YASHIKA SINGH

102 1/2 Western Avenue, Floor 1, Jersey City, NJ 07307

+1(646)242-3188 • yashika6499@gmail.com • linkedin.com/in/yashika-singh-944700195 • yashika-portfolio.streamlit.app

Summary

Innovative and driven AI/ML Engineer with 3+ years of hands-on experience building end-to-end solutions in Machine Learning, Generative AI, NLP, and Data Engineering. Adept at designing real-world applications—from AI-driven interview platforms to healthcare diagnostics—powered by deep learning, LLMs, and scalable data pipelines. Holds a filed patent for a novel authentication system, reflecting a strong foundation in problem-solving and originality. Passionate about pushing the boundaries of AI to craft impactful, human-centric solutions that drive automation, enhance decision-making, and bridge the gap between research and real-world use.

Education

Yeshiva University, Katz School of Science and Health, New York, NY
Master of Science in Artificial Intelligence, Current GPA: 3.7/4.0 — Expected May 2025
Galgotias College of Engineering And Technology, Greater Noida, India
Bachelor of Technology in Computer Science, CGPA: 3.5/4.0 — Aug 2021

Work Experience

AI/ML Engineer Intern, Product Manager Accelerator, Boston, MA (Remote) — Jan 2025 - Present

- Develop a video-based, LLM-powered AI interviewer using Generative AI, NLP and Computer Vision for Product Managers, integrated with a virtual avatar and API-enabled scripts to simulate real-world interview experiences. Scheduled for market release in May 2025, with paid customers onboarded.

- Utilize cutting-edge conversational LLMs (e.g., Anthropic's Claude) along with advanced prompt engineering techniques to generate natural, human-like follow-up questions, enhancing the realism of AI-driven interviews and ensuring a more authentic and dynamic interview experience.

- Collaborate with cross-functional teams to enhance product features, including product sense case studies and theme-based interviews.

Data Scientist, Axtria, Noida, India — Apr 2022 - Dec 2023

- Developed a K-Means clustering model using pharmaceutical engagement data to segment HCPs (Healthcare Professionals) based on behavioral patterns, enabling stakeholders to optimize marketing strategies, personalize outreach, and enhance overall HCP engagement.

- Developed and maintained ETL pipelines using Matillion, SQL and Python.

- Automated complex data processing workflows using DAGs in Apache Airflow, reducing manual intervention and enabling autonomous operations in dynamic environments.

- Integrated MLOps best practices with CI/CD pipelines to automate AI model training, deployment and monitoring on Databricks, ensuring scalability and reliability.

Data Analyst, Axtria, Noida, India — Aug 2021 - Apr 2022

- Optimized automated pipelines in Azure Data Factory and Amazon Apache Airflow, reducing data processing errors and improving overall efficiency.

- Supported in maintenance and advancement of the company's product Axtria Datamax, a cloud-based data management product.

- Created interactive dashboards and data visualizations using Power BI and Tableau to deliver insights and support decision-making for internal stakeholders.

Patent

Patent: PEHCHAAN - An Authentication System and Methods for Social Media Accounts (Filed on Nov 8, 2020, Application No. 202011048750) for effective reduction of fake social media accounts.

Projects

Automatic Teaching via Vision-Language RAG (Ongoing): Develop a Vision-Language RAG model integrating advanced AI, visual data, and textual information, leveraging open-source models like Whisper and Llama 3.2 11B Vision-Instruct Model to enhance interactive learning and enrich the educational experience.

Drowsiness Detection for Drivers: Built a CNN-based, real-time system using TensorFlow to detect driver fatigue and improve road safety with validation accuracy around 95%.

Detection of Cardiomegaly in Dogs: Created a deep learning model using custom CNN for X-ray analysis and automating veterinary diagnostics with an accuracy of 76.75%, surpassing the benchmark set by VGG16 (75%).

Skills and Certifications

Programming Languages: Python, C, C++, JavaScript, SQL

Developer Tools/Cloud: VS Code, Amazon Apache Airflow (DAGs), Azure Data Factory, AWS, Google Cloud Platform (GCP), Git, TensorFlow, Streamlit, Tableau, Microsoft Excel, PowerBI, PyTorch, Computer Vision, Anaconda, Jupyter, Google Colab, Pandas, Scikit-learn, Matplotlib, Data Pipelines, Databricks

Skills: Data Analysis, Machine Learning, Deep Learning, Natural Language Processing (NLP), Convolutional Neural Networks (CNN), Automatic Speech Recognition (ASR), Advanced Analytics, Predictive Modeling, Statistics, Mathematics, Exploratory Data Analysis (EDA), Whisper, Large Language Models (LLM), Prompt Engineering, Generative AI, Retrieval-Augmented Generation (RAG), MLOps, CI/CD Pipelines, MySQL, PostgreSQL, Extract-Transform-Load (ETL), HTML5, CSS, Technical Writing and Documentation

Trainings: Advanced Certificate Program in Data Science - upGrad, Machine Learning A-Z™: Hands-on Python in Data Science - Udemy, Deep Learning Specialization - Coursera, Python Certificate - HackerRank, C++ - Coding Ninjas, JavaScript Essentials - Udemy, 50 Days of DSA Python Data Structures Algorithms LEETCODE

Soft Skills: Problem-Solving, Critical Thinking, Adaptability, Collaboration, Communication, Leadership, Attention to Detail, Time Management, Innovation