Shriyansh Singh

+1 930 333 5141 | shriyansh.singh24@gmail.com | linkedin.com/in/shriyansh-bir-singh

SUMMARY

Machine Learning Engineer with expertise in building and deploying ML models and data pipelines using Python, TensorFlow, Spark, and AWS. Proven ability to drive business impact through scalable, cloud-based solutions and automation. Passionate about optimizing the entire ML lifecycle to deliver high-performance, reliable outcomes.

PROFESSIONAL EXPERIENCE

Machine Learning Intern
Apr 2024 – Present

Hyphenova AI
Los Angeles, California

- Developed and deployed NLP models using Python and TensorFlow to filter content with 30% higher accuracy, directly increasing user satisfaction by 15% through Agile-based iteration and feedback loops.
- Built and automated ML pipelines with Apache Spark and AWS (S3, Glue), cutting data processing time by 40% and facilitating real-time model deployment and monitoring.
- Improved predictive accuracy of models by 25% using advanced feature engineering and optimization of Random Forest and deep learning algorithms, enhancing brand-creator matching.
- Streamlined data preprocessing and feature selection to enhance training efficiency, achieving 15% better accuracy for underrepresented categories by applying SMOTE and custom resampling methods.
- Converted data science prototypes into production models using AWS SageMaker and Lambda, implementing CI/CD pipelines to automate deployment and governance processes.
- Established model monitoring and governance using real-time validation scripts, boosting data reliability by 30% and reducing errors by 20%, ensuring consistent model performance.
- Managed the full ML lifecycle, including data collection, preprocessing, deployment, and monitoring using AWS and Python, ensuring scalable and reliable ML operations.
- Implemented advanced ML algorithms, such as deep neural networks and LLMs, to tackle complex datasets and enhance decision-making accuracy.
- Automated ML deployment workflows on AWS using CI/CD pipelines and Terraform, cutting deployment time and minimizing manual errors.

Data Analyst Intern
May 2022 – Oct 2022

Enterprise Business Technologies Pvt. Ltd
Mumbai, India

- Analyzed large datasets using SQL and Python to drive business strategies, increasing project completion rates by 25% and client satisfaction by 20%.
- Optimized data models and automated reports in Power BI, boosting data accuracy by 15% and cutting report generation time by 40%, facilitating quicker decision-making.
- Built predictive models using regression analysis in Python, improving forecast accuracy by 18% and driving a 10% increase in quarterly revenue.
- Collaborated with cross-functional teams to embed data analytics into decision-making, advancing data-driven business strategies and project outcomes.
- Created automated data validation pipelines in Python, enhancing data integrity across reporting systems and reducing manual errors by 20%.

EDUCATION

Indiana University Bloomington
Indiana, United States

Master of Science in Data Science
Aug 2023 – May 2025

SKILLS

Programming Languages: Python, Java, C++, Scala
Machine Learning Frameworks: TensorFlow, PyTorch, Scikit-learn, XGBoost, Keras
Data Engineering and Big Data: Apache Spark, Hadoop, AWS (S3, EC2, Glue, Lambda, SageMaker), Databricks, Kafka, Airflow, Terraform
Cloud and Automation: AWS (SageMaker, Lambda, Glue), Azure, GCP, Docker, Kubernetes, CI/CD Pipelines, Infrastructure as Code (Terraform)
NLP and Deep Learning: NLP Techniques, LLMs, Transformers, Hugging Face
Data Analysis and Visualization: Pandas, NumPy, Power BI, Tableau, D3.js
Databases: **SQL** (MySQL, PostgreSQL, Microsoft SQL Server), **NoSQL** (MongoDB, Cassandra, Neo4j)
Development Tools: Git, Bash, Jupyter Notebooks, VS Code, IntelliJ

PROJECTS

NLP-Based Content Filtering System | *Python, TensorFlow, NLP, AWS*
Nov 2023 – Dec 2023

- Enhanced text classification and entity recognition with LLMs and pre-trained embeddings, improving filtering accuracy of inappropriate content by 40%