Desh Iyer

(+91) 90589-88842 | desh.iyer@gmail.com | Desh Iyer | deshiyer
Alahalli, Anjanapura, Bengaluru, KA 560062

Executive Summary

Driven computer science graduate with a B.Tech in Computer Science with valuable experience in scientific distributed computing garnered through the Belle II experiment at KEK, Japan. Proven expertise in collaborative research environments, optimizing grid infrastructure for global experiments. Seeking to contribute to CERN's world-class research, leveraging skills in grid computing, optimization, and interdisciplinary collaboration.

Skills

Programming Languages: Python, R, Mojo, Scala, C/C++, Lua, Rust, SQL (MySQL), JavaScript, HTML/CSS
Frameworks: TensorFlow, PyTorch, Node.js, Flask, FastAPI, Burn, Processing, p5.js
Developer Tools: Linux, NVIM, Linode, Git, Docker, VS Code, LaTeX
Grid Computing: CI/CD, Apache Spark, Apache Kafka, Apache Hadoop, GitLab
AI/ML: Ollama, llama.cpp, Transformers, Pandas, NumPy, Matplotlib, Seaborn, TensorFlow, Keras, PyTorch, Sklearn, nltk
Languages: English (proficient), French (intermediate), Tamil (native), Kannada (native), Hindi (proficient)

Experience

Belle II Grid Computing and R&D Collaborator 27/03/2023 - Present
High Energy Accelerator Research Organisation (KEK) *Remote (Tsukuba, Japan)*
BelleDIRAC

- Impacted development of the BelleDIRAC distributed framework for the Belle II grid and gbasf2 for the command-line client, amplifying users' experience to submit jobs to the Belle II grid
- Enhanced components handling **job submission, data transfer, and parsing and logging dataset objects** using the DIRAC API
- Optimized dataset logging by making it **20% faster to relay data transfer diagnostics** to the client by parsing dataset objects in JSON
- Improved the subsequent versions of BelleDIRAC by **decreasing job submission failures by 15%**, handling diagnostic data better, and unit testing components using *pytest*

Shifts

- Monitored the Belle II data production system, **accumulating 25+ hours of shifts** on the *B2MonitoringSystem* web app
- **Investigated and logged the performance of grid sites** in data transfer, data processing, and MC (Monte Carlo) production resources

Machine Learning Engineer 01/06/2023 - 01/09/2023
IBM *Bengaluru, India*

- Engineered a data pipeline to pre-process music data from the *Google Magenta MAESTRO* dataset in MIDI format and generate music with a custom generative adversarial network (GAN)
- Yielded an **average log likelihood of -2, an Inception Score (IS) of 4.2, and a Fréchet Inception Distance (FID) of 68** by *batch normalization* of the training data and *dropout* as a regularization technique to boost the similarity between distributions
- Incorporated model inference into an API hosted on an IBM cloud server and **achieved a 62% increase in model response time and precision** over the previous generation music generation model on the cloud

Projects

***What's Up Doc?* - Language Agnostic Code Documentation** *Python, Ollama, LLMs, TensorFlow, Transformers, Git*

- Designing a command-line tool named *whats-up-doc* to **automatically generate documentation** of a codebase using open-source large language models, *Ollama* and *LangChain*
- Fine-tuning the *codellama-2-7b model* on the *CodeLlama-2-20k* and *CodeSearchNet* datasets for Python, C++, and Java to produce a language-agnostic model that **achieves a baseline BLEU score of 32.2**
- Advancing code generation through exploring LLM quantization and pruning methods

***Splice and Dice* - Text Summarization with Transformers** *Python, Transformers, PyTorch, Streamlit, Git*

- Designed a custom T5 transformer architecture to **summarize academic and news articles in a web application** named *splice-and-dice*
- Fine-tuned the T5 model on the *Kaggle News Summary* dataset to **achieve a ROUGE score of 0.58**, translating to a **7% increase in ROUGE precision** compared to the pre-trained model on the *cnn_dailymail* dataset
- Deployed the fine-tuned T5 transformer model on a *Streamlit* web application, providing an intuitive and interactive interface for text summarization.

***Bulletin* - Project Management in Neovim** *Neovim, Lua, Linode, Shell Scripting, GitLab*

- Created a plugin for the command-line IDE Neovim during a college group project to **keep track of action items, bugs, and short-hand documentation directly in the Neovim development workspace**
- Integrated a module to **parse compiler warnings and errors** and log file paths automatically to the workspace's *bulletin*
- **Recorded a 66% surge in team productivity** from having the *bulletin* built into the editor, allowing us to **complete project development 3 weeks ahead of schedule**

Education

UPES Dehradun, India

B.Tech in Computer Science & Engineering (Hons.) Spl. AI/ML, GPA: 8.51 01/08/2020 - 01/05/2024 (Expected)*

Indian School Muscat Muscat, Oman

CBSE Senior School Certificate Examination Class XII, Aggregate: 95.6% 01/04/2020

Indian School Muscat Muscat, Oman

CBSE Secondary School Examination Class X, Aggregate: 92.8% 01/08/2018