ONKAR EKNATH SHELAR

os9660@rit.edu | +1 (971) 447-7003 | https://www.linkedin.com/in/onkarshelar/ | https://github.com/Onkar2102

OBJECTIVE

Driven Machine Learning enthusiast, poised to apply Advanced Analytics and Machine Learning expertise to develop cutting-edge AI solutions, while embracing challenges and driving impactful outcomes. (Seeking Summer 2024 Co-op/Internship)

EDUCATION

Rochester Institute of Technology | Rochester, NY **(Expected) May 2025**
Master of Science in Data Science **GPA: 4.00/4.00**
Relevant Courses: Software Construction, Database Design and Implementation, Foundations of Data Science and Analytics

Centre for Development of Advance Computing (CDAC) | Pune, India **April 2022**
Post-Graduate Diploma in Big Data Analytics **GPA: 3.91/4.00**
Relevant Courses: Linux Programming and Cloud Computing, Practical Machine Learning, Advanced Analytics using Statistics, Python and R programming, Data Collection and DBMS, Big Data Technologies, Data Visualization, Object Oriented Programming with Java

Savitribai Phule Pune University | Pune, India **April 2021**
Bachelor of Engineering in Mechanical Engineering **GPA: 3.21/4.00**

SKILLS

- **Programming Languages:** Python, SQL, Apache Spark, Java
- **Data Tool:** Git, Azure Databricks, Azure Data Factory, Azure Data Lake Storage, Azure Synapse, Azure Logic App
- **Data Management and Analytics:** Data Wrangling and Orchestration, Data Modeling, DBMS, Statistical Analysis, Power BI
- **Machine Learning Techniques:** Data Mining, Supervised Learning, Unsupervised Learning, ML model optimization, Algorithm Development

PROFESSIONAL EXPERIENCE

***Associate Data Engineer*, Fractal Analytics** **May 2022 – May 2023**

- Led end-to-end development of data pipelines, integrating diverse data sources such as SharePoint, Blob Storage, and SAP into ADLS Gen2/Azure Synapse Analytics.
- Designed and managed dynamic Linked Services, Datasets, and Pipelines in Azure Data Factory, utilizing triggers for streamlined data orchestration.
- Implemented Medallion architecture within Databricks, optimizing data orchestration processes and boosting overall efficiency.
- Improved dashboard performance in Databricks by customizing KPIs for Power BI, reducing 2 hours of daily computation time.
- Enhanced pipeline stability and data accuracy in Databricks by proactively monitoring and debugging existing Notebooks and creating dynamic notebooks for quality checks and time frame calculations.

***Production Coordinator*, Sandeep Plastic Industries** **July 2018 - March 2020**

- Orchestrated production scheduling, trained 11 new team members, and managed equipment maintenance, while spearheading initiatives to elevate production efficiency and reduce downtime.
- Fostered continuous improvement by optimizing processes and collaborating with vendors for equipment maintenance and repairs, enhancing operational workflow.

***Mechanical Engineering Intern*, Sun-Axis Industries** **July 2017 - June 2018**

- Supervised production schedules and coordinated work orders, while diagnosing machinery malfunctions and ensuring product quality standards.

TECHNICAL CERTIFICATES

- **Microsoft Certified:** Azure Data Engineer Associate (Dp-203) | Power BI Data Analyst Associate (PL-300) | Azure Data Fundamentals (DP-900) | Azure Fundamentals (AZ-900)
- **Databricks:** Databricks Accredited Lakehouse Fundamentals

RELEVANT PROJECTS

Title: Machine Learning Algorithms from Scratch (Rochester Institute of Technology) **Aug 2023 - Dec 2023**

Tools: Decision Tree, Ada Boost, Logistic Regression, K-means Clustering, Naïve Bayes, Nearest Neighbor, Genetic Algorithm
Contribution: Engineered multiple ML algorithms without relying on advanced libraries like Scikit-Learn. This project showcased a deep understanding of underlying ML principles and algorithmic efficiency.

Title: Job Posting Authenticity Analysis (Rochester Institute of Technology) **Dec 2023**

Tools: Statistical Analysis, Frequency Encoder, TFID Vectorization, Random Forest, Model Optimization, Text Analysis
Contribution: Developed a robust model to distinguish between real and fake job listings, enhancing recruitment process integrity. This involved intricate data-driven techniques for accurate classification and analysis.

Title: Business Insights on Post-COVID Taxi Systems (CDAC) **Jan 2022 – April 2022**

Tools: Flask, Linear Regression, XGBoost, AWS S3, Jupyter Notebook, Power BI
Contribution: Conducted an in-depth post-COVID analysis for New York City taxi system, focusing on profit optimization for drivers. The project involved developing a real-time analytical tool providing insights on peak times and locations for maximizing earnings.