Sandeep Chahal

AI Scientist | ML Engineer | Data Scientist

Email: chahal.sdp@gmail.com | Phone: +1 (224) 391-2179 | Chicago, IL, USA - 60654 (Open to relocation)
LinkedIn: linkedin.com/in/sandeepchahal | GitHub: github.com/chahalsandeep | Portfolio: chahalsandeep.github.io/portfolio

Professional Summary

AI/ML engineer with 5+ years of experience in architecting scalable systems and working with large, high-dimensional datasets with multi-modal systems (images, text, time series) in distributed environments. Strong foundation in model evaluation, pipeline optimization, and end-to-end MLOps. Proven track record driving cross-functional initiatives and collaborating across engineering, product, and analytics teams. Currently building decision-support tools using LLMs and Generative AI. Specialized in aligning user facing ML systems with real-world business impact.

Technical Skills

Programming Languages & Scripting: Python, C++, MATLAB, SQL, Bash.

Machine Learning & CV Libraries: Scikit-learn, PyTorch, TensorFlow, OpenCV, HuggingFace, LangChain, LangGraph, Prompt tuning, Vector Store, XGBoost, Pandas, NumPy, SciPy, ONNX, FFmpeg.

MLOps & Pipelines: MLflow, Apache Airflow, DVC, Streamlit, Jupyter Notebooks, A/B Testing, Databricks, PySpark, Prometheus (familiar), Docker, Kubernetes.

Cloud & Infrastructure: AWS (S3, EC2, SQS, DynamoDB, EKS, Step Functions, Lambda), GCP (Vertex AI, Colab, Dataproc, Cloud Storage), Snowflake.

Visualization: Matplotlib, Seaborn, Plotly, Tableau, Power BI, Snowflake Dashboards.

Robotics & Simulation: ROS, OpenAI Gym, Raspberry Pi, NUC, Phidgets, GPS, IMU, Sonar, Bluetooth, Wi-Fi, Camera.

Professional Experience

Freelance AI/ML Consultant — Remote
AI/ML Consultant — 08/2024 – Present

- **Designed** near real-time semantic and sentiment-driven feedback triage system for customer support using LLMs (LLama, OpenAI) and clustering algorithms to dynamically prioritize issues based on urgency, theme, and historical resolution patterns.
- **Developed** a predictive analytics pipeline to estimate manpower and automotive resource needs in urban areas based on dynamic population behavior, mobility trends, and external events.
- **Built** a machine learning framework for predicting baseball player performance using historical tournament data; evaluated Decision Trees, Random Forests, and transformer models for comparative accuracy.

Sift Science — San Francisco, CA
Senior Data Scientist → Data Scientist — 01/2022 – 08/2024

Senior Data Scientist

- **Led** cross-product revenue insights project by modeling behavioral signals across fraud, payments, and account protection uncovering $2M+ ARR opportunity and driving 100% adoption.
- **Architected** hybrid agentic (Multi Agent - LLM Chatbot) retrieval system (Snowflake + Confluence) to enable structured/unstructured query support increasing platform usage by 5×.
- **Optimized** real-time supervised learning pipeline with public data and verification signals improving fraud model precision by 15% and saving $1.2M+ quarterly.
- **Managed** model monitoring, deployments, workflow, customer escalations and coordinated cross-functional releases across engineering, product, and analytics teams.

Data Scientist

- **Built** multilevel and multistage adpaptive sampling pipeline with stages like class balancing, undersampling, oversampling, etc. improving model performance by 20% and reducing customer complaints by 60%.
- **Developed** statistical comparison framework using distributional metrics and hypothesis testing for explainable model benchmarking for XGBoost, RandomForest, etc. instead of using surrogate metrics like ROCAUC, F1, etc.
- **Defined** metrics, success, and performance KPIs; contributed to release guidelines, OKRs, and product roadmap planning.

Stats Perform Chicago, IL
AI Scientist II → Senior ML Engineer → ML/CV Engineer 07/2019 - 12/2021

AI Scientist II

- **Led** development of a computer vision system for sports player tracking and jersey recognition using ResNet50, R-CNN, U-Net, OpenPose, OCR, and optical flow, improving accuracy by ~20%.
- **Enhanced** tracking architecture by integrating sparse/dense optical flow with Kalman filtering for better coverage and stability.
- **Supported** scalable deployment using AWS Step Functions, Docker, and Kubernetes to automate long-form video processing.

Senior ML Engineer

- **Led** the design, development, and optimization of a modular MLOps platform supporting diverse model types across classification, regression, and clustering tasks.
- **Architected** an end-to-end ML pipeline integrating MLflow, Cortex, and Neptune on AWS, improving model release efficiency and reducing deployment time by 60%.
- **Worked** cross-functionally to define workflows and CI/CD automation for model lifecycles.

ML/CV Engineer

- **Developed** modules of an OCR-based video recognition system and time series and Analysis using SSD/YOLOv3 to process long-form footage, reducing compute/storage by ~30%.
- **Applied** frame reduction and clustering techniques with PCA and FFNNs (downsampling 60fps to 15–30fps) to increase inference throughput.
- **Managed** experimentation and model optimization workflows using MLflow and ONNX-based deployment.

Fetch Robotics *05/2018 - 08/2018*
AI Intern San Jose, CA

- **Improved** robotic QA process by designing an automated testing framework using computer vision techniques, increasing testing uptime by 50% and reducing manual verification.
- **Developed** CV-based detection scripts (using barcode, QR code, Aruco, AprilTag etc.) for real-time monitoring of robot tasks, contributing to faster release validation and deployment cycles.

SignalAware *06/2017 - 08/2017*
Predictive Analyst Intern Sunnyvale, CA

- **Built** Bluetooth and Wi-Fi–based localization algorithms to enable outdoor device tracking for connected mobile devices.
- **Engineered** predictive models to analyze user movement and forecast next likely locations, enabling advanced mobility insights for marketing and analytics use cases.

Education

Master of Science in Computer Science - University of Texas at Arlington *2019*
Research: Generating an Adaptive Path Using RRT Sampling and Potential Function with Directional Nearest Neighbor. Combined the advantages of sampling-based global planning with local navigation using potential-field functions. Designed a hybrid control composition scheme similar to null-space control to reduce re-planning, incorporate runtime constraints, and enable adaptive, efficient path generation.
Notable Courses: Algorithms, Data Analysis & Modelling Techniques, Artificial Intelligence, Computer Vision, Robotics, Unmanned Vehicle Systems.
Bachelor's in Information Technology - JNTUH *2015*

Academic Experience

- **ML Researcher** at LearnLab, UTA - Conducted research on predictive solutions for assistive technology.
- **Teaching Assistant** for AI course, UTA - Assisted in course development and student mentoring.
- **System Administrator**, UTA - Managed IT infrastructure and user support systems.

Soft Skills

- Cross-functional collaboration and team leadership
- Technical communication and storytelling
- Problem-solving and innovation
- Project management and strategic planning
- Stakeholder management and cross-functional coordination