HITESH NARAYANA

+1 (213)68-12369 • hiteshna@usc.edu • linkedin.com/in/hiteshnarayan/ • github.com/Hit07 • hit07.github.io/

EDUCATION

UNIVERSITY OF SOUTHERN CALIFORNIA — Los Angeles, CA
Master of Science in Computer Science — *December 2025*

GPA: 3.5/4.0 **,** Coursework: **Data Science Machine Learning,** Analysis of **Algorithms**, **Database Systems,** Information Retrieval

BANGALORE INSTITUTE OF TECHNOLOGY — Bangalore, India
Bachelor of Technology in Information Science and Engineering — *July 2023*

GPA: 9.0/10.0, Coursework: Data Structures and Algorithms**,** Machine Learning, Artificial Intelligence, **Big Data**

SKILLS

Languages: **Python** (Proficient), C/C++(intermediate), Java(moderate), R(basic)

Machine Learning: pandas, NumPy, scikit-learn, statsmodels, matplotlib, seaborn, MATLAB, PyTorch

Big Data Tools: Apache Spark (PySpark, Dash), Hadoop **Databases**: MySQL, MongoDB, PostgreSQL,

Cloud & Tools: AWS, Azure, Docker, GCP, Git, Flask, RESTful APIs, Microsoft Office Suite, Tableau**,** Power BI

PROFESSIONAL EXPERIENCE

USC Information Science Institute — Los Angeles, CA
Research Assistant , HUMANS Lab (HUmans MAchines Networks Society) | *Web Scraping, ETL, Apache Spark, Scikit-learn.* — *July 2024 - Present*

- Collaborated with a team of researchers to explore the effects of social media discourse on public opinion and hate in the upcoming 2024 election led under **Prof. Emilio Ferrara.**
- Developed machine learning models (**K-Means Clustering**) to predict user behaviour and sentiment shifts using X data.
- Engineered an **ETL** pipeline utilizing **Apache Spark**, processing over **100+GB** and achieving a notable increase in efficiency that facilitated quicker data analysis for strategic decision-making.

Genpact — Bangalore, India
Business Analyst | *SQL, Power BI, DAX, ETL, SAS, Excel, Tableau* — *July 2023 – December 2023*

- Optimized SQL queries by rewriting complex joins and indexing critical tables; achieved a **30%** reduction in query runtime which enhanced data retrieval speed for **5+ reports** generated weekly.
- Created an automated reporting utilizing Power BI, which enabled real-time monitoring of payroll trends, reduced turnaround time for report delivery.

PROJECTS

Food Vision: Multiclass Image Classification using TinyVGG | Deep Learning | *Python, PyTorch, CUDA, torchvision*

- Built an **deep learning** image recognition system for food classification, handling **100K+ data points**; integrated feedback loops that improves learning efficiency.
- Engineered a custom CNN model (TinyVGG) that surpassed the baseline accuracy by **20%,** enabling precise categorization of diverse food items and improving classification speed by **15%** during testing phases.
- Improved model performance through hyperparameter tuning techniques , achieving a **15%** increase in accuracy while cutting training time by **40%** using automated scheduling.

Human Activity classification -body sensor data | 7 activities (walking, cycling, etc...) | *Scikit-Learn, Python, pandas*

- Conducted data slicing and advanced feature extraction on **time-series data**, employing RFE, **PCA**, and P-tests to isolate relevant features
- **Tuned hyperparameter** via CV for multi-class Ridge, Lasso and Naive Bias(gaussian) models, achieving peak accuracy of 94% in Lasso
- Implemented **SMOTE** to address class imbalance, boosting overall OVR classification accuracy by 12%.

Full Stack Blog with Analytics | ***Flask, SQLite, Bootstrap, KPI***

- Designed a responsive blog application using Flask for enhanced user experience; facilitated post creation/editing capabilities.
- Developed and implemented data visualization tools that enhanced the tracking of key metrics; integrated CKEditor and Gravatar systems into user profiles, enabling personalization features.

Predictive Modelling Using Multivariate Regression | Pandas, Scikit-Learn, Matplotlib, Seaborn

- Developed a predictive algorithm for real estate prices data; achieving peak R-squared score; validated through **(F-test, t-tests).**
- Implemented **log transformation** techniques, resulting in a 20% increase in model accuracy by effectively reducing prediction errors.
- Conducted comprehensive exploratory data analysis (**EDA**) and feature engineering to enhance model performance and interpretable

PUBLICATION

[1] "*E-Voting System Using Blockchain Technology and Homomorphic Encryption"* published in *International Journal for Research in Applied Science & Engineering Technology (**IJRASET**), Volume 11, Issue VII, July 2023*. [DOI: 10.22214/ijraset.2023.54573]