**Profile**

Piyush Singh

Phone Number: (+91) 9205973514 Email ID: piyushsingh50105@gmail.com Address: 4C/2082, Vasundhara, Ghaziabad, 201012 LinkedIn: <https://www.linkedin.com/in/piyush-singh-328302289> GitHub: <https://github.com/PiyushVIT346>, Portfolio: [https://piyushvit346.github.io/Portfolio](https://piyushvit346.github.io/Portfolio/)

Aspiring AI engineer and CSE (AIML) student at VIT Bhopal, driven by a passion to solve real-world problems through intelligent systems. Throughout my engineering journey, I've built ML and NLP projects, participated in national hackathons, and led collaborative teams, all while exploring the depths of artificial intelligence, web technologies, and innovation. My goal is to contribute to impactful AI solutions that improve lives and push the boundaries of what technology can achieve. Constantly learning, building, and growing to become a well-rounded tech leader of tomorrow.

# Education

## B. TECH CSE AIML, VIT BHOPAL (2023-27) CGPA: 9.05

## 12th CBSE Board – 89.6%, Vanasthali PublIc School, Ghaziabad

10th CBSE BOARD – 96.4%, VANASTHALI PUBLIC SCHOOL, GHAZIABAD

# Skills: Machine Learning, GenAI, LLM, Advance RAG, LangChain, LangGraph, LangSmith Monitoring, MCP, Transformer/ HuggingFace, CrewAI, n-8-n, Ollama, Neural Network, Natural Learning Process, Computer Vision, Data Analyst, Data Science, DVC, MLOPS, MLFlow, Deployment using Vercel and Docker, API generation, FastAPI, Flask, Continuous Integration and development, Web Development

**Languages:** Python, C++, Java, JS, ReactJS, HTML, CSS

**Research Paper:**

* **EMG Signal Disease Detection** (Published)
* **Skin Disease Detection** (Pending to be published)
* **EEG Signal** (Pending to be published).

**Project**

**Water Potability Detection: MLOPS and CI/CD**

[**https://drive.google.com/drive/folders/1cnoZm1GkztMcnmQ2WjtaQMtvZVMBbnpo?usp=sharing**](https://drive.google.com/drive/folders/1cnoZm1GkztMcnmQ2WjtaQMtvZVMBbnpo?usp=sharing)

* Collected and processed raw data using web scraping techniques; structured data into clean CSV format with the help of Large Language Models (LLMs) for parsing and transformation.
* Performed hyperparameter tuning, model training, and versioning using MLflow, with proper model registry, tracking, and monitoring integrated into the pipeline.
* Developed and tested RESTful APIs using FastAPI for serving machine learning models and ensuring robust backend integration.
* Implemented Continuous Integration (CI) workflows using GitHub Actions for automating code testing and deployment.
* Conducted experiment tracking and version control through DagsHub for collaborative ML development.
* Built and compared multiple ML model versions, evaluating performance through accuracy, F1-score, and other key metrics**.**

**Advance RAG**

[**https://drive.google.com/drive/folders/199rmgvYn6VDyycDeIL\_QnxZ5JEwHd8Ki?usp=sharing**](https://drive.google.com/drive/folders/199rmgvYn6VDyycDeIL_QnxZ5JEwHd8Ki?usp=sharing)

* Implementing Hybrid RAG Search (using keyword search + Chroma DB retrieval + vector search) using LangChain and HuggingFace transformer and also having memory.
* Implementing Reranking (rank\_bm25, cross encoder and cohere).
* Implement Lost in middle phenomena and overcoming using merger and compression retriever.
* Implement RAG Fusion to combining results from multiple query variations.
* Flash Reranking as it ensures the most useful ones guide the generation.
* Implementing Context Compression Retriever, Self Retriever, Parent-child Retrieval, Sentence Window Retriever and HyDE Retriever.

**Kidney tumor detection and MLOPS**

[**https://github.com/PiyushVIT346/kidney-tumor-classification-mlops**](https://github.com/PiyushVIT346/kidney-tumor-classification-mlops)

* Developed a VGG16‑based deep learning model (~94% accuracy) for CT‑based kidney tumor detection.
* Designed and implemented a full MLOps pipeline using DVC (data/model versioning), MLflow/DagsHub (experiments), and GitHub Actions (CI/CD), ensuring end-to-end reproducibility.
* Containerized the model and deployed it via a Flask API, enabling scalable inference in production-like environments.

**Trained Model for CT-Scan, X-Ray and Brain tumor**

[**https://github.com/PiyushVIT346/health-hackathon**](https://github.com/PiyushVIT346/health-hackathon)

[**https://drive.google.com/drive/folders/1HgC4aK16dUtKjlNqlo3c3U4bvXzW4QYl?usp=sharing**](https://drive.google.com/drive/folders/1HgC4aK16dUtKjlNqlo3c3U4bvXzW4QYl?usp=sharing)

* Trained a deep learning model to classify medical documents and scans, accurately identifying whether an image was an X-ray, CT scan, or Brain MRI using image classification techniques.
* Enhanced model performance by applying Convolutional Neural Networks (CNN) and transfer learning with ResNet-50, improving accuracy on complex medical image datasets.
* Developed a chatbot for disease detection that interacts with users through symptom-based input, identifies potential illnesses using NLP, and provides basic first-aid suggestions and care tips.
* Built a personalized food timetable generator that considers the user's health history (e.g., past diseases), age, dietary preferences (veg/non-veg), and chronic conditions to generate nutritionally appropriate meal plans.

**Certified Courses:**

* **IBM GenAI Course**
* **Microsoft DP-900**
* **LangChain, LangGraph and Ollama-RAG Udemy**
* **Kaggle - Intermediate ML**
* **Computer Vision – GreatLearning**
* **DSA – Simplilearn**
* **Learn Advance C++ - Simplilearn**

**Badges:**

* **Leetcode** – leetcode pandas, leetcode 50 days code, leetcode November batch
* **Hackerrank** – Problem Solving, CPP (silver), Python (Gold), Java(silver)
* **AWS**

**Experience:**

**-LearnFlu Internship and Training** (September 2024 – February 2025): Completed a 4-month internship and training program at LearnFlu focused on Python programming, machine learning, statistics, and data analysis. Gained hands-on experience with data preprocessing, feature engineering, model development, and evaluation using libraries like Pandas, NumPy, Matplotlib, and Scikit-learn. Built and deployed multiple machine learning models for real-world business problems. Key projects included Iranian Churn Prediction using classification algorithms, Online Retail Customer Segmentation using RFM and clustering, Real Estate Price Prediction with regression models, and Wholesale Customer Segmentation through unsupervised learning. Strengthened understanding of business analytics, customer behavior, and predictive modeling for decision-making in real-world applications.

**-CDC Summer Internship** (June 2025 - July 2025): Completed 2 certified course

1. IBM GenAI: Learned to create ChatBot, NLP concepts and Handon practice on Watsonx IBM cloud. Explored Watsonx.ai for building and fine-tuning LLMs using foundational models hosted on the IBM Cloud platform.
2. Microsoft DP-900: Learned basic core concept of Cloud and Azure Platform and deploy projects on Azure. Worked with SQL, NoSQL, and Azure ML tools for model deployment. Learned data storage, governance, and analytics with Synapse & Power BI.

**Technical Core member at Freelancing Club** (January 2025): learned leadership, organized events and work Collaboratory.

**Achievements:**

* Participated in Adobe Hackathon.
* Participated in google AI Agent Hackathon.
* Participated in Health Hackathon.
* Participated Flipkart Hackathon.