MAHARSHI SHUKLA

Data Engineer and Analytics — Data Analysis, Pipeline Development, & Machine Learning

maharshishukla19@gmail.com (236)-862-6576 Vancouver LinkedIn GitHub Portfolio

SKILLS

- **Programming Languages:** SQL, Python, Java, Bash/Shell Scripting
- **Databases and Query Languages:** MySQL, PostgreSQL, MS SQL Server, MongoDB, SQLite, T-SQL, PL/SQL
- **Data Engineering and Integration:** Apache Spark, Kafka, Airflow, Azure Data Factory, ELT Frameworks
- **Core Skills:** Machine Learning and AI, Cloud Engineering, CI/CD and DevOps, Data Visualization and Business Intelligence, Version Control and Collaboration

WORK EXPERIENCE

Teaching Assistant – Neural Networks and Deep Learning **September 2024 – December 2024**
Northeastern University, British Columbia

- Facilitated deep learning tutorials for students, teaching core concepts like backpropagation, and neural network optimization using Python and TensorFlow. This practical exposure resulted in a 20% increase in student comprehension.
- Delivered interactive workshops on advanced topics like transfer learning and model fine-tuning, demonstrating how to apply pre-trained models to new datasets. These sessions led to a 15% increase in students' ability to deploy models.
- Supervised weekly lab sessions, providing students with hands-on experience in developing neural networks from scratch. Offered individualized guidance, resulting in a 25% reduction in debugging time and an increase in model accuracy.
- Evaluated lab assignments and projects, offering feedback on neural network architecture choices, optimization techniques, and debugging practices. Helped students improve the respective models' performance by an average of 18%.

Data Engineer **May 2021 – September 2022**
Shree Drashti Infotech, India

- Engineered and optimized Power BI reports for clients by creating advanced data models and performing ETL transformations. This effort resulted in a 10% increase in the customer base through enhanced decision-making capabilities.
- Refined and deployed an interactive, real-time marketing dashboard using SQL-based ETL pipelines, integrated with cloud storage. This dashboard enabled the marketing team to quickly access up-to-date insights, boosting marketing.
- Optimized data ingestion pipelines for real-time and historical data processing in a cloud environment, implementing robust error handling and query optimization. This decreased data latency by 20%, enabling faster and reliable insights.
- Implemented data quality assurance measures by creating automated testing scripts for ETL processes. These measures ensured 99% accuracy in data reporting and eliminated discrepancies in client reports, leading to improved trust.
- Worked closely with the data science team to identify new data sources and integrate the data into existing pipelines. This integration led to a 30% increase in the amount of actionable data available for analysis and decision-making.

PROJECTS

Machine Learning-Driven Cryptocurrency Trading System **September 2024 – December 2024**

- Implemented machine learning algorithms, including XGBoost and Random Forest, to predict trends and optimize strategies. Achieved impressive prediction accuracies of 82% and 80% respectively, demonstrating a robust model.
- Designed an interactive, and agile dashboard for real-time monitoring of trading performance. Integrated sentiment analysis using the VADER algorithm to analyze market sentiment, which enhanced predictive accuracy.

Health Risk Analysis using Smoking and Drinking Dataset **September 2023 – December 2023**

- Facilitated exploratory data analysis (EDA) on a large dataset of over 991,000 records to analyze correlations between smoking, drinking, and diverse set of health indicators. Utilized Azure cloud services for efficient data processing.
- Built interactive, visually rich Power BI dashboards to represent the data analysis results. These dashboards provided actionable insights by visualizing correlations between health factors like age, BMI, and blood pressure.

Bone Fracture Detection Using ML and DL Algorithms **August 2021**

- Experimented with numerous machine learning models and architectures to identify the most effective approach. After thorough evaluation, the DNN architecture being selected as the best-performing model for this task.
- Published a review paper on the project in the AUT-AUT International Research Journal, this article being presented as a step towards the use of artificial intelligence for the improvement of healthcare care.

EDUCATION

MS in Data Analytics Engineering **Sep 2023 - Dec 2024**
Northeastern University, Vancouver, BC

BE in Computer Engineering **August 2017 – May 2021**
Gujarat Technological University, India