Prateek Goswami

+91 79767 24248 prateekgoswami93@gmail.com LinkedIn GitHub Portfolio

Summary

Computer Science & AI undergraduate with hands-on experience in NLP, deep learning, and end-to-end ML pipelines. Built AI-powered chatbots, sentiment analysis systems, computer vision models, and multimodal architectures. Proficient in Python, HuggingFace Transformers, data preprocessing, feature engineering, and model evaluation. Comfortable working across the full data science lifecycle — from raw data to deployed solution.

Education

JK Lakshmipat University Jaipur, Rajasthan
B.Tech in Computer Science & Artificial Intelligence — CGPA: 9.09/10 *2023 – Present*

IIT Gandhinagar (Exchange Semester 3) Gandhinagar, Gujarat
Coursework: Computer Networks & Security, Foundations of AI

IIT Jammu (Exchange Semester 5) Jammu, J&K
Coursework: Deep Learning, Computer Vision, Artificial Intelligence, DBMS, Statistics

Experience

AI / ML Engineer Intern Remote
Patidar Agriculture — AgriTech (Agricultural Machinery) *Jun 2024 – Jul 2024*

- Built and deployed an NLP-powered chatbot for the company website to handle product queries and machinery recommendations — covering full pipeline from data preparation and preprocessing to deployment.
- Developed computer vision models for agricultural disease detection using image processing and predictive models for data-driven decision support; performed feature engineering, model training, and evaluation using Python, OpenCV, and Scikit-learn.
- Delivered end-to-end ML solutions integrated into a live production website; maintained clear technical documentation of model decisions, experiments, and outcomes.

Skills

Languages & Tools: Python, SQL, JavaScript, C, GitHub, Docker, Jupyter Notebook, Google Colab, Databricks

Machine Learning: Scikit-learn, PySpark ML, MLflow, Feature Engineering, Model Evaluation, Model Optimization, Predictive Modeling, Statistics

Deep Learning & Neural Networks: TensorFlow, PyTorch, Keras, CNN, LSTM, Attention Mechanisms, Transfer Learning

NLP: HuggingFace Transformers, BERT, Sentiment Analysis, Text Classification, TF-IDF, Chatbot Development, Tokenisation

Computer Vision: OpenCV, YOLOv5, Mediapipe, Object Detection, Image Captioning, VQA

Data Science: Pandas, NumPy, Matplotlib, Seaborn, EDA, Data Cleaning, Data Preprocessing, Data Manipulation, Data Visualization, Analytical Thinking

AI / LLM: LLM APIs (Claude, OpenAI, Groq), LangGraph, Prompt Engineering, RAG Pipelines, Agentic Workflows, Tool Use

Projects

AI Outreach Agent | *Python, LangGraph, FastAPI, Llama 3, SQLite, Google Sheets API, Google Drive API, RAG*
GitHub | Live Demo

- Built and deployed a LangGraph-based AI agent with tool-calling workflows that scrape company websites, score resume-company fit, generate personalized outreach emails, and evaluate outputs through hallucination detection and quality-scoring pipelines.
- Tested and debugged agent outputs to improve grounding and reduce hallucination-prone responses.
- Designed a human-in-the-loop review workflow with retrieval-based Style Memory, enabling future generations to adapt from previous user edits and approved responses.
- Integrated SQLite, Google Drive API, and Google Sheets API for persistent tracking of outreach history, evaluation metrics, approval decisions, and generated communications across the agent lifecycle.

Visual Question Answering (VQA) System | *PyTorch, ResNet/ViT, BERT, HuggingFace Transformers* | GitHub

- Built a multimodal neural network combining image feature extraction (ResNet/ViT) with BERT-based question encoding to generate natural language answers; evaluated on the VQA v2 benchmark.
- Designed image-text fusion architecture and analysed accuracy across open-ended and binary question types to guide model optimisation.

Real-Time Anti-Cheating Proctoring System | *Python, YOLOv5, Mediapipe, TensorFlow, OpenCV* | GitHub

- Built a real-time multi-model inference system detecting gaze deviation, head pose, hand gestures, and unauthorised devices; logged behaviour events with timestamped alerts for automated exam monitoring.
- Orchestrated multiple neural network models in a single pipeline, applying computer vision and predictive inference under real-world deployment constraints.

Customer Churn Prediction Pipeline | *PySpark, Databricks, MLflow, Scikit-learn*

- Built an end-to-end binary classification pipeline on Databricks using PySpark for large-scale data processing and Spark SQL for EDA across 7,000+ customer records.
- Performed feature engineering via StringIndexer and VectorAssembler; trained Logistic Regression and Random Forest models; tracked experiments with MLflow. Achieved AUC-ROC of 0.83 and 78.4% accuracy.

Image Captioning Model | *TensorFlow, CNN-LSTM, Bahdanau Attention, Flickr30k*

- Developed a CNN-LSTM captioning pipeline with Bahdanau attention on Flickr30k; improved BLEU-4 score through attention integration and systematic hyperparameter tuning.
- Implemented custom data loaders and training loops; visualised loss and accuracy curves using Matplotlib for experiment tracking.

Lip Sync Generation / Audio-Visual Synchronisation | *Python, Wav2Lip, Deep Learning, FFmpeg*

- Fine-tuned a Wav2Lip-based deep learning architecture to synchronise lip movements in video with arbitrary audio; built a preprocessing pipeline covering face detection, mouth-ROI extraction, and audio-video alignment.

Sentiment Analysis Pipeline | *Python, HuggingFace Transformers, Scikit-learn, NLP*

- Built an end-to-end NLP pipeline for multi-class sentiment classification; performed data cleaning, preprocessing, EDA, and feature extraction before model training.
- Compared TF-IDF + classical ML baselines with fine-tuned HuggingFace transformer models; evaluated using F1-score and confusion matrix; documented experiment results across runs.

Leadership & Achievements

Organising Head — College Hackathon JK Lakshmipat University
Led a cross-functional team of 200+ volunteers across logistics, scheduling, and event execution *2026*

- Coordinated sponsor outreach, participant onboarding, and real-time problem resolution across all hackathon tracks.

Teaching Assistant — C Programming JK Lakshmipat University
Selected based on academic performance and programming expertise to support 60+ students *2025*

- Conducted doubt sessions, graded assignments, and covered core C programming, debugging, and problem-solving techniques.

Dean's Honour List JK Lakshmipat University
Recognised for outstanding academic performance across consecutive semesters