

ACHHUTA NAND JHA

Darbhanga, Bihar | +91-8439131459 | theanjha@gmail.com | [LinkedIn](#) | [GitHub](#) | [LeetCode](#) | [Portfolio](#)

PROFESSIONAL SUMMARY

Final-year B.Tech Computer Science student with hands-on production experience in Machine Learning, Deep Learning, NLP, and LLM integration. Proven ability to build and deploy production-grade AI systems in fast-paced product environments. Industry experience at Microsoft, IBM, Shell India, and Vodafone Idea across AI/ML and Data Analytics domains. Recognized at AI-NIRMAN 2026 (3rd Runner-Up, MeitY Startup Hub) and shortlisted under Startup Bihar Government Seed Fund initiative. Strong foundations in DSA, system design, and applied ML research.

TECHNICAL SKILLS

Languages: Python, Java

Machine Learning & Deep Learning: PyTorch, TensorFlow, Keras, Scikit-learn, NumPy, Pandas, SciPy; supervised & unsupervised learning, neural networks, model training & evaluation, hyperparameter tuning

Natural Language Processing: Hugging Face Transformers, TF-IDF, text classification, sequence modeling, tokenization, NLP pipelines, LLM fine-tuning & integration

Generative AI & LLMs: Google Gemini API, Prompt Engineering, multi-task LLM systems, structured output generation

Computer Vision: YOLOv8, MediaPipe, OpenCV, object detection, pose estimation, facial landmark detection, real-time inference pipelines

Data Analysis & Visualization: Pandas, NumPy, SciPy, Matplotlib, Seaborn; exploratory data analysis (EDA), statistical analysis, data cleaning & preprocessing, feature engineering, SQL-based data querying & aggregation

Computer Science Fundamentals: Data Structures & Algorithms (Arrays, Trees, Graphs, Heaps, Hashing, DP, Greedy, Backtracking), OOP, Design Patterns, Basic System Design, REST API Design, Authentication & Authorization (JWT, RBAC)

Web Development: React.js, Next.js, Node.js, Express.js, MongoDB, PostgreSQL, RESTful APIs

Tools: Git, GitHub, Docker, SQL, MongoDB

Cloud & Deployment: AWS (EC2, S3), Azure (AI Services), Render

PROFESSIONAL EXPERIENCE

AI & Cloud Intern | Edunet Foundation (in collaboration with IBM) July 2024 - August 2024 *Remote, India*

- Analyzed **243-record Algerian Forest Fires dataset** (11 features: temperature, humidity, wind, FWI indices) using Python, Pandas, Seaborn, and Matplotlib.
- Evaluated and compared 4 regression models using R^2 scoring — Linear Regression achieved $R^2 = 0.9830$ (best); documented performance gaps across Lasso (0.9492), Ridge (0.9825), and ElasticNet (0.8687).
- Applied structured evaluation criteria to identify and remove low-signal variables (DC, BUI); documented rationale for feature exclusion to reduce model ambiguity and improve consistency.
- Deployed trained model on **IBM Watson Studio** using AutoAI; exposed predictions via REST API endpoint on IBM Cloud.
- Developed a **Flask web app** with HTML/CSS frontend for real-time FWI prediction; integrated **IBM Watson X virtual assistant** (portfolio bot) for conversational user interaction.

Artificial Intelligence & Data Analytics Intern | Shell India Markets Pvt. Ltd. × AICTE Apr 2025 - May 2025 *Remote, India*

- Developed a healthcare prediction model on diabetic patient data using Python — achieving 92% model accuracy
- Performed data preprocessing including missing value handling, normalization, and feature selection
- Applied and compared ML algorithms including Random Forest and SVM to build and optimize predictive models

AI & ML Intern | Edunet Foundation (in collaboration with Microsoft) Apr 2025 - May 2025 *Remote, India*

- Developed NLP-based SMS spam classifier using TF-IDF and Random Forest achieving 98% classification accuracy.
- Designed and implemented end-to-end ML pipeline including preprocessing, feature engineering, model training, and performance evaluation.

PROJECTS

Fluency AI — AI-Powered Career Accelerator Platform | [GitHub](#) | [Live](#)

Jan 2025 - Present

Tech Stack: Python · FastAPI · PyTorch · YOLOv8 · MediaPipe · Gemini 2.5 · ImageKit · Next.js · MongoDB

- Built a **multi-modal real-time behavioral analysis system** using **MediaPipe Face Mesh (468 landmarks) + Pose (33 landmarks)** to score 5 dimensions live — Confidence, Eye Contact, Posture, Smile, Stress Level — streamed via WebSocket

from a **FastAPI + CPU PyTorch** backend; when poor posture or high stress is detected, **ImageKit generates contextual corrective illustrations** (posture guides, breathing techniques) delivered instantly to the user

- Built a **multi-agent AI GD simulation** where 3–7 role-specific agents (Initiator, Analyzer, Challenger, Moderator, etc.) generate contextually distinct responses via **Gemini 2.5 Flash** with role-aware prompting and unique per-agent TTS voices — simulating real company GD rounds (Meta, Google) at configurable intensity levels (Calm / Moderate / Aggressive)
- Developed a **YOLOv8 proctoring pipeline** detecting cheating events (phone usage, multiple persons, frame absence) during live sessions — with async frame-dropping and backpressure handling for production stability on memory-constrained free-tier infrastructure
- Integrated **Gemini 2.5 Flash/Pro** as multi-task LLM backbone for structured answer evaluation (5-score JSON output), STT error correction (Hindi-English code-mix), interview question generation, personalized study plan creation, and ATS resume scoring — all serving active production users
- Engineered an **NLP-based ATS Resume Scoring engine** with 8 scoring dimensions (keyword match, skills coverage, format, experience relevance, readability) including keyword gap analysis and actionable improvement suggestions against target JDs
- Built **GD behavioral analytics** capturing per-participant speaking time, interruptions, filler word frequency, pace (WPM), and volume stability — aggregated into 7 scoring dimensions for automated post-session AI evaluation

Diamond Price Prediction | [GitHub](#) | [Live](#)

Tech Stack: Python, Pandas, scikit-learn, Flask

- Audited 193,573-record diamond dataset for quality and consistency; analyzed skewness, kurtosis, correlation heatmaps, and distribution gaps to identify key price drivers.
- Detected and removed ~8,500 outliers via Z-score; documented label noise and distribution gaps (price skew: 1.60, carat skew: 0.99) with structured rationale for each exclusion decision.
- Applied taxonomy tagging to ordinal categories (cut, clarity, color) using domain-specific annotation guidelines — ensuring consistent label mapping across 193K+ records.
- Trained 4 regression models (Linear, Lasso, Ridge, ElasticNet); Ridge achieved **R² of 93.7%** and RMSE ~1,014.
- Deployed Flask app on Render with Docker and modular ML pipeline (ingestion → transformation → training).

CERTIFICATIONS & ACHIEVEMENTS

- Microsoft Applied Skills – Azure AI Language (NLP) | Microsoft | [View Credential](#)
- GfG 160 – 160 Days of Problem Solving (22-Week Structured DSA Program), GeeksforGeeks | [View Credential](#)
- Programming Fundamentals using Python (66 Hours), Infosys Springboard | [View Credential](#)
- Database Management System (55 Hours), Infosys Springboard | [View Credential](#)
- Software Engineering (8 Hours), Infosys Springboard | [View Credential](#)

ACHIEVEMENTS

Startup Bihar — Government of Bihar Seed Fund Initiative

- [TechSeva](#), an AI-powered home appliance service automation platform targeting Tier-2/3 cities, was officially shortlisted and pitched before government expert panels — demonstrated live with business model validation, AI-based technician auto-assignment system, dynamic pricing engine, and scalability roadmap

3rd Runner-Up — AI-NIRMAN 2026 | 36-Hour GenAI Startup Hackathon

Apr 2026

GLA University × MeitY Startup Hub | Investor Jury from Indicorn Angels, Soonicorn Ventures & Faad Capital

- Secured **3rd Runner-Up** at AI-NIRMAN 2026 by pitching **Fluenzy AI** — a production-ready AI interview & GD simulation platform — competing against university-level AI startups; received strategic feedback on fundability, product clarity, and go-to-market from investors at Indicorn Angels, Soonicorn Ventures, and Faad Capital.
- Built & shipped **Fluenzy AI** as founding engineer — production-grade AI interview prep SaaS featuring 9 training modules, FAANG company-specific tracks, live 1:1 peer matching, GD simulation with 7 roles, ATS resume scoring engine, and real-time analytics dashboard tracking 7 performance dimensions (Confidence, Grammar, Body Language, Pace, Communication, Accuracy, Filler Words) across 40+ sessions.
- Conducted enterprise-grade **load testing using k6** — processed **56,623 requests** at **1,594 concurrent users** achieving **0% HTTP failure rate** and **93.5 req/s throughput**, validating platform reliability under production-scale traffic — demonstrating structured QA validation approach.

EDUCATION

B.Tech in Computer Science and Engineering

GLA University, Mathura | CGPA: 7.39/10.0

2023 - 2027