Phaneendra Ganji

phaneendraganji95@gmail.com | +1 (561) 759-7282 | Washington, USA | LinkedIn | Portfolio

Summary

AI/ML Engineer with 4+ years of experience designing and deploying end-to-end machine learning, data engineering, and cloud-based AI solutions across finance, technology, and enterprise domains. Skilled in developing predictive models, recommendation systems, and NLP-driven applications using Python, R, SQL, TensorFlow, PyTorch, and Scikit-learn. Proficient in building ETL pipelines, automating MLOps workflows, and deploying scalable solutions with Spark, Airflow, Docker, Kubernetes, and cloud platforms (AWS, Azure, GCP). Adept at transforming complex datasets into actionable insights with Tableau, Power BI, and advanced analytics to drive innovation, improve decision-making, and deliver measurable business impact.

Technical Skills

- **Methodologies, Programming & Databases:** SDLC, Agile, Waterfall, Python, R, SQL, Go, Java, C, C++, MATLAB, Scala, MySQL, Oracle, Firebase, MongoDB, Redshift, Visual Studio Code, PyCharm, Google Colab
- **Libraries & Frameworks:** TensorFlow, PyTorch, Keras, Scikit-Learn, XGBoost, NumPy, Pandas, SciPy, OpenCV, NLTK, SpaCy, Regex, Matplotlib, Seaborn, Ggplot, ResNet-50, Flask, FastAPI, Django, RASA
- **Machine Learning & AI:** Supervised & Unsupervised Learning, Regression, Classification, Clustering, Logistic Regression, Random Forest, Decision Trees, Deep Learning (CNN, RNN, Autoencoders, GANs), NLP, Time-Series Forecasting, LLMs, OpenAI
- **Generative AI & Advanced Techniques:** RAG, LangChain, Agents, Vector DB, LoRa/QLoRa, Prompt Engineering
- **Data Engineering & Tools:** Apache Spark, Apache Airflow, Hadoop, Snowflake, Hive, Alteryx, ETL, Data Warehousing, Data Cleaning, Data Wrangling, Data Mining, Pattern Extraction, EDA
- **MLOps & DevOps:** MLflow, Docker, Kubernetes, Git, GitHub, GitHub Actions, Jenkins, Terraform, CI/CD Pipelines
- **Visualization & BI:** Tableau, Power BI, Microsoft Excel, Matplotlib, Seaborn, Ggplot, Data Visualization, KPI Dashboards, Google Analytics
- **Cloud Platforms & Statistical Analysis:** Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform (GCP), Hypothesis Testing, A/B Testing, Statistical Modeling

Professional Experience

AI Engineer, *iLink Digital Inc*   Jan 2025 – Present | Washington, USA
- Implemented anomaly detection using Microsoft Fabric and integrated an RCA chatbot with Azure OpenAI, streamlining IT operations and improving incident response time by 35%.
- Built a Flask-based invoice processing solution with Azure SQL and Blob Storage, reducing manual effort by 40% and enabling scalable document management.
- Containerized applications with Docker, deployed via Azure Kubernetes Service (AKS), and automated CI/CD pipelines, ensuring 98% uptime and consistent, error-free deployments.
- Engineered a GPT-4o–driven OCR pipeline, achieving 93% extraction accuracy and 94% categorization accuracy across diverse invoice layouts, enhancing automation efficiency.
- Partnered with business stakeholders to align AI prototypes with client objectives, enabling adoption of production-ready solutions that improved financial workflows and operational decision-making.

Cloud Solution Engineer, *Yogaya Inc*   May 2024 – Dec 2024 | Washington, USA
- Integrated an Alexa skill with AWS backend (S3, DynamoDB, API Gateway, Lambda), delivering seamless voice-enabled experiences for end users.
- Designed and implemented a scalable AWS data architecture, supporting rapid growth while maintaining high availability and fault tolerance.
- Optimized AWS resource management with auto-scaling and monitoring, reducing operational costs by 20% without sacrificing performance.
- Enhanced deployment reliability by implementing CI/CD best practices and automating infrastructure provisioning for faster, error-free releases.

Lead Machine Learning Engineer, *Omdena*   Jan 2024 – Mar 2024 | Washington, USA
- Directed a team of six to design and develop a Mental Health Chatbot using LLama2 and Mistral, achieving a 20% improvement in patient recommendation accuracy.
- Enhanced chatbot intelligence by integrating a RAG pipeline with domain-specific data sources, improving context relevance and response quality.
- Implemented RASA framework with voice assistance features, boosting usability and accessibility for diverse user groups and delivered a production-ready LLM solution within 3 months, showcasing rapid prototyping and deployment capabilities in an Agile environment.

AI Engineer, *TCS*   Jul 2019 – Jul 2022 | Hyderabad, India

- Developed advanced predictive modeling systems for credit risk scoring and fraud detection, applying Random Forest, Gradient Boosting, and XGBoost, improving default prediction accuracy by 21% for banking clients.
- Processed and enriched large-scale unstructured financial transaction logs using NLP techniques like tokenization, lemmatization, and sentiment analysis, enhancing downstream fraud detection and anomaly detection pipelines.
- Designed and automated ETL workflows using Apache Spark and Airflow to integrate data from 15+ sources, reducing manual preparation time and improving model training reliability by 40%.
- Conducted systematic A/B testing across multiple ML algorithms and hyperparameter sets, identifying optimal models for loan approval predictions and investment portfolio analysis with measurable business value.
- Created interactive Tableau and Power BI dashboards showcasing fraud detection metrics, customer segmentation patterns, and credit scoring models, enabling senior management to make faster, data-driven investment decisions.
- Partnered with compliance and data governance teams to ensure predictive modeling pipelines adhered to GDPR and PCI-DSS standards, mitigating regulatory risks and strengthening client data security frameworks.
- Containerized ML applications using Docker and deployed them on Kubernetes clusters, improving scalability, reducing deployment errors, and cutting model rollout timelines from weeks to just days.
- Collaborated with cross-functional teams of data scientists, analysts, and financial consultants to integrate AI insights into credit approval workflows, reducing loan processing time by 25% across enterprise clients.

Projects

Mental Health LLM Chatbot

- Led the development of a Mental Health Chatbot leveraging LLMs to analyze social media impact on mental health.
- Utilized RAG, Mistral LLM, Chroma DB, and Streamlit for chatbot interface development, deployed using Huggingface.
- Employed LangChain framework throughout the project.
- Delivered a user-centric chatbot capable of providing tailored mental health recommendations

Enhanced Image Gallery Flask Web App

- Enabled multi-factor authentication using Firebase Authentication to enhance data security.
- Developed a user-friendly image gallery with PIL for metadata extraction and integrated Git + Google Cloud Run for CI/CD.
- Ensured secure image storage using Google Cloud Storage with signed URLs for scalability.

Education

Master of Science, *Florida Atlantic University*   Aug 2022 – May 2024 | FL, USA
Data Science and Analytics

Bachelor of Engineer, *Sphoorthy Engineering College*   Jul 2016 – Jul 2020 | Hyderabad, India
Computer Science