Vanshika Gupta

vanshika421@gmail.com | +91 9611591432 | LinkedIn: vanshika421 | GitHub: vansjyo

Education

University of Illinois at Urbana-Champaign (UIUC) *2021 - 2023*
Master of Science (M.S. Thesis), ***Specialization****: Mathematics and Computing* *GPA: 3.84 / 4.00*

National Institute of Technology Karnataka (NITK) *2016 - 2020*
Bachelor of Technology (B. Tech.), ***Specialization****: Mathematics and Computing* *GPA: 8.96 / 10.00*

Work Experience

Software Developer - AI/Backend, *Enmovil AI* *Sept 2023 - Present*

- Built an LLM based multimodal chatbot with **Voice** interface for generating insights, charts, and custom dashboards from customer data using natural language queries. Implemented **Retrieval Augmented Generation (RAG)**, leveraging **Qdrant** vector database and **Langroid** for knowledge base indexing
- Applied prompt engineering techniques and fine-tuning to optimize LLM interactions. Incorporated tool-calling to execute actions
- Currently developing a **multi-agent orchestration** framework integrating agents associated to different products using **LangGraph**, to build a unified AI assistant
- **Managed a team of 9** as Product Manager, driving the product lifecycle from ideation to launch, collaborating with cross-functional teams, and aligning solutions with business objectives. Engaged directly with high-value clients (Maruti, Nestle, Mahindra) to develop POCs and secure key deals
- Developed core product features and algorithms for ML and combinatorial optimization problems, including vehicle routing with resource constraints, bin packing, and forecasting (PyTorch, Pandas, Google OR-Tools)
- **Led the architecture and development** of a configuration-driven **multi tenant** backend model for SaaS platform from scratch, reducing development effort for customer POC and onboarding by 10x (Pydantic, MongoDB, Python, FastAPI, AWS Cloud, REST APIs, Docker)
- Migrated multiple python scripts to a full-fledged centrally-hosted ML-Optimizer **microservice** consumed by multiple projects within the company using flask. Used MongoDB and AWS S3 to manage tasks, run configurations and generate output
- Developed scheduler based pipeline for ingestion of customer data using **Celery** and **Apache Kafka**

Machine Learning Engineer, *Lowe's India* *July 2020 - Aug 2021*

- Built and deployed scalable modules using **Python scripting** for automating reporting (ETL) workflow, including **big data processing** (Hadoop, Teradata), querying database (SQL) and upholding data quality using tests (reduced 20 hours/week)
- Built **time-series forecasting model** like ARIMA, LSTM to predict weekly inventory sales using H2O open source ML Ecosystem
- Developed root-cause analysis tool to mitigate inventory deficit risk by correlating 100s of data points and applying specific business domain knowledge to automatically flag the reasons for deficit

Data Scientist, *IIT Mandi* *May 2019 - December 2019*

- Led research for detecting *clutters* (material changes) on geospatial Mars Subsurface Signal (GPR) data using **computer vision** based **semantic segmentation** pipeline (Keras, PyTorch, TensorFlow)
- Achieved 70% balanced accuracy with **UNet architecture** based **Deep learning** CNN framework on data with 0.0005% imbalance; Built parallel routines for signal denoising, multimodal registration, feature generation and lazy loading to handle large data ingestion

Software Developer Intern, *Carabiner Technologies* *Dec 2018 - Jan 2019*

- Developed various interactive front-end image editing tools for annotating images for ML applications using Canvas
- Wrote back-end feature for batch processing images for their product (NodeJS, Angular)

Machine Learning Intern, *Indian Academy of Sciences* *May 2018 - July 2018*

- Employed Spectral **Clustering**-based **unsupervised Dimensionality Reduction** to reduce size of satellite image with 400 bands to 30 bands with only 1% loss in multiclass cropland **classification** accuracy, determined using **Support Vector Machines** (MATLAB)
- Compared various dimension estimation/feature selection approaches like Principal Component Analysis to determine best approach

Selected Publications

- **Gupta, V.**, Gupta, S. K., & Kim, J. **Deep Learning Based Automated Discontinuity Detection and Reconstruction in Subsurface Environment of Mars: A Case Study of SHARAD Observation**, *Applied Sciences 10*(7), 2279 (2020)
- **Gupta, V.**, Gupta, S. K.,& Shetty, A. **Fractal-based supervised approach for dimensionality reduction of hyperspectral images**, *Computers & Geosciences*, 193, 105733. (2024)
- **Gupta, V.**, Gupta, S. K., & Shukla, D. P. **Optimal Selection of Bands for Hyperspectral Images Using Spectral Clustering**, In *International Conference on Recent Trends in Image Processing and Pattern Recognition* (pp. 288-304). Springer (2018)

Technical Skills

- **Programming**: Python, C++, SQL, MATLAB, R, JavaScript
- **Other Skills**: LangGraph, Langchain, Qdrant (Vector DB), Keras, TensorFlow, PyTorch, OpenCV, Docker, MongoDB, Pandas, Teradata, Git CI/CD
- **Courses**: Convex Optimization for Machine Learning, Graph Networks, Algorithms, Financial Computing, Advanced Stochastic Processes, Game Theory & Fair Division, High-Frequency Trading Technology, Statistical Analysis, Deep Learning, Signal Processing, Data Structures

Other Projects

- **Modelling resource allocation problems with fairness and Pareto Optimality** Designed two-sided market matching-based game-theoretic algorithm using Mixed Integer Linear Programming. Conducted computational analysis with statistically generated inputs (C++, Gurobi)
- **Portfolio Risk Optimization Bond Immunization** compute immunized portfolio of Fixed Income bonds using linear programming based convexity maximization and duration matching; performed stress testing (Gurobi solver, C++)
- **Automatic Detection of Wireless Financial Trading Networks (HFT)**
  – Developed graph search-based algorithm to identify and analyze wireless HFT links and estimate their latency/path length using NetworkX
  – Created clean, standardized schema for efficient information retrieval from unstructured database; Incorporated Integration tests (Vagrant)
  – Performed address normalization using Google Maps Geocoding API to match SEC broker address and path endpoint (SQL, Python)
- **Algorithms for morphological extraction of Scanning Electron Microscope (SEM) Fiber Images**
  – Ranked **3**rd/218 in *APEER Image Processing Contest 2020* organized by **ZEISS** for developing modules for processing fiber images
  – Built end-to-end background correction, fuzzy clustering-based image segmentation, estimation of orientation, diameter and length of individual fiber using connect components approach (Python, OpenCV, MATLAB)

Additional Experience & Achievements

- **Open Source Contributions:** Resolved Back-end and Front-end issues in **Mozilla** Firefox-DevTools Profiler and **Public Labs**, Developed Python3 compatible code for morphological image analysis library **Quanfima**
- **Graduate Teaching Assistant** at UIUC for 3 advanced undergraduate courses like Deterministic Optimization with 100+ students.
- Received **Spot Award** twice for exceptional performance at Lowe's
- Awarded **KVPY** (Kishore Vaigyanik Protsahan Yojana) Fellowship (2015)
- Recipient: **Indian Academy of Sciences** (IAS'17) Research Fellowship, **Merit Scholarship** for Department Rank **1**st/88