



Akhil Negi  
Manager  
Data & Analytics

396-B, Pocket-2,  
Mayur Vihar, Phase-1  
New Delhi  
+91-9555452237  
writetoakhilnegi@gmail.com

Education

G.L. BAJAJ INSTITUTE OF  
TECHNOLOGY  
Bachelor of Technology  
(Electronics and Communication)  
Graduated May 2012

ASN Sr. Sec. School  
Senior Secondary (CBSE)  
Graduated March 2007

Skills

Programming language

Python:	★★★★★
Tools	
Jupyter	★★★★★
Anaconda	★★★★★
Pycharm	★★★★★
VSCode	★★★★★
ArcGIS	★★★★★
QGIS	★★★★★

DL/ML Stack

DL algorithms	★★★★★
Pytorch	★★★★★
Fastai	★★★★★
Transformers	★★★★★
Keras	★★★★☆
Pyspark	★★★★☆
Pandas	★★★★★
ArcGIS Python API	★★★★★
GDAL	★★★★☆
Langchain	★★★★☆
Llama Index	★★★★☆
Unstructured	★★★★☆

Cloud platforms:

AWS	★★★★★
Azure	★★★★☆
GCP ML APIs	★★★★☆

OS / Virtualization:

Linux	
Windows	
Docker	

Other tools:

ArcGIS Pro	
Redis Cache	
MS Project Server	
Sharepoint 2013	
SQL	

Interests

Trail running, hiking.  
Deep learning evangelism.

Experience

EY Global Delivery Services  
Manager – Data and Analytics  
December 2021 – Present

**Generative AI Center Of Excellence setup:**  
Successfully set up a fully operational Generative AI Center of Excellence.  
Implemented a strategic roadmap for generative AI applications, resulting in the development and deployment of key Generative AI projects.

Core Responsibilities:

- Established Core workstreams: GenAI training and enablement, Innovation, Generative AI factory, GenAI SOPs and best practices.
- Spearheaded the development of distinct workstreams within the COE, involved in shortlisting nominations, and defined the core responsibilities and strategic vision for each workstream.
- Strategic Planning and Execution: Developed a comprehensive two-month backlog for each workstream, ensuring alignment with long-term business goals and facilitating smooth initial operations.
- Resource Allocation and Funding: Managed prioritization for Generative AI initiatives, focusing on the development and delivery of prioritized use cases to maximize return on investment and enhance operational efficiencies.

**Led the design and development of MVPs with generative AI models, directly impacting key areas such as:**

- Enterprise knowledge management and discovery, enhancing information accessibility and improving decision-making.
- Optimizing contact center operations through AI-driven insights, boosting customer service and operational efficiency.
- Application leveraging RAG, LLM Agents and tools:
  - ESG benchmarking
  - Code documentation and translation.
  - NL query for structured data.
  - Document comparison and disclosure validation.
  - Web search based cataloging tool.

Core Responsibilities:

- Helping customers with analyzing their current processes and identifying transformation areas.
- Requirement analysis, discovery, drafting high-level and low-level application design.
- Going through the latest state of the art research.
- LLM and framework selection.
- Training data preparation.
- Pre-processing pipeline design and development.
- Evaluation pipeline design and development.

Tech stack: Azure OpenAI, Azure app service, Opensource LLMs (Ilama, mistral, etc.), whisper, Azure STT/TTS, Azure AI search , Langchain, LangSmith, LlamaIndex, peft, evaluate, ragas, promptfoo, vLLM, bitsandbytes, tiktoken, Lamini, huggingface-transformers, chromaDB, redis, FAISS, Streamlit, FastAPI.

**Implementation of a transactional fraud detection strategy for a leading UK bank:**  
Worked as a Sr. Data Scientist for customization and implementation of transaction fraud detection strategy for a large UK bank

Core Responsibilities:

- Data feed validation, mapping and transformation.
- Data preparation for model selection and training.
- MVP and production model building
- Building fraud strategy rules for augmenting model decisions.
- Shadow model monitoring and production deployment.
- Live model monitoring.

Tech stack : LightGBM, XGBoost, Random Forest, Python, Pyspark, JupyterLab, GCP DataProc, Hadoop

**Lead the development of a Data Monetization platform for a major power and utility sector client:**

- Analyzed power consumption data for millions of customers across different regions to come up with high quality reliable insights.
- Contributed to the data monetization strategy by developing high quality spatial trends and predictions to be monetized as Data Insight reports and data enrichment API.
- Leveraged spatial analytics, countrywide power consumption data (over a decade), powerful regression models to generate different insightful reports:
  - Up and coming neighborhoods.
  - Real-estate development trends.
  - Commercial and residential development trends.
  - Geographic socio-economic indicators.

Tech stack: QGIS, SAP BW, Python, JuPyter, Folium, Streamlit, Pytorch, scikit-learn, GDAL, Shapely, GeoPandas

**Lead the development and implementation of a Document AI solution for a significant client in the Chemical (procurement) sector for transforming their raw material procurement pipeline:**

Core Responsibilities:

- Lead the project from conceptualization to deployment, ensuring the solution effectively streamlined procurement processes.
- Data ingestion pipeline design and development on Azure.
- Database modelling and UX design.
- Led the training data labelling effort with 6 junior DS and experts chemists from the client team. Output: 50k labelled documents.
- Optimized the labelling data process by developing AI assisted labelling approach.
- Drove model pipeline development. As part of which 16 classification and ner models were trained and iterated to achieve acceptable performance.
- Setting up a feedback loop to retrain models with user feedback.

Tech stack: Azure logic apps, Azure functions, Azure language studio, Azure Form-recognizer, Python, Huggingface-transformers, Doccano, CosmosDB

**Contributed to business development efforts by:**

- Initiating and executing multiple Natural Language Processing (NLP) and Computer Vision (CV) POCs for clients across different industries, showcasing the practical benefits of AI technologies.
- Utilizing technical and project leadership skills to foster company growth and widen its client base with innovative AI solutions.

**ESRI New Delhi R&D center (Environmental Systems Research Institute)**

Data Scientist 2

May 2019 – November 2021

- Played a pivotal role in the ArcGIS Python API team, integrating cutting-edge deep learning models for processing imagery, unstructured text, and 3D point cloud data.
- Served as Product owner for the Python API text module, leading agile practices and facilitating the seamless development and delivery of product enhancements.
- Drove the research and development of new datasets, spearheading the training of production-scale deep learning models on multi-GPUs. These models were primed for publication on the ArcGIS Living Atlas, showcasing a blend of technical acumen and strategic project direction.
- Assumed the role of NLP specialist, overseeing the curation and maintenance of the text sub-module within the Python API, ensuring the delivery of robust NLP capabilities.
- Identified and pursued opportunities for enhancing the existing product ecosystem with ML/DL technologies, demonstrating a strategic approach to product development and innovation.
- Led the deployment of ready-to-use models as web services on cloud infrastructure, facilitating the operationalization of deep learning models for broader accessibility and impact.
- Collaborated across teams to embed deep learning solutions within various products:
  - Few examples of DL inclusion:
    - Enhancing search and discoverability in ArcGIS Online.
    - Implementing solutions for the redaction of personally identifiable information from public documents, thereby elevating privacy and security measures.
- Championed the Python API and the team's innovations through presentations at Webinars, ESRI Developer Summits, and user conferences, effectively communicating the value and applications of our work to a wider audience.

**Tata Consultancy Services (Life Sciences- TEG)**

Data Scientist

Jan 2017– Mar 2019

- Engaged with stakeholders to comprehensively understand their processes, laying the groundwork for targeted improvements.
- Analyzed the existing workflows to pinpoint opportunities for ML/NLP applications, aiming to enhance operational efficiency.
- Collaborated closely with fellow developers to extract and prepare the necessary data for analysis and modeling.

- Developed and enforced data labeling guidelines, overseeing the manual labeling efforts carried out by my team for training data preparation.
- Outlined the technical infrastructure needs, leading the setup of Virtual Machines (VMs) on the Google Cloud Platform to support our computing requirements.
- Architected and implemented a data cleansing and preprocessing pipeline, ensuring data quality and readiness for model training.
- Conducted thorough research on cutting-edge deep learning architectures and academic papers, adapting state-of-the-art solutions to meet our specific project needs.
- Took charge of model development, from initial building and training to fine-tuning and establishing deployment pipelines, ensuring models are ready for production use.
- Provided ongoing support to the team, troubleshooting and resolving challenges encountered during the model training process.
- Focused on team capability enhancement by hiring and mentoring new ML talents, fostering a skilled and knowledgeable workforce.

**Tata Consultancy Services (Ameriprise Financial)**

Sharepoint developer/Project Server SME

Mar 2013 – Sep 2016

- Design, develop, and manage site content and capabilities and provide daily onsite maintenance of the existing SharePoint portals.
- Design intelligent form-based applications (InfoPath) as per business requirements.
- Work with Project Managers, providing accurate development estimates as requested.
- Ensure systems change management is carried out in accordance with agreed standards and procedures on Service Now.
- Walkthrough of the design and code to the Production support team and Change management team.
- Reporting of status to the client through weekly meetings.
- Act as Project Server SME for Project Management Group.
- Assist PMs with the maintenance of project pages in SharePoint.