Vikas Prajapati

AI/ML Engineer

+91 9161589883 | vp191198@gmail.com | linkedin.com/in/vikas1998 | github.com/VikasPrajapati1998

Summary

Results-driven AI/ML Engineer with 1+ years of industry experience specializing in computer vision and NLP solutions. Demonstrated expertise in optimizing deep learning models (reducing inference time by 40% while maintaining accuracy) and building scalable production systems with Django and FastAPI. Passionate about developing accessible AI applications that solve real-world problems across judicial, security, and enterprise domains.

Experience

Software Developer Oct 2024 – Present

Houston System *Greater Noida, Uttar Pradesh*

- **Architected** a real-time face detection and recognition system with multilingual chatbot integration, achieving 95% recognition accuracy across diverse lighting conditions and user demographics.
- **Implemented** robust Role-Based Access Control (RBAC) systems using Django RestAPI and FastAPI with secure JWT authentication, reducing unauthorized access attempts by 80%.
- **Developed** a comprehensive case management system for the Punjab and Haryana High Court, streamlining judicial workflows and reducing case processing time by 35%.
- **Established** data-driven project management using JIRA and Excel-based reporting, enabling accurate sprint planning and stakeholder communication.

Software Engineer Oct 2023 – April 2024

Multicoreware Inc. *Chennai, Tamil Nadu*

- **Spearheaded** end-to-end software lifecycle management for high-availability systems, reducing time-to-market by 20% through improved cross-functional coordination.
- **Engineered** state-of-the-art AI/ML solutions using MaxViT, YOLOv5, and Mask R-CNN models, optimizing inference performance by 40% using AIMET techniques.
- **Constructed** resilient Python backend services with Flask and FastAPI that supported 10,000+ concurrent users for a major East Coast client's Angular application.
- **Cultivated** high-performance team culture through Agile coaching and code reviews, reducing bug rates by 25% and improving overall code quality metrics.

Projects

Visitor Management System | *Python, Django, ORM, API, MySQL, Git* Feb 2025 – Present

- Developing an enterprise-grade Visitor Management System with Django and MySQL, implementing JWT-based RBAC authentication system supporting 3 distinct user roles.
- Creating dynamic employee availability scheduling that optimizes visitor meeting allocation and reduces scheduling conflicts by an estimated 60% .
- Implementing intelligent visitor identification system that reduces registration time by 75% for returning visitors while maintaining data integrity.
- Designing comprehensive admin controls for employee management with hierarchical permission structures compliant with organizational security policies.

Automatic Number Plate Recognition | *Python, DL, FastAPI, MySQL, Git* Nov 2024 – Jan 2025

- Engineered an ANPR system achieving 97% accuracy in diverse environments using YOLOv5, enhancing traffic monitoring capabilities for law enforcement agencies.
- Executed comprehensive data pipeline handling 50,000+ vehicle images across various lighting conditions and camera angles.
- Integrated the deep learning model with FastAPI backend, enabling <200ms response time for real-time recognition requirements.
- Optimized model deployment for edge devices, reducing computational requirements by 30% while maintaining recognition accuracy.

Court Guidance Case Management System | *Python, Tkinter, MySQL, Git* Oct 2024 – Oct 2024

- Designed an accessible kiosk-based Court Case Guidance System supporting users with diverse disabilities, increasing judicial system accessibility for underserved populations.
- Implemented multi-modal interaction methods with 98.5% uptime and average response time under 2 seconds.
- Enabled real-time case information retrieval through multiple identifier types, reducing information access time by 70% compared to manual processes.
- Optimized the system for public legal environments with enterprise-grade security protocols protecting sensitive case data.

Testorent (Qualcomm) | *Python, Deep Learning, CNN*, Git, Bitbucket, Ubuntu Oct 2023 – April 2024

- Optimized deep learning models across multiple frameworks while maintaining accuracy loss within <1%, enabling deployment across Qualcomm's heterogeneous hardware.
- Implemented advanced quantization techniques (INT8, INT16, FP16) using both post-training and quantization-aware training methods, reducing model size by 75%.
- Leveraged Qualcomm's AIMET toolkit to perform high-fidelity quantization and layer-by-layer sensitivity analysis, enabling efficient compression with minimal accuracy degradation.
- Engineered quantization-aware strategies and calibration methods validated through automated regression testing, ensuring production-ready model performance.

Education

PG-Diploma in Big Data Analytics
Center of Development of Advanced Computing, Bangalore, Karnataka *March 2023 – Aug 2023*

Bachelor of Technology in Electronics and Communication Engineering
Dr. Rammanohar Lohia Avadh University, Ayodhya, Uttar Pradesh *July 2018 – Oct 2022*

Technical Skills

Programming Languages: Python, R, C/C++, JavaScript, SQL, Shell Scripting
AI/ML Frameworks: PyTorch, TensorFlow, Keras, scikit-learn, ONNX, AIMET, Torchvision, OpenVINO
Web Frameworks: Django, FastAPI, Flask
Data Processing: NumPy, Pandas, SciPy, Matplotlib, Seaborn
Computer Vision & NLP: OpenCV, Torchvision, SpaCy, NLTK
Databases: MySQL, PostgreSQL, MongoDB
DevOps & Cloud: AWS (EC2, S3, RDS), Docker, Git, Bitbucket, CI/CD pipelines
Development Tools: VS Code, PyCharm, JupyterLab, Google Colab, Kaggle, Roboflow
Core Competencies: Deep Learning, Computer Vision, NLP, System Design, API Design, Agile/Scrum