Shivam Mishra

+18123610960 | mishras0297@gmail.com | Portfolio | GitHub | LinkedIn

EDUCATION

Indiana University, Bloomington, United States   **Aug 2022 – May 2024**

Master of Science In Data Science | GPA: 3.7 / 4

Applied Algorithms, Advance Database Concepts, Data Mining, Deep Learning Systems, Applied Machine Learning, Data Analytics

PROFESSIONAL EXPERIENCE

Kelley School of Business, Bloomington, United States of America   **Jan 2024 – Present**

Research Assistant | Skills: Python, PySpark, SQL, Tableau

- Conducted a comprehensive overhaul and analysis of datasets exceeding 1 terabyte, uncovering pivotal insights that redefined research approaches and led to a notable 30% increase in operational efficiency.
- Spearheaded the design and deployment of over 10 sophisticated Tableau dashboards, catalyzing a 40% increase in decision-making efficiency and bolstering cross-departmental collaboration with impactful, real-time data visualizations.

Rocket Central, Detroit, Michigan, United States of America   **May 2023 – Aug 2023**

Data Engineer Intern – Data Intelligence | Skills: AWS S3, Glue ETL, Lambda, Python, PySpark, Tableau

- Engineered and optimized ETL pipelines using AWS Glue, PySpark achieving cost savings of $10,000 and adeptly handling datasets exceeding 5 terabytes.
- Processed and analyzed over 5 TB of customer data with PySpark, driving insights through Tableau dashboards.
- Engineered robust CI/CD pipelines utilizing Jenkins, CircleCI, and Git to automate and ensure the reliability of job deployments.

Indiana University, Bloomington, United States of America   **Aug 2022 – May 2023**

Data Analyst Research Assistant, O'Neill School | Skills: Python, PySpark, SQL, Tableau, Snowflake

- Crafted and deployed Python data wrangling scripts to extract over 3 terabytes of insurance data from various online sources, integrating seamlessly with Snowflake.
- Executed meticulous data operations, encompassing the cleaning, aggregating, transforming, and visualizing of over 2 million data points from multiple health insurers, significantly enhancing research precision and operational workflow.

Ericsson, Bangalore, India   **Jun 2021 – Aug 2022**

Software Developer – Cloud Container Distribution | Skills: Kubernetes, Docker, Open Stack, Python

- Effectively oversaw the management of containerized applications within Ericsson products through Kubernetes and OpenStack deployments and improved the software reliability by 30 %.
- Initiated and directed the automation of version updates across comprehensive YAML and JSON files using Python and Golang, which enhanced process productivity by 80%.

Blackrock, Mumbai, India   **Jul 2019 – Jun 2021**

Data Analyst – Data and AI | Skills: Python, PySpark, SQL, Hadoop, Snowflake, Tableau

- Built and oversaw data pipelines for client portfolios, moving over 5 million daily data points to BlackRock's database, improving risk analytics efficiency by 20%.
- Crafted a Python and SQL-based file monitoring tool to streamline data handling from various index vendors, achieving a 60% increase in process efficiency.
- Utilized Spark and SQL for in-depth analysis of production errors in client reports, achieving a 40% reduction in data discrepancies and significantly boosting data quality and reliability.

ACADEMIC PROJECTS

Stocks Recommendation System

- Developed an Recurrent Neural Network powered stocks recommendation system, utilizing sentiment analysis and user input parameters to guide beginners in achieving their investment goals efficiently.

Smart Production and Logistics System

- Analyzed production machine data (power, voltage, temperatures) to detect anomalies using a K-Nearest Neighbors model. Awarded IoT hackathon victory for successful implementation.

TECHNICAL SKILLS

- **Programming languages:** C, C++, Java, Perl, Python, JavaScript, SQL, Bash
- **Libraries/Frameworks:** Pandas, Scikit-learn, ReactJS, PySpark, PyTorch
- **Databases/Cloud:** MongoDB, Postgres, MySQL, Sybase, Cassandra, AWS S3, Glue ETL, Lambda, Route 53, Athena
- **Miscellaneous:** Tableau, Snowflake, Kubernetes, Docker, Hadoop, HTML, CSS