Ayush Agarwal

ayush.ag.05@gmail.com | 9112647721 | Linkedin=>ayush-agarwal-261041215 | Github=>ayush-agarwal-0502

Education

IIT BHU Varanasi - BTech in Electronics Engineering (ECE) (CGPA = 9.04) 2020 – 2024

Experience

DATA SCIENTIST AT INFO EDGE (NAUKRI) - May 2024 – Jan 2025

- **Reels Notification pipeline** - Constructed **ETL Pipelines** for **Data Engineering** to deliver notifications to **6,00,000+ users** achieving **5x DAU** (Daily avg Users) on Naukri reels. Integrated data from multiple AWS and Mongo DBs, translating complex business logic into code computing dynamically computed user pool.
- **News Recommender System** - Developed **Real Time Recommender System** for Naukri Techminis, a news platform similar to Inshorts, delivering personalized content to **40,000+ users** with response time **<150ms**, using Approximate Nearest Neighbors (**ANN**) based techniques like **HNSW**.
- **Notification Generation using GenAI** - Used **Prompt Engineering** over **Llama** (8B) to automate notification content generation, leveraging video data (title, transcription(using Whisper AI)). Designed **guardrails**(profanity checks, length check etc) to ensure quality content, processing 400+ videos with zero manual effort.
- **Sallie - Sales Intelligence AI (Hackathon) -** Ideated and prototyped an innovative AI tool to empower sales teams with features like call summarization, salesperson evaluation, and client-specific data extraction (using RAG) from Naukri's internal databases. Integrated speech diarization.
- **Metrics Mailer and Dashboard -** Developed a metrics mailer system and Streamlit dashboard to visualize and track key performance indicators (KPIs) like CTR, delivery rate etc for monitoring the user journey funnel for Reels notifications, automating insights delivery to the team.
- **Other Projects -** Meme Scrapper, Reels Recommender System

ASIC DESIGN (HARDWARE ENGINEERING) INTERN AT NVIDIA - May 2023 – July 2023

- **Tools:** Perforce (Version Control System), Unix, gVim, Verdi (ASIC File Management), JasperGold, System Verilog
- **SFV SLCG deployment - Deployed, Executed and Documented** "SLCG SFV" workflow over **3** submodules in USB IP to root out clock gating bugs. Discovered and **debugged 1** critical RTL **bug** in the chip.
- **Netlist Linting over USB IP -** Performed **Rule Based Data Analytics** over **8000** Netlist Linter outputs thus assuring **best quality practices** in Netlist. Using this Data analytics, **optimized Design** via RTL Coding to enhance Clock Tree structure.

Honours and achievements

- **SPECIAL MENTION - GYMKHANA AWARD:** Among **top 20 students across IIT BHU** in technical proficiency
- **2ND POSITION IN RECOGNIZANCE PRASTUTI IIT BHU :** Machine Learning Event. (Made U-Nets)
- **3rd POSITION IN CASSANDRA UDYAM IIT BHU :** Data Science Event (Invoice payment duration prediction model)
- **6th POSITION IN IDEEAVOLT ,IIT ROORKEE :** Idea Pitching Competition.(Made Innovative Electronic Door) (Link)
- **COMPLETED MLWare TECHNEX'22 :** Computer Vision event (Made Stained Cell Counter) (Link)
- **Hobbies :** Guitar, Photography, Dance, Reading, Travelling

Skills and Technologies

Skills: Recommender Systems (Information Extraction), Data engineering, Generative AI (GenAI), Prompt Engineering, Data Science, Machine Learning (Deep Learning, Computer Vision), Electronics, Robotics, Software (SDE), Product Management

Technologies: Python, Mongo DB, Kafka, Aerospike, Flask APIs, SQL, Cron, Git, Confluence, Kubernetes, Jenkins, Postman, Gunicorn, Requests, BeautifulSoup, C++, Tensorflow, Pandas, Numpy, Scikit Learn, Matplotlib, Seaborn, Excel, Power BI

Position of Responsibility (POR) - Digisim Coordinator - Udyam 22 (ECE Dept Fest)

- **Skills Demonstrated:** Time Management, Communication Skills, Mentorship, Innovation, Leadership.
- **Spearheaded** the Digisim event, overseeing end-to-end execution, including workshops and problem statements. Conducted hands-on sessions for **50+ juniors**, mentoring them in PCB design and fostering a culture of backend VLSI.
- Designed **innovative** real-life digital electronics problems and organized a final challenge and **ensured seamless event execution** through strong **leadership and mentorship**.

Projects -

Credit Card Fraud Detection -

- **Exposure:** Logistic Regression, SVM, KNN, Financial Data Analysis, Data Science, Risk Analytics (Link)
- Conducted Exploratory Data Analytics (**EDA**), **Data Visualization** and processing over highly imbalanced dataset.
- Utilized Machine Learning (ML) algorithms such as **KNN, Logistic Regression and SVM** achieving classification recall - 94%.
- Evaluated and compared model performances on the basis of **F1 Scores, Confusion Matrix and ROC Curve**.

Starbucks Customer Segmentation -

- **Exposure :** PCA, K-means, t-SNE, cluster analysis, market segmentation, Unsupervised learning (Link)
- Processed 3 datasets containing over 300,000 entries, ensuring data cleanliness and merging wherever necessary. Performed EDA, feature engineering, and applied PCA to reduce data dimensionality.
- Employed **K-Means Clustering** to cluster 17,000 customers, selecting the optimal number of clusters based on **Silhouette score** and Inertia/Sum of Squared errors value. Inferred the primary demographic traits of each cluster based on characteristics of each member.
- Utilized t-SNE to perform **cluster analysis** and determine the most effective approach for delivering promotional offers.

Optical Character Recognition (OCR)-

- **Exposure:** Computer Vision, Machine Learning, Deep Learning, Transfer Learning, **CNN** (Convolution Neural Networks) (Link)
- Employed **OpenCV** for word and character segmentation; trained the model on the Standard OCR Dataset using TensorFlow.
- Developed an OCR pipeline using **Transfer Learning** (ResNet50) to classify printed characters and digits with 98% accuracy.

Auto-Adjusting Short Straddle Quant Bot | Options Trading Strategy -

- **Exposure:** Quantitative Finance, Delta Hedging, **Options & Derivatives**, Algorithmic Trading, Python (Link)
- Coded an automated **short straddle** strategy with **delta hedging** using real-time position adjustment logic to adapt to price swings and preserve profitability.
- **Visualized P&L chart via GIF** demo, simulating how the strategy responds in real-time to market movement.

Sign Language Translation using GMM -

- **Exposure:** Gaussian Mixture Models (GMM), Mediapipe, Data Collection (Link)
- **Created a custom dataset** by collecting ASL sign language data using **Mediapipe** and processing hand landmarks with PCA to visualize the alphabets .
- Trained a **Gaussian Mixture Model (GMM)** for gesture classification and applied a 99% confidence threshold for accuracy.

Background Remover for Photos -

- **Exposure:** UNet, Semantic Segmentation, Image Processing (Link)
- Developed a background removal tool using **UNet for semantic segmentation**, trained to detect humans in images and remove the background.
- Implemented custom loss functions to address class imbalance, improving foreground segmentation accuracy in the model.

Restaurant Finder Dashboard -

- **Exposure:** Power BI, Data Visualization, Dashboard Development, Data Analysis (Link)
- Built an **interactive Power BI dashboard** to filter and recommend top-rated restaurants in Bangalore based on user preferences such as cuisine type, region, dine-in or delivery, and table booking requirements.
- Integrated real-time **data visualization** to provide users with detailed restaurant recommendations, average expenditure estimates, and cost range insights for two people.

ConsultBot – RAG based LLM for Case Study Practice -

- **Exposure:** Python, OpenAI, Docker, Pathway, **Retrieval Augmented Generation (RAG)**, GenAI, LLM, Streamlit (Link)
- Developed a RAG-powered chatbot using Pathway to enable interactive case study practice for consulting, analytics, and product management interviews, incorporating IIM casebooks. Built a user-friendly **Streamlit dashboard** for easy interaction, allowing users to practice case studies.
- Ranked among the top in the LLM Bootcamp, earning a certificate and prizes for the impactful application built.

Voice-Controlled Mecanum-Based Forklift (VOCMEF) -

- **Exposure:** Mechatronics, IoT, Robotics, Machine Learning, NLP, Speech Detection, Arduino, Proteus, CAD (Link)
- Ideated and developed VOCMEF, a voice-controlled forklift with Mecanum wheels and robotic arms to optimize stability and reduce human injuries in material handling.
- Integrated **Speech Detection** using Fourier Transform for voice command recognition, leveraging **Natural Language Processing (NLP)** for efficient control commands.
- Ranked 6th out of 250+ ideas at the Ideeavoltz IIT Roorkee competition and 3rd in the Robotics Conclave at Technex, IIT BHU.