Nilesh Pandey

pandeynilesh325a@gmail.com | 9026468481 | LinkedIn | GitHub | Hugging Face

Technical Skills

- **Programming & Databases:** Python, Rust, C++, Java, SQL, MySQL, PostgreSQL, MongoDB, VectorDB, ChromaDB
- **Analytics & Visualization:** Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, Plotly, Tableau (basic), Power BI (basic)
- **Machine Learning & AI**: Deep Learning, NLP, Transformers, LLMs, GANs, Reinforcement Learning, Forecasting Models, Statistical Analysis, Hypothesis Testing
- **Frameworks & Tools:** TensorFlow, PyTorch, Keras, FastAI, Flask, FastAPI, Streamlit, LangChain, SentenceTransformers, Hugging Face, vLLM, AirLLM
- **Big Data & Processing:** PySpark, batch processing, vector search
- **Networking:** HTTP, HTTPS, REST APIs, Flask, FastAPI, client-server architecture, TCP/IP
- **Operating Systems:** Linux, Windows
- **Tools & Platforms:** Git, GitHub, Hugging Face, Google Cloud, AWS
- **Soft Skills:** Documentation, teamwork, project planning

Projects

Twitter-Based Mental Health Crisis Detection | 04/2025 | Link

- Built a **distributed NLP pipeline** to collect **42 M+ tweets**, analyze, and visualize across multiple stages (data collection with **Selenium, NLTK, and RegEx,** preprocessing with **VADER, sentiment classification, visualization with Plotly**). **Handled large-scale data** using **batch processing** and **parallel NLP pipelines**, simulating production-scale distributed workflows.

AI-Based Symptom Detector | 11/2024 | Link

- **Architected a distributed AI system** combining a **fine-tuned Mistral-7 B-Instruct** LLM with **vector search (FAISS)** and **embedding pipelines** (**Sentence Transformers**), **deployed on cloud infrastructure** for scalable symptom detection.Designed modular components (**NLP preprocessing, model inference, vector DB retrieval**) communicating via **APIs and Python** services, simulating real-world distributed system behavior.

Web-Based Facial Authentication with Anti-Spoofing System | 08/2023 | GitHub link

- Developed a web-based facial authentication system using **TensorFlow, OpenCV, and MediaPipe** to replace traditional CAPTCHA and password logins. Integrated **Flask-PyMongo and Flask-Bcrypt for secure, real-time authentication over HTTPS,** ensuring **encrypted client-server communication.** Employed **SSL/TLS protocols** to safeguard biometric data in transit and **prevent spoofing attacks.** Achieved **over 97% recognition accuracy,** enhancing both security and user experience.

Achievements

Grand Finalist & 1st Place in Healthcare Category, IIC 2024

- **Secured a Top 10 position out of 764 teams** in the grand finals of the **International Innovation Challenge.** Designed and **deployed an AI-driven medical assistant within 48 hours** during an intense hackathon. Fine-tuned Mistral 7B LLM model as part of a **Retrieval-Augmented Generation (RAG) system for accurate symptom detection and diagnosis.**

1st Runner-Up – Dark Pattern Hackathon 2023

- Built an ML-driven system to **detect fake reviews across 100+ Google Play Store apps.** Scraped app data using **BeautifulSoup** and **applied NLP techniques** to uncover deceptive patterns, promoting transparency and user trust.

Education

- **B.TECH:** Vellore Institute of Technology - Bhopal, **CGPA:** 7.43, **From:** 07/2022 **To:** 07/2026.

Certifications

- **Certified Ethical Hacker (CEHv12)** -by EC-Council | **Certification Number : ECC9680215374**
- **Generative AI** - by IBM | **Link**
- **Data Science: Transformers and Natural Language Processing** -by Udemy | **Link**
- **Bits and Bytes of Networking** -by Shounak Saha - Coursera | **Link**
- **Data Science Professional** -by Oracle | **Currently Pursuing**