NAVIN KUMAR M

+919884285774 ⋄ Chennai, Tamil Nadu, India

personal.mnk@gmail.com ⋄ linkedin.com/in/mnk-navin ⋄ github.com/NavinKumarMNK ⋄ megnav.com

EXPERIENCE

Software Engineer - Machine Learning   Remote, India
Full Time - B1 @ Branch International   Jun 2024 - Present

- Contributing significantly to the ongoing efforts to scale **ML Services** and facilitating improvements in the **Model Deployment** process.
- Conducted intensive **QA testing** to evaluate data integrity across various forms and developed an algorithm for **Credit Report** creation from historical files.

Intern @ Branch International   Jan - May 2024

- Engineered over **200 unique features** from credit reports, enhancing model performance by increasing the **AUC by 10%**.
- Implemented a cash flow and balance classification and extraction model, which tripled coverage and achieved a perfect **AUC score of 1.00**.

Machine Learning Engineer   Remote, India
Intern @ Gutsy Innovation   May - July 2023

- Developed real-time **Face Detection & Recognition System** using **Yolov7 & ArcFace**.
- Implemented the AI service on **DeepStream SDK**, with integration of **Azure IoT Hub** and **Qdrant**.

SKILLS

- **Frameworks:** PyTorch, Ray, AWS Cloud, Docker, Kubernetes, Qdrant, FastAPI, Django
- **Languages:** Python, PostgreSQL, Go, Rust, C++
- **Technical Expertise:** Credit Risk Modeling, Time-series Forecasting, Computer Vision, NLP, LLM, Linux, Cloud & DevOps

EDUCATION

B.Tech in Computer Science & Engineering, Artificial Intelligence & Machine Learning Specialization.
Vellore Institute of Technology, Chennai
Cum.GPA : 8.92   Sep 2020 - May 2024

Relevant Coursework: Machine Learning, Deep Learning, Data Structures, Algorithms, Operating Systems, Business Intelligence, Security, Networks, Databases, Linear Algebra, Statistics, Software Engineering.

PROJECTS

AI Learning Management System   (github)

- Built **AI-Learning-Platform**, an **LLM-RAG pipeline** that acts as a guide and resolves students' doubts.
- Deployed the product on-premise in **IBM ppc64le** architecture & scaled using **Docker Swarm**.
- Utilized **vLLM** for model inference, and integrated **Qdrant** with **Langchain** for the RAG pipeline. The server was implemented using **Django**, with **PostgreSQL** and **Cassandra** as the SQL and NoSQL databases.

AI Malware System   (blog) (github)

- Developed a **Cross-Platform Application** that uploads files to an AI-driven backend for malware classification, with results visualized real-time in **Azure PowerBI**.
- Employed **CoAtNet Transformer** model pre-trained on 2 million malware binaries, and conducted advanced feature analysis using **LSTM & LightGBM** based on logs from **Cuckoo Sandbox** and Header Information.

Benetech - Making Graphs Accessible   (github)

- Visual Reasoning from charts and plots to make them accessible for **Visually Impaired People**
- Utilized **EfficientNetv2, Matcha (Pix2Struct)** for analyzing **bar, line, and dot plots**, and **Faster-RCNN** for **scatter plot derendering**.