

# SOWJANYA BOJJA

6692974674 • simplysowj@gmail.com • LinkedIn • GitHub • Website • LLM Portfolio • Medium

Data Science and AI/ML professional with expertise in generative AI, machine learning models, data analysis, and big data engineering. Skilled in Python, SQL, Kafka, PySpark, and visualization tools. Proven ability to implement end-to-end ML solutions, enhance algorithmic efficiency, and collaborate effectively in team environments. Passionate about innovation and deploying scalable AI systems.

## Technical skills

**Programming:** Python, SQL, Java, R

**Machine Learning:** Regression, SVM, Decision Trees, XGBoost, LightGBM, Clustering, Hyperparameter Tuning

**Deep Learning:** CNNs, RNNs, LSTMs, GANs, GPT, BERT, T5

**NLP:** Tokenization, Stemming, Lemmatization, TF-IDF, Word2Vec, BERT

**Big Data:** PySpark, Kafka, Data Pipeline Development, Nosql

**Data Visualization:** Tableau, Power BI, Matplotlib, Seaborn, Excel

**Tools:** TensorFlow, PyTorch, Scikit-learn, Docker, JupyterLab, Streamlit, Autogen

**Cloud:** Google Colab, Microsoft Azure, AWS SageMaker, AWS AutoGlue

## Certifications

Data Science, IBM Registered Business Partner – June 2022

Multi Agent Ai Systems with CrewAI and Autogen and DeepLearning.AI – Nov 2024

Generative AI and Learnathon Genai, GUVI Geek Networks, IITM Research Park – Sep 2024

Data Visualization with Tableau, SQL, Advanced Excel, Internshala – Oct 2023 to Nov 2023

Google Data Analytics – Dec 2024 to Jan 2025 [link]

Microsoft Azure AI/ML Engineering – Jan 2025

Microsoft Power BI Data Analyst

Generative AI for Data Science [link]

Intro to GenAI [link]

AWS Technical Essentials (AWS Skill Builder)

Statistics for Data Science, coursera

GenAI for Data Analytics [link]

## Education

Master of Science in Data Science **Certificate**

Chandigarh University, India | Jul 2022 - Aug 2024

GPA: 8.0

Key Coursework: Python, Calculus, Applied Probability, Data Analysis & Visualization, SQL, Machine Learning, Deep Learning, NLP, Cloud-Native Development, Java, Data Structures, Business Analytics, Data Mining, and Data Engineering.

Bachelor of Technology in Electronics and Communications

Jawaharlal Nehru Technological University, India | Jan 2002 - Jan 2006

## Experience

### Gen AI Intern | MENTORNESS

Developed the ATS Resume Expert app using generative AI. Conducted research on emerging AI technologies to enhance development.

**ML Intern | Bharat Intern**

Built and deployed machine learning models to predict customer behavior. Evaluated model performance using accuracy, precision, and recall metrics.

**Data Science Intern | Cognifyz Technologies**

Applied ML algorithms to large datasets for classification and regression tasks. Conducted exploratory data analysis to identify patterns and trends.

**Data Analyst Intern | Internshala Trainings**

Developed data analysis projects using SQL, Tableau, Python, and Excel. Created visualizations and dashboards to track business metrics.

**Data Science Research Intern | Innomatics Research Labs**

Applied advanced ML algorithms and BERT vectorization to develop models. Utilized MLflow and Prefect for experiment tracking and model management. Deployed solutions on Heroku and GitHub.

**Systems Engineer | Tata Consultancy Services**

Developed PL SQL stored procedures, functions, and packages. Designed database objects (tables, views, triggers) to meet business requirements. Tools: TOAD, PL/SQL Developer.

**Career Break | Feb 2014 - Jan 2021**

Managed full-time caregiving for a family member, honing organizational, time management, and multitasking skills.

**Projects****Gen AI & LLM Projects:**

- Rag with Graph Database [link](#) , Rag with VectorDB [link](#) , Multi-Agent AI Systems with CrewAI [link](#)
- Data Balancing with Gen AI (Credit Card Fraud Detection with GAN, PCA, TensorFlow)[link](#)
- ATS Resume Expert App [link](#) , Fraud Detection Using Big Data (Kafka & PySpark)[Link](#)
- Personal Portfolio using LLM [link](#)
- Image Captioning with OCR and GenAI, [Link](#) , Integrated OCR and Generative AI

**Web & Cloud Applications:**

- Azure Speech-to-Text Web App [link](#), ML with PySpark: Recommender System [link](#), Semantic Segmentation on SageMaker [link](#)
- Image Classification with AWS AutoGluon [link](#), LLMops, Personal Portfolio using LLM [link](#)

**NLP Projects, [Link](#)****Java & ML Projects:**

- Microservices with Spring Boot, Quora Duplicate Question Detector (End-to-end MLOps) [Link](#)
- Quora Duplicate Question Detector, [Link](#)

**Data Analysis:**

- Data Analysis Portfolio (SQL, Excel, Tableau, Python, PowerBI) [Link](#)
- Sentiment Analysis and EDA with Databricks, [Link](#) ,Classified and analyzed sentiment in text data

**Awards & Publications**

- **Letter of Recommendation**, Innomatics Research Labs (Nov 2022)
- **Fraud Detection with Gen AI & Big Data Engineering**, Medium, Jan 2024 [link](#)
- **Multimodal Image Captioning with OCR and LLM**, Medium, May 2024 [link](#)