

# Harshal Patil

716-750-9045 | [hpatil2@buffalo.edu](mailto:hpatil2@buffalo.edu) | [linkedin.com/in/harshalsp0011](https://www.linkedin.com/in/harshalsp0011) | New York, USA

## EDUCATION

- SUNY Buffalo** New York, USA  
*MS Engineering Science (Data Science)* Aug. 2024 – Present
- Modern Education Society's College of Engineering** Maharashtra, India  
*Bachelor of Engineering in Electronics and Telecommunication* Aug. 2018 – April 2022

## SKILLS SUMMARY

- Programming Languages:** Python, SQL, C#, R, C++, JavaScript, HTML/CSS, Data structures
- Frameworks:** TensorFlow, Blazor, .NET, React, Camunda Workflow, Entity Framework
- Libraries:** Pandas, NumPy, Matplotlib, Scikit-learn, Keras
- Big Data & Distributed Computing:** Apache PySpark, Hadoop, HDFS, MapReduce, Spark SQL
- Cloud Platforms:** Microsoft Azure, AWS
- Tools & Platforms:** Azure, Git, Tableau, Docker, Power BI, VS Code, snowflake, Visual Studio, Matlab, Swagger, Postman, Office, Excel, R-Studio, Jupyter Notebook, PostgreSQL
- Soft Skills:** Leadership, Time Management, Team Collaboration, Communication, Attention to Detail

## EXPERIENCE

- Assistant Manager (Senior Software Developer Engineer)** Sept 2022 – July 2024  
*Reliance Jio Platforms LTD* New Mumbai, India
  - Database Optimization:** Improved MySQL database performance through indexing strategies and query profiling, enabling faster data retrieval and a 20% improvement in query efficiency.
  - Automation & Data Pipeline Optimization:** Built Python scripts and automated ETL pipelines, maintaining a production database with 100,000+ records and reducing server downtime by 10 hours/month.
  - API Development for Data Access:** Developed RESTful APIs to enable seamless access to internal data systems, contributing to improved interoperability across business units.
  - Semantic Modeling:** Collaborated with cross-functional teams to structure domain-level data models, improving discoverability and enabling integration with metadata tools.
  - Workflow Integration:** Used Camunda BPMN to design and implement scalable process workflows, automating critical operations and increasing operational efficiency.
  - Project Management and Quality Assurance:** Led agile sprints, maintained release schedules, and coordinated testing efforts to ensure deliverables met business and performance requirements.

## PROJECTS

- Job Market Analysis (PySpark, HDFS, MLlib, Big Data Pipeline):** Built a scalable data pipeline to analyze LinkedIn's 1.3M job records using PySpark and Hadoop. Performed distributed EDA, automated data cleaning (Parquet), and implemented ML models for job classification, skill demand prediction, and clustering. Persisted models and outputs to HDFS for reproducibility and future access.
- Job Portal Database System (SQL, ER Modeling, Query Optimization):** Designed and implemented a relational database to manage job listings, applications, and candidate profiles. Created normalized schemas, optimized search queries for skill-based filtering, and enforced data integrity across entities. Demonstrated scalability and accuracy for real-world hiring workflows.
- Database Operations Web Application (SQL, .NET, REST API):** Developed a modular full-stack application enabling dynamic CRUD operations on a relational database via RESTful APIs. Employed SQL for optimized query execution and utilized Entity Framework with C# to enforce schema integrity, ensure transactional consistency, and streamline backend logic.
- Hybrid Forecasting Pipeline (ARIMA + LSTM):** Designed and deployed a hybrid time series forecasting architecture integrating ARIMA and LSTM models. Implemented scalable data preprocessing workflows, dynamic model selection logic, and performance-weighted ensembling to enhance predictive accuracy across multiple temporal patterns.

## CERTIFICATIONS

- Python for Data Science and Applied Analytics
- Foundations of Cloud Computing and Scalable Data Systems
- Full-Stack Web Development with HTML, CSS, and JavaScript
- Entity Framework and Database-Driven Application Design
- Building Front-End Interfaces with React for Data Products