

# Kshitij Singh

+91-8088184297 | [kshitij.tech06@gmail.com](mailto:kshitij.tech06@gmail.com) | [Linkedin](#) | [Github](#) | [Portfolio](#)

## EDUCATION

### Dayananda Sagar College of Engineering

*Bachelor of Engineering in Computer Science and Engineering (Data Science)*

Bengaluru, Karnataka

2023 – 2027

### Air Force School

*XII (CBSE - 93%)*

Bengaluru, Karnataka

2023

## PROJECTS

### Visual Language Translation & Typesetting Engine | *Python, PyTorch, OpenCV, PaddleOCR, LaMa* [GitHub](#)

- Engineered end-to-end CV+NLP pipeline automating manga translation through 4-stage architecture: text detection (YOLOv5), OCR extraction (PaddleOCR - 94% CER), inpainting (LaMa model), and context-aware translation, reducing manual effort by 85%
- Optimized text detection achieving 0.92 F1-score on custom manga dataset through transfer learning and post-processing filters; integrated multiple translation backends (Google Translate API, DeepL) with quality scoring to select optimal output
- Implemented smart inpainting with LaMa deep learning model achieving 28.5 dB PSNR for background reconstruction, preserving art style integrity while enabling clean text replacement
- Deployed production-ready system with FastAPI backend, batch processing for 50+ page volumes, and comprehensive evaluation framework measuring detection IoU, OCR accuracy, and translation BLEU scores

### Customer Intelligence System: CLV & Churn Prediction | *Python, LightGBM, XGBoost, SHAP, Streamlit* [GitHub](#)

- Engineered end-to-end customer intelligence System combining LightGBM for CLV prediction ( $R^2$  0.93) and XGBoost for churn classification (AUC 0.95), processing 500K+ transactions across 4,300 customers with 32 engineered behavioral features.
- Applied SHAP analysis for model interpretability, identifying recency and monetary value as top predictive drivers, enabling targeted retention strategy recommendations for high-value at-risk customers.
- Built an interactive Streamlit dashboard enabling real-time customer segmentation with automated feature engineering from raw transaction data.

### Full-Stack Ecommerce Platform with ML Recommendations | *Node.js, Express, PostgreSQL, React* [GitHub](#)

- Developed full-stack ecommerce application with Node.js, Express, PostgreSQL, and Prisma ORM, implementing secure REST APIs for authentication, product management, cart operations, and order processing
- Built robust backend with JWT authentication, role-based authorization, product CRUD, dynamic cart updates, stock validation, and order lifecycle handling
- Integrated React frontend with backend APIs featuring product catalog, detail pages, cart UI, and admin dashboard with image upload via Multer; designed scalable architecture using modular controllers and Prisma migrations

### CNN-Based Digit Recognition System | *Python, TensorFlow/Keras, Streamlit* [GitHub](#)

- Architected 6-layer CNN achieving 99.5% accuracy on MNIST through systematic experimentation with batch normalization, dropout regularization ( $p=0.5$ ), and Adam optimizer, outperforming baseline MLP by 12%
- Built interactive Streamlit web application with drawable canvas interface, displaying real-time predictions with confidence scores and class probability visualizations using cached model for instant inference
- Engineered preprocessing pipeline normalizing user input (grayscale conversion, bounding box detection, center-of-mass alignment,  $28 \times 28$  resizing) ensuring consistency with training data distribution

## TECHNICAL SKILLS

**Languages:** Python, C++, JavaScript, SQL, R

**Machine Learning:** TensorFlow, Keras, PyTorch, scikit-learn, XGBoost, LightGBM, SHAP, Optuna

**Deep Learning:** CNNs, Transfer Learning, Model Optimization, Computer Vision, NLP

**Data Science:** Pandas, NumPy, Matplotlib, Seaborn, Feature Engineering, A/B Testing

**MLOps & Deployment:** Streamlit, FastAPI, Docker, Model Versioning, REST APIs

**Frameworks & Tools:** React, Node.js, Express, Git, Linux, AWS

**Databases:** PostgreSQL, MongoDB, MySQL