Parit Kansal

paritkansal121@gmail.com | (+91) 639-895-0353

EDUCATION

HARCOURT BUTLER TECHNICAL UNIVERSITY
Bachelor of Technology in Computer Science & Engineering
May 2025 | Kanpur, India
CGPA: 8.3/10.0

DEWAN PUBLIC SCHOOL
Senior Secondary Education (Class XII)
May 2020 | Hapur, India
Percentage: 94%

DEWAN PUBLIC SCHOOL
Secondary Education (Class X)
May 2018 | Hapur, India
Percentage: 98%

PROFESSIONAL PROFILES

GitHub: **github.com/ParitKansal**
LinkedIn: **linkedin.com/in/paritkansal**

TECHNICAL SKILLS

PROGRAMMING LANGUAGES
• Python • C++ • C

AREAS OF EXPERTISE
• Data Science • Machine Learning
• Deep Learning • Computer Vision
• Natural Language Processing
• Graph Neural Networks
• Few-shot VLM Training

TOOLS & TECHNOLOGIES
• LangChain • MySQL • Power BI
• Microsoft Azure • Large Language Models • Web Scraping • CI/CD
• Retrieval-Augmented Generation
• Workflow & Agent •LangGraph

FRAMEWORKS
• TensorFlow • PyTorch • FastAPI

PROFESSIONAL SKILLS
• Team Collaboration • Analytical Problem Solving
• Technical Communication

CERTIFICATIONS

TensorFlow: Advanced Techniques
Coursera & DeepLearning.AI | June 2024

Machine Learning with Python
freeCodeCamp | April 2024

PROFESSIONAL EXPERIENCE

XELPMOC DESIGN AND TECH LTD. | Machine Learning Scientist
January 2025 – Present | Hyderabad, India

- **Live Store Analysis:** Contibuted in builting a CCTV-based real-time store monitoring system for a major global fast-food chain, analyzing **live RTSP feeds** to detect **customer entry/exit**, **staff presence at counters**, and triggering alerts when customers wait too long unattended. Implemented automatic detection of **unclean tables** after customer departure and **dustbin cleanliness**, generating manager alerts. Deployed using a **microservice-based backend** for continuous live processing.
- **Intelligent Document Understanding:** Fine-tuned **DONUT** for conditional key–value extraction from insurance claims, achieving **98% field-level accuracy**. Developed a **Gemma-3-4B (4-bit VLM)** automated judge to validate predictions against the **original document image**, reaching **96.67% document-level accuracy**. The production pipeline processes **100,000+ documents monthly**.
- **Integrated Lead Scoring Model:** Built an event-level scoring pipeline to predict post–site visit likelihood by integrating WhatsApp, web, and audio interactions. Engineered cumulative behavioral features for each event and trained a **ModernBERT** model on interaction transitions, using its outputs as features in a final **XGBoost** model, achieving **83% recall** in the top 13% events.
- **Web Visit Scoring:** Designed an end-to-end scoring pipeline to address a 2+ month feedback delay. Analyzed **8M+ sessions** using **XGBoost**, achieving **88% session-level recall** and **90% visitor-level recall**.
- **Multi-Document Detection:** Developed a synthetic document generation pipeline to train a **Mask R-CNN** model for multi-document detection, achieving **97.1% AP@[IoU=0.50:0.95]**.

ACADEMIC PROJECTS

COMPARATIVE ANALYSIS OF RECOMMENDATION ALGORITHMS | Machine Learning
June 2024
Benchmarked **SVD, SVD++, KNN, RBM, and AutoRec** models on **RMSE, Hit Rate, and Novelty**, providing a comprehensive performance comparison.

QUESTION PAIR SIMILARITY DETECTION | NLP
October 2023
Developed a **duplicate question detection model** on the **400K+ Quora dataset** using **Optuna-based hyperparameter tuning**, achieving **83.7% accuracy**.

RESEARCH PROJECTS

Few-Shot Layout Detection (In Progress)
Building a few-shot document layout detection model that generalizes layout understanding from a limited set of annotated exemplars. For instance, when a user provides bounding boxes around headings in a few sample images, the model learns the visual and spatial characteristics of those regions and predicts the corresponding heading boundaries in unseen documents.