complex railway stations
- Designed and implemented a station simulator to model train routing dynamics and evaluate scheduling algorithms, reducing computational overhead by 98% compared to traditional methods
- Created a Q-Learning agent that outperformed baseline algorithms, reducing train delays by up to 60% compared to deterministic scheduling approaches
- Validated algorithms using real data from Kanpur Central Station (150+ daily trains, 10 platforms), demonstrating significant improvement in handling delayed train scenarios

Parallel Programming Jan 2022 - May 2022
CS433 Assignment - Under Prof. Mainak Chaudhuri

- Parallel OpenMP implementation for TSP and linear equation solver
- Implemented barriers and locks using OpenMP and POSIX threads
- Implemented GPU based parallel matrix vector product and Gauss-Seidel Iterative solver

Cybersecurity in Additive Manufacturing June 2021 - Aug 2021
NYU SURP - Under Prof. Nikhil Gupta

- Researched threat vectors and viable PUFs like microstructure, volume fraction as security feature for AM.
- Attended lectures on state of the art technology in Additive Manufacturing and Cybersecurity.
- Worked on identification of faulty G-code files and embedded security codes in CAD designs.
- Reverse Engineered CAD designs from representations like STL, Point Clouds, Multi-view.

Falsification of Cyber-physical Systems using Deep Reinforcement Learning Jan - April 2021
CS659 Course Project

- Falsification of Simulink Models using Deep Reinforcement Learning
- Used Double DQN and A3C algorithms for Falsification from ChainerRL library.
- Used S-Taliro for robustness calculation.

- Re-implementation of Falsification of Cyber-physical Systems using Deep Reinforcement Learning paper by T. Akazaki et. al.
- Code: https://github.com/ghostktjMactavish/CS659_Project

Parallelization of Lid Driven Cavity flow using OpenMP Jan - April 2021
PHY690 Course Project

- Wrote Lid-Driven Cavity flow simulation in python and C++ for 2D and 3D case.
- Parallelized the Cavity flow simulation using OpenMP, OpenACC.
- Performed strong and weak scaling analysis for varying grids and processors
- Code: https://github.com/ghostktjMactavish/PHY690Code

Simple baseline for citation context purpose and influence classification June 2020
Joint Work with Shubhanshu Mishra

- Used Random Forests, Gradient Boosted Trees and Multi-Layered Perceptrons for identifying citation context.
- Achieved 2nd position in influence task and 3rd position in the purpose task
- Systems Paper Published for 8th International Workshop on mining scientific publications, 3C Shared Task.

AI-based Monitoring and Response System for Hospital Preparedness towards COVID-19 in Southeast Asia April 2020 - June 2020
With IITK Consulting Group

- COVID-19 monitoring and response system to identify the surge in the volume of patients at hospitals and shortage of critical equipment like ventilators in South-east Asian countries
- Paper Published in HSYS 2020 Track of International Conference of Machine Learning 2020 Workshop.

Trolling, Aggression and Cyber-bullying Content Identification March 2020 - April 2020
Joint work with Shivangi Prasad and Shubhanshu Mishra

- Used fine-tuning of multi-lingual transformer neural network with joint training training on different languages and joint labelling of multiple tasks. Trained Models available at huggingface.co/socialmediaie
- Our approach came 2^{nd} and 4^{th}in 2 sub-tasks and 3^{rd} in 4 of the 6 sub-tasks.
- Code: github.com/socialmediaie/TRAC2020

Hate Speech Identification June 2019 - May 2020
Joint work with Shubhanshu Mishra

- Used fine-tuning of multi-lingual transformer neural networks, multi-task models and Back Translation.
- Among the 8 shared tasks, our solution won the 1^{st} place for 3 tasks. Additionally, it was within the top 5 for 7 of the 8 tasks, being within 1% of the best solution for 5 out of the 8 sub-tasks.
- Published in the Proceedings of the 11th annual meeting of the Forum for Information Retrieval Evaluation. Extension of this work has been submitted to "Springer Nature Computer Science" journal.
- Invited for an oral presentation at Forum for Information Retrieval Evaluation 2019
- Code: github.com/socialmediaie/HASOC2019

Open World Classification using Adversarial Auto-Encoders May 2019 - July 2019
Summer Internship Project (SRFP) under the guidance of Prof. Nikhil R. Pal (ISI Kolkata)

- Worked on Open World Problem of handling data which does not belong to any known class of the classifier.
- Experimented with various types of Generative deep learning algorithms like Vanilla GAN, W-GAN and Adversarial Autoencoders to develop a method to detect open world samples using the MNIST dataset.
- Code: github.com/ghostktjMactavish/Adversarial-Autoencoders

SKILLS

Programming Languages	C, C++, Python, MATLAB, Julia
Libraries and Softwares	Keras, Pytorch, JAX, Fusion 360, P5.js

RELEVANT COURSES

Vibration Control(ME354)*	Parallel Programming	HPC and ML
Modern Control Systems	Finite Element Method*	Non-Linear Vibrations*
Numerical Methods*	Engineering Drawing (ME251)*	Machining Dynamics*
Autonomous Cyber-Physical Systems*	Computational Fluid Dynamics*	ML for Signal Processing*

*: Received Top Grade (**A**) ^{o}: Ongoing

ONLINE COURSES

Machine Learning by Andrew Ng, Stanford University on Coursera.	(Course Certificate)
RL Specialization, University of Alberta, AMII on Coursera	(Course Certificate)
Sequence Models, Deeplearning.ai on Coursera	(Course Certificate)

EXTRA-CURRICULAR ACTIVITIES AND ACHIEVEMENTS

Selected for attending Eastern European Machine Learning Summer School	2022
Presented Decision Transformers paper for EEML RL Journal Club discussion	2022
Selected for attending Eastern European Machine Learning Summer School	2021
Attended HPC Workshop on Materials and Mechanics	2021