Raj Kumar Maurya

📞 +91-9044883800 @ therajmaurya@gmail.com

 GitHub   in LinkedIn 🌐 Website ✹ Certifications 🎖 Badges

WORK EXPERIENCE (∼6 YEARS)

- **Senior Data Scientist, Ernst & Young Global Delivery Services (EY GDS)**   *(July 2022 - Present)*
  - Responsible for building end-to-end **Generative AI, NLP, and Computer Vision** solutions
  - **Managed** teams of up to **10 professionals** across diverse sectors including pharmaceuticals, automobiles, and real estate
  - Engaged with stakeholders to analyze business opportunities and propose strategic approaches and solutions
  - **Fine-tuned** FLAN-T5, Falcon-40B, etc. open source LLMs using techniques like **Quantization, LoRA, QLoRA,** etc.

- **Data Scientist, Diwo (DataFactZ) Pvt. Ltd.**   *(July 2018 - June 2022)*
  - Responsible for building end-to-end natural language processing solutions for the **Diwo ASK Platform**
  - **Led and guided teams of up to 3 professionals** in platform initiatives and engagements spanning retail, insurance, and finance domains
  - Worked with **cross-functional teams** to gather data, domain knowledge, handle dependencies, perform overlapping tasks, and deploy models in production
  - Worked on model explainability and trust aspect for client-specific projects
  - Performed **causal inference analysis** using the What-If Tool to find out several impact points

- **Data Science Intern, Tidyquant Pvt. Ltd.**   *(Jan 2018 - Apr 2018)*
  - Implemented various Machine Learning algorithms for tasks like Regression, Classification, Clustering, Time series analysis, and Association rules mining using Python for a wide range of problems on masked data
  - Used the concept of **transfer-learning, ensembling** (stacking, bagging, boosting), feature importance, feature selection, PCA, etc. to enhance model performance

EDUCATION

- **Master of Technology (Data Science)**
  Computer Science and Engineering   *July 2022 – June 2025*
  Indian Institute of Technology (IIT), Hyderabad   *Ongoing*

- **Bachelor of Technology**
  Information Technology   *July 2014 – June 2018*
  Dr. A.P.J. Abdul Kalam Technical University, Lucknow   *Percentage: 78.1*

INDUSTRY PROJECTS

- **EY: Infrastructure-as-Code (IAC) Script Generation [GenAI]**   *(Sep 2023 - Present)*
  - Enhanced IAC compliance and adoption, reducing scripting time from weeks to days; developed a **conversational chat interface** for personalized script generation; implemented validation and mitigation mechanisms
  - Utilized **RAG**-based approach, signature matching, slot-filling, and multi-stage few-shot prompting
  - Achieved over **50% reduction in manual** IAC provisioning effort and improved overall adoption

- **EY: Photo-review Automation [Computer Vision]**   *(Jul 2022 - Sep 2023)*
  - Enhanced customer experience on a resale car company's website by accurately associating and reviewing vehicle images, **minimizing overall review effort**
  - Led development of an end-to-end solution using transfer learning and pre-trained models for precise car image classification
  - Employed Text and Object Detection Models, achieving a **92% reduction in manual photo review** effort
  - Integrated and deployed 16+ models including **VGG16, ResNet, YOLOv5, YOLOv8, SVM,** etc., using CI/CD in Azure DevOps

- **EY: Functional Testing Automation [GenAI]**   *(Apr 2023 - Oct 2023)*
  - Automated end-to-end functional testing framework using GenAI with Human-in-the-loop; generated **intelligent test scenarios and test cases** based on custom web pages
  - Developed a GenAI-powered Synthetic Data Generation Tool for test cases, Web-page Element Locator Extractor Tool, and integrated a chat interface for fine-tuning outputs
  - Utilized **Selenium scripts generated by GenAI** for automated functional testing; enabled controlled updates to existing test cases based on partial web page changes
  - Expected to **enhance testing efficiency by 30-50%**

- **Diwo: Related Questions Suggestion [NLP]** *(Sep 2021 - May 2022)*
  - Reduced the barrier for users to interact with the system by suggesting cold start questions and context-aware questions during an active session
  - Built **context-aware prediction** of questions that the user might be interested in asking next based on their past interaction, behavior, and questions
  - Conceptualised and designed the end-to-end architecture and approach from scratch
  - The modeling also encompasses **user permission level security** in post-processing and multi-use case aspects in pre-processing with the help of ontology.
  - The **click rate** of the suggestions was close to **30% across use cases**.

- **Diwo: Named Entity & Action Recognition [NLP]** *(Dec 2020 - Sep 2021)*
  - Transitioned the unreliable manual ontology-based custom logic with a more reliable machine learning-based entity and action recognition
  - Enhanced the existing rule-based implementation and made it intelligent using **transfer learning of NER from the Spacy** model for complex entities
  - Worked on user state management and conflict resolution of entities
  - **Improved the overall intelligence** of the system w.r.t named entity and action recognition

- **Diwo: Command Handling [NLP]** *(Jul 2020 - Dec 2020)*
  - Substituted the keyword match-based query classification with an ML-based approach
  - Built a model to classify commands from non-command questions and then to **classify the type of commands** and then extract relevant entities from the question based on its intent-action pair
  - **Fine-tuned BERT model** for sub-command classifications and built BERT embeddings-based MLP model for binary classification

- **Diwo: Category Expansion/Rationalisation [Core ML]** *(Jan 2020 - Jun 2020)*
  - **Introduced new products** based on the performance of past products of a similar category and predicted its **revenue potential** along with cannibalization impact
  - Used floor-set information to break down the data and extracted similarity features in the regression model for baseline prediction and to study cannibalization

- **Diwo: Paid Rate Prediction for quotes offered by insurer [Core ML]** *(Nov 2018 - Dec 2019)*
  - Predicted whether a customer will accept the quote offered or not within the specified periods, to **streamline the marketing strategy** and target the potential customer
  - Built the model, handled imbalance and **performed causal inference** for the client to find the list of KPIs affecting the quote acceptance

TECHNICAL SKILLS

- **Languages**: Python, SQL, C/C++

- **Data Science**: Statistics, Machine Learning, Deep Learning, Natural Language Processing, Computer Vision, MLOps

- **Generative AI (GenAI)**: Large Language Models (LLMs), Attention & Transformers, Stable Diffusion, OpenAI GPT, Langchain, VectorDB, LLM Agents, Prompt Engineering, Chunking Strategy, LLM Context Management, Memory Management, and State Management, Retrieval-augmented Generation (RAG), LLM Finetuning, LLMOps, Model Optimization for GPUs

- **Tools/Libraries**: Pandas, Numpy, Scikit-learn, PyTorch, Keras, Spacy, NLTK, Matplotlib, Plotly, Flask, Streamlit

- **Others**: MySQL, Git, LaTeX, Jira, Confluence, Docker, AWS, Azure, HuggingFace, MLFlow, HTML/CSS

ACHIEVEMENTS AND ACTIVITIES

- Received multiple **Spot awards** at EY. Recognised as the **best employee** of the month twice at Diwo

- Among 300+ participants worldwide to complete **IBM Quantum Computing Challenge** 2020 (Advanced)

- Participated in the Smart India **Hackathon** organized by MHRD, AICTE, India (2018)

- Authored and **published the research paper "Machine Learning in Big Data Analytics"** as the first author in the International Journal of Recent Research Aspects (IJRRA ISSN: 2349-7688)

- Member of college society – Big Data Centre of Excellence (2016-18) [**Technical Head** during 2017-18]

- **Organised and mentored workshops** on Big Data Science technologies for three consecutive years at BDCOE