

## MEGHNA REDDI

(973) 906-9462 | [mreddi2001@gmail.com](mailto:mreddi2001@gmail.com) | [/in/meghnareddi](https://www.linkedin.com/in/meghnareddi) | [Github](https://github.com/meghnareddi) | Jersey City, NJ

### PROFESSIONAL EXPERIENCE

#### Maryland Department of Planning | Data Engineer – Baltimore, MD

03/2024 – Present

- Engineered an ETL pipeline to extract, standardize American Community Survey(ACS) data from Census API into MS SQL Server, designed reusable components adopted in 2+ datasets, reducing processing time by 40%; utilized GitLab for orchestration and version control
- Developed an interactive Power BI dashboard to visualize population projections, incorporating Population Dynamics KPIs; replaced static tables on the Maryland Planning website, enhancing data accessibility for 5+ internal planning teams

#### New Jersey Equity in Commercialization Collective | Data Research Analyst – Newark, NJ

01/2024 – 12/2024

- Enhanced data quality for innovation trend analysis by reducing false positive rate (FPR) by 60% in university identification using regex and fuzzy matching, processing 13GB of USPTO PatEx data in a Data Lakehouse
- Enabled 40% increase in women inventors receiving financial support by creating advanced Power BI visualizations with Dataflow Gen2, transforming complex patent data into actionable insights for non-technical stakeholders, improving resource distribution
- Improved data retrieval by 11% through API integration (RESTful web services), validating gender information from LinkedIn and university web pages, ensuring higher accuracy and reliability in demographic analysis

#### New Jersey Institute of Technology | Teaching Assistant – Newark, NJ

01/2024 – 12/2024

- Progressed assignment scores by 15% through 6 lab exercises in RStudio on CNNs, Data Analysis and Large Language Models (LLMs)
- Conducted workshop on pandas, seaborn and statistical models for 30+ students, improving project quality and reproducibility

#### Zenoti India Pvt Ltd | Data Analyst – Hyderabad, IND

07/2022 – 07/2023

- Designed and implemented 10+ custom DAX KPIs in Power Query to analyze undertrained customer accounts, providing region-specific insights to PMs, leading to targeted improvements that increased course completion rates by 2 courses per month
- Reduced churn by 12% in key accounts by developing a BI dashboard (Google Analytics connector) to track customer activity and revenue trends, contributing to a \$20,000 quarterly revenue boost through seasonal churn mitigation
- Built 4+ Excel reports with Pivot Tables and VLOOKUP as data sources for Power Automate flows in a Scrum-based Agile environment, saving 3 hours/week by automating employee reporting, boosting operational efficiency

### ACADEMIC PROJECTS

#### AI-Powered Document Processing and QA System with Azure

01/2025 – 02/2025

- Developed an Azure data processing system using AI Vision to extract text from 30GB+ of documents in Blob Storage, streamlining access to processed data for the team; Integrated AI Language Service for sentiment analysis, NER, categorizing customer reviews
- Deployed a Question Answering system using Azure AI Language, reducing response time by 35% through knowledge base integration

#### Bias Detection and Mitigation in LLM Generated Text

10/2024 – 12/2024

- Developed a bias detection model using AIF360 to analyze LLM outputs, assessing fairness metrics across demographic factors
- Evaluated GPT-2 and BERT using the Crows-Pairs dataset, analyzing fairness disparities across key demographic factors such as gender
- Reduced bias by 10% using Adversarial Debiasing, improving model trustworthiness and fairness in AI-generated content

#### Training Optimization Algorithms in Neural Networks

02/2024 – 05/2024

- Designed a layer-wise & end-to-end neural network training framework, implementing Coordinate Descent (CD), SGD for optimization
- Achieved 97.8% test accuracy with CD, 97.06% with SGD after 15 epochs, outperforming baseline models on high-dimensional datasets
- Analyzed minimum norms across training approaches, comparing CD, SGD to identify lowest norms and assess model generalization

#### Predictive modeling & Time Series Forecasting on Divvy Bicycle Sharing System

10/2023 – 12/2023

- Designed an ETL pipeline using AWS Glue to preprocess 25M+ records stored in S3 data lake optimizing data transformation
- Performed exploratory data analysis (EDA) using Athena, visualizing ride duration and user behavior trends With QuickSight
- Trained and fine-tuned time series models (ARIMA, VAR, Prophet) on AWS SageMaker reducing forecasting error by 20%, improving demand prediction and resource management
- Deployed models using Docker and established CI/CD pipelines with GitHub Actions, ensuring reproducibility & continuous integration

### TECHNICAL SKILLS

**Machine Learning:** Linear & Logistic Regression | Decision Tree | SVM | PCA | Random Forest | KNN | K Means

**NLP & Deep Learning:** NLTK | Spacy | Llama2 | BERT | GPT 3.5 | CNN | RNN | GAN | LSTM | Transformers

**Statistical Methods:** A/B Testing | Time Series Analysis (ARIMA, VAR, SARIMAX) | Hypothesis Testing | Minitab

**Data Analysis:** Predictive Modeling | Data Mining | Feature Engineering | Data Wrangling

**Programming Languages:** Python (Sklearn, Polars, TensorFlow, Keras, Pandas, Numpy, PyTorch, Seaborn) | R (ggplot) | SQL

**AWS Services:** EC2 | S3 | Sagemaker | Lambda | EMR | SQS | Redshift | Rekognition

**Azure Services:** Azure AI/ML | AI Vision | AI Language Service | Data Lake Storage | Databricks | Data Factory | Synapse Analytics | Dataflow Gen2 | Data Pipelines | SQL Database

**Data Visualization & Business Intelligence:** Power BI (MS Fabric) | Tableau | RShiny | Quicksight

**Database & Automation:** Oracle | MS SQL Server | MySQL | Docker | Gitlab | Power Automate

**Project Management:** JIRA | Confluence | Agile | Scrum

**Certifications:** [AWS Certified Cloud Practitioner](#) | [Generative AI with Large Language Models](#) | [Google Data Analytics Professional](#)

### EDUCATION

New Jersey Institute of Technology, Newark, NJ – Master of Science (M.S.) – Data Science, Concentration in Statistics – GPA: 3.8

Mahatma Gandhi Institute of Technology, Hyderabad, IND – Bachelor of Technology (B.Tech.) – Computer Science – GPA: 3.5