

Kumar Saurav Jha

Arlington, VA | 510-962-0224 | kumarsaurav.jha@gwu.edu |
<https://www.linkedin.com/in/kumar-saurav-jha-063a63168/> | <https://kumarsauravjha.github.io/>

EXPERIENCE

George Washington University, Research & Teaching Assistantship Aug 2024 - current

- Working under Prof Hazim Shatnawi on a novel GAN-driven knowledge distillation framework for real-time object detection on edge devices [🔗](#); and creating a curriculum for “Generative AI and Transformers” for the department; also TAing for *Visualization of Complex Data* under Prof Reza Jafari

Deloitte India LLP, Consultant Oct 2022 - July 2023

- Reduced pipeline downtime by 25% and lowered annual repair costs by \$ 1.2M by designing predictive ML models (classification and anomaly detection) for *Shell Global Solutions Inc.*,
- Leveraged NLP and Computer Vision techniques to analyze large volumes (~500k) of equipment failure reports and technician notes, extracting critical insights and automating root-cause analysis.

Accenture Solutions Pvt Ltd, Data Analyst Sep 2019 - Oct 2022

- Implemented automated data validation process using Power BI for Cargill Inc., eliminating 2-hour daily manual effort by directly integrating SQL scripts into dashboard and enabling real-time tracking of system replication metrics.
- Automated SAP S/4HANA tasks like Shipping Point Creation and Fixed Asset Management, using SQL scripts for *Cargill Inc.*, reducing manual efforts by 35% and cutting processing time by 50%.
- Designed payment inquiry gateway for *Nationwide Mutual Insurance*, integrating automated text and email alerts, leading to 20% increase in customer engagement and satisfaction.
- Worked closely with business stakeholders to define data requirements, improving the accuracy of KPIs and providing actionable insights for operations team.

PROJECTS

World Bank: Transport Cost Modelling for Sub-Saharan African Countries [🔗](#) Jan 2025 - ongoing

- Working with the World Bank team (Transport Global Knowledge Unit) to develop a calibrated transport choice model for efficient mineral transportation, better infrastructure planning, and future investments.

Author Profiling on Blogpost Corpus - Academic [🔗](#) Nov - Dec 2024

- Implemented age-group prediction using blog text, author gender, and occupation on a Fully Connected Neural Network, conducting a comparative analysis of static vs. contextual embeddings.

Highlight Reel: Automated Sports Content Generation - Independent [🔗](#) Oct - Nov 2024

- Created a pipeline for Volleyball highlight reel generation using OpenCV, metadata-based game liveness filtering, and custom features like speed & acceleration; Achieved 90% recall with Logistic Regression

Demand Prediction and Strategic Customer Segmentation - Academic [🔗](#) Oct - Dec 2023

- Used XGBoost model for daily/ weekly demand prediction for a supermarket; reducing overstock costs and enabling targeted marketing by segmenting customers (KMeans) based on spending behavior.

TECHNICAL SKILLS

Languages: Python, R, SQL, CQL (Cypher Query Language), HTML, CSS

Databases: MySQL, MongoDB, and Neo4j

Tools: Excel, Tableau, Power BI (PowerQuery, DAX), PyTorch, Tensorflow, Scikit-Learn

Product Development: Agile Methodology, HP ALM, MVP model, Git, and Flask

Certification: IBM Machine Learning Specialization, AWS Certified Developer - Associate

EDUCATION

The George Washington University

Master of Science, Data Science [🔗](#)

Washington, DC
expected May 2025

University of Pune

Bachelor of Engineering, Electronics and Telecommunication

India
May 2019