# DURGA PRASAD N

dprasad.datas@gmail.com | 203679-1183 | in

## PROFESSIONAL SUMMARY

Experienced **AI/ML Data Scientist Engineer** with over 5 years of success in designing and deploying **machine learning models**, **LLMs**, and **Generative AI solutions** across **healthcare, finance, and insurance domains**. Proficient in Python, TensorFlow, PyTorch, Scikit-learn, LangChain, and AWS SageMaker, with hands-on expertise in **NLP**, **computer vision**, and **predictive modeling**. Delivered impact-driven projects including **HIPAA-compliant virtual assistants**, **fraud detection engines**, and **personalized recommender systems** that boosted operational performance by up to **30%**. Skilled in **MLOps**, building scalable pipelines with Docker, Kubernetes, CI/CD, and deploying production-grade models across AWS, Azure, and GCP. Strong collaborator with a focus on aligning technical solutions to business goals while ensuring compliance with **FCRA, HIPAA, and GDPR**.

## SKILLS

- **Programming & Scripting:** Python, R, SQL, Go, Scala
- **Machine Learning & AI:** Supervised & Unsupervised Learning, Regression, Classification, Clustering, Time Series Forecasting, Deep Learning, Ensemble Models, Transfer Learning, Recommender Systems, LLMs *(GPT, BERT, RoBERTa)*, OpenAI API, Prompt Engineering, LangChain
- **NLP & GenAI:** Named Entity Recognition (NER), Sentiment Analysis, Topic Modeling, Text Summarization, Chatbots, PHI De-identification, BioBERT, SpaCy, NLTK, Hugging Face Transformers
- **Frameworks Libraries:** TensorFlow, PyTorch, Scikitlearn, Keras, XGBoost, LightGBM, OpenCV, MLflow, Pandas, NumPy, SciPy
- **MLOps:** FastAPI, Docker, Kubernetes, CI/CD Pipelines, MLflow, GitHub Actions, Jenkins, Model Monitoring, Model Drift Detection
- **Cloud Platforms:** AWS (S3, SageMaker, Lambda, Bedrock, ECS, Glue, Redshift, CloudWatch), Azure (ML Studio, Data Factory), Google Cloud Platform (GCP)
- **Big Data & Pipelines:** Apache Spark, PySpark, Hadoop, Airflow, Snowflake, ETL Pipelines, Data Lakes
- **Databases & APIs:** PostgreSQL, MongoDB, MySQL, NoSQL, FHIR, REST APIs, OpenSearch
- **Data Visualization & BI Tools:** Power BI, Tableau, Plotly, Matplotlib, Seaborn, Excel

## EXPERIENCE

**Cigna Healthcare** – Senior AI/ML Data Engineer                    **Bloomfield, CT • 10/2023 - Current**

- Designed and deployed a **HIPAA-compliant Generative AI platform** on **AWS** using **Bedrock** and **SageMaker**, reducing manual claim triage time by **30%** and boosting patient satisfaction scores by **25%**.
- Built a scalable **RAG pipeline** integrated with **FHIR-compliant EHR systems** and **OpenSearch** to provide **real-time, personalized patient insights** across digital touchpoints.
- Engineered **PHI de-identification workflows** using **AWS Glue**, **KMS**, and **IAM-based access control**, ensuring **100% compliance** with **HIPAA** and **HITECH** regulations.
- Developed a **fraud detection engine** using **XGBoost** and **deep learning models**, achieving **92% precision** and **89% recall** on real-time claim datasets.
- Created a **model drift detection framework** using **Kinesis**, **CloudWatch**, and **MLflow**, enabling **automated alerts and retraining** for evolving clinical models.
- Collaborated with **cross-functional teams** to streamline **MLOps workflows**, integrating **CI/CD** with **SageMaker Pipelines** and improving deployment time by **40%**.

**Wells Fargo** – AI Data Scientist                    **Phoenix, AZ • 08/2022 - 09/2023**

- Developed and deployed a **credit default prediction model** using **XGBoost** and **LightGBM**, improving **AUC-ROC by 28%** and enhancing **risk segmentation** for underwriting decisions.
- Built **explainable ML pipelines** using **SHAP** and integrated them with **model governance workflows** to ensure compliance with **FCRA**, **Basel III**, and **GDPR** regulations.
- Engineered 75+ **advanced financial features** from **transactional behavior** and **credit utilization trends** using **PCA** and **UMAP**, boosting model performance across loan products.
- Created a **hybrid recommendation engine** combining **collaborative filtering** and **BERT-based sequence models** to deliver **hyper-personalized financial insights**, increasing **engagement by 30%**.
- Partnered with **compliance, product, and engineering teams** to align **AI solutions** with **customer experience goals**, **regulatory standards**, and secure deployment on **AWS infrastructure**.

**Hexaware Technologies** – Data Scientist                    **Chennai, India • 06/2019 - 11/2019**

- Analyzed large-scale **policyholder datasets** using **SQL** and **Python (Pandas, NumPy)**, identifying key variables that influenced **customer renewal behavior**.
- Built and deployed **classification models** using **Scikit-learn** and **XGBoost** for **churn prediction** and **customer segmentation**, leading to a **20% improvement** in retention targeting.
- Conducted **A/B testing** and statistical analysis with **Statsmodels** and **SciPy** to measure the impact of **personalized pricing strategies**, contributing to a **15% increase** in renewal rates.
- Designed interactive **dashboards** in **Tableau** and **Power BI** to visualize KPIs such as **churn risk**, **policy lapse trends**, and **renewal success rates** for business stakeholders.
- Collaborated with **product and marketing teams** to translate insights into actionable strategies, supporting successful rollout of targeted campaigns and optimized pricing policies.

## TECHNICAL PROJECTS

**Fraud Pattern Detection Engine for Logistics Operations** *(Apache Spark, XGBoost, SQL, Power BI, Redshift)*
- Built anomaly detection models using Spark and XGBoost to identify high-risk delivery zones, behavioral outliers, and synthetic accounts; enabled fraud scoring in near real-time and improved detection precision by 35%.

**Clinical Question Answering System Using Retrieval-Augmented Generation (RAG) (LangChain, OpenAI, Amazon Bedrock, FHIR, OpenSearch)**
- Developed a scalable RAG pipeline integrated with FHIR-compliant EHR data to generate context-aware medical responses; enhanced information retrieval for care coordinators and reduced query resolution time by 40%.

**Credit Risk Scoring & Behavioral Segmentation Platform** *(LightGBM, SHAP, AWS SageMaker, PCA, UMAP, PostgreSQL)*
- Built explainable ML pipelines to classify credit default risk and segment borrowers based on financial behavior; improved risk prediction AUC by 28% while meeting regulatory transparency standards.

**LLM-Powered Virtual Health Assistant (Amazon SageMaker JumpStart, FastAPI, BioBERT, AWS Lambda, DynamoDB)**
- Designed and deployed a HIPAA-compliant chatbot that processed real-time clinical conversations and achieved 94% NER accuracy for symptoms, diagnoses, and medications; reduced support workload by 30%.

## EDUCATION

**Master of Science**: **University of New Haven |**                                        West Haven, CT • 12/2023
Data Sciences
**Bachelors of Technology**, Vignan University                                        06/2016 - 05/2020 | Guntur, India
Computer Science
**Post-Graduation Diploma**, IIIT Bangalore                                        01/2020 - 02/2021 | India
Data Science | West Haven, USA 06/2016 - 05/2020 | Guntur, India