ABHISHEK KUMAR

Hoboken, New Jersey, 07307, +1(201)-234-7692, akumar47@stevens.edu, LinkedIn

EXPERIENCE SUMMARY

- Expertise in **data science, machine learning, and big data analytics**, with hands-on experience in **Python, R, NLP, Tableau, PySpark, and TensorFlow** for scalable data-driven solutions.
- Adept at **analyzing large datasets, designing predictive models, and implementing scalable solutions** using **Google Cloud Dataproc, PySpark, TensorFlow, and MLLib**.
- Proven ability to manage end-to-end projects, from **data extraction and processing to model deployment and evaluation**.
- Certified **Databricks Associate Developer for Apache Spark 3.0 (2025–2027)**, demonstrating proficiency in distributed computing and big data frameworks.
- **Graduate Student Ambassador at Stevens Institute of Technology**, assisting over 1200+ prospective students with academic and career guidance.
- **Teaching Assistant for Deep Learning and Mathematical Foundations of Machine Learning**, supporting students in understanding advanced AI and ML concepts.
- Completed MS in **Computer Science** from **Stevens Institute of Technology**. Recognized as a **Provost Scholar.**

SKILLS

Technologies and Tools: Python, C++, Java, Javascript, SQL, R, Microsoft Excel, MongoDB, Jupyter Notebook, Google Colab, MySQL, MATLAB, AWS, GCP

Machine Learning and NLP: LangChain, HuggingFace, Transformers, Supervised and Unsupervised Machine Learning models and their optimization, TensorFlow, PyTorch, Keras, Databricks, Spark, Pyspark

PROFESSIONAL EXPERIENCE

Teaching Assistant, Stevens Institute of Technology **Jan 2025 - Till Date**
Hoboken
Deep Learning and Mathematical Foundations of ML

- Assisted students in understanding **machine learning algorithms, deep learning architectures, linear algebra and optimization techniques**.
- Provided hands-on guidance in **coding deep learning models** like **transformers, LSTM, RNN, CNN, Q-Nets using Python, TensorFlow, PyTorch, and NumPy**.
- Reviewed and graded **assignments, projects, and exams**, ensuring **accuracy and conceptual clarity** in student submissions.
- Debugged student code and optimized **ML models**, helping troubleshoot issues related to overfitting, convergence, and performance tuning.
- Developed **supplementary materials, tutorials, and Jupyter notebooks** to reinforce complex ML concepts.

Student Ambassador, Stevens Institute of Technology **July 2024 - Feb 2025**
Hoboken
Schaefer School of Engineering

- Acted as the primary representative for Stevens Institute's Schaefer School of Engineering, effectively addressing queries from over **1200 prospective students**, providing guidance on program specifics, admissions processes, and campus resources.

Subject Matter Expert, Chegg India Pvt Ltd **Aug 2020 - Aug 2023**
India
Computer Science

- Provided **detailed solutions** to complex **computer science, machine learning, and data structures** problems for global students.
- Created **well-documented, high-quality solutions** adhering to academic integrity and best practices in programming.
- Resolved **hundreds of queries** across topics like **data structures, computer networks, web-development database management, C++, Python, Java, NodeJS,** and **Machine Learning**

PROJECTS

New York Citibike demand prediction using PySpark **Dec 2024**

- Applied regression models using Spark on Google Dataproc to forecast Citibike demand using nine years of historical data stored in Google Cloud Storage. Implemented Linear Regression, Random Forest, and Decision Tree algorithms with MLLib to uncover ridership patterns. Incorporated weather data, lockdown periods, and holidays, improving the model's R^2 from 0.38 to 0.74. Designed an end-to-end data pipeline that automated data ingestion from Google Cloud Storage, performed preprocessing and feature engineering in Spark, and seamlessly integrated MLLib for model training and evaluation, ensuring scalability and efficiency in large-scale data processing.

Movie genre trend prediction using reddit discussions **April 2024**

- Leveraged the Pushshift API and Selenium to scrape, collect and analyze Reddit posts and comments totaling about 35k posts. Developed a comprehensive engagement score by systematically integrating over 35,000 Reddit posts' likes and comments with sentiment analysis via VADER. Generated a genre score by calculating the frequency of each genre in discussions normalized by post length. Achieved an R2 of 0.69, indicating solid predictive performance by Decision Tree Regression

Rocker-bogie robot with semi-autonomous navigation **May 2022**

- Designed a 6 wheeled rover–robot based on NASA's rocker-bogie mechanism for stair and obstacle navigation. Implemented YOLOv5 for obstacle detection and swarm optimization, reducing costs significantly.

EDUCATION

Stevens Institute of Technology, USA **Aug 2023 - May 2025**

Master of Science in Computer Science **GPA 3.78**

Coursework: Deep Learning, Big Data Technologies, Machine Learning, Web Mining, Database Management Systems, Mathematical Foundations of Machine Learning

Indian Institute of Information Technology, India **Aug 2018 - May 2022**

Bachelor of Technology, Electronics and Communication Engineering **GPA 3.10**

Coursework: OOPs with C++, Computer and Communication Networks, Python for Big Data, Embedded and Intelligent Systems, IT Project Management, Geospatial Technologies and Applications