MOUKTHIKA GUNAPANEEDU

moukthikagunapaneedu@gmail.com ⋄ linkedin.com/in/moukthika-gunapaneedu

303-332-1899 ⋄ Boulder, Colorado

SUMMARY

Data Science graduate student with **2+ years of experience** in data engineering, analysis, and machine learning. Skilled in developing **scalable data pipelines, automation workflows, and statistical/ML models** that drive decision-making. Proficient in Python, SQL, and cloud platforms, with a focus on building reliable, production-ready systems from complex datasets.

SKILLS

- **Programming & Data Science:** Python, Pandas, NumPy, Scikit-learn, TensorFlow, PyTorch, OpenCV, SQL (MySQL, PostgreSQL), R, Bash/Shell scripting

- **Data Engineering & Big Data:** ETL/ELT (dbt, Airflow), AWS (S3, Redshift, Lambda), GCP (BigQuery, Pub/Sub), Azure, Spark, Hadoop, Snowflake, MongoDB, NoSQL

- **Modeling & Analytics:** Regression, Classification, Clustering, PCA, Time Series, Predictive Modeling, Forecasting, A/B Testing, Statistical Analysis, Hypothesis Testing, NLP (SpaCy, NLTK, Transformers)

- **MLOps & Automation:** Docker, Kubernetes, CI/CD (GitHub Actions, GitLab CI, Jenkins), REST/GraphQL APIs, Flask, FastAPI, Streamlit, MLflow, EvidentlyAI, Workflow Automation (Airflow, Prefect)

- **Visualization & Reporting:** Tableau, Power BI, Matplotlib, Seaborn, Plotly, Dash, Excel (Pivot Tables)

- **Collaboration & Documentation:** Git, GitHub, Confluence, Jira, Trello, Google Workspace, Microsoft Office

EXPERIENCE

Data Engineer - Metadata   Jan 2025 - Present
University of Colorado Boulder   *Boulder, CO*

- Improved metadata accuracy for 3,500+ NOAA tsunami microfilm records with automated Python workflows.

- Reduced duplication errors by 30% by designing and implementing an automated image deduplication workflow.

- Increased pipeline adoption by creating reusable documentation and guides for cross-functional research teams.

Database Intern   Jan 2024 - Jul 2024
Studio Zenoya   *India*

- Integrated 100K+ JSON records into analytics pipelines by extracting and validating data from multiple APIs.

- Cut manual workload by 40% by developing reusable automation scripts for data transformation.

- Boosted MongoDB query efficiency by 25% through optimized indexing and aggregation pipeline tuning.

Data Scientist   Sep 2023 - Dec 2023
Solar Secure Solutions   *India*

- Reduced fraud false positives by 20% by developing ML models with Scikit-learn and XGBoost.

- Improved data readiness by 35% by designing ETL workflows that streamlined processing of structured and unstructured data, enabling faster analysis for data science teams.

- Built interactive dashboards to visualize fraud risks and highlight transaction patterns for stakeholders.

EDUCATION

Master of Data Science, *University of Colorado Boulder*   2024 – 2026
Relevant Coursework: Machine Learning, Data Mining, Statistics and Applications, Big Data Architecture, Neural Networks and Deep Learning (NN/DL), Natural Language Processing (NLP), Information Visualization