Devasheesh Mishra

+91 93197 72072 | devasheeshmishra4@gmail.com | linkedin.com/in/devasheesh-mishra/ | github.com/devasheeshG
huggingface.co/devasheeshG

Professional Summary

- **Technical Lead** at GeeksforGeeks (GFG), conducting **machine learning workshops** and organizing **hackathons** for 200+ students
- Developed advanced projects including **GPT-2 model implementation** and **Automatic Speech Recognition System** using **PyTorch**, **Transformers**, and **FastAPI**
- Proficient in **Python**, C/C++, SQL, **PyTorch**, **Transformers**, **GCP**, **AWS**, Azure, **Docker**, and Git with multiple **hackathon wins** and **ML certifications**

Skills & Interest

Programming Languages: Python, C/C++, SQL

Frameworks & Libraries: PyTorch, Transformers, TRL, Pydantic, FastAPI, Pandas, NumPy, Matplotlib

Cloud & DevOps: Google Cloud Platform, AWS, Azure, Docker

CI/CD: GitLab CI, GitHub Actions

Databases: PostgreSQL, MongoDB, Vector Databases (Milvus, Pinecone)

Machine Learning/AI: Large Language Models (LLMs), Model Training, Fine-tuning

Version Control: Git, GitHub

Testing: Unit Testing, Integration Testing, Pytest

Projects

GPT-2 from Scratch | *PyTorch, Transformers*

- Implemented a decoder-only GPT-2 model (135M parameters) trained on 2B tokens using PyTorch and Transformers
- Utilized Distributed Data Parallelism (DDP) with 2x Tesla P40 GPUs, reducing training time to 19 hours
- Optimized performance with Flash Attention 2, Gradient Accumulation, and Mixed-precision training
- Developed custom tokenizer with 51K vocabulary size, trained on 100K tokens with over 50,000 merges

Automatic Speech Recognition System | *PyTorch, Transformers, FastAPI, WebSockets*

- Fine-tuned 2x Whisper models (750M& 1.1B parameters) for Hindi using 10,000 hours of audio from the Gram Vani Dataset
- Achieved Real-Time Factor (RTF) of 0.3 and reduced inference latency to 200-300ms for 30s audio chunks
- Implemented data cleaning, Voice Activity Detection (VAD), and increased decoding heads for improved performance
- Developed low-latency server using FastAPI and WebSockets, incorporating speculative decoding and Medusa technique

Work Experience

GFG Technical Lead April 2024 – Present

SRM University *Delhi-NCR, IN*

* Conducted comprehensive machine learning workshops for 80+ students, teaching essential ML and DL principles and applications, which enhanced participants' practical skills and resulted in a 25% improvement in course project quality
* Organized and led 10+ GeeksforGeeks (GFG) workshops across SRM, facilitating doubt clearance sessions that enhanced coding proficiency for over 200 students, resulting in a 40% increase in coding competition participation

GFG Core Tech Member October 2022 – April 2024

SRM University *Delhi-NCR, IN*

* Contributed as a Core Member of GFG, organizing the Phoenix Hackathon, where I achieved 1st place
* Organized and led Hack-Innovate, a two-day hackathon during the tech fest, attracting over 300 participants and showcasing 50+ innovative solutions, significantly enhancing event participation and engagement by 40%

Achievements

Phoenix Hackathon: Secured 1st position (2022)

Code-A-Thon: 1st runner up (2022)

InnoSprint D4 Hackathon: 1st runner up (2023)

Live-Project Competition: 2nd runner up (2023)

SIH: 1st runner up at college level (2023)

Certifications

DeepLearning.AI: Neural Networks and Deep Learning

Google: Transformer Models and BERT Model

Udemy: Machine Learning - Fundamental of Python Machine Learning

Udemy: PyTorch for Deep Learning Bootcamp

Education

SRM Institute of Science and Technology Delhi NCR, IN

Bachelor of Technology in Computer Science, Specialization in AI and ML; CGPA: 8.0 *May 2022 – May 2026*

- **Relevant Coursework:** Data Structures and Algorithms, Machine Learning, Deep Learning, Natural Language Processing, Database Systems, Artificial Neural Network, Computer Networks