Jyotirmay Khavasi

(+91) 9987540788 | jyotirmay.khavasi@gmail.com | Jyotirmay Khavasi | theory-in-progress

EDUCATION

Vishwakarma Institute of Information Technology **2020 – Present**
B.Tech in Artificial Intelligence and Data Science, **CGPA: 9.3** *Pune, India*

EXPERIENCE

Wolters Kluwer **January 2024 – Present**
Data Science Intern *Pune, India*

- Optimized production-deployed Vision Encoder-Decoder and Layout Transformer models through batch classification and data augmentation techniques resulting in a 4% improvement across all metrics.
- Utilized Hugging Face Optimum to convert models to ONNX format for a 40% reduction in evaluation time for clients.
- Working on object detection and segmentation, with a focus on table detection, employing transformer models and OpenCV to achieve precise and robust results.
- Automated tasks tailored to meet specific business needs, leveraging clustering algorithms to streamline workflows and enhance productivity.

Google Summer of Code @PyTorch-Ignite **May 2023 – September 2023**
Open Source Contributor *Remote*

- Created a template for Reinforcement Learning using DQN and Advantage Actor Critic algorithms, configuring code to efficiently utilize parallel processing for spawning multiple environments and handling various Reinforcement Learning tasks, using interchangeable and flexible components of TorchRL.
- Enhanced CI/CD pipelines with GitHub workflows, enabled Docker containerization, and wrote comprehensive contribution guidelines. Also, improved the configuration by centralizing YAML attributes, allowing command-line overrides, and helped in integrating with Google Fire and Hydra.
- Implemented code refactoring with Vue.js tags, enabling template inheritance and resulting in more than 1000-line reduction across templates.
- Building a few-shot learning template with a custom sampler, following PyTorch's data loading, training, evaluation, and metric integration pipeline.

HCL Technologies **September 2022 – March 2023**
Research Intern *Pune, India*

- Worked on a NLP Model which extracts Rules from given Text. Software converts and Processes Given Text to output the Mathematical Expressions of the Rules.
- Worked on custom BERT architecture which is used for Named Entity Recognition with a Recall of 88%.
- Employed Random Forests for Classification of a given sentence into 'Rule' or 'not Rule' with 92% Accuracy.
- Created Custom Parse Trees which Convert the Rule from the BERT model into a Logical Mathematical Expression.

SKILLS

Languages: C++, Python, SQL.
Developer Tools: Git, Docker, Vim, Shell, LaTeX, VS Code.
Technologies: MySQL, PyTorch, Machine Learning, Image Processing, Deep Learning, Natural Language Processing, Data Science, MongoDB, CUDA.
Soft Skills: Communication, Analytical Thinking, Teamwork, Problem Solving.

PROJECTS

Opensource Contributor to PyTorch-Ignite *PyTorch, Deep Learning, Neural Networks, Torchvision*

- Active Contributor to the PyTorch Ignite Opensource Library which helps in training and evaluating neural networks.
- Wrote Event Filters for the Engine and Doctests for different Metrics in the Contrib Module

TODO App API *Docker, Dockerfile, Flask, API, REST*

- Built a REST API app exposing the end points. Users can log in and make API Calls. Containerized the application and deployed it using Docker

Movie Ticket Booking System *DJango, MySQL, WebApp, HTML*

- Developed a WebApp using Django in Python. Implemented Backend Connection using MySQL Database. Integrated the frontend with the Backend seamlessly. Integrated Role based login and booking.

Querying with LLMs *LangChain, OpenAI, Q&A*

- Prompt Engineering to Fetch relevant content with Langchain
- Evaluating LLMs for document similarity on consumer-grade GPUs.

RELEVANT COURSEWORK/MOOC

Discrete Mathematics
Probability and Statistics
Data Structures
Artificial Intelligence
Operating Systems
Database Management
Software Engineering
Cloud Computing

INF8245E - Machine Learning: *Chandar Research Lab:* Math Intensive, Rigorous introduction to the field of Machine Learning. Use of ML algorithms, how and when the algorithms work.