# Preethu Nath Manjunath

Boulder, CO 80304 | +1 (303) 472-5880 | prma3214@colorado.edu | github.com/preethu-manjunath | preethu-manjunath.github.io/github-portfolio/ | linkedin.com/in/preethu-nath-78274417a/

### **EXPERIENCE**

### Data Consultant, KOBEYO - Basalt, CO, USA

Dec. 2024 - Present

- Leading AI/ML development and managing a team of four to build a scalable, autonomous talent-matching platform using Python, Supabase, and LLMs like Anthropic Claude and Apollo AI, which identifies job lead opportunities from web data, enriches them with skill tags, and matches seekers to companies based on preferences and location with minimal human intervention.
- Designed and deployed interactive dashboards using Tableau and Holistics, resulting in a 35% reduction in manual effort and saving users 3hours/daily
- Performed geospatial job market analysis using Python and SOL, enriching the dataset to improve database coverage by 40%.
- Automated data ingestion pipelines, streamlining ETL workflows and reducing query execution time by 50% and streamlining decision-making.

## Data Science Intern, CodeModulo Technologies Pvt. Ltd. - Bengaluru, India (Remote)

Jul. 2024 - Sep. 2024

- Developed a business intelligence dashboard for a retail client using Power BI, integrating sales, inventory, and customer data to enable data-driven decisions, resulting in a 15% increase in sales through optimized inventory management and proactive stock alerts.
- Enhanced chatbot performance by integrating LLMs such as LLaMA and Claude, improving text summarization, and reducing query resolution time by 30%, while enabling more contextual and human-like responses and boosting customer service efficiency.
- Built and deployed NLP pipelines using Python, SQL, spaCy, and NLTK, and leveraged AWS services including SageMaker, Glue, S3, and Lambda to create scalable model training and deployment workflows with automated retraining and streamlined model-to-UI integration collaboratively.

#### Graduate Research Assistant, Leeds School of Business, University of Colorado Boulder - Boulder, CO, USA

Jul. 2024 - Sep. 2024

- Developed LLM-powered models using GPT-3.5, Hugging Face Transformers, LangChain, and OpenAI API to analyze and predict patterns in human behavior, generating actionable insights that advanced research findings.
- Automated data collection workflows through web scraping with Selenium and BeautifulSoup, reducing manual efforts by 60%.

### Data Scientist (Consultant), Deloitte Consulting USI - Bengaluru, India

Jan. 2022 - Jul. 2023

- Automated NLP pipelines using Python, spaCy, and NER models to process 1,000+ daily news alerts, achieving 90% accuracy and reducing manual review time by 2 hours/day through real-time risk identification and event detection.
- Received Spot Award for designing and deploying an AI-driven risk event tracking system using AWS Glue and SageMaker, boosting global supplier monitoring efficiency by 30%, supporting faster risk mitigation decisions.
- Built proof-of-concept LLM applications for document automation, conversational AI, and stable diffusion image generation, using Retrieval-Augmented Generation (RAG) and cloud-hosted models to enhance business intelligence workflows.
- Engineered a fully automated data ingestion pipeline using AWS Lambda, S3, and Step Functions, eliminating reliance on third-party vendors and manual collection processes, resulting in a 15% reduction in operational costs and more reliable, on-demand reporting for stakeholders.

# Data Scientist (Associate Software Engineer), Hexaware Technologies - Chennai, India

Oct. 2020 – Dec. 2022

- Developed a custom Named Entity Recognition (NER) model using Python and spaCy for a major retail client, achieving 70% accuracy and enabling faster, more reliable extraction of key information from unstructured customer data.
- Built and deployed a loan default prediction model with XGBoost, improving recall to 93%, which significantly enhanced the client's risk assessment process and supported better decision-making for credit approvals and optimized data pipelines.

# SKILLS

Programming Languages: Python, R, HTML/CSS, Java, C/C++

Machine Learning & NLP: Scikit-learn, XGBoost, TensorFlow, NLP (spaCy, NLTK), Hugging Face Transformers, Statistical Analysis, A/B Testing, RAG Cloud & DevOps: AWS(Lambda, ECS, EC2, API Gateway, CloudWatch); Azure(MLOps, Virtual Machines, WebApps); DevOps(Docker, CI/CD) Databases & Analytics: SQL (PostgreSQL, Snowflake), MongoDB, Supabase, Power BI, Tableau(Prep, Desktop, Online, Server)

Other Tools: Git/GitHub, VS Code, DataRobot, Dataiku, JIRA, Flask, FastAPI

#### PROJECTS

## ADHD Assistance App – Your Personal Focus Sidekick

- Developed a responsive web and mobile-friendly app to support adults with ADHD in daily productivity and task management.
- Integrated **LLMs** with **voice input functionality** to deliver personalized productivity suggestions and **hands-free task management**, enhancing accessibility and user engagement.
- Designed with an accessibility-first approach, including blind-friendly features and attention-enhancing video modules for improved focus.

# **AI-Powered Perpetrator Identification in Road Accidents**

- Developed an ML-based system to analyze accident footage, automatically identify potential perpetrators using computer vision and behavioral patterns.
- Applied YOLOv5 for real-time object detection and OpenCV for motion tracking and trajectory analysis.
- Incorporated spatiotemporal pattern recognition to assess vehicle behavior and detect anomalies pre- and post-incident.
- Generated automated summary reports to assist law enforcement with faster, evidence-backed incident evaluations.

### **EDUCATION**

# University of Colorado, Boulder, USA

Aug. 2023 - May 2025

Masters in Data Science

# CERTIFICATION

- Microsoft Certified: Azure Fundamentals

- Microsoft Certified: AI Fundamentals

- Google Cloud Certified (Coursera): MLOps Fundamentals

- DataRobot Certified: DataRobot for Data Scientist