

Paritosh Singh

(213) 820-2207 | paritoshsinghrahara@gmail.com | [LinkedIn](#)

EXPERIENCE

USC Center for AI in Society

Research Fellow

Los Angeles, CA

January 2022 - Present

Project funded by the Department of Homeland Security

- Developed a Python program for identifying vulnerable water pipes in Los Angeles, leveraging earthquake fault lines and liquefaction areas to bolster resilience for critical services.
- Devised resilient water infrastructure plan for LADWP (LA Department of Water and Power), determining top 3 KPIs via integer programming and network flow modeling.
- Optimized existing code under the lab to enhance pipeline replacement planning by up to 40%.

Solidigm

Software Engineer (Machine Learning)

Rancho Cordova, CA

May 2022 – May 2023

- Developed a machine learning pipeline to automate Failure Analysis of NAND Solid State Drives (SSD).
- Performed feature extraction on 500 GBs of raw data to create a processed data frame compatible with ML libraries.
- Implemented code optimization and parallelization techniques, reducing runtime from 1 hour to less than 20 minutes.
- Designed pipeline automated debugging of faulty SSD drives and reduced the product development life cycle by 30%.

USC Brain-Body Dynamics Lab

Robotics Intern

Los Angeles, CA

September 2021 – March 2022

Project funded by the Defense Advanced Research Projects Agency (DARPA)

- Designed a real-time multi-channel Data Acquisition (DAQ) System for a quadruped robot using Arduino and Python.
- Spearheaded the implementation of cross-learning practices across the lab by initiating and curating software design documentation in LaTeX, benefiting over 30 team members.
- Built DAQ system reduced the development costs by 60% compared to existing solutions.

CSIR - Central Scientific Instruments Organisation

Machine Learning Research Fellow

Chandigarh, IN

January 2021 - July 2021

- Developed machine learning models to identify causes of Anemia in Indian women.
- Attained 80% accuracy using Python with XgBoost, Decision Tree, Random Forest, and SVM algorithms.
- Programmed robust data preprocessing on a large-scale Demographic Surveys dataset (800k+ Rows, 2000+ Columns) with Python and Vaex, enhancing data readiness for ML model training.
- Enabled insightful health analysis by predicting 20 crucial factors causing Anemia.

PROJECTS

Automating W2 Forms Processing

LayoutLM Transformer, AWS Textract OCR

- Designed a multimodal (Text and Image) deep learning end-to-end pipeline for automating W2 forms processing.
- Achieved SOTA accuracy using LayoutLM architecture with 99% classification across all 26 labels.

Named Entity Recognition on Patient Notes

DeBERTa and RoBERTa Transformer, NLTK Library

- Utilized RoBERTa, and DeBERTa architectures for annotating phrases in patient notes.
- Achieved 88% F-1 score with machine learning models for accurate phrase annotation.
- Exceeded USMLE Clinical Skill Step 2 Examination's benchmark performance.

EDUCATION

University of Southern California

Los Angeles, CA

Master of Science in Computer Science (Artificial Intelligence) / GPA: 3.9

May 2023

Panjab University

Chandigarh, IN

Bachelor of Engineering in Computer Science & Engineering (with honors) / GPA: 3.8

July 2021

SKILLS

Python (TensorFlow, PyTorch, Transformers, OpenCV, Scikit-learn, Pandas, NumPy, spaCy, NLTK, Flask, Selenium, Vaex, Gurobi), R, C++, JavaScript, SQL, NoSQL, MapReduce, Spark, PowerBI, Tableau, Git, Linux, Docker, GCP.