MEGHNA REDDI

973.906.9462 | mreddi2001@gmail.com | /in/meghnareddi | Github | Jersey City, NJ

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

TECHNICAL SKILLS

Machine Learning: Linear & Logistic Regression | Decision Tree | SVM | PCA | Random Forest | KNN | K Means | ARIMA | VAR | SARIMAX
Natural Language Processing & Deep Learning: NLTK | SpacY | Llama2 | BERT | GPT 3.5 | CNN | RNN | GAN | LSTM | Transformers
Data Analysis & Statistical Methods: Regression Analysis | Hypothesis Testing | A/B Testing | Time Series Analysis | Statistical Testing | Minitab | 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 | IAM | Quicksight | EMR | SQS | Lambda | Redshift | Rekognition
Azure Services: Azure AI/ML | AI Vision | AI Custom Vision | AI Language Service | Azure Data Lake Storage | Databricks | Fabric | Lakehouse | Data Factory | Synapse Analytics | Dataflow Gen2 | Data Pipelines | Notebooks | SQL Database
GCP Services: BigQuery | Cloud Storage
Data Visualization & Business Intelligence: MS Power BI | Tableau | Looker | RShiny | Quicksight
Database & Automation: Oracle | MS SQL Server | MySQL | Docker | Git | Power Automate | Power Apps
Project Management: JIRA | Confluence | Agile | Scrum
Certifications: AWS Certified Cloud Practitioner | Generative AI with Large Language Models | Google Data Analytics Professional

PROFESSIONAL EXPERIENCE

New Jersey Equity in Commercialization Collective | Data Research Analyst – *Newark, NJ* **01/2024 – 12/2024**

- Enhanced data quality by reducing FPR by 60% in university name identification using regex and fuzzy matching in Python, processing USPTO PatEx data (13GB) in Data Lakehouse
- Identified over 3000 male & female inventors using Fabric Python notebook based on inventor names across 8 NJ institutions
- Increased inventor visibility by 20% by building a Power BI dashboard using Dataflow Gen2, analyzing gender demographics
- Built a scalable ETL pipeline in Data Factory, streamlining data collection, organization, and data transformation for evolving datasets
- Leveraged SQL Analytics Endpoint to generate actionable insights, supporting NJECC stakeholders in decision making
- Improved data retrieval by 11%, API integration (RestFul web services), validating genders from LinkedIn and university web pages

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 a data visualization & statistical modeling workshop using Python, teaching data wrangling and interpretation practices

Zenoti India Pvt Ltd | Data Analyst – *Hyderabad, IND* **07/2022 – 07/2023**

- Improved customer training strategies by designing 10+ DAX KPI s in Power Query to analyze undertrained customer accounts
- Increased biweekly course completions by 200% by delivering region-specific insights to PMs, enabling targeted improvements
- Reduced churn by 12% in key accounts by developing a Power BI dashboard (Google Analytics connector) to monitor customer activity, churn rates, and revenue trends; Boosted quarterly revenue by $20,000 via data-driven retention strategies preventing seasonal churn
- Saved 10+ hours by developing Power Automate flows, to extract and format employee reports weekly using Python and HTML
- Developed 10+ custom reports in Excel, leveraging Pivot Tables and VLOOKUP to analyze business performance metrics
- Collaborated with cross-functional Agile Scrum teams to refine user stories, align development with business goals, ensuring timely execution of testing and feature rollouts

ACADEMIC PROJECTS

End to End Visual and Textual Processing & Support System using Azure AI **01/2025 – 02/2025**

- Developed an AI driven data processing system using Azure AI Vision enhancing text extraction, searchability, and metadata generation
- Built an image classification model with Azure AI Custom Vision to categorize images, optimizing content management workflows
- Integrated Azure AI Language Service for sentiment analysis, NER, and key phrase extraction, enabling insights from 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**

- Designed a bias detection pipeline using AIF360 for fairness metrics, analyzing LLM outputs for biases across demographic factors
- Evaluated Llama2, BERT, and other pre trained models using CrowS Pairs dataset, quantifying bias levels and fairness disparities
- Conducted multilingual bias analysis, focusing on Indian languages to address cultural disparities in AI generated text, improve inclusivity
- Enhanced AI model trustworthiness and mitigated harmful stereotypes, contributing to improved NLP fairness and performance

Training Optimization Algorithms in Neural Networks **02/2024 – 05/2024**

- Implemented a Python pipeline for layer wise & end to end neural network training, for Coordinate Descent(CD) and SGD for optimization
- Achieved a test accuracy of 97.8% with CD and 97.06% with SGD after 15 epochs, demonstrating robustness in high dimensional datasets
- Analyzed parameter norms to uncover local minima variations, improving understanding of model generalization, optimization behavior

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