Bhavana Vippala

bhvi9697@colorado.edu — (303) 521-8141 — BOULDER, CO, USA
GitHub: bhavanareddy19 — **LinkedIn:** bhavanareddy19 — **Portfolio:** Website

Education

- **University of Colorado Boulder** — Boulder, CO
  Master of Science in Data Science; CGPA: 3.88/4 — *2024–2026*
- **CMR College of Engineering & Technology** — Hyderabad, India
  Bachelor of Technology in Computer Science; CGPA: 8.3/10 — *2020–2024*

Experience

- **AI/ML Software Engineer Intern** — May,2025 – Present
  PM Accelerator — *Florida,US*
  - Cut LLM latency by 40% with RESTful Flask APIs; scaled to 500+ req/day at 99.8% uptime.
  - Automated ETL for 10k+ docs/mo to JSON on Databricks (Spark/Delta), boosting throughput by 35%.
  - Integrated ChatGPT + custom ML with MLflow on Databricks; +30% perf and (-25%) token spend across 1k+ daily interactions.
  - Built DS pipelines (MiniLM embeddings; DBSCAN/K-Means GPU-cosine; BART summaries) on Databricks; added JWT+OTP, rate limiting, and Redis caching.

- **Python Developer** — May 2022 – May 2023
  Swecha Organization — *Hyderabad, Telangana*
  - Developed a Python-based application for real-time bus tracking and arrival time predictions for **200+** buses, significantly enhancing commuter convenience.
  - Implemented advanced geolocation algorithms, improving the accuracy of bus arrival times by 25%, increasing user reliance and satisfaction.
  - Engineered the integration of live GPS data with a team of **4** developers, enabling real-time tracking for **200+** buses across the city and improving operational efficiency while enhancing user experience for daily commuters.

Projects

Virtual Vogue: Deep Learning for Realistic Fashion Try-On — PyTorch, GANs, Diffusion, CV:

- Built a coarse-to-fine VITON pipeline: encoder–decoder coarse generator + mask-guided non-parametric warping + refinement/blending, using pose heatmaps and person segmentation for geometry-aware garment transfer.
- Beat strong baselines (PRGAN, CAGAN, CRN) on 2,032 test pairs: **SSIM +5–8%**, **IoU +5–17 pp**, Inception Score **3.22**; deployed on **AWS** (EC2/S3) with real-time ∼**30 ms** inference at 256 × 192.
- Ran ablations (NoWarp/NoRefine/NoPose/NoSeg) showing warp as most critical (**IoU 17 pp** without it); user study (30 raters, 100 trials/baseline) preferred our outputs **88–97%** for realism/fit.

Echolab — Feedback→Hypothesis Mining (Python, NLP, Supabase, Databricks):

- Our AI-powered platform eliminates the manual bottleneck between customer voice and product experimentation, delivering **end-to-end ¡48 hr** hypotheses that teams can immediately test and validate.
- Ingested multi-source tickets into Supabase/PostgreSQL; standardized text/meta and embedded 100% of records using `sentence-transformers/all-MiniLM-L6-v2` with cosine-sim vector search.
- Auto-clustered themes via hybrid DBSCAN + K-Means on **Databricks**; deduped near-duplicates and produced cluster briefs with `facebook/bart-large-cnn`, each with representative examples and citations.
- Cut analysis cycle time from **3–4 weeks to ¡48 hours** and lifted hypothesis throughput **3–5×** via auto triage + clustering + summarization; RAG grounded hypotheses in UX patterns and exported test-ready cards to GrowthBook/Jira.

PLATE-TO-HEALTH: A Global Nutritional Journey — Python, Machine Learning, GCP:

- Led a project using the Global Dietary Database on **GCP** (BigQuery/Vertex AI) to model global dietary patterns; employed advanced regression, achieving $R^2 = 1.000$.
- Analyzed dietary data across **N** demographics/**M** regions, improving data-driven insights for public health interventions.
- Enabled precise dietary trend forecasting, setting strong accuracy benchmarks with MAE = 30.5 and RMSE = 25.0.

Technical Skills

- **Programming**: Python, R, Java, JavaScript, C/C++
- **ML/AI & NLP**: PyTorch, TensorFlow, scikit-learn, XGBoost, Transformers, LLMs/RAG
- **Data/Analytics & Tools**: SQL, Pandas, NumPy, Spark, Hadoop, Tableau, Power BI, Git
- **Frameworks & Backend**: Flask, Django, Node.js, React, Angular
- **Cloud & Databases**: AWS (EC2, S3), GCP, Databricks; PostgreSQL, Supabase, MongoDB, MySQL, Snowflake/BigQuery
- **Certifications**: Full Stack Data Science and AI — Naresh I Technologies (2023)

Research Papers

- **Cardiac Arrest Prediction in Newborns**: Published in *IJRASET* (Paper ID: IJRASET59408)
- **Public Opinion Detection: Sentiment Analysis and Data Visualization**: Published in *IJRAR.org* (Reference Number: IJRAR$_2$80114)