

Sanket Navnath Janger

Binghamton, NY | (609) 302-1913 | sanketjanger10@gmail.com | [linkedin.com/in/sanket-janger/](https://www.linkedin.com/in/sanket-janger/)

EDUCATION

Binghamton University, State University of New York, Thomas J. Watson College of Engineering and Applied Science

Masters of Science in Computer Science, Artificial Intelligence track

Expected May 2026

Relevant Coursework: Design Analysis and Algorithm, Programming Language, Design Pattern

Savitribai Phule Pune University, K. K. Wagh Institute of Engineering Education and Research

Bachelors in Computer Engineering

June 2024

Honors: Data Science

Relevant Coursework: Data Science and Visualization, Machine Learning and Statistics, Big Data Analytics, Discrete Mathematics, Database Management System, Cloud Computing

TECHNICAL SKILLS

Programming Languages and Query Languages: Python, SQL

Database Software : MySQL, Apache Spark, Relational Database System

Frameworks: NumPy, Pandas, Scikit-learn, Matplotlib, Seaborn

BI Tools : Tableau, Power BI, MS Excel, Qlik

AWS Services: EC2, S3, Elastic Beanstalk, VPC, Route53, IAM, DynamoDB

Cloud Frameworks: Amazon Web Services(AWS), GCP(Google Cloud Platform)

Additional: Git, Github co-pilot, Statistical Analysis, MS Office

INTERNSHIP EXPERIENCE

Emerging Technologies, Full Stack Intern | India

February 2023 – March 2023

- Developed an automated ID card generation system for students and staff using C#, JavaScript and ASP.NET framework
- Created a web interface for instant card generation through photo and data upload, implementing full-stack functionality with ASP.NET to handle front-end interactions and server-side processing
- Transformed manual card creation into a streamlined digital process, reducing generation time from days to minutes while gaining hands-on experience in building end-to-end web applications

ACADEMIC PROJECTS

Distributed Machine Learning Pipeline with Hadoop and Spark | Independent Project

August 2024 – September 2024

- Built an end-to-end data processing pipeline leveraging Hadoop and Spark for real-time machine learning model deployment
- Implemented real-time data processing using Spark Streaming and PySpark MLlib, integrating Random Forest and XGBoost algorithms for model development and prediction
- Delivered a scalable distributed computing solution that handles big data processing in real-time, demonstrating practical expertise in building production-ready ML pipelines

Retail Analytics Dashboard: Sales & Inventory | Independent Project

June 2024 – July 2024

- Created an interactive retail analytics dashboard using Power BI to optimize sales tracking and inventory management for store operations
- Developed data pipeline using SQL for database management and Python for data cleaning, implementing features to analyze sales patterns and predict future trends
- Delivered a real-time visualization solution that empowered stores with actionable insights for inventory control and sales forecasting, enhancing decision-making capabilities

Exploratory Analysis of Bank Stock Price Data | Independent Project

September 2022 – October 2022

- Performed comprehensive analysis of historical bank stock prices to uncover market trends and price movement patterns
- Applied statistical techniques and data visualization tools for stock performance analysis, focusing on market drivers and financial indicators
- Generated actionable investment insights by identifying key market conditions and risk factors affecting bank stock performance

Customer Segmentation | Independent Project

February 2023 – March 2023

- Developed a data-driven customer segmentation system to identify distinct purchasing behavior patterns and target groups
- Implemented K-Means clustering using Python, leveraging Pandas for data manipulation and Scikit-Learn for customer behavior analysis
- Automated the segmentation process, transforming manual analysis into an efficient system while gaining practical expertise in data science techniques

Enhancing ML Model through Dynamic Features Selection | Group Project

September 2023 – April 2024

- Created an automated feature selection system for healthcare data prediction, improving overall model efficiency by ~ 35%
- Implemented OSFS and DOSFS algorithms for optimal feature selection, integrating PCA for dimension reduction that cut feature space by ~ 60% while maintaining data integrity
- Boosted model performance metrics significantly, accuracy by ~ 18%, F1-score by ~ 22%, and AUC-ROC by ~ 15% across SVM, Logistic Regression, and Decision Tree models

LEADERSHIP EXPERIENCE

Event Lead, Rotaract Club of Nasik Youth Force,

October 2023 & March 2024

- Organized and ran NYFPL Season 3&4 cricket tournament, handling everything from team management to event operations
- Team up with different club committees to ensure smooth event planning and execution
- Handled event promotion on Instagram and brought in sponsors to support the tournament

Core Committee Member, MIBCS (Machine Learning, IoT, Blockchain, Cyber Security)

August 2022 - August 2023

- Led the Scavenger hunt of Synapse 3.0., the largest technical event in our college
- Developed complex riddles in morse codes with a team, enhancing participant engagement and challenge level