Vishal Vaka

+1-872-258-6648 | vishalvaka@gmail.com | LinkedIn | GitHub | Portfolio

Summary

Recent M.S. Computer Science graduate energised by the challenge of turning cutting-edge machine-learning research into polished, customer-facing products. Skilled in the full ML lifecycle—from framing problems and curating data to training Bayesian-optimised deep-learning and large-language-model pipelines in PyTorch/Hugging Face, then shipping them with cloud-native MLOps, rigorous testing, and CI/CD automation. Demonstrated success building real-time optimisation systems, scalable recommendation prototypes, and intuitive analytics dashboards that make complex models usable for non-experts. A collaborative, fast learner who thrives in cross-functional teams and is driven to solve high-impact problems with creativity, empathy, and engineering excellence.

Education

University of Illinois at Chicago | Master of Science in Computer Science | GPA - 3.78/4.0 *Aug 2023 – May 2025*
Coursework: Advanced Machine Learning, Artificial Intelligence, Parallel Processing, Natural Language Processing

Osmania University, Hyderabad | Bachelor of Technology in Computer Science and Engineering *Jul 2017 – May 2021*
Coursework: Machine Learning, DBMS, BigData, Python App Programming, Data Science, Data Analytics

Experience

- **Machine Learning Researcher | PyTorch, BoTorch, GPyTorch, Lab Streaming Layer, GitHub Actions, Jupyter** *Feb 2024 - Aug 2024*
  University Of Illinois at Chicago, Rehab Robotics Lab Chicago, USA
  - Built real-time LSL to BoTorch pipeline streaming gait / metabolic data; GPU-optimised Bayesian loops ran 25 % faster.
  - Designed multi-objective Rank-Weighted GP ensemble with Chebyshev-EHVI, doubling Pareto coverage and halving convergence iterations.
  - Refactored RGPE into pip package; added 2-D benchmarks, Matplotlib diagnostics, Jupyter workflows, automated hyper-parameter sweeps.
  - Integrated PyTest and GitHub Actions CI/CD, lifting nightly build stability to 100 % across more than 70 commits.
  - Resolved clock-sync preprocessing bugs and wrote dev docs, enabling reproducible multi-GPU experiments for incoming researchers.
- **Business Technology Analyst | Salesforce, Apex, Jenkins, GitHub, Jira, Confluence, Apttus CPQ, CI/CD** *Sep 2021– Aug 2023*
  Deloitte USI Hyderabad, India
  - Automated CPQ deployments via Jenkins-GitHub pipeline; Apex unit tests and quality gates push code from QA to PROD.
  - Generated multilingual quotes by crafting Apttus X-Author templates and email microservice, slashing global-entity sales cycles 30%.
  - Implemented Apex price-list engine with country-specific rules; cut Belgium staging overrides 40% and improved pricing accuracy.
  - Achieved 95% release stability through exhaustive test classes, telemetry logging, and Jira-Confluence traceable CI documentation.

Projects

- **Biomedical NER Comparison Suite | Python, PyTorch, HuggingFace Transformers, BioBERT, LoRA, CRF, Gradio, YAML** [GitHub]
  - Benchmarked BioBERT fine-tuning vs LoRA, CRF, adapter fusion; reached 0.88 F1 using 98 % fewer parameters.
  - YAML runner executes triple-seed sweeps, logs latency + VRAM, finishes reproducible training in 15 min on a T4.
  - Published HuggingFace Spaces Gradio app—live token-level NER predictions with F1, runtime, memory dashboards.
  - Structured src package, PEP-8 lint, pinned requirements; MIT-licensed repo is CI/CD-ready for production ML workflows.
- **NLP Sentiment-&-Coherence Benchmark Toolkit | scikit-learn, Gensim Word2Vec, spaCy, Stanford CoreNLP, TF-IDF, SVM, Git LFS** [GitHub]
  - Benchmarked TF-IDF, Word2Vec, handcrafted linguistic features across SVM, Logistic, MLP; macro-F1 rose from 0.57 to 0.71.
  - Added spaCy POS-diversity and conjunction-count extractors, boosting stylistic recall nine percent on social-sentiment data.
  - Integrated CoreNLP coherence metrics (TTR, LSA, content overlap) enabling real-time dialogue-quality dashboards for human-in-loop review.
  - Versioned models with Git LFS; CLI offers live classification, YAML sweeps, PEP-8 lint, CI-tested reproducible workflows.
- **Siamese DenseNet Fake-Face Detection | TensorFlow, Keras, DenseNet, Contrastive-Loss, AWS SageMaker, CUDA, OpenCV** [GitHub]
  - Built Siamese DenseNet with contrastive loss; hit 99.5 % accuracy, 0.0017 loss on 100 k GAN-vs-real faces.
  - Sampled PGGAN, StyleGAN, StarGAN pairs; kept more than 97 % N-way one-shot accuracy at 1.9 s inference.
  - Optimised mixed-precision SageMaker runs; 64-image batches cut epoch time from 6 h to 1.5 h without accuracy drop.
  - Deployed Tkinter-OpenCV GUI; delivers sub-second single-face verdicts and 60 % accuracy on multi-face images.
- **Distributed 2-D FFT Convolution Accelerator | C, MPI, OpenMP, FFT, PBS/SLURM, HPC** [GitHub]
  - Parallelised 512 × 512 2-D FFT with MPI Scatter/Gather and OpenMP, achieving 7.8 × speedup over serial baseline.
  - Implemented communicator splitting and custom Scatterv/Gatherv, sustaining 92 % strong-scaling efficiency across 32 CPU cores.
  - Optimised cache-friendly transposes and SIMD complex-multiply kernels, maximising FLOPS for vision and transformer-attention convolutions.
  - Packaged Makefile and PBS scripts for reproducible SLURM runs, blueprinting distributed gradient-sync workflows for ML.

Skills

- **Programming Languages:** Python, C, C++, Java, Apex, JavaScript, HTML, CSS
- **ML/DL & Vector Retrieval:** PyTorch, TensorFlow, Keras, scikit-learn, HuggingFace Transformers, BoTorch, GPyTorch, FAISS, LoRA, CRF
- **Data Processing & Visualization:** NumPy, Pandas, Matplotlib, Seaborn, NLTK, spaCy, OpenCV, YAML
- **AI/ML:** LLMs, CNN, RNN, LSTM, Gaussian Processes, Bayesian Optimisation, Siamese Networks, Contrastive Learning, TF-IDF, Word2Vec, Sentiment Analysis, NER
- **Data Science:** Data Mining, Text Mining, Topic Modelling (LDA, LSA), Feature Engineering, Hyper-parameter Sweeps, Reproducible Pipelines
- **Natural Language Processing:** Sarcasm Detection, Named Entity Recognition, Semantic Analysis, Coherence Scoring
- **Tools & Platforms:** AWS SageMaker, Git, GitHub, Jenkins, GitHub Actions, Git LFS, Docker, Gradio, HuggingFace Spaces, Lab Streaming Layer, MPI, OpenMP, CUDA, SLURM/PBS, Salesforce (CPQ, Flows), Jira, Confluence

Affiliations

- **Vice Chair, Decent Work & Economic Growth** | Hyderabad Youth Assembly (Season IX) – Street Cause - Hyderabad Sep 2019 – Feb 2020
  - Promoted from Delegate to Vice Chair; mentored 20+ delegates executing SDG 8 projects.
  - Co-led orphanage and government-school outreach; installed equipment, taught environmental lessons, impacting 200+ students.
  - Designed arts-and-crafts micro-enterprise workshop; empowered 30 orphans to create and sell products for income.
  - Managed budget, vendors, and reports; ensured 100 % on-time delivery and transparent fund utilisation.