Aayush Jaiswal

aajais@iu.edu | linkedin.com/in/jaiswal-aayush | github.com/aayush9400 | Google Scholar | medium.com/@jaayush12

Education

Indiana University Bloomington — Aug 2022 – May 2024
Master of Science in Data Science — *GPA: 3.60*

Manipal University Jaipur — Aug 2017 – May 2021
Bachelor of Technology in Computer Science — *GPA: 3.70*

Technical Skills

Programming Languages: Python, R, Java, C/C++, SQL
Frameworks: TensorFlow, PyTorch, Keras, Numpy, Pandas, ScikitLearn, spaCy, W&B, OpenCV
Tools: Git, Docker, Kubernetes, Grafana, Prometheus, PostgreSQL, BigQuery, Elastic Search, Kafka, Spark, dbt, Bash, Cronicle, Flask, MLFlow, AWS, GCP
Areas of Expertise: Deep Learning, Data Engineering, MLOps, Computer Vision, Recommendation Systems, Agile

Experience

Graduate Assistant — Aug 2023 – May 2024
Indiana University — *Bloomington, IN*

- Enhanced research capabilities by developing Generative AI models for 3D brain image synthesis, collaborating with cross-functional stakeholders (Doctors, Researchers, ML Experts)
- Mentored 150+ students in Big Data Applications and Data Representation courses, focusing on Java, Spark, Hadoop, Data Structures and Algorithms, Distributed & Cloud Computing, and AWS.

Data Scientist / Machine Learning Engineer - 1 — Jan 2021 – Mar 2022
Trell, Sequoia Backed Social Media Platform — *Remote*

- Enhanced user engagement by 20% and retention by 45% by deploying a multilingual short video recommendation system for 1M+ users, collaborating with cross-functional stakeholders to improve feed quality and drive satisfaction.
- Optimized operational efficiency by 30% through automating model training and deployment using AWS, Prometheus, Grafana, and W&B.
- Reduced data retrieval time by 25% and improved data quality by 10% by optimizing real-time data pipelines and cloud database efficiency with GCP Big Query and DBT.
- Defined success metrics and prioritized solutions by creating an MIS dashboard for RecSys projects based on user metrics.

Research Assistant — Jan 2020 – Jul 2020
Manipal University — *Jaipur, India*

- Published 2 papers with 700+ citations by researching computer vision algorithms for COVID-19 detection. Created an end to end application accessible to doctors to get predictions with low latency.

Data Science Intern — Jun 2019 – Jul 2019
Airtel (Leading telecom provider of India) — *Delhi, India*

- Increased marketing campaign effectiveness by 10% by collaborating with the marketing team and developing user behavior-based clustering algorithms.

Projects

Multi-Lingual Short Video Recommendation System | *Python, TensorFlow, AWS, GCP, SQL, Docker, MongoDB, MLFlo*

- Implemented a custom two-tower recommendation system including retrieval and ranking resulting in 50% faster retrieval times.
- Used collisionless embeddings to adapt to changes in user preferences and behavior in real-time
- Hosted the model on AWS EC2 and developed a fully automatic retraining pipeline using AWS Lambda for tackling data drift. Medium Article

Real-time data pipeline using DBT | *Python, SQL, GCP, DBT*

- Demonstrated expertise in SQL for crafting complex queries and aggregations, reducing data access times by 15% and significantly enhancing data analytics efficiency across multiple cloud platforms.

3D Conditional Diffusion Model | *Python, PyTorch, TensorFlow, GPU, Slurm, W&B, Generative AI*

- Implemented a 3D Vector Quantized AutoEncoder, Generative Adversarial Network (VQ VAE, VQ GAN), diffusion models for synthetic brain image generation and accurate latent representation, under Dr. Eleftherios Garyfalladis

Rapid SARS-COV-2 Detection | *Python, TensorFlow, OpenCV, AWS, Flask API, EC2*

- Achieved 96% testing accuracy with minimal false negatives for COVID-19 detection from X-ray and CT scans.
- Published in Taylor & Francis and Springer, advised by Dr. Dilbag Singh and Dr. Manjit Kaur.

Virtual coding assistant for DiPY | *Python, Transformers, Vector Databases, RAG, LangChain, LLM*

- Created a code assistant using the DiPY codebase.
- Implemented LangGraph RAG pipeline using local LLM for prompt engineering.

Brain Image Segmentation using Vision Transformers (ViT) | *Python, PyTorch, Hugging Face*

- Implemented a 3D CNN neural network (U-Net) with an attention mechanism to identify tumor regions in MRI images, achieving 95% accuracy.

Monitoring for Cloud-Native Environments | *AWS, Prometheus, Grafana, Kubernetes, PromQL, S3*

- Implemented scalable monitoring on AWS, enhancing system reliability, uptime, and performance insights by over 30%. Medium Article

Publications

3D Conditional Diffusion Model for Synthetic Brain Image Generation TBD

Aayush Jaiswal, JongSung Park, Eleftherios Garyfalladis *2024*

- Developed 3D VQ-VAE and VQ-GAN models using Python and PyTorch, achieving high-quality synthetic brain images with improved latent representation.
- Contributed to advancements in medical imaging, enhancing diagnostic capabilities and research methodologies.

Classification of the COVID-19 Infected Patients Using DenseNet201 Based Deep Transfer Learning Taylor & Francis

Aayush Jaiswal, Neha Gianchandani, Dilbag Singh, Vijay Kumar, Manjit Kaur *2020*

- Achieved 96.25% accuracy and AUC of 0.97 in classifying COVID-19 from chest CT scans using DenseNet201 and deep transfer learning techniques.
- Utilized Python and TensorFlow to provide a faster, non-invasive alternative to RT-PCR for COVID-19 diagnosis, significantly reducing false positives and negatives.

Rapid COVID-19 Diagnosis Using Ensemble Deep Transfer Learning Models from Chest Radiographic Images Springer

Aayush Jaiswal, Neha Gianchandani, Dilbag Singh, Vijay Kumar, Manjit Kaur *2020*

- Designed ensemble models using Python and TensorFlow, achieving 98.25% accuracy in differentiating COVID-19, viral, and bacterial pneumonia from chest X-rays.
- Validated on two datasets, models demonstrated high sensitivity and specificity, enabling effective large-scale COVID-19 screening.