Vikas Kumar

Research Engineer – Generative AI

vikasaps.17@gmail.com +91-7860010819 vizzard110 theSquaredError

Experience

- **Research Engineer at TCS Research** Aug 2024 - Present
  - **Scalable LLM Training**: Fine-tuned a 7 B-parameter Mixtral variant using **DeepSpeed ZeRO-3** on a 4-GPU node; leveraged **LoRA** adapters and **bf16** precision to cut memory footprint and keep the pipeline ready for multi-node expansion.
  - **RLHF & DPO Pipeline**: Built an end-to-end feedback loop that reduced finetuning cost by 38 % through parameter-efficient tuning.
  - **Task-Planning Research**: Integrated **Reflexion** and Retrospective Memory in AlfWorld, raising task success from 46 % to 56 %.

Publications

- **"Emergence of Recursive Language through Bootstrapping and Iterated Learning", in AAMAS 2025 (top-tier AI conference) (First Author)**

Projects

- **LLM-Powered Research Assistant Web App**: **(Ongoing)**
  - **Tech Stack: Python, Streamlit, FastAPI, HuggingFace Inference API, Docker, GCP**
  - Built a full-stack Streamlit application powered by Mixtral-8x7B-Instruct for research question answering.
  - Implemented Retrieval-Augmented Generation (RAG) by embedding uploaded documents and querying them via FAISS to inject relevant context into LLM prompts.
  - Enabled multi-turn chat using HuggingFace Inference API and streaming responses from backend to frontend.
- **Enhancing task planning in LLMs**: **(Ongoing)**
  - **Environment: AlfWorld, HuggingFace Transformers, Reinforcement Learning**
  - Integrated Reflexion (LLM self-feedback) and Retrospective Memory (trajectory refinement) into AlfWorld agents.
  - Achieved +10% improvement in task success by combining verbal RL with retrospective grounding.
- **Emergent Communication in multi-agent system using deep reinforcement learning**: **(Research Project)**
  - **Environment: Custom Language Game, PyTorch, Decentralized policy gradient, LSTMs**
  - Simulated multiple agents evolving grounded, recursive, and compositional language from scratch.
  - Used decentralized policy gradients to train agents for spatial navigation via symbolic coordination.
- **A reinforcement learning agent for crypto trading (Link)**: **(AI challenge: Inter IIT)**
  - **Tools: Binance API, SAC, PPO, PyTorch, FinRL**
  - Developed an automated crypto trading system for Binance, incorporating data retrieval, preprocessing, and cleaning.
  - Formulated the trading task as an MDP and implemented SAC and PPO algorithms for optimal decision-making.
  - Designed contrastive datasets from failure cases to improve sequential reasoning accuracy.

Education

- **Indian Institute of Technology, Tirupati** Tirupati, IN
  MS (by Research) Computer Science; CGPA: 8.2 *Jan 2021 - 2024*
- **APJ Abdul Kalam Technological University** Lucknow, IN
  BTech in Computer Science; CGPA: 8.03 *Aug 2016 - July 2020*

Technical Skills

- Programming Languages: Python, C++, C, SQL
- DL Frameworks: PyTorch, Transformers, HuggingFace, LangChain, DeepSpeed
- Distributed Training: ZeRO, model/pipeline parallelism
- Data & Visualization: Pandas, NumPy, Matplotlib
- DevOps: Docker, FastAPI, Streamlit, GitHub Actions
- Reinforcement Learning: OpenAI Gym, nnabla-rl, Ray, MuJoCo
- Other Tools: LaTeX, VS Code, Jupyter, Git
- Finetuning: LoRA, DoRA, DPO, RLHF

HONORS AND AWARDS

- Achieved 7th place in the ZeltaLabs Untrade Crypto Trading Challenge at Inter IIT Tech 12.0.
- Recipient of Research Assistant Scholarship at Indian Institute of Technology, Tirupati by MHRD, Government of India