

# NIKHIL KUMAR SHARMA

Data Scientist  
Reliance Jio

+91-9968-944-825 | Bengaluru | [nikhil.sharma1294@gmail.com](mailto:nikhil.sharma1294@gmail.com) | [LinkedIn](#) | [Portfolio](#)

PROFESSIONAL SUMMARYTOTAL EXPERIENCE: 7 years

Innovative Data Scientist with **7+ years** of expertise in **Generative AI, NLP, and Large Language Models (LLMs)**. Proven track record of deploying scalable **AI solutions (chatbots, STT/TTS systems)** at Reliance Jio, Cognizant and Capgemini, increasing workflow efficiency by 70% - 80% while maintaining the quality and accuracy. Proficient in **Python, PyTorch, Azure/GCP, and open-source LLMs (LLAMA, Mistral)**. Passionate about transforming complex data into actionable insights for **telecom, banking & finance and automotive industries**.

## SKILLS AND TOOL SETS

- **Languages:** Python, SQL, Pyspark.
- **Generative AI:** OpenAI, LLAMA, Mistral, Langchain, PyTorch, Tensorflow, Gemini, HuggingFace, SST and TTS.
- **Machine Learning:** Supervised & Un-supervised, NLP, Clustering, Deep Learning, Scripting and Automation.
- **Cloud & DevOps:** Azure AI, GCP, Docker, Jenkins, CI/CD.
- **Databases:** Chroma DB, Faiss, Qdrant, Pinecone, MySQL, PostgreSQL, MongoDB.
- **Data Visualization:** Power BI, Matplotlib, Seaborn, Plotly, DASH.
- **Tools & Software:** Azure AI services, GCP, Jupyter Lab, VS code, Anaconda, Spyder.
- **Domain:** Telecom, Banking & Finance, Automotive and Network Security.

## EXPERIENCE

■ RELIANCE JIODATA SCIENTISTOct 2021 - Present

- **Generative AI | LLM based Solutions**
  1. Developed a **RAG-based chatbot using OpenAI GPT, fine-tuning GEMMA and LLAMA models** to enhance Jio Roaming Consultants' turnaround time by 70%.
  2. Designed a custom **STT** solution by **fine-tuning Whisper-tiny and Whisper-large-v3 models**, enhancing transcription accuracy, reducing latency, and eliminating external dependencies to increase revenue.
  3. **TTS (Text-to-Speech) for Jio Translate Platform**
    - 3.1. Built a high-quality **TTS** using **Fish TTS model**, ensuring natural speech in Hindi and cutting external service costs.
    - 3.2. Engineered a custom pipeline for **Jio Translate Platform** utilizing the **Parler-TTS, MetaVoice** and **RVC** model, enabling personalized voice synthesis with high speaker similarity.
    - 3.3. Integrated **RE** (reverberation effects) and **UVR** (Ultimate Vocal Remover) for voice separation and noise reduction, improving audio clarity across environments.
- **Network Monitoring and Optimization Framework**
  1. Implemented **network monitoring dashboards** to identify anomalies across India for proactive identification of areas where network issues are present.
  2. **Connected Components**  
Clustering of telecom affinity data on the geospatial region to get the data segregated cluster to work on e-Tilt optimization.
  3. **e-Tilt Optimization (cvxOpt)**  
Implemented a custom in-house developed optimization algorithm to optimize the e-tilts of the cells on the network tower with the sole aim to improve the user's network experience.

- **Computer Vision - Hand gesture recognition with face validation**  
Build 5 class hand gesture classification model with face validation. Along with pipeline to generate the hand and face data for model training.
  - Google's MediaPipe is to generate features of hand classification.
  - Face images are used to train a separate Sequential deep learning model of face validation.

**Tech Stack Used:** LangChain, Transformers, Gen AI, HuggingFace, Pyspark, spectral clustering, Jupyter Lab, data transformation, HDFS, Git, Kepler-maps, Deep Learning, Neural Networks, etc.

<b>COGNIZANT</b>	<b>DATA SCIENTIST CONSULTANT</b>	<b>Sep 2020 – Oct 2021.</b>
------------------	----------------------------------	-----------------------------

**Client:** Mercedes-Benz Research & Development India (MBRDI) – Global Software Engineering Platform (GSEP)

- Created NLP models for spelling correction and chatbot performance dashboards.
- Developed price simulation dashboards for Mercedes-Benz using Azure & Plotly Dash.
- Automated Data Extraction Pipeline with Jenkins and Postgres SQL | Cron Job triggered daily at 00:00

**Tech Stack Used:** Python, NLP, Plotly, Dash, SQL on Azure, Azure Services, Azure Devops, Spello Language Model (based on SymSpell-Fuzzy Search), Postgres dB, Docker, Network error handling.

<b>CAPGEMINI</b>	<b>DATA SCIENTIST CONSULTANT</b>	<b>Nov 2018 – June 2020.</b>
------------------	----------------------------------	------------------------------

**Client:** ANZ – Bank Guarantee for Trade & Finance – Digital Transformation

- Built document classifiers & NER models for ANZ Bank, automating form processing by 80%.
- AI enabled automation of Bank Guarantee Generation.
  1. **Document classification** of Bank Documents (ex. BG, Request Forms, ASIC, CLAA, Email, Others).
  2. **Entity extraction** from different documents like Applicant Name, Beneficiary Name, ABN/ACN Numbers, Issue Date, Expiry Date, Purpose etc. from Request forms, Bank Guarantees and Emails.
  3. **Automated insertion of extracted entities** (Applicant/Beneficiary Name, ABN/ACN, Issue/Expiry Date) into ANZ Bank Guarantee Templates and client documents ensuring compliance and accuracy.

**Processes:** Abby + Data Cap OCR, XML Parsing, Text Cleaning, data preparation, Text Analytics, Feature Engineering, Algorithm Selection, data tagging, model building, model training, and model validation.

**Tech Stack:** Spacy model/LSTM + Rule Engine + Regex + Logic-based solution to automate the form-filling process by the Maker and Checker that reduced the time required by 80% on an average.

**Capgemini AI-COE**

- Contributed to Cognitive Document Processing (CDP) platform with OCR & NLP capabilities.  
OCR, Text Extraction, Text Classification, Language Detection, Natural Language Generation.
- **PII Redaction**  
Developed a tool for automated PII masking in legal documents, ensuring compliance with GDPR.

<b>ADAPTIVE SECURITIES</b>	<b>PYTHON DEVELOPER</b>	<b>April 2018 – Nov 2018.</b>
----------------------------	-------------------------	-------------------------------

- Built a **web scraping** pipeline and CLI tools output parser to **extract critical domain intelligence** (DNS records, server fingerprints), enabling identification of DDoS attack vectors, finding vulnerabilities and mitigation strategies to fortify network defences.
- Developed and tested different kinds of **DDoS** prevention and attack programs/frameworks.

<b>EDUCATION</b>		
------------------	--	--

Graduation	HMRITM, Delhi (B. Tech)	2012 – 2016
12 <sup>TH</sup>	Kendriya Vidyalaya, Delhi (CBSE)	2011 – 2012
10 <sup>TH</sup>	Kendriya Vidyalaya, Delhi (CBSE)	2009 – 2010