AASTIK

8076945930 ◇ aastik.solanki99@gmail.com ◇ Portfolio ◇ linkedin.com/aastik ◇ github.com/aastik

EDUCATION

MTech in CSE - Indian Institute of Technology (IIT), Bombay *Aug'21 - Jun'23*
Specialization: *Machine Learning and Artificial Intelligence* - CPI 9.5

EXPERIENCE

Senior Machine Learning Engineer, *Samsung Research India* *(Aug'23 - Present)*
Semantic Search using Vision Language Model, in collaboration with University of Cambridge

- Transitioned from tag-based to natural language search using Vision Language Models (VLM) on edge devices, improving search relevance for **10 million+ flagship devices**.
- Experimented with a novel training regime of subset selection of negative pairs via submodular selection in SIGLIP for image-text alignment pretraining for **11 percent faster training**.
- Proposed new metrics (Recall@80p) for Image retrieval task, and optimized VLM pipelines on edge devices, reducing model size by **16 percent and inference time by 25 ms per inference.**

NEXIN:**N**egative **Ex**clusion **In**ference

- Proposed and implemented **NEXIN**, a novel preprocessing technique for exclsuion search queries, **improving result accuracy by over 85% for** ΔU.
- Authored and submitted the **NEXIN paper to EMNLP 2024**, covering 100+ hours of research and development, playing a pivotal role as the main contributor and **secured a patent**.

TestSuite

- Developed a model evaluation tool(TestSuite) used by 5+ teams, **cutting evaluation time by 50**%.
- Owned full life cycle of the project, from initial development to ongoing maintenance, ensuring its continuous improvement and alignment with evolving team requirements, delivering 5 critical evaluation features.
- Deployed a tool to evaluate 7 VLM pipelines on **60+ datasets, saving 200+ man-hours/month**.

PROJECTS

LLM based Multi-Agent Medical Assistant *(Nov'24-Dec'24)*

- Built AI Medical Assistant with 3 LLM agents for edge devices at the LLM Agents Hackathon (Berkeley RDI).
- Integrated smartwatches for real-time monitoring for 4+ vitals, medication reminders, and intelligent scheduling.
- Designed a privacy-first architecture with secure, on-device processing and modular AI optimization.

Multitask federated Learning (GitHub) *(Aug'22-Jun'23)*

- Led a **joint team of 6 from IIT Bombay and Massachusetts Institute of Technology (MIT)** on federated learning in healthcare, published results in **CODS-COMAD 2024**.
- Experimented with class skew for 5 different federated learning algorithms.

Know your Tweeters (GitHub) *(Jul'21-Nov'21)*

- Implemented Django-based platform with **87 % accuracy** in detecting bots and black market users on Twitter.
- Automated data collection for **1,500+ accounts** and identified **65 attributes** to identify bots via Twitter API
- Published in ComPE 2021, winning **Session's Best Paper Award**, and introduced 2 confidence indexes for account trustworthiness.

SKILLS

Languages: Python, C/C++, Java, Rust, Kotlin
Frameworks: TensorFlow, PyTorch, Keras, Pandas, NumPy, Scikit-learn, ONNX, PyTorch Lightning, MLOps
Others: Django, Flask, LLM finetuning, AI agent modeling

PATENT & PUBLICATIONS

- Published paper **Framework for Co-distillation Driven Federated Learning to Address Class Imbalance in Healthcare** in **CODS-COMAD 2024**.
- Published paper **Machine Learning-Based Identification of Collusive Users in Twitter Stream** in **ComPE 2021** and won the **Session's best paper award**.
- Filed a patent for **NEXIN**, introducing novel metrics and benchmarks for exclusion-based image-text retrieval.