

# Ms. Harshali Sunil Narkhede

harshan2@umbc.edu  
443 653 0703

New York City, New York

GitHub ID- HarshaliN11  
LinkedIn -www.linkedin.com/in/harshali-n

## EDUCATION:

**University of Maryland-Baltimore County** - Master of Professional Studies, Data Science.

2022- 2024

**Savitribai Phule Pune University** - Bachelor of Engineering, Instrumentation & Control

2014-2018

## SKILLS:

**Skills & Technologies:** Python, R Programming, Scala, SQL, RDBMS, NoSQL, LINUX, Java, Network Protocols, Scala, Data Analysis, Data Warehousing, Data Mining, Snowflake, Databricks, Apache Hadoop, Apache Spark, PySpark, Hive, Apache Pig, Apache Kafka, Apache Airflow, Azure Synapse, Azure Fabric, Azure Data Factory, AWS Glue, AWS Redshift, AWS S3, GCP, Regression, Classification, Clustering, Time Series Analysis, MLOps, Feature Engineering, Jenkins

**Tools:** Tableau, Microsoft Power BI, Talend, Zookeeper, Sqoop, Docker, SSIS, SSAS, SSRS, GitLab CI/CD, GitHub

**Frameworks:** Apache Airflow, ETL Pipelines, Scikit-Learn, TensorFlow, PyTorch.

## WORK EXPERIENCE:

**University of Maryland-Baltimore County** - Data Engineer

September 2022 – Current

- **Data Modeling & Deployment:** Architected and implemented new data models and processes, managing the Blumen-Compansol Database with 99.8% accuracy, reducing data retrieval time by 20%.
- **Data Analysis & Insights:** Conducted data analysis to generate business insights, leading to a 30% increase in college application submissions by leveraging Python.
- **Data Quality & Expertise:** Built expertise in data management strategies, ensuring high data quality and effective tracking of student progress.
- **Data Reporting:** Designed comprehensive reports on student performance and program effectiveness using SQL and Power BI, enhancing report generation efficiency by 40%.

**Gauri Ltd, Pune, India** - Data Engineer

April 2021 - July 2022

- **ETL & Data Optimization:** Enhanced SSIS packages and optimized Azure Synapse SQL queries and SSIS pipelines, reducing data processing time by 25%.
- **Data Analysis & Visualization:** Created Power BI reports with DAX calculations, improving data insights accuracy by 30% and connected them to OLAP cubes for interactive data analysis.
- **AWS Integration:** Designed and implemented data pipelines using AWS Glue and AWS Redshift, improving data processing efficiency by 20%.
- **Troubleshooting & Debugging:** Systematically debugged SSIS packages and PySpark scripts, ensuring smooth data pipeline operations.
- **Collaboration:** Worked closely with cross-functional teams to understand data requirements and ensure data accuracy.

**GeoSensing and Imaging Consultancy, Kerala, India** - Data engineer

October 2020 - March 2021

- **Data Preprocessing & Accuracy:** Improved dataset accuracy by 50% through Python preprocessing techniques and leveraging Apache Spark.
- **Model Development:** Utilized TensorFlow for deep learning model development and evaluation, enhancing model accuracy.
- **Data Pipeline Management:** Streamlined data ingestion and processing pipelines using Python and Apache Kafka reducing processing time by 25%.
- **Quality Assurance:** Reduced false positives by 35% through thorough script revisions and comprehensive testing.

**Parishram Enterprises, Pune, India** – Data Engineer

June 2018 - September 2020

- **IoT Data Integration:** Directed the integration of IoT devices, significantly reducing data discrepancies by 40%.
- **ETL Workflow Management:** Managed dataflow and preprocessing in ETL workflows, leading to faster data retrieval time by 30%.
- **Data Quality Improvement:** Implemented Python-based data quality controls, improving accuracy by 20%.
- **Visualization:** Built Power BI reports with DAX calculations for efficient analysis and visualization of IoT sensor data.

## PROJECTS:

### Factors Impacting Crop Yield Prediction

- Predicted agricultural production accuracy to 20% by enhancing models with data segmentation and ensemble methods like Gradient Boosting and Random Forest.

### YouTube Trending Videos for Marketing Strategies.

- Applied targeted advertisement strategies to drive sales growth and leveraged data visualization for an insightful analysis of trending videos by category.

### Business Analysis and Future Predictions

- Used Time Series Analysis (FB Prophet algorithm) for future prediction.
- Developed a user interface with Python Flask and performed comprehensive analysis using a Tableau dashboard.

### Remote Monitoring of Track Grinding Machine

- Established real-time monitoring to prevent errors and breakdowns by 30%, transmitting data to a designated client. Also, early fault detection minimized machine downtime, enhanced product quality, and reduced scrap rates.

## CERTIFICATIONS:

**Microsoft Certified Azure Data Engineer (DP-203)** – Certification number: I133-3218

**Centre for Development of Advanced Computing, Pune, India. (PG-Diploma in Big Data Analytics)**