

# Sudhanshu Kumar

631-739-3555 / [sudhanshu.kumar@stonybrook.edu](mailto:sudhanshu.kumar@stonybrook.edu) / [linkedin.com/in/sudhanshu-kumar7k7](https://www.linkedin.com/in/sudhanshu-kumar7k7) / [github.com/sudhanshu7k](https://github.com/sudhanshu7k)

## EXPERIENCE

### Research Aide (Data Engineer)

Stony Brook Medicine

May 2024 – Present

Stony Brook, NY

- Analyzed clinical N3C COVID-19 data nationwide to **predict Long COVID trends** and pinpoint COVID sensitive demographic areas.
- Engineered models with ICD-10 and SNOMED codes to **explore autoimmune disease** effects on Long COVID.
- Leveraged statistical and **machine learning methods** to deliver actionable insights for healthcare stakeholders.

### Software Engineer

Harman International (Samsung)

Oct 2020 – Aug 2023

Bangalore, India

- Spearheaded projects integrating Harman's digital cockpit with machine learning, utilizing **random forests** and **gradient boosting** for electric vehicle (EV) route optimization, leading to a **40%** improvement in route accuracy and a **20%** reduction in travel time.
- Developed the Stress-Free Routes feature using **time series forecasting** and **k-means clustering** for real-time traffic and charging station data, resulting in a 24% decrease in user-reported stress levels during journeys
- Deployed **convolutional neural networks (CNNs)** and **recurrent neural networks (RNNs)** for driver behavior analysis, including aggressive driving detection and drowsiness monitoring, enhancing driver safety through personalized feedback

### Software Engineer Intern

Birla Group of Industries

May 2018 – July 2018

Kota, India

- Built a **purchase propensity model** to identify high-propensity leads for focused lead prospecting. This model prioritized leads, resulting in an **18%** increase in sales over a period of 3 months.
- Examined data on active clients and created visualizations to inform marketing tactics and product strategies. Facilitated data-driven decision-making processes, resulting in a **40%** decrease in time-to-insight and a **25%** increase in overall data utilization efficiency.

## PROJECTS

### Colorizing Black-and-White Images with GANs /

- Constructed a GAN-based model for colorizing black-and-white images, using a U-Net with ResNet18/34 backbones.
- Applied advanced image augmentations(Albumentations, TTA) to improve training robustness and testing accuracy.
- Experimented with different optimizers for Generator and Discriminator, significantly enhancing color accuracy and realism.

### Financial Chatbot with Retrieval-Augmented Generation (RAG) /

- Engineered a Python-based financial chatbot using **RAG** to provide accurate and current financial insights.
- Integrated advanced web scraping with semantic search models like **all-mpnet-base-v2** for information retrieval and response generation.
- Leveraged **Qwen/Qwen-14B-Chat-Int4** model for high-quality response generation, enhancing chatbot's ability to deliver precise advice.

### HISAB Fintech Application /

- Designed a Fintech mobile app with integrated machine learning algorithms to enable MSMEs to track finances, manage accounts, and optimize inventory. Leveraged **AI** to build features such as **predictive stock alerts** and **automated order generation** based on sales trends.
- Implemented advanced data analytics to identify unique **market trends** and **demand patterns**, leading to the successful **commercialization** of the product to a **Singapore-based NBFC**.

## TECHNICAL SKILLS

**Languages:** Python, SQL, R, Java, C/C++, Bash/Shell, JavaScript

**Frameworks:** TensorFlow, PyTorch, Scikit-learn, Keras, FastAPI, Flask, Streamlit

**Developer Tools:** Git, Docker, Kafka, AWS, GCP, Tableau, Power BI, MLflow, Spark, MongoDB

**Libraries:** Pandas, NumPy, Matplotlib, Seaborn, XGBoost, LightGBM, Plotly, NLTK, SpaCy, Hugging Face Transformers, OpenCV, StatsModels, NetworkX

## EDUCATION

### Stony Brook University

Master of Science in Data Science

Stony Brook, NY

Aug. 2023 – May 2025

[Distributed Systems, Data Science, Machine Learning, NLP, Statistical Learning]

### Vellore Institute of Technology

Bachelor of Technology in Computer Science

Vellore, India

Jul. 2016 – May 2020

[Data Structures, Algorithms, Operating Systems, Network Security, Database Management]