Smit Patel

AI/ML Engineer| **Location:** IL | **Mobile**: 224-251-0424 | **Email:** smitpatel0049@gmail.com | GitHub

SUMMARY:

AI/ML Engineer with 3+ years of experience designing, developing, and deploying scalable machine learning and deep learning solutions across healthcare and analytics domains. Skilled in Python, TensorFlow, and AWS, with proven expertise in NLP, computer vision, and predictive modeling. Experienced in building end-to-end MLOps pipelines using Docker, MLflow, and CI/CD, improving model reliability and deployment efficiency. Adept at translating complex data into actionable insights through Tableau, Power BI, and Plotly, driving measurable business impact and process automation.

SKILLS:

Programming: Python (NumPy, Pandas, Dask), SQL, R, Java, C++
Machine Learning & AI: Supervised & Unsupervised Learning, Ensemble Methods (XGBoost, LightGBM), CNNs, RNNs, LSTMs, Transformers, Generative AI (LLMs, BERT, GPT Models, Diffusion Models), Recommender Systems, Predictive Analytics, Time-Series Forecasting, Autoencoders
Artificial Intelligence Specializations: Natural Language Processing (Text Classification, NER, Sentiment Analysis, Question Answering), Computer Vision (Image Classification, Object Detection, Segmentation)
MLOps & Deployment: Docker, Kubernetes, MLflow, Kubeflow, TensorFlow Serving, Weights & Biases, CI/CD (Jenkins, GitHub Actions, GitLab CI), Model Serving (Flask, FastAPI, REST APIs, gRPC), Monitoring (Prometheus)
Cloud Platforms & Databases: AWS (SageMaker, Lambda, EC2, S3), GCP (AI Platform, BigQuery, Vertex AI), Azure (Azure ML, Cognitive Services); PostgreSQL, MySQL, MongoDB, Redis, Snowflake
Visualization & Analytics: Matplotlib, Seaborn, Plotly, Tableau, Power BI
Collaboration & Tools: Git, GitHub, GitLab, Bitbucket, Jupyter Notebook, VS Code, PyCharm

EDUCATION:

Master's in Data Science **Nov 2023**
DePaul University, Chicago, IL
Bachelor of Computer Engineering **May 2021**
Charotar University of Science and Technology

EXPERIENCE:

Cigna, IL | June 2023 – Current | AI/ML Engineer

- Designed and deployed an ML pipeline using Python (Pandas, Scikit-learn) and XGBoost to predict patient readmission risks, improving hospital resource allocation and reducing readmission rates by 15%.
- Developed NLP models using BERT and GPT-based transformers for medical text extraction (diagnoses, medications, and procedures) from unstructured EHR data, reducing manual data entry time by 40%.
- Built anomaly detection models using unsupervised learning (Isolation Forest, Autoencoders) and integrated them into Cigna's claim system via Flask APIs, reducing fraudulent claim cases by 20%.
- Engineered a time-series forecasting framework using LSTMs and Prophet for predicting patient health deterioration trends, enabling proactive care interventions.
- Streamlined model lifecycle using MLflow, Docker, and AWS SageMaker; implemented CI/CD with GitHub Actions and Prometheus-based monitoring for real-time model health tracking.
- Designed interactive dashboards using Tableau and Power BI for executives to visualize patient outcomes, cost efficiency, and ML model KPIs across business units.

Infinite Infolab, India | July 2020 - Aug 2021 | ML Engineer

- Built and optimized supervised learning models for customer churn prediction using Python and LightGBM, improving retention campaign targeting accuracy by 22%.
- Developed personalized recommender systems using collaborative and content-based filtering, increasing user engagement and cross-sell conversions by 18%.
- Created and fine-tuned CNN-based image classification and object detection models with TensorFlow and PyTorch, achieving 92%+ model accuracy on real-world datasets.
- Designed robust time-series forecasting solutions using LSTMs, ARIMA, and Prophet for sales prediction, improving forecast reliability by 15–20% and enabling data-driven inventory planning.
- Deployed end-to-end ML pipelines with Docker and FastAPI for production-grade inference, reducing model latency by 35% and supporting scalable performance under heavy traffic.
- Conducted A/B testing and hyperparameter optimization using Bayesian search, improving model precision and recall metrics by 10% on average.
- Delivered actionable insights via interactive dashboards and EDA visualizations in Plotly and Seaborn, helping business teams accelerate decision-making and identify emerging trends.