

Saransh Kumar

College Park, MD | +1 (240) 413 1734 | saranshk@terpmail.umd.edu | [linkedin.com/in/saransh-kr](https://www.linkedin.com/in/saransh-kr) | github.com/saranshkr

PROFESSIONAL EXPERIENCE

Data Scientist | Python, SQL, AWS, Big Data technologies

Apr 2022 - Aug 2024

Samsung Research & Development (R&D) Institute - Delhi, India

- Analyzed over **500 TB** of viewership and advertisement data from the global Samsung SmartTV network, identifying **trends** and user behavior patterns that improved ad relevance by **20%** for over **10 million** users.
- Conceptualized and implemented end-to-end **data pipelines** for live **big data** streams with **Python, SQL**, and cloud technologies, reducing data processing and analysis time by **30%** and enabling real-time insights across **12+** markets.
- Designed a unified **dashboard** for customizable **data visualization** across **130+** services to communicate key insights and actionable **recommendations** to stakeholders, driving data-informed decision-making that led to a **15%** rise in user satisfaction.
- Collaborated with cross-functional teams to build **predictive** machine learning **models**, optimize content **fingerprinting** and recognition by **10%**, and reduce cloud environment costs by **\$200,000** per month.
- Researched and **pioneered** the adoption of new tools and technologies in the Cloud department, increasing the efficiency of **data analytics** workflows by **40%** and saving **60+** manhours per week across teams.

Junior Data Scientist | Java, SQL, Hybrid recommendation systems

Jul 2021 - Apr 2022

Wipro Limited - Bangalore, India

- Developed data-driven features for a major client's **e-commerce** platform intended for **B2B** sales, integrating **APIs** and **databases** to enhance content personalization, boosting user engagement by **15%**.
- Programmed and optimized machine learning models for **data processing**, reducing load times by **30%** and improving real-time accuracy of **recommendation systems** by **10%** on the web application.
- Applied **Agile** practices and utilized **Git** for version control to manage projects, ensuring timely deployment of data models and resolving **1000+** code quality issues to enhance system reliability.

Research Data Science Intern (Capstone) | Computer Vision, Deep Learning

Jan 2021 - May 2021

National Institute of Technology Silchar, India

- Led a team to design a 3-step **Automatic License Plate Recognition** model with raw input images focused on use cases involving license plates with non-standard designs and dimensions.
- Compared performance of pre-built and custom-tuned **neural network** architectures to achieve a final accuracy of **96%** with a modified version of **MobileNet**, including distorted and blurred images and license plates with special characters or symbols.

EDUCATION

Master of Science in Data Science

Aug 2024 - May 2026

University of Maryland, College Park, Maryland, USA

- Relevant coursework:** Probability and Statistics, Data Science, Machine Learning, Data Modeling, Data Presentation, Big Data Systems, Communication in Data Science, Deep Learning, Natural Language Processing (NLP), Computer Vision

Bachelor of Technology in Electronics and Communication Engineering

Jul 2017 - May 2021

National Institute of Technology, Silchar, Assam, India

- Relevant coursework:** Calculus, Object Oriented Programming, Linear Algebra, Probability and Statistics, Random Variables and Random Processes, Data Structures and Algorithms, Neural Networks and Fuzzy Logic

SKILLS

- Data Science & Machine Learning:** Data Analytics, Data Visualization, Data Engineering, Predictive Analytics, Business Intelligence, Business Analytics, Deep Learning, Computer Vision, Natural Language Processing (NLP), Generative AI, Data Structures, Statistics, Mathematics, Regression, Decision Trees, Database Management, Hyperparameter Tuning, Cloud Computing, Containerization
- Languages & Tools:** Python, SQL, Java, C++, Convolutional Neural Networks (CNN), Large Language Models (LLM), Tensorflow, Keras, PyTorch, OpenCV, SciKit-learn, Pandas, NumPy, Matplotlib, Apache Spark, ETL, MongoDB, MySQL, Apache Airflow, AWS, Docker, CI/CD, Data Pipeline, Linux, Bash, Snowflake, Microsoft Excel, Tableau, Jupyter Notebook, Microsoft PowerBI, MATLAB
- Other Skills & Tools:** Data Transformation, Data Strategy, Advanced Analytics, Predictive Models, Business Models, Compliance, Jira, GraphQL, DevOps, Kubernetes, NoSQL, Microservices, API, HTML, CSS, JavaScript, React, Web Design, Backend, Confluence, Project Management, Collaboration, Communication Skills, Analytical Thinking, Problem-Solving, Decision Making, Attention to Detail

PROJECTS

Real-time Bitcoin Analytics Pipeline with LangChain & Neo4j

May 2025 - Jun 2025

LangChain, Graph Data Modeling, Graph Analytics, Streamlit, NLP, Cypher, ETL

- Designed a real-time analytics pipeline for Bitcoin data, visualizing trends and **2000+** simulated graph relationships with **Streamlit**.
- Integrated **LangChain** with Mistral (via Ollama) to convert user questions into explained **Cypher** queries using few-shot prompting.

ClimaSense: AI-Driven HVAC Optimization using Reinforcement Learning

Apr 2025 - May 2025

Energy Systems Modeling, Time Series Forecasting, FastAPI, Sensor Fusion, Model-Based Control

- Trained a **SAC agent** in a custom Gymnasium simulation to reduce HVAC energy use by **81%** across **20,000+** episodes.
- Deployed real-time control with **8-hour** forecasts, reducing daily HVAC runtime by **3.6 hrs** with **90%** alignment to predicted conditions.

Ocular Disease Analysis and Recognition using Multimodal Data

Sep 2024 - Dec 2024

Statistical Analysis, Hypothesis Testing, Predictive Modeling, Hyperparameter Tuning, Transfer Learning

- Developed a **Convolutional Neural Network** model on top of the **VGG16** architecture, achieving a final accuracy of **90.39%**.
- Conducted **data analysis** on eye fundus images and tested **5+** hypotheses for ocular disease prediction.