Abhishek Reddy Gorla

abhireddy2748@gmail.com | +1 (720)3333891 |LinkedIn| Github

EDUCATION

University of Colorado Boulder   **Aug 2023 - May 2025**
Master of Science in Data Science, CGPA: 3.80/4   Boulder, CO
Coursework: Statistical Methods and Applications, Data Mining, Machine Learning, Neural Networks, Data-Center Scale
Computing, Information Visualization, Natural Language Processing, Computer vision

Ramaiah Institute of Technology   **Aug 2019 - May 2023**
Bachelor of Engineering in Electronics and Communication, CGPA: 3.21/4   Bangalore, India

SKILLS

- **Programming Skills:** Python, R, SQL, Web Scraping, A/B Testing, Statistical Modeling, Hypothesis Testing, Relational Databases, RestfulAPI, Modular coding.
- **Data Science and Machine Learning:** Tensorflow, Scikit-learn, Pandas, NumPy, Matplotlib, Seaborn, Plotly, ggplot2, LLMs, ML Ops, Open CV, Hugging Face, Fine-tuning, Deep Learning (Certified), Machine Learning (Certified), EDA, LangChain, LangGraph, NVIDIA NIM.
- **Data Engineering:** Snowflake, AWS, Git, Data Cleaning, Preprocessing, Feature Engineering, ETL, CI/CD, Docker, Kubernets.
- **Tools & Utility**: Jupyter Notebook, RMarkdown, API, R Studio, VS Code, Microsoft Power BI, Snowflake Snowsight, Notion.
- **Disciplinary Skills:** Prompt Engineering, Mathematical Reasoning, Data Visualization, Technical Proficiency, Data Storytelling.

PROFESSIONAL EXPERIENCE

Value Health Inc.   **Jun 2025 – today**
Business Analyst Intern   New York, USA
- Wrote clear and detailed user stories to capture functional and non-functional requirements for application development.
- Researched, gathered, and evaluated data sources; performed data cleaning and preparation to ensure high-quality inputs for application workflows.
- Collaborated on developing and improving an in-app chatbot, including defining intents, designing conversational flows, and refining responses to enhance user experience.

Pucho Inc.   **Sep 2022 – Oct 2022**
Data Science Intern   Bangalore, India
- Built a deep learning model to pull handwritten and printed text from scanned files, improving digitization accuracy by 35% and reducing manual effort.
- Delivered interactive plots for quick insights, cutting reporting time by 50% and enabling faster decisions.
- Engineered a Convolutional Neural Network (CNN) model to detect handwritten content in images with 98% accuracy, flagging 5,000+ images weekly and accelerating data validation processes by 60%.
- Automatically pushed OCR results to Snowflake via AWS Lambda triggers on S3 uploads, enabling real-time analytics dashboards for managers and eliminating a full day of manual data hand-offs each week.

Get It Done   **Jun 2022 – Aug 2022**
Deep Learning Intern   Bangalore, India
- Completed a robust facial recognition system using a Siamese Network with TensorFlow and OpenCV, augmenting training data with noise to mimic real-world conditions and achieving 96% match accuracy.
- Collaborated with cross-functional teams to integrate the facial recognition system into existing security frameworks, contributing to the enhancement of overall security measures.
- Containerized the inference service with Docker and deployed it on AWS ECS Fargate, scaling to 200 + concurrent video streams without downtime.
- User access logs streamed through Kinesis Firehose to Snowflake, where scheduled Snowpark jobs generated daily compliance reports, cutting audit preparation time by 70 %.

PROJECTS

Robust EEG Signal Classification for Motor Imagery Tasks in BCI
- Tackled signal instability and artifact interference in EEG-based BCI systems to improve motor imagery classification reliability.
- Enhanced signal quality through advanced filtering and temporal segmentation of EEG data, boosting classification model accuracy by 15% and improving neural signal interpretability.
- Designed and implemented neural network architectures, including DeepConvNet, EEGNet, and ShallowConvNet, achieving 90% accuracy in classifying motor imagery tasks from EEG data for BCI applications.

F1 Race Q&A Generation using Transformer Models
- Engineered a natural language QA system by converting structured F1 race results into contextualized narratives and generating question-answer pairs.
- Fine-tuned transformer models (RoBERTa, Flan-T5) on the custom dataset using JSONL formatting for input-output alignment in supervised learning.
- Achieved competitive performance with chrF and ROUGE-L scores of 0.4 and 0.5 respectively after targeted hyperparameter optimization.
- Automated the end-to-end data pipeline including preprocessing, Q-A generation, model training, and evaluation.

Dynamic SQL Generator for JSON Extraction
- Developed a scalable SQL generation engine for automated multi-level JSON extraction in Snowflake using Snowpark.
- Automated complex query generation, reducing manual effort and accelerating development by 95%.
- Optimized query performance with adaptive batching, operator handling, and schema caching—cutting execution time by 70% and memory usage by 30%.
- Boosted real-time analytics efficiency via metadata integration and low-memory schema generation.