

HITESH NARAYANA

Los Angeles, CA | (213) 681-2369 | hiteshna@usc.edu | linkedin.com/in/hiteshnarayan/ | hit07.github.io/ | github.com/Hit07

SKILLS

Languages: Python (proficient), SQL(proficient) C/C++ (intermediate), Java(moderate), Scala(moderate)
Machine Learning: PyTorch, SQL, Pandas, Numpy, Scikit-Learn, HuggingFace, NLTK, Keras, TensorFlow
Cloud & Tools: Apache Spark, Hadoop, AWS, Azure, GCP, SAS, GCP, Git, Docker, Fast API, MongoDB, SQL, Power BI, Excel

PROFESSIONAL EXPERIENCE

University of Southern California Los Angeles, CA
Research Assistant, HUMANS Lab under Prof. Emilio Ferrara July 2024-Present

- Engineered multi-threaded scraper and Apache Spark pipeline for distributed data processing, processing 23M+ posts for election discourse analysis using HPC for efficient execution.
- Designed a scalable ETL pipeline to process 15GB+ of Twitter/X data, leveraging data collection for predictive analytics and sentiment analysis (NLP, TF-IDF) to identifying key political narratives and user behavior patterns.

Genpact Bangalore, India
Business Analyst July 2023-December 2023

- Designed payroll trend reporting dashboards using Power BI for key KPI metrics, leveraging Azure data services to cut down report delivery time by 40% and enabling real-time insights for strategic decision-making.
- Implemented event-driven ETL pipelines on AWS Glue, enabling serverless queries for automating data extraction, transformation, and loading processes to enhance operational efficiency by 30% for a SaaS employee data.

EDUCATION

University of Southern California Los Angeles, CA
Master of Science in Computer Science January 2024-December 2025
Coursework: Data Science Machine Learning, Information Retrieval, Analysis of Algorithms, Database Systems

Bangalore Institute of Technology Bangalore, India
Bachelor of Engineering in Information Science and Engineering August 2019-July 2023
Coursework: Data Structures and Algorithms, Machine Learning, Artificial Intelligence, Big Data
CGPA: 9.0 / 10.0

PROJECTS

FitGPT: AI-Powered Fitness Coach, HackSC'24 SoCal Runner-up

- Developed a React frontend with Flask backend, integrating Firebase and OpenAI GPT-4o-mini for personalized workout recommendations.
- Implemented advanced prompt engineering and Elasticsearch to process (VO2 max, walking asymmetry) and historical workouts, enhancing recommendation accuracy by 20% and enabling real-time fitness insights.

Food Vision: Multiclass Image Classification using CNN

- Engineered custom CNN model (TinyVGG) for food classification, achieving 90% accuracy on 100K+ images; optimized training pipeline with CUDA and automated scheduling, reducing training time by 40% and improving classification speed by 15%.

Human Activity Classification - Body Sensor Data

- Orchestrated ML pipeline for sensor-based classification, attaining 94% accuracy across 7 activities.
- Applied feature selection via data slicing and feature extraction on time-series data using (RFE, PCA, P-Tests); Implemented SMOTE with CV-tuned (L1 & L2 regularization) to address class imbalance, boosting overall accuracy by 12%.

PUBLICATIONS

- [1] Balasubramanian, A., Zou, V., Narayana, H., You, C., Luceri, L., & Ferrara, E. (2024). "A Public Dataset Tracking Sial Media Discourse about the 2024 U.S. Presidential Election on Twitter/X." arXiv preprint arXiv:2411.00376
- [2] Narayana, H. et al. (2023). "E-voting System Using Blockchain Technology and Homomorphic Encryption." International Journal for Research in Applied Science & Engineering Technology (IJRASET), 11(7), pp. 54573-54582. DOI: 10.22214/ijraset.2023.54573

LEADERSHIP

- Collaborated in a team of 3 to develop FitGPT, securing runner-up position at HackSC's SoCal Tech Week hackathon.
- Mentored two graduate students on course selection, research opportunities, and career development at USC.