Shivansh Srivastav

Data Scientist | AI Engineer

shivanshthedeveloper@gmail.com | +91- 7985763648 | linkedin.com/in/shivansh-srivastav | shivansh408.github.io/portfolio/

PROFESSIONAL SUMMARY

Data Scientist with experience in building data-intensive applications with proficiency in **AI** and, **neural architectures** such as CNNs, RNNs, Encoder-Decoders, and generative models including GANs and GPT. Hands-on experience with **Cloud platforms**, especially AWS, and Azure, and adeptness with **Container orchestration tools** like Docker and Kubernetes. Acquaintance with **agile** development practices and worked on **version control utilities**, such as Git, and GitHub. Mastery of programming languages, notably **Python, C++ and Java** with Stellar communication and presentation abilities.

WORK EXPERIENCE

May 2024 – Present

AI Engineer | Neural Niti | Delhi, INDIA

- Led high-impact projects in the healthcare, management and Automotive industries utilizing AI for Factory digitalization and automation, designed products for US and International market for Hospitals.
- Managed a team of professionals, overseeing research, data analytics, and advancing projects from initial concept to development, resulting in a 25% improvement in project efficiency and 60% more investment.

Sep 2022 – Apr 2024

Open-Source Developer | PyTorch | Delhi, INDIA

- Added self-attention graph pooling, dense GCN layer, and edge weights to Graph-Conv: layer, and image classification, torchvision and optim.
- Ported GPU efficient model, and implemented transformation for images.
- Performed unit tests to ensure correctness, and fixed multiple bugs and issues within community norms.

Feb 2021 – Apr 2021

Technical and Data Analyst | 1k Entrepreneurs | Bengaluru, INDIA

- Maximized the processing speed of the website by 30% using analytical plugins.
- Envisioned a 5-member cross-functional team of TiE Bangalore to format unstructured data, analyzed data to provide valuable insights and market trend analytics.
- Produced a suite of meaningful reports to help Directors make important decisions, and mentored new interns to increase productivity by 18%.

EDUCATION

2021 – 2023

Master of Computer Applications, Major in Artificial Intelligence | Jain University, Bengaluru

2017 – 2020

Bachelor of Computer Applications, Major in Information Technology | Jain University, Bengaluru

SKILLS

Data Science/Machine Learning/Deep Learning: Python, Data Visualization, Feature Engineering, Supervised/Unsupervised ML algorithms, BERT, Transformers, LSTM, CNN, RNN, GANs, Neural Networks, Computer Vision, OpenCV, PIL, EDA, Feature Selection, A/B testing, Data Augmentation, Distributed Model Training, Stochastic Gradient Descent (SGD), Adam, AdamW, FSDP, Deepspeed Zero
Mathematics for ML & DL: Linear Algebra, Probability, Advanced Statistics, Calculus, Matrices
Python Packages & Frameworks: Scikit-learn, TensorFlow, PyTorch, NumPy, Pandas, SciPy, Keras, Beautiful Soup, PySpark, OpenCV, Pillow
MLOps Tools: CI/CD, DVC, MLflow, Tf-extended, Jenkins
Programming Languages: Python, C, C++, JavaScript
Databases: MySQL, MongoDB, PostgreSQL, Redis, Milvus, Neo4j, SQL, NoSQL
Cloud Deployment: AWS, Azure, AWS EC2 and S3, Heroku, Text-to-Speech and Speech-to-Text services.

PROJECTS

Automated HRM with Video Interview Platform — Project Report

Project Brief:

- Developed an automated HRM system integrating a video interview platform to streamline candidate assessment resulting in a 40% reduction in recruitment time and a 30% increase in assessment accuracy.

Results/Deliverables:

- Reduced recruitment time by 40% using advanced analytics, with Milvus DB for fast similarity searches and AWS cloud for scalable infrastructure.
- Increased assessment accuracy by 30% by integrating machine learning algorithms and leveraging Llama for prompt engineering, and employing computer vision techniques for Resume parsing.
- Developed a robust video interview platform with Backend development, AWS, and advanced APIs, utilizing text-to-speech and speech-to-text technologies for enhanced candidate interaction.

Credit Assessment for Predicting Defaulters — Project Report

Project Brief:

- Analyzed and processed the American Express dataset to prioritize applications, accurately predicting healthy customers (0) and potential defaulters (1).

Results/Deliverables:

- Conducted clustering, k-fold cross-validation, and autoencoder-based anomaly detection to preprocess and enhance data quality for credit risk analysis.
- Implemented advanced machine learning models, including Random Forest, SVM Regressor, XGBoost, ANN, and Logistic Regression, achieving a prediction accuracy of 96.8% in identifying potential defaulters.
- Developed and fine-tuned model features to optimize performance in credit assessment, contributing to more accurate risk prediction and decision-making.

Diagnosing cancerous & non-cancerous breast tumor — Project Report

Project Brief:

- Predicted the malignancy of breast tumors by analyzing 569 breast tissue samples, applying feature synthesis techniques to accurately classify tumors as cancerous (malignant) or non-cancerous (benign).

Results/Deliverables:

- Plotted model performance graphs for various ML algorithms, demonstrating a 27% improvement, and achieved a prediction accuracy of 99.1% with Logistic Regression classifier for diagnosing breast tumours.
- Utilized AUC-ROC and Precision-Recall curves to evaluate model performance for distinguishing between cancerous and non-cancerous cases.