linkedin.com/in/akashsara/
akashsara.github.io/

Akash Saravanan

akashsara@outlook.com
+1 587-936-6416

Work Experience

Machine Learning Scientist Dec 2023 - Present
- Applied Research Intern Jan 2023 - Dec 2023
Georgian Partners *Edmonton, Canada*
  - Led applied AI projects using LLMs (RAG & data extraction) & analytics for 10+ portfolio & pipeline companies.
  - Developed solutions for 10+ applied research problems across 7+ companies using NLP/LLMs (RAG, fine-tuning, prompt engineering), computer vision (image2text), graph ML (classification), & traditional ML (XGBoost).
  - GenAI Bootcamps (Technical Head): 82 companies, 350+ participants, 55+ projects (15+ in-production) over 3 iterations. Developed majority of technical content, led sessions on RAG, RLHF, & prompt engineering.
  - Created a public repository containing guides & tutorials for LLM reasoning, retrieval-augmented generation (RAG), model alignment, image-to-text models and text-to-image (diffusion) models.
  - **Tech Stack:** Python, PyTorch, LangChain, Transformers, GCP, Git, LLMs (GPT, Mistral), Vector DBs

- **Graduate Research Fellow** May 2021 - Dec 2022
University of Alberta *Edmonton, Canada*
  - Research on AI for video games (computer vision, reinforcement learning) and debiasing language models (NLP).

- **Machine Learning Engineer** Aug 2019 - Nov 2020
Mad Street Den (Vue.ai) *Chennai, India*
  - Created a machine learning model to extract keywords from 37+ million retail products.
  - Developed & deployed solutions for classification & entity extraction problems using language models like BERT.
  - Boosted precision of a rule-based classification system by 15% using ML & reduced codebase latency by 40%.
  - Implemented Word2Vec across a dataset of 2+ million retail product descriptions.
  - **Tech Stack:** Python, PyTorch, Tensorflow, Keras, Transformers, Django, Javascript, AWS, GCP, Git

Education

- **University of Alberta** Jan. 2021 – Dec. 2022
Master of Science (Thesis) in Computing Science; CGPA: 3.75/4.0 *Edmonton, Canada*
  - **Thesis:** Visualizing Characters and Evaluating their Balance in Competitive Video Games.

- **Anna University (Sri Venkateswara College of Engineering)** Jun. 2015 – Apr. 2019
Bachelor of Engineering in Computer Science and Engineering; First Class. *Chennai, India*
  - **Thesis:** Natural Language Generation using Generative Adversarial Networks (Awarded grant of INR 10,000)

Skills

- **Languages & Databases:** Python, MySQL, SQLite, MongoDB, Vector Databases (QDrant, LanceDB), Markdown
- **Frameworks & Libraries:** PyTorch, Tensorflow, Keras, Transformers, LangChain, NumPy, Pandas, scikit-learn
- **Tools & Technologies:** Git, LaTeX, AWS (Sagemaker, EC2, S3, Redis) GCP (VertexAI, Compute Engine)

Publications

- **A Framework for Predicting the Impact of Game Balance Changes through Meta Discovery:** First author. Under review.
- **FineDeb: A Debiasing Framework for Language Models:** Co-first author. AI4SG Workshop, AAAI 2023.
- **Pixel VQ-VAEs for Improved Pixel Art Representation:** First author. EXAG Workshop, AIIDE 2022.
- **Facial Emotion Recognition using Convolutional Neural Networks:** First author. AICV 2018.

Projects & Open Source Work

- **[Open Source] Multimodal Toolkit:** Primary maintainer. Refactored codebase, added tests, resolved 30+ issues.
- **Homebrew Helper:** Developed & deployed a Discord bot with database connectivity for online role-playing games.
- **[Open Source] poke-env:** Identified & fixed several bugs, added example code.