

Yash Kishor Sanap

571-631-9400 | ysanap@gmu.edu | linkedin.com/in/yash-sanap/ | github.com/yashsanap14

EDUCATION

George Mason University

Master of Science in Computer Science GPA: 3.83/4

Fairfax, VA

Aug. 2024 – Expected May 2026

Mumbai University

BE of Computer Engineering CGPA: 9.34/10

Mumbai, India

Aug. 2018 – May 2022

EXPERIENCE

Data Science Intern

City Of Virginia Beach IT

Jun 2025 – Aug 2025

Virginia Beach, VA

- Analyzed multi-month Sentinel-5P satellite imagery using Python geospatial libraries (GeoPandas, Rasterio) to map NO₂ pollution across Virginia; identified seasonal emission trends (spring–summer dips) and regional hotspots (Arlington vs. Virginia Beach), delivering actionable insights for environmental monitoring and public health
- Built an automated end-to-end image segmentation pipeline for reservoir water-level monitoring: integrated automated dataset labeling (Roboflow + Grounding DINO) and trained YOLOv11+SAMv2 models, achieving 95% detection precision and 90% mAP; solution scales to 7K+ images and minimizes manual annotation effort
- Implemented time-series forecasting with Nixtla's TimeGPT to predict coastal water levels from historical trends; delivered accurate short-term forecasts aligned with USGS gauge readings, enabling proactive flood-risk management and timely community alerts
- Orchestrated scalable data workflows on cloud infrastructure: leveraged AWS SageMaker Studio Lab and Coiled (Dask) to process NOAA National Water Model forecasts, enabling distributed analysis and reproducible hydrologic insights to support scientific research and emergency planning

Data Scientist

Capgemini Technology Services

Aug 2022 – Jun 2024

Mumbai, India

- Developed and optimized end-to-end data pipelines using Google Colab, Python, and SQL to process, analyze, and visualize large datasets; improved model inference speed and accuracy through workflow optimization
- Designed and deployed cloud-native machine learning solutions on AWS (EC2, S3, Lambda, RDS) for real-time analytics, reducing operational costs by 30% while ensuring scalability and security
- Built and managed REST APIs to serve ML model predictions, enabling seamless microservice integration and enhancing data-driven decision-making across applications
- Implemented NLP and LLM-powered conversational AI systems, collaborating with UX teams to create intuitive, AI-driven chatbot experiences; applied Agile methodologies to streamline development cycles

PROJECTS

Air Pollution Analysis with Sentinel-5P

Personal Project

Jan 2025 – May 2025

Virginia, USA

- Built Python pipeline (GeoPandas, Rasterio, Matplotlib) to analyze NO₂ satellite data
- Automated retrieval, cloud masking, and visualization of seasonal/regional pollution patterns

Total Water Level Forecast Extraction

City of Virginia Beach – VBDEM Support

Jul 2025

Virginia Beach, VA

- Created Python pipeline to extract NOAA NWM v3.0 NetCDF TWL forecasts for Atlantic–Gulf region
- Automated retrieval and processed data with Xarray + Coiled for distributed analysis

TECHNICAL SKILLS

Languages: Java, Python, C/C++, SQL (Postgres), JavaScript, HTML/CSS, R

Frameworks: React, Node.js, Django, Express.js

Developer Tools: Git, Docker, TravisCI, Google Cloud Platform, VS Code, Visual Studio, PyCharm, IntelliJ, Eclipse

Libraries: pandas, NumPy, Matplotlib, rasterio, Geopandas