# Dr. Venkatesha Nayak B

**Artificial Intelligence & Machine Learning | Healthcare & Applied AI**

+91 9480560737 | venkateshnayakb23@gmail.com | Bengaluru, India

Portfolio | GitHub | LinkedIn

## SUMMARY

AI engineer with a clinical background, focused on building machine learning applications and healthcare analytics solutions. Passionate about applying data and AI to improve healthcare technology.

## EDUCATION

**PGDM - AI & Data Science in Healthcare** — **Jul 2025 - Present**

*Indian Institute of Health Management Research, Bengaluru (IIHMR-B)* — *Expected 2027*

**Bachelor of Ayurvedic Medicine and Surgery (BAMS)** — **Oct 2018 - Oct 2024**

*SDM Institute of Ayurveda and Hospital, Bengaluru*

## CORE SKILLS

- **Programming:** Python, SQL, JavaScript / TypeScript, R
- **Machine Learning:** scikit-learn, XGBoost, LightGBM, TensorFlow, PyTorch, SHAP, classification, regression, time-series forecasting, semi-supervised learning
- **Generative AI & LLMs:** LLM application development (Llama-3 via Groq), prompt engineering, RAG, inference cost optimisation
- **Data & Geospatial Engineering:** Pandas, GeoPandas, PyArrow / Parquet, PostgreSQL, MongoDB, AWS S3 (boto3), ETL pipelines
- **Backend & MLOps:** FastAPI, Docker, Git, GitHub Actions (CI/CD), MLflow, DVC, Weights and Biases, Prometheus, Grafana
- **Frontend & Visualisation:** React, Vite, Leaflet, Recharts, Streamlit, Power BI, Tableau
- **Healthcare Domain:** Clinical assessment, diagnosis and treatment planning, EMR data, NFHS survey data, Multidimensional Poverty Index (Alkire-Foster)

## EXPERIENCE

**AI Engineer Intern** — **Bengaluru**

*Data, Evidence & Action Lab (DEA Lab) | Full-time internship* — *Apr 2026 - Jul 2026*

- Engineered an AI analytics platform on National Family Health Survey (NFHS-5) data: a 3-layer Python pipeline on **AWS S3 (boto3)** that filters DHS microdata and computes **14 survey-weighted child-malnutrition indicators**, with incremental (ETag-based) processing and golden-output regression tests for full reproducibility.
- Computed the **India National Multidimensional Poverty Index (Alkire-Foster method)** and validated it against NITI Aayog's official methodology - **national MPI 0.067 vs 0.066, with all 36 states matching the official ranking (correlation 0.997)**.
- Built the lab's flagship **"Precision MPI"** - a semi-supervised geospatial ML system that estimates multidimensional poverty for **~589,000 villages across 31 states and 620 districts**.
- Engineered **39 census-derived village features** from SHRUG data, linked NFHS survey clusters to villages using **GeoPandas**, and applied **graph-based label propagation** to generate calibrated village-level estimates (r ~ 1.0 vs district survey MPI) - validated with spatial cross-validation and served at scale through a Leaflet-canvas / CDN-chunked map.
- Deployed explainable **child-undernutrition (CIAF) risk-prediction models** for two age cohorts - benchmarked XGBoost, LightGBM, Random Forest and Logistic Regression with SMOTE and threshold tuning, selected the production model by ROC-AUC, explained it with **SHAP**, and served live predictions through a **FastAPI service on Render**.
- Shipped the organisation's **public-facing React (Vite) analytics application** - 9 lazy-loaded pages, interactive Recharts dashboards and **Leaflet choropleth maps** - with an automated S3-to-frontend sync and a GitHub Actions **CI/CD pipeline** (unit and Playwright E2E tests, Lighthouse) deploying to Vercel.
- Integrated a **generative-AI insight layer** (Llama-3 via Groq) that auto-generates public-health policy summaries for maps and dashboards, using engineered prompts and content-hash caching for deterministic, low-cost inference.

**Ayurvedic Physician** — **Bengaluru**

*Sapthashwa Ayurvedic Health Care Center | Full-time* — *Feb 2025 - Jul 2025*

- Ran OPD consultations and supervised Panchakarma therapies for 6-7 patients daily - delivering personalised treatment plans, monitoring outcomes and maintaining structured EMR / follow-up documentation.

**Ayurvedic Intern** — **Bengaluru**

*SDM Institute of Ayurveda and Hospital | Clinical internship* — *Oct 2023 - Oct 2024*

- Completed OPD and IPD clinical training in patient assessment, diagnosis, patient education and EMR support.

## PROJECTS

**TrialScope | Clinical Trial Competitive Intelligence Platform |** Link **Jun 2026**

- Disease-first, built end to end: a Python ETL pipeline ingests ClinicalTrials.gov (API v2) data into a **6-table PostgreSQL schema (Supabase)**, a **FastAPI** backend serves aggregated KPIs, trial momentum, drug-mechanism classes, sponsor leaderboards and geography, and a **Next.js + TypeScript** dashboard visualises them with choropleth maps - refreshed daily via a GitHub Actions cron and deployed on Render and Vercel.
- Designed a **config-driven module architecture** so a new disease area is added with data/config only and zero new code - validated by onboarding a second therapeutic area (Alzheimer's) with no code changes.

**Prakriti AI | Clinical Assessment Platform |** Link **Jan 2026**

- Two-part Ayurvedic assessment: 22 weighted questions identify the lifelong constitution (Prakriti) and 8 more read the current state (Vikriti). The plan focuses on a single dosha, prioritising one that is currently imbalanced or that the live season aggravates, with seasonal diet and lifestyle guidance, cutting manual effort by 65-70%. Built in React 19 / Vite, deployed on Netlify.

**Logistics Delay Prediction | End-to-End MLOps Pipeline |** Link **Nov 2025**

- Early / on-time / late classification: a LightGBM model served via **FastAPI + Streamlit**, containerised with Docker, versioned with DVC and Weights and Biases, automated through GitHub Actions CI/CD, and monitored with Prometheus + Grafana.

**Hospital Admissions Forecasting | Healthcare Time-Series Modelling |** Link **Oct 2025**

- 1,064 days of low-volume, irregular admissions; benchmarked naive (MAE ~2.16), moving-average, weekday/weekend and SARIMA (MAE ~3.56) models, and selected a 7-day adaptive rolling-window model on robustness and interpretability for intermittent, zero-heavy data rather than lowest MAE alone.

## CERTIFICATIONS

- **Project Management for Managers - NPTEL SWAYAM:** 12-week Elite certification, Top 5% rank.
- **Six Sigma - NPTEL SWAYAM:** 12-week certification in DMAIC.
- **Data Literacy Fundamentals - DataCamp.**

## LANGUAGES

**Konkani, Kannada** - Native | **English** - Full Professional | **Hindi** - Conversational | **Tamil, Telugu, Malayalam** - Elementary