A. SURYA KAUSTHUB

Bengaluru, India | +91 93912 46119 | kausthubsurya@gmail.com | Linkedin | Github

SUMMARY

Final-year CS student specializing in **ML/DL/LLMs**. Build **end-to-end** pipelines (data → training/tuning → evaluation → **deployment**) with **PyTorch/TensorFlow**/sklearn. Strong **MLOps**: **MLflow** tracking, **FastAPI** microservices, **Docker/Kubernetes** CI/CD; emphasize **measurement & ablations** (p95/p99 latency, accuracy–latency–memory trade-offs). Experience in **NLP** (Transformers, seq2seq, RAG) and pragmatic **CV**; optimize inference with **ONNX**/quantization and profiling. Proficient in **Python**, **C++**, **Java**; solid math foundation (linear algebra, probability, optimization).

EDUCATION

- **Bachelor of Technology in Computer Science**   CGPA: 8.3/10 (Top 10%)
  Amrita Vishwa Vidyapeetham, Bengaluru   *2022 – 2026*

TECHNICAL SKILLS

- **Languages:** Python, Java, C++, SQL
- **Data Analytics:** Pandas, NumPy, Matplotlib, Scikit-learn, Excel (Advanced), Power BI
- **Modeling:** Linear Regression, Classification, Time-Series Forecasting, LSTM
- **Tools & Platforms:** Jupyter, Git, Prometheus, Firebase , PyTorch, Android Studio
- **Concepts:** Object-Oriented Programming, Design Patterns, Distributed Systems, Socket Programming, SQL Server
- **Testing:** Test Automation, Debugging, Profiling, Unit Testing, Manual and Automated QA

EXPERIENCE

- **Technical Co-Founder**   Nov 2023 – Present
  Digg Deeperr — Micro-learning Platform (Remote)
  - Architected **Next.js 14 + Firebase** app with real-time **WebSocket** coding arena; containerized and deployed on **GKE**.
  - Delivered >**99.9% uptime** for **500+ users** using **readiness/liveness probes**, resource limits, and **blue–green** rollouts; executed zero-downtime schema changes and tracked error budgets against SLOs.
  - Hardened security with **RBAC** and **JWT**; automated **AES-256** backups (**RTO < 60s, daily RPO**).
  - Optimized performance with **Prometheus/Grafana** dashboards and **eBPF** profiling; added CDN/edge caching and hot-path tuning to sustain ≤ 150 ms p95 latency (met **98.7% SLO**).

PROJECTS

- **MorseVerse: TinyML vs Cloud Morse Translator** — ESP32, TFLite, PyTorch   2024
  - Built dual translators: **ESP32 TinyML (int8 2-layer LSTM)** vs **PyTorch LSTM** cloud on **40k** sequences.
  - Edge: **77% acc** (**F1 76%**) at **0.12s p95**; Cloud: **89%** at **80.9s** ($\sim$**670$\times$ faster** on-device).
  - Designed **TTP223B** tap input + debounced timing; exported to **TFLite int8** for real-time ESP32 inference.
  - Ablated window sizes/thresholds; documented **accuracy–latency** trade-offs for assistive/emergency/IoT use.
- **Low-Resource Telugu–Hindi Translation LLM** — PyTorch, Fairseq, BART, ONNX   2024
  - Dataset: **21.4k** pairs; subword vocab **16k (Hindi) / 32k (Telugu)**.
  - LSTM (TPU) scaling study **10–100%** data; compared schedulers/regularization.
  - Transformer: **val loss 6.48 vs 7.63** (LSTM), **4$\times$ faster** convergence (**10 vs 40** epochs).
  - Implemented **beam search + length penalty**; **BLEU/chrF** eval; exported **ONNX** for inference.
- **Smart Health Tracker App** — Android, Java, Firebase   2024
  - Android app for **GPS activity** & health; **Firebase Auth/Firestore** with offline caching & sync.
  - Applied **MVVM** (ViewModel, LiveData) and **Room** for local persistence.
  - Optimized RecyclerView diffing/I/O; $\sim$**30% faster** p95 navigation; smooth UI on >**1,000** records.
  - Added background workers & permission handling; improved reliability under intermittent network.

CERTIFICATIONS & ACHIEVEMENTS

- **CodeChef 4 (1901)** & **Codeforces Expert (1775)** — Top $\sim$2–5%. [CC] [CF]
- **IBM: ML with Python (Honors)** — Jul 2024. [Cert]
- **IBM: Deep Learning with Keras** — Jul 2024. [Cert]

RESEARCH PUBLICATIONS

- **SecureBidPro: A Robust Online Auction System** — IEEE Conference, 2023.
- **Efficient ML Models for Telugu–Hindi Translation** — INCOFT 2025.
- **Quantum-Secured Alerts in Smart Grids** — IEEE ICET 2025.
- **Smart Health Tracker with GPS** — CRM 2025.

LEADERSHIP & MENTORING

- Conducted Distributed Systems Bootcamp (100+ participants)
- Mentored juniors in Android app development and code optimization