

SHRIYANSH SINGH

+1 930 333 5141 — shriyansh.singh24@gmail.com — LinkedIn — GitHub

EDUCATION

Indiana University Bloomington	Aug 2023 – May 2025
Master of Science in Data Science	Indiana
Relevant Courses: Information Visualization, Data Mining, Applied Machine Learning, Statistics, Big Data Applications, Cloud Computing, Graph Analytics, Applied Database Technologies, Intelligent Systems	

PROFESSIONAL EXPERIENCE

Machine Learning Intern	April 2024 - Present
Hyphenova AI	Los Angeles, California

- Led the end-to-end development of NLP-based content filtering algorithms using Agile methodologies, achieving a 30% increase in brand-creator matches and 15% improvement in user satisfaction.
- Collaborated with cross-functional teams to design and implement advanced Random Forest models with enhanced feature selection techniques, resulting in a 25% boost in campaign performance.
- Deployed scalable machine learning models using AWS SageMaker, ensuring seamless integration with existing systems and enabling real-time analytics that drove a 40% reduction in data processing time.
- Enhanced model accuracy by 15% for underrepresented categories by applying SMOTE resampling techniques, improving the system's reliability in handling data imbalance.
- Integrated real-time data quality checks and monitoring with custom validation scripts, leading to a 30% increase in data reliability and a 20% reduction in analytics errors.
- Optimized and streamlined data pipelines using Apache Spark and AWS, reducing latency and supporting a 3x increase in data throughput for high-velocity data streams.
- Implemented continuous model monitoring and automated retraining pipelines to adapt models to evolving data patterns, maintaining a high level of accuracy and performance over time.

Machine Learning Intern	May 2022 - Oct 2022
Enterprise Business Technologies Pvt Ltd	Mumbai, India

- Developed a customer segmentation model using unsupervised learning techniques (K-means clustering, PCA), which improved targeted marketing strategies and led to a 20% increase in customer engagement.
- Optimized an existing predictive analytics pipeline by integrating XGBoost for better model accuracy, driving a 15% increase in forecasting precision and aiding business decisions.
- Collaborated with data engineering teams to design and deploy machine learning models on Azure ML, reducing model deployment time by 30% and enabling quicker time-to-market for new features.
- Automated data preprocessing workflows with Python and Pandas, reducing manual effort by 50% and allowing for more efficient data handling and faster model training.
- Analyzed customer behavior with time series forecasting techniques to identify key trends, leading to a 10% increase in revenue through data-driven decision-making.

PROJECTS

Fraud Detection Using Graph Neural Networks

- Developed a graph neural network (GNN) model using GraphSAGE to analyze complex relationships between entities, achieving a 35% increase in fraud detection accuracy compared to traditional methods.
- Engineered graph-based features utilizing NetworkX and Neo4j to model entity interactions, improving the model's ability to identify anomalous patterns indicative of fraudulent behavior.
- Deployed the model on Google Cloud Platform (GCP), leveraging GCP's scalable architecture to handle large, dynamic graphs with millions of nodes and edges, enabling real-time fraud detection.
- Implemented an anomaly detection system that continuously monitors transaction networks, reducing false positives by 20% and providing actionable insights to the fraud investigation team.
- Optimized GNN training using distributed computing on GCP, reducing model training time by 40%.

Document Summarization and Keyword Extraction

- Implemented BERT-based extractive summarization using Hugging Face Transformers to condense lengthy documents, improving information retrieval efficiency by 50% for large corpora.

- Developed a keyword extraction pipeline using SpaCy and NLTK, achieving 85% precision in identifying relevant phrases, enabling quicker data indexing and searchability.
- Integrated the summarization and keyword extraction system with Elasticsearch, providing scalable search capabilities for millions of documents and reducing search times by 60%.
- Deployed the system on TensorFlow with TensorFlow Serving for real-time summarization and keyword extraction, supporting large-scale document processing with minimal latency.
- Fine-tuned the BERT model to optimize performance for specific document types, resulting in 30% higher summarization accuracy compared to baseline models.

Object Detection and Classification in Satellite Images

- Developed an object detection model using YOLOv5 to accurately identify and classify objects in satellite images, achieving a 95% accuracy rate.
- Enhanced model performance by integrating OpenCV for image preprocessing, improving detection accuracy in low-resolution images by 20%.
- Deployed the model on AWS SageMaker, utilizing its scalable infrastructure to process terabytes of satellite imagery, enabling real-time monitoring and classification.
- Leveraged Google Earth Engine for geospatial data processing and model deployment, supporting large-scale analysis for urban planning and disaster management.
- Optimized the object detection pipeline for parallel processing on cloud, reducing inference time by 35%.

Image Classification for Autonomous Vehicles

- Designed and implemented an image classification model using the YOLOv3 framework to detect and classify road objects, achieving 98% accuracy in identifying pedestrians, vehicles, and traffic signs.
- Optimized real-time inference using TensorFlow Serving on Google AI Platform, reducing latency by 25%.
- Enhanced model robustness through extensive data augmentation techniques in Keras, leading to a 15% improvement in detection accuracy under various conditions.
- Deployed the classification system on Google AI Platform with autoscaling capabilities, ensuring the model could handle fluctuating data loads from real-time video feeds.
- Implemented a model monitoring and feedback loop that continuously updates the model with new data, maintaining high accuracy as environmental conditions evolve.

NLP-Based Sentiment Analysis for Social Media

- Developed a sentiment analysis model using the BERT architecture, achieving 90% accuracy in classifying social media posts into sentiments.
- Fine-tuned the BERT model using domain-specific data to improve sentiment prediction accuracy by 20% for niche industries.
- Deployed the sentiment analysis model as a RESTful API using FastAPI, enabling real-time sentiment analysis for large-scale social media monitoring.
- Scaled the model deployment using TensorFlow Serving, supporting thousands of simultaneous API requests with minimal latency.
- Integrated the sentiment analysis tool with social media platforms, providing actionable insights for brand reputation management and customer engagement.

Demand Forecasting for Retail

- Designed and trained an LSTM-based model for demand forecasting, achieving a 25% reduction in forecast error compared to traditional models.
- Implemented advanced time series preprocessing techniques in Pandas, enhancing model accuracy.
- Deployed the forecasting model on Azure ML, leveraging cloud-based scaling to process large volumes of sales data.
- Optimized model performance using hyperparameter tuning in TensorFlow, leading to a 15% increase in forecast accuracy.
- Developed an automated model retraining pipeline to adapt the LSTM model to new data, ensuring continued accuracy over time.

Predictive Maintenance in Manufacturing

- Developed a predictive maintenance model using time series analysis and anomaly detection algorithms, reducing unexpected equipment failures by 40%.
- Integrated IoT sensor data into the model pipeline using Apache Kafka, enabling real-time monitoring of equipment health.
- Deployed the model on AWS SageMaker, ensuring scalable and reliable predictions across multiple manufacturing sites.
- Implemented data preprocessing and feature extraction techniques in Scikit-learn to enhance model accuracy.
- Automated the model update process with continuous learning capabilities, allowing the system to adapt to new data patterns.

Personalized Healthcare Recommendations

- Designed a personalized recommendation model using collaborative and content-based filtering techniques, improving patient treatment adherence by 30%.
- Integrated NLP techniques to analyze unstructured medical records, extracting key insights for personalized treatment plans.
- Deployed the recommendation system on Kubernetes, ensuring scalable and secure delivery of treatment plans.
- Enhanced model interpretability by incorporating SHAP values, allowing healthcare providers to understand treatment recommendations.
- Optimized the system for real-time performance using PyTorch and TensorFlow, enabling timely delivery of personalized treatment plans.

SKILLS

- **Machine Learning Frameworks:** TensorFlow, PyTorch, Scikit-learn, XGBoost, Keras, Hugging Face Transformers
- **Data Processing:** Pandas, NumPy, Apache Spark, Dask, Pyspark
- **Model Deployment:** TensorFlow Serving, TorchServe, AWS SageMaker, Google AI Platform, Azure ML, Docker, Kubernetes, FastAPI
- **Cloud Platforms:** AWS (S3, EC2, Lambda, SageMaker, Redshift), GCP (AI Platform, BigQuery, Dataflow), Azure (ML, Synapse Analytics)
- **MLOps and Infrastructure:** Terraform, Docker, Kubernetes, MLflow, Apache Airflow
- **Big Data Technologies:** Hadoop, Spark, Kafka, Apache Hive, Apache Flink
- **Data Storage:** NoSQL (MongoDB, Cassandra, Neo4j), SQL Databases, Data Lakes
- **AI Techniques:** NLP, Computer Vision, Reinforcement Learning, Deep Learning, Transfer Learning, Model Interpretability (SHAP, LIME)
- **DevOps for ML:** CI/CD Pipelines, Jenkins, GitLab CI, ML Pipelines (Kubeflow, TFX)