Shreyansh Kumar AI/ML Engineer

shreyansh.developer07@gmail.com | (617) 259-0101 | USA | LinkedIn | GitHub

Summary

AI/ML Engineer with 3+ years of strong expertise in machine learning, deep learning, and data analytics. Skilled in building scalable models, data pipelines, and cloud-based deployments. Proficient in Python and advanced ML frameworks, with experience across the full model lifecycle. Adept at cross-functional collaboration, automation, and delivering data-driven solutions to support business decision-making and improve operational efficiency.

Technical Skills

- **Programming & Scripting**: Python (OOP, scripting), SQL, R, API integration (Flask, FastAPI)
- **Data Analysis & Visualization**: Pandas, NumPy, Dask, Polars, Matplotlib, Seaborn, Plotly
- **Statistical Modeling & Math Libraries**: SciPy, Statsmodels, probability, linear algebra, calculus, optimization
- **Machine Learning & Deep Learning**: scikit-learn, XGBoost, LightGBM, CatBoost, Random Forest, SVM, KNN, Decision Trees, clustering (K-Means, DBSCAN), PCA, UMAP, t-SNE, TensorFlow, PyTorch, Keras, CNNs, RNNs, LSTMs, GRUs, Autoencoders, Transformers, GANs
- **Natural Language Processing (NLP):** spaCy, NLTK, Hugging Face Transformers, BERT, GPT, RoBERTa, tokenization, text classification, sentiment analysis, RAG, LangChain, prompt engineering
- **Generative AI & LLMs**: OpenAI API, LLM fine-tuning, AI agents, zero/few-shot learning
- **Time Series Forecasting**: ARIMA, SARIMA, SARIMAX, VARMAX, Prophet, RNNs for sequence modeling
- **Model Optimization & Evaluation**: Hyperparameter tuning (Grid Search, Random Search, Bayesian Optimization), cross-validation, ROC-AUC, F1-score, precision/recall
- **Data Engineering & Pipelines**: Apache Spark, Kafka, dbt, Airflow, ETL/ELT, relational embeddings, rolling aggregations, feature stores
- **Cloud Platforms & MLOps**: AWS (S3, SageMaker), Azure (Data Factory, Azure ML, AKS), GCP; Docker, Kubernetes, MLflow, CI/CD, model drift detection, monitoring with Grafana
- **Computer Vision:** OpenCV, YOLO, Detectron2, image segmentation, object detection, ViT (Vision Transformers)
- **Tools & DevOps**: Git, GitHub, GitLab CI/CD, VS Code, Jupyter, Colab, Postman, Streamlit, Tableau, Power BI

Professional Experience

AI/ML Engineer, *Brex* 10/2024 – Present | Remote, USA

- Worked on an AI-Powered Expense Intelligence System, collaborating with product managers, finance teams, and data engineers to deliver personalized spend insights, fraud detection, and budget optimization tools for corporate clients.
- Collaborated with data engineering to collect, preprocess, and store transactional, merchant, and category data on AWS S3. Used Python and Pandas for data cleaning, normalization, and feature extraction, resulting in a high-quality dataset with 98% completeness for model development.
- Build reinforcement learning models in TensorFlow and OpenAI Gym to optimize real-time budget recommendations and anomaly alerts. Designed reward functions centered around spend efficiency, policy compliance, and user engagement, achieving an F1-score of 0.87 in fraud detection.
- Implemented feature engineering pipelines including user segmentation via clustering, temporal spend patterns, and merchant behavior encoding. Leveraged AWS SageMaker for scalable training and hyperparameter tuning, improving fraud AUC by 12% and budget recommendation accuracy.
- Applied Bayesian optimization to fine-tune model parameters and alert thresholds. Ran extensive A/B tests with client cohorts, collaborating with product and UX teams to drive a 25% increase in usage of smart spend insights and a measurable uplift in client satisfaction.
- Validated models for regulatory compliance and operational robustness in coordination with risk and compliance teams. Partnered with MLOps to containerize models using Docker and deploy them via Kubernetes. Delivered RESTful APIs through Flask for seamless integration into Brex's web and mobile platforms.

Data Scientist, *Nexdigm* 08/2022 – 06/2023 | Gurgaon, India

- Led development of a Python-based time series tool using ARIMA, SARIMA, SARIMAX, VARMAX, and RNN models, reducing exploratory forecasting time by 70% through efficient automation and streamlined analysis workflows.
- Automated model selection and hyperparameter tuning, running 50+ iterations per dataset, which significantly reduced project timelines from 3 weeks to just 4 days, accelerating delivery cycles for multiple time-sensitive client projects.
- Enhanced model accuracy by 12% using iterative tuning, advanced feature engineering, and performance optimization, enabling more reliable predictions and data-driven decision-making for stakeholders across various industries and use-case domains.
- Boosted client engagement by 15% through timely delivery of predictive insights, leveraging automated time series forecasting outputs to support strategic planning, drive value, and strengthen ongoing relationships with enterprise customers.

ML Engineer, *Barclays* 01/2021 – 12/2021 | Remote, India

- Built a credit risk scoring system using ensemble models and Temporal Graph Convolutional Networks, integrating behavioral and demographic features to enhance borrower segmentation, boost AUC by 29%, and improve high-risk borrower recall by 40%.
- Engineered an end-to-end data pipeline using Azure Data Factory and Apache Spark on Azure Databricks, processing over millions daily transactions with 99.5% reliability, supporting large-scale data ingestion and real-time transformation needs.
- Developed custom T-GCNs with PyTorch Geometric and DGL to model dynamic borrower-lender relationships. Leveraged Azure Machine Learning for distributed training, automated orchestration, and efficient experiment tracking across multiple compute environments.
- Applied advanced feature extraction methods like rolling-window aggregations, repayment patterns, and graph-based embeddings. Optimized latency by 35% while maintaining accuracy using time-series cross-validation, backtesting, and hyperparameter tuning with Optuna.
- Deployed models on Azure Kubernetes Service using Docker and integrated CI/CD pipelines. Created Grafana dashboards to monitor model drift, prediction confidence, and risk alerts, ensuring robust production performance and explainability.

Education

Master of Science in Applied Data Analytics 09/2023 – 01/2025
Boston University, Boston, MA, USA
Bachelor of Technology (B.Tech), Computer Science 2019 – 2023
Bennett University Greater Noida, Uttar Pradesh, India

Certificates

- AWS Academy Cloud Foundations
- Microsoft Certified: Azure AI Fundamentals
- Introduction to Data Science
- Machine Learning with Python