Vansh Raja

12C, B-Wing, Whitespring, Rivali Park, Magathane, Mumbai-400066

Ph: +91 7021954565; E-mail: vanshraja32@gmail.com, Site: vanshraja.me

ACADEMIC QUALIFICATION

Bachelor of Engineering in Computer Science with specialization in AI and ML, Vellore Institute of Technology, Bhopal; Aggregate/CGPA: 8.87; Jul'25

INTERNSHIPS

Data Analyst Intern, PreProd Corp, Bangalore **Nov '23 – Jan '25**

PreProd Corp is a company that empowers individuals and businesses to build ethical, sustainable AI through up-skilling, consulting, and product innovation, fostering collaboration and advancement in AI technology.

- Designed and **implemented robust MLOps (Machine Learning Operations) pipeline applications** by streamlining the **deployment using DVC (Data Version Control)** and monitoring of machine learning models with **MLFlow**
- Contributed to **creating open-source resources on databases,** and other **machine learning tools** for the **PreProd Opnsrc initiative** to fostering community learning and engagement, by democratizing industry level machine learning projects like **MLFlow, Pyspark and algorithms like RAG, GAN, Mobilenet**, etc.
- **Led 6 teams** in developing **30+ open-source projects** (for most of the new industrial libraries like PySpark, Pandas, Polars, CV, RAG, GANs) as **Team Leader and Scrum Master,** conducting weekly sprint meetings and guiding teams achieve the goals as set by the company, promoting MLOps best practices, and hands-on support to enhance learning outcomes

Tools Used: PowerBI, Tableau, Apache Kafka, Spark, PySpark, Scikit-Learn, DVC, SQLite, MongoDB, MLFlow, Flask, Pandas, NumPy, and Pickle

ACADEMIC PROJECTS

Title: Food Demand Forecasting
Duration: Nov'23 – Feb'24
Team size: 7 members
Summary: Designed a time-series model to forecast food demand with high accuracy. Implemented an MLOps pipeline to enhance project efficiency. The model significantly improved food demand prediction accuracy, optimizing inventory management and reducing food wastage.
Tools Used: Kafka, MLFlow, DVC, Scikit-Learn, Flask

Title: Calpred
Duration: Nov'23 – Jan'24
Team size: 3
Summary: Developed a calorie prediction framework to accurately estimate caloric expenditure from workout data, leveraging advanced machine learning techniques and interactive visualization. Designed a robust system with a user-friendly interface to provide real-time calorie tracking and insights. The system delivered accurate, real-time predictions of caloric expenditure, empowering users to monitor workout performance, track fitness progress, and make data-driven decisions to meet fitness goals
Tools Used: Flask, SQLite3, Plotly, Dash, Pandas, NumPy, Matplotlib, XGBoost, TPOT, Seaborn, Gunicorn, HTML CSS

Title: CozyReader
Duration: Apr'24 - May'24
Summary: Created an AI-powered bedtime storyteller that crafts personalized bedtime stories based on a child's preferences and specified parameters, like custom characters, themes, environments, and custom locally hosted text-to-speech voices. The system generates cozy, fully AI-produced audio narrations,

offering a warm and engaging storytelling experience. Delivered user-friendly storytelling platform that offers personalized audio narrations with customizable elements, enhancing bedtime routines by making storytelling interactive, engaging, and uniquely tailored to individual preferences.
Tools Used: Google Generative AI API, Various TTS models from HuggingFace, CouchDB, Python, Streamlit

AWARDS & ACHIEVEMENTS

Awarded the **1st place** in the Electronics Technical Quiz at the XPOTECH event organized by VITronix, VIT Bhopal, Feb'24

WORKSHOPS

Attended the Industry AI Workshop on MLOPS Fundamentals, conducted by Krishnav Dave of PreProdCorp held at Bangalore, Oct'24

CERTIFICATIONS

- "Privacy and Security in Online Social Media" by NPTEL, Apr' 24
- "Artificial Intelligence and Machine Learning", by Google Developers and SmartInternz, Nov'23
- "Applied Machine Learning in Python" by University of Michigan, Feb' 23

TECHNICAL SKILLS

- **Languages:** Python
- **Data Science/ML:** NumPy, Pandas, Polars, Scikit-learn, TensorFlow, PySpark, Dask, H2O AutoML, PyCaret, TPOT
- **Web & MLOps:** Flask, Streamlit, FastAPI, MLflow, Docker
- **GenAI:** Langchain, OpenAI/Google AI APIs, Groq, Ollama
- **Data Engineering**: Hadoop, Kafka, SQL, PostgreSQL, Cassandra, MongoDB, Redis, Neo4j, CouchDB
- **Cloud & Viz:** Oracle Cloud, Matplotlib, Seaborn, Plotly, Tableau
- **Tools:** Git, Agile/Scrum
- **OS:** Windows, macOS, Linux

EXTRACURRICULAR ACTIVITIES

- Assisted in designing, organizing, and executing the "Buildathon" Hackathon, as **co-organizer** to develop an AutoML application, promoting innovation and practical skill-building, PreProd Corp, VIT Bhopal, Aug '24
- VITBMUN 24' **Vice Chair UNSCW**: Participated as a Vice Chair on the UNSCW committee to successfully guide the delegates to pass a draft resolution on the agenda: "Reproductive Rights Under Threat: Backlash Against Abortion Access Globally"
- PWSkills Ambassador: Helped in organising and promoting technical skills in the VITB Campus.