

ANUSHA UDAYAKUMAR

(716)-939-5651 | anusha.ukumar14@gmail.com | [LinkedIn](#) | [Portfolio](#) | New York

SUMMARY

Master's student in Data Science with hands-on experience in Python, SQL, and machine learning, specializing in data-driven applications and AI solutions. Proficient in statistical analysis, A/B testing, and predictive modeling, with a focus on enhancing user experience and solving complex problems. Strong collaboration skills in cross-functional teams, leveraging data science to drive innovation, product development, and growth. Skilled in defining project scope, communicating strategies, and implementing best practices for scalable, high-performance data systems.

EDUCATION

State University of New York at Buffalo, NY

(Master of Science in Data Science)

Aug2023-Dec2024

CGPA: 3.8/4(till now)

Vidyaavardhaka College of Engineering (affiliated to VTU, Belgaum)

2015-2019

(Bachelor of Engineering in Computer Science)

CGPA: 8.4/10

TECHNICAL SKILLS

- **Programming Languages:** Python (Pandas, NumPy, Sklearn), SQL, R, MATLAB, Scala, SAS, C++, Java, REST APIs, Go.
- **Machine Learning:** Scikit-Learn, TensorFlow, PyTorch, XGBoost, K-Means, Random Forest, Decision Trees, SVM, Keras.
- **Statistical Modeling:** Regression, Hypothesis Testing, ANOVA, , Time Series Forecasting, Statistical Methodologies.
- **Data Analysis & Visualization:** Matplotlib, Seaborn, ggplot2, Power BI, Tableau, Plotly.
- **Big Data Tools & Application:** Apache Spark, Hadoop, HDFS, Kafka, Hive, Kubernetes, Docker.
- **Data Management:** SQL (Advanced Queries), PostgreSQL, MySQL, Snowflake, Databricks, CI/CD, Git, GitHub.
- **Cloud Platforms:** AWS (SageMaker, Lambda), Google Cloud (AI Platform, BigQuery), Azure (ML Studio, Synapse).
- **Collaboration & Communication:** Agile, Scrum, JIRA, Cross-functional Teamwork, Data Storytelling.
- **Security & Compliance:** Data Encryption, GDPR Compliance, RBAC, Secure API Integration
- **Process Optimization:** Automated Testing, A/B Testing, Performance Tuning, Infrastructure Monitoring, Process Improvement

PROFESSIONAL EXPERIENCE

Data Analyst Research Intern, Verusen, Atlanta, USA

May2024-Aug2024

- Implemented **OpenAI's** GPT-4 with **prompt** engineering for supply chain recommendations, improving response time to 2s.
- Engineered **Python** modules in **Databricks** to optimize data analysis for EDA for **finance** and **accounting** workflows.
- Developed **FastAPI** - based microservices to enhance real-time data flow, utilizing **Postman** for **API testing**.
- Optimized **PostgreSQL** queries using **indexing**, **partitioning**, and **query tuning**, reducing execution time by 40%.

Graduate Student Assistant, University at Buffalo, USA

Jan2024-May2024

- Fine-tuned biomedical data using **XGBoost**, improving classification accuracy by 25% on a 100,000-sample dataset.
- Designed **CNN-LSTM** models for patient outcome prediction, forecast accuracy improved by 20% on time-series data.
- Automated **ETL** pipelines with **Pandas** and **Spark**, reducing data preprocessing time from 4 hours to 2 hours

Data Engineer, Tata Consultancy Services, India

Jul2019-Jul2023

- Optimized **ETL** pipelines in **Snowflake**, using **Python**, **Teradata**, **SQL**, **Informatica**, reducing data extraction costs by 60%.
- Automated **CI/CD** pipelines with **Azure Pipelines** and **GitHub**, accelerating deployment cycles and integration processes
- Boosted data processing by 37%, refining **Informatica** workflows and optimizing **Spark** jobs for distributed data processing.
- Applied **Agile** methodologies in **Scrum** to lead digital transformation and adaptive planning in software development cycles.
- Enhanced product quality by 49% through data analytical, defect analysis, and visualization using **Power BI** and **Databricks**.
- Increased customer satisfaction via collaborative **Tableau** dashboards, leveraging **Azure RBAC** for secure data access control.
- Conducted **ad-hoc** statistical methodologies, leveraging **A/B testing**, **KPIs** to enhance infrastructure performance monitoring.
- Automated metric workflows using optimized **SQL stored procedures**, cutting operational costs by \$93.9k annually.

ACADEMIC PROJECTS

AI-Powered Legal Document Summarizer | Python, PyTorch, Hugging Face Transformers, GPT-4, Flask

April2024

- Built an AI legal document summarizer using GPT-4 and Hugging Face Transformers, leveraging fine-tuning and NER for real-time, accurate summaries. with a focus on scalable NLP solutions.

F1 Data-Driven Foresights | Python, Scikit-Learn, Real-time Analysis

Dec2023

- Spearheaded a detection system for real-time F1 race data monitoring using Isolation Forest and One-Class SVM.

Predictive Modeling for IMDb Ratings | Python, SQLite, Scikit-Learn

Sept2023

- Engineered a relational SQLite database for large-scale cinematic data and deployed a Random Forest model in Python.

NLP-Based Virtual Assistant | Python, NLP, Google Dialogflow

Oct2021

- Formulated a cross-platform virtual assistant with Python and Dialogflow and published a paper on speech recognition in the International Journal of Innovative Science and Research Technology.

ACHIEVEMENTS AND CERTIFICATIONS

- Received 'Pinnacle Award' and 'Team Player Award' at TCS for performance and ability to collaborate.