Phanindra Kumar Mulamreddy
Data Scientist (AI/ML Engineer)
+1 (667) 464-0071 |mulamreddyp7@gmail.com|

SUMMARY

With 3 years of experience in data engineering and AI/ML engineering, I specialize in building scalable cloud-based solutions and intelligent systems. Leveraging AI/ML frameworks like TensorFlow, PyTorch, and scikit-learn, I have developed advanced systems for medical image analysis, recommendation engines, and predictive analytics. My expertise includes designing and optimizing data pipelines for efficient ETL workflows and real-time processing, as well as implementing robust MLOps practices to ensure seamless deployment, monitoring, and lifecycle management of machine learning models on platforms such as AWS and GCP. I have achieved significant reductions in model size and latency through techniques like quantization, pruning, and A/B testing. In healthcare IT, I have integrated systems adhering to HL7 standards, ensured HIPAA compliance, and developed secure APIs to manage sensitive medical data. My experience in Agile environments includes leading cross-functional teams, driving iterative development, and delivering scalable, high-quality solutions that align with business objectives.

SKILLS

- **Programming/Libraries**: Python (Pandas, NumPy, Scikit-learn, TensorFlow, Pytesseract, Hugging face Jupyter), PySpark
- **Databases**: SQL, MongoDB, Oracle, MySql
- **Machine Learning**: Predictive Modeling, Supervised and Unsupervised Learning, Anomaly Detection, Feature Engineering, LLM's
- **Algorithms:** KNN, Regression (Linear, Logistic, Multiple), Naive Bayes, Random Forest, SVM, NLP, K- Means
- **Cloud Platforms**: AWS (Glue, Redshift, SageMaker, Lambda, Athena, S3, EMR, Kinesis, Firehose, IAM)
- **Big Data & Workflow Automation:** Apache Airflow, Spark, Hadoop, Kafka, ETL/ELT Pipelines
- **Data Engineering:** Data Extraction, Data Validation, Dimensional Modeling, Data Warehousing
- **Data Visualization**: Tableau, Looker Studio, Power BI
- **Project Management**: Workflow orchestration, Agile, Software Development Life Cycle, Work Breakdown Structure, Slack, Jira
- **Version Control Tools**: Git, GitHub, GitLab, Bitbucket

CERTIFICATES:

- AWS Machine Learning Specialty
- AWS Data Engineer Associate

WORK EXPERIENCE:

Data Scientist (AI/ML Engineer) | Broadcom | March 2023 – Present

- Conducted statistical analysis and developed predictive models, reducing feature engineering workflows' runtime by 20% and improving machine learning model training efficiency.
- Extracted and analyzed large datasets using PySpark, SQL, and AWS tools, generating actionable insights that led to a 15% increase in product performance and revenue optimization.
- Designed and implemented GPT-like large language model prototypes, achieving 10% efficiency gains over baseline transformer models and enabling scalable deployment for real-world applications.
- Applied transfer learning and fine-tuning techniques to LLMs, improving performance for domain-specific tasks like conversational AI, sentiment analysis, and document summarization.
- Conducted prompt engineering experiments to optimize LLM outputs, saving 15% in content generation costs.
- Leveraged Kafka to process over 10,000 events/second, building scalable and fault-tolerant data pipelines that reduced data latency by 25% and supported real-time analytics.
- Implemented natural language processing (NLP) pipelines for text classification, sentiment analysis, and recommendation systems.
- Stayed up-to-date with the latest AI/ML trends and technologies, integrating emerging techniques like deep learning, reinforcement learning, and transfer learning into projects.
- Applied statistical and machine learning techniques like regression, clustering, classification, and deep learning to drive business outcomes.
- Developed a cloud-native MLOps platform on AWS for scalable AI/ML deployment and management using Python, TensorFlow, and Kubernetes, ensuring 99.9% availability while handling petabytes of medical imaging data.
- Utilized TensorFlow and PyTorch for deep learning analysis of medical images, achieving 90% accuracy on 100,000 X-ray/MRI scans. Conducted A/B testing to compare model performance for data-driven decision-making.

- Employed advanced optimization techniques such as quantization and pruning (TensorFlow-Model-Optimization) to reduce ML model size by 60% without sacrificing accuracy.
- Established a scable data ingestion pipeline on AWS (S3, Glue, Athena) for processing and storing terabytes of medical data from diverse sources, enduring reliability and scalability.
- Utilized pre-built machine learning algorithms (scikit-learn, XGBoost) for predictive analytics on medical data, resulting in an 18% improvement in diagnosis accuracy.
- Coordinated cross-functional teams to resolve critical system issues, achieving 99.9% uptime and reducing 50% incident resolution time.

Programmer Analyst | Cognizant, India | February 2021 – June 2022

- Developed an AI-driven recommendation engine for a leading e-commerce conglomerate with Deloitte, using microservices architecture with Java and Spring Boot.
- Implemented advanced machine learning algorithms to personalize product recommendations based on user behavior and preferences, leading to a 35% reduction in response times.
- Employed advanced AI techniques like NLP and collaborative filtering to enhance product recommendations, leading to a 20% increase in customer engagement and conversion rates.
- Preparation of Understanding documents, creating knowledge repository for the application process and end to end flow docs.
- Implemented reinforcement learning techniques to optimize product recommendations over time, allowing the system to adapt and improve based on user feedback and interactions.
- Integrated Elasticsearch and Apache Kafka for real-time data processing, enabling the recommendation engine to dynamically adapt to changing user interactions and market trends.
- Perform test estimation, test planning, requirement analysis, test design, test execution, defect management and test closure activities.
- Utilized unsupervised learning algorithms such as clustering and dimensionality reduction to segment customers based on browsing and purchasing behavior, enabling more targeted and personalized recommendations.
- Conduct daily review meetings for the entire team to track and streamline the workflow.
- Leveraged machine learning models for sentiment analysis and trend forecasting to provide actionable insights to the e- commerce client, enabling data-driven decision-making and strategic planning.

Operations Research Analyst | TATA Capital India | April 2020 – January 2021

- Utilized SQL to query and analyze large datasets, generating actionable insights that supported process optimization and strategic decision-making.
- Assisted in identifying operational inefficiencies and recommending process improvements to enhance performance.
- Assisted in developing resource allocation models by performing data analysis with SQL, contributing to a 15% improvement in project delivery timelines.
- Used tools like Excel, Python, R, and statistical software to perform quantitative analysis and modeling.
- Designed and maintained dashboards to monitor key performance indicators (KPIs) using SQL and data visualization tools, enabling real-time decision-making for operations teams.
- Operated financial market research, collected data and performed analysis by using statistical software
- Performed sensitivity analysis and risk assessment using SQL and statistical techniques to evaluate the impact of variable changes on business outcomes, enhancing decision-making accuracy under uncertainty.

EDUCATION

- **Master's in Data Science** from University of Maryland Baltimore County, Baltimore, MD – USA.
- **Bachelor's in Mechanical Engineering** from VR Siddhartha Engineering College - INDIA.