

# SURYANSH KHARE

Varanasi, Uttar Pradesh

+91-7376914547 — kharesuryanshkhare@gmail.com — in LinkedIn — GitHub

## SUMMARY

AI/ML undergraduate with a strong command of ML architectures and backend systems, sharpened through internship experience and exposure to industry practices. Focused on solving real-world problems and delivering measurable value.

## EDUCATION

|                                                                    |                         |
|--------------------------------------------------------------------|-------------------------|
| Pranveer Singh Institute of Technology, Kanpur                     | 2023 – 2027             |
| B.Tech – Artificial Intelligence and Machine Learning – CGPA: 7.64 | Kanpur, Uttar Pradesh   |
| Kendriya Vidyalaya BHU, Varanasi                                   | 2022 – 2023             |
| Senior Secondary – Percentage: 83.2%                               | Varanasi, Uttar Pradesh |
| Kendriya Vidyalaya BHU, Varanasi                                   | 2020 – 2021             |
| Secondary – Percentage: 92.3%                                      | Varanasi, Uttar Pradesh |

## EXPERIENCE

|                                                                                                                                                                                                                                                                                                                                                                                                               |                                |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------|
| Junior ML Engineer                                                                                                                                                                                                                                                                                                                                                                                            | FEB 2026 (Industry ML Program) |
| IBM (International Business Machines) - PBEL                                                                                                                                                                                                                                                                                                                                                                  | Kanpur, Uttar Pradesh          |
| <ul style="list-style-type: none"><li>Reduced inference latency by 35% and preprocessing time from 40 ms to 18 ms using vectorized NumPy and FastAPI optimizations.</li><li>Optimized transactional pipeline for ML models processing 5M+ records, reducing preprocessing time by 40% using vectorized NumPy and memory optimizations over row-wise transformations and repeated timestamp parsing.</li></ul> |                                |
| ML Analyst Intern                                                                                                                                                                                                                                                                                                                                                                                             | OCT 2025 – DEC 2025            |
| Geeks for Geeks                                                                                                                                                                                                                                                                                                                                                                                               | Remote Job                     |
| <ul style="list-style-type: none"><li>Developed a traffic flow prediction pipeline improving model accuracy optimizing Trees and max depth hyperparameter tuning using historical road sensor data and time-based features, improving congestion prediction accuracy by 18% approx.</li></ul>                                                                                                                 |                                |

## TECHNICAL SKILLS

**Languages:** Python, JavaScript, C, Java, SQL, PostgreSQL  
**Frameworks:** Node.js, Express.js, MongoDB, Mongoose, JWT, FastAPI, TensorFlow, Transformers, FAISS, Llama  
**Core Skills:** Data Structures, Algorithms, Dynamic Programming, Graph Theory, Advanced Data Structures

## PROJECTS

|                                                                                                                                                                                                                                                                                                                                                                                            |               |                     |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------------|
| Fire detection & prediction system                                                                                                                                                                                                                                                                                                                                                         | GitHub - Live | AUG 2024 – DEC 2024 |
| <ul style="list-style-type: none"><li>Published a Springer LNEE research paper on real-time fire detection using YOLOv8n, improving baseline accuracy from 86.3% to 97.2% through manual architectural enhancements. [978-3-032-27342-0, EACE 2025]</li><li>Developed a real-time fire detection pipeline using YOLOv8, OpenCV, and Python for object detection.</li></ul>                 |               |                     |
| Project Nexus(LLM from Scratch)                                                                                                                                                                                                                                                                                                                                                            | GitHub - Live | OCT 2025 – DEC 2025 |
| <ul style="list-style-type: none"><li>Built a Node.js/Express backend with JWT authentication to route requests between frontend and model, using MongoDB (Mongoose) for temporary and permanent memory storage.</li><li>Architected an end-to-end Transformer-based LLM from scratch with a modular architecture, allowing easy customization and extension of core components.</li></ul> |               |                     |
| Exam Preparation Platform                                                                                                                                                                                                                                                                                                                                                                  | GitHub - Live | June 2026 – Present |
| <ul style="list-style-type: none"><li>Designed a Node.js backend with JWT-secured routes to handle requests, storing user results and recommendations in a MongoDB database.</li><li>Built backend services for document upload, RAG retrieval, and real-time AI responses using Llama.</li></ul>                                                                                          |               |                     |

## ACHIEVEMENTS & CERTIFICATIONS

- Secured Rank 1 in Semi-Finals and won Tech Expo (2023–2026) consecutively three times
- Secured Pre-Placement from TCS among 40 Projects for Kissan llama
- Published Conference Papers on Fire Detection (International Conference 2025) and PET (Pancreatic Endocrine Diseases) Early Diagnosis
- Oracle Certified Generative AI Professional – Oracle