SANJANA RAJESH PISAL

sanjana@itjobworks.com | +1 217-986-0387 | Springfield, IL | LinkedIn

SUMMARY

Data Scientist and AI/ML Engineer with 4 years of experience delivering scalable machine learning, deep learning, and generative AI solutions across Healthcare, IT Services and IT Consulting and E-Commerce. Proficient in Python, R, SQL, TensorFlow, PyTorch, and Spark, with expertise in building cloud-native pipelines on AWS and GCP, developing predictive models, and fine-tuning LLMs for real-world applications. Skilled in fraud detection, risk scoring, NLP, and anomaly detection, with a strong track record of improving efficiency, compliance, and business outcomes. Adept at MLOps, Agile delivery, and cross-functional collaboration to deploy secure, production-ready AI systems that drive measurable impact.

EDUCATION

Master of Science in Data Analytics | **University of Illinois Springfield, Springfield, USA** May 2025

SKILLS

Methodologies: SDLC, Agile, Waterfall, A/B Testing, Experimental Design
Languages: Python, R, SQL, MATLAB, HTML, CSS
Frameworks & Libraries: TensorFlow, PyTorch, Keras, Scikit-Learn, NumPy, Pandas, Matplotlib, Flask, Transformers Architecture
Big Data & ETL Tools: PySpark, Apache Spark, Hadoop, MapReduce, HDFS, Spark Streaming
Databases: MySQL, PostgreSQL, SQL Server 2008, MongoDB
Data Visualization: Tableau, Power BI, Excel (Pivot Tables, VLOOKUP)
Machine Learning & Deep Learning: Linear/Logistic Regression, Clustering, SVM, PCA, Random Forest, Boosting, Lasso, Ridge, CNN, RNN, Fine-tuning & Transfer Learning
Generative AI & LLMs: Foundation Models (GPT, LLaMA, Claude), Prompt Engineering, Tokenization (BPE, WordPiece), Reinforcement Learning from Human Feedback (RLHF)
Data Management: S3, Glue
Tools & APIs: Postman, REST APIs, JSON Parsing, Git, Jenkins
IDEs: Jupyter Notebook, PyCharm, Visual Studio Code
Soft Skills: Problem-Solving, Communication, Collaboration

EXPERIENCE

Centene Health, USA | Data Scientist **Jan 2025 – Present**

- Developed and deployed ML models (Logistic Regression, Gradient Boosting, CNNs) in Agile healthcare environments for fraud detection, patient readmission prediction, and personalized care recommendations, reducing data processing time by 25% and improving patient outcomes.
- Engineered secure AWS-based data pipelines (Glue, Redshift, S3) for real-time claims analytics, clinical workflow optimization, and EHR data integration; delivered Tableau/SSRS dashboards, improving query latency and healthcare reporting efficiency by 30–40%.
- Designed and optimized PostgreSQL and Oracle databases to support claims processing, patient records, and provider network management with compliance-focused architecture (HIPAA, PHI) and performance-tuned queries.
- Built AI/ML solutions using SageMaker, Bedrock, and LangChain for disease risk prediction, patient demand forecasting, and NLP-driven clinical note summarization, improving decision support and healthcare service delivery.
- Integrated structured and unstructured healthcare data with RAG models and deployed GenAI chatbots using Bedrock for patient self-service, provider search, and personalized care guidance, increasing digital health engagement by 20%.
- Automated end-to-end model deployment pipelines with CodePipeline, Docker, and GCP Kubernetes, ensuring scalable, compliant (HIPAA, HITRUST), and audit-ready analytics workflows aligned with healthcare IT and enterprise data standards.

HCL Tech, India | AI/ML Engineer **Mar 2020 – Jul 2023**

- Designed and implemented deep learning models (CNNs, RNNs, LSTMs, Transformers) for anomaly detection, customer behavior prediction, and real-time personalization in e-commerce and IT service delivery, achieving 90%+ accuracy and reducing downtime for enterprise clients.
- Built scalable ML pipelines using AWS SageMaker, EC2, Lambda, Redshift, and S3 for large-scale data ingestion, demand forecasting, and transaction analytics, leveraging ARIMA, Prophet, and advanced time-series models to improve forecasting by 25%.
- Developed XGBoost-based risk scoring engines and NLP pipelines (spaCy, BERT) for unstructured e-commerce data such as product reviews, logs, and service tickets, automating decision workflows and improving operational efficiency by 20%.
- Led Agile AI/ML initiatives, partnering with consulting and software engineering teams to integrate ML solutions into enterprise platforms, reducing data processing latency by 15% and accelerating release velocity for global clients.
- Mentored junior engineers on MLOps best practices including CI/CD (Git, Jenkins), automated retraining pipelines, feature engineering, and deployment using AWS CloudFormation and Infrastructure-as-Code (IaC).
- Fine-tuned LLMs (GPT, Hugging Face, LangChain) for automated log summarization, intelligent ticket classification, and customer support chatbots, enhancing service desk productivity and reducing mean-time-to-resolution (MTTR).
- Containerized and deployed AI/ML models using Docker, Kubernetes, and Terraform across AWS and GCP, ensuring secure, scalable, and cloud-native deployment for IT consulting and e-commerce clients.
- Applied TensorFlow, PyTorch, and Keras for predictive analytics, fraud detection, computer vision (document/ID verification), and intelligent automation, enabling compliance with enterprise validation and regulatory workflows.

CERTIFICATIONS

- **Microsoft pl900**
- **Microsoft pl300 in progress**
- **Generative AI fundamentals-databricks**
- **Data visualisation using plotly -coursera**
- **Machine Learning pipeline using azure ml studio- coursera**