Wasimuddin Salar Fathimullah

+1 (618) 434-5330 | wasimsalar.f.m@gmail.com | LinkedIn | GitHub | Carbondale, IL

Summary

Machine Learning Engineer with 5 years of experience designing, deploying, and optimizing end-to-end AI/ML solutions across academia, startups, and enterprise. Skilled in Natural Language Processing, federated learning, predictive modeling, and anomaly detection, with expertise in Python, PyTorch, TensorFlow, and GCP. Skilled at building real-time data pipelines, scalable ML models, and cloud-native deployments using Docker, BigQuery, and Databricks. Strong background in MLOps, data engineering, and stakeholder collaboration, with proven success in improving business outcomes through actionable insights. Certified Pega Senior System Architect (CSSA), Google Data Analytics, and Deep Learning specialist.

Skills

Languages & Libraries: *Python (NumPy, Pandas, Scikit-learn, Matplotlib, Seaborn), R, C*
ML & AI: *PyTorch, TensorFlow, Keras, Vision Transformers, Federated Learning, Hugging Face (NLP), Deep Learning (CNNs, RNNs, LSTMs, Transformers), Recommender Systems, Anomaly Detection*
Data Engineering & Cloud: *BigQuery, Google Cloud Platform, Docker, FastAPI, Flask*
Developer Tools: *Git, VS Code, MATLAB, Jupyter, Qualtrics*
Certifications: *Pega [CSA, CSSA], Google Data Analytics [Coursera],* Machine Learning *[Coursera]*, Deep Learning *[DataCamp]*

Professional Experience

Software Engineer

Uktha Global Inc. | Atlanta, GA | April 2025 – Present

Responsibilities:

- Developed deep learning recommendation systems with collaborative filtering and embeddings, increasing customer engagement by 20%.
- Built NLP models with Hugging Face Transformers for text classification and sentiment analysis, deployed via FastAPI + Docker in production.
- Partnered with data engineers to migrate batch ML models into near real-time pipelines on Databricks, cutting processing time from hours to minutes.
- Implemented model monitoring dashboards (Grafana + custom Python scripts) to track drift, accuracy, and latency in production.
- Optimized ML training workflows on GCP (Vertex AI, BigQuery), improving model retraining speed by 40%.
- Collaborated with cross-functional teams to integrate ML APIs into core business applications, enabling data-driven personalization at scale.

Environment: Python, PyTorch, Hugging Face, Databricks, GCP, Vertex AI, BigQuery, FastAPI, Docker, Grafana

Volunteer Technical Assistant (ML Focus)

Southern Illinois University Carbondale | Carbondale, IL | July 2024 – April 2025

Responsibilities:

- Modeled 1,000+ housing records using Pandas & Scikit-learn, predicting utilization patterns and reducing vacancy by 15%.

- Built predictive dashboards with Matplotlib & Seaborn, enabling leadership to allocate housing resources more efficiently.
- Implemented survey response analysis with NLP (TF-IDF + Logistic Regression), boosting engagement and feedback quality by 20%.
- Automated ETL pipelines for housing and student datasets using Python & Airflow, reducing manual reporting time by 50%.
- Deployed ML-driven dashboards on Flask + Docker, ensuring scalability and reproducibility for university-wide adoption.
- Conducted model explainability analysis (SHAP, LIME) for administrative decision-making, ensuring trust in ML outcomes.
- Collaborated with IT staff to maintain data governance and compliance standards while handling sensitive student data.

Environment: Python, Scikit-learn, Pandas, Airflow, Flask, Docker, SHAP, LIME, Matplotlib, Seaborn

Graduate Assistant (ML Engineer & Data Analyst)

Southern Illinois University Carbondale | Carbondale, IL | August 2022 – May 2024

Responsibilities:

- Leveraged Applied NLP sentiment analysis (BERT, Hugging Face) on 10,000+ student feedback entries, reducing housing vacancy rates by 10%.
- Engineered predictive models with Scikit-learn & TensorFlow, improving student retention prediction with 90% accuracy.
- Unified Navigate and Advocate platforms with automated triage ML workflows, cutting issue resolution time by 50%.
- Mentored 10+ graduate assistants on ML practices, data visualization, and reproducibility, raising overall project success.
- Built interactive dashboards in Flask to visualize academic performance metrics, used by administrators for strategic planning.
- Conducted A/B testing on engagement interventions, identifying strategies that improved inclusivity metrics by 12%.
- Published findings on ML-driven student housing solutions in internal research symposiums.

Environment: Python, TensorFlow, Scikit-learn, BERT, Hugging Face, Flask, Pandas, Matplotlib

Assistant System Engineer (SWE)

Tata Consultancy Services | Chennai, India | May 2021 – June 2022

Responsibilities:

- Automated operational reports with Python + Pandas, improving reporting efficiency by 25% and client satisfaction by 15%.
- Collaborated with data science teams to develop ML anomaly detection models (Isolation Forest, Autoencoders) for IT monitoring.
- Reduced application downtime by 40% through predictive maintenance models integrated into client systems.
- Deployed ML pipelines using Docker + Pega, ensuring production readiness with a 95% Guardrail score.
- Designed classification models for incident triage, reducing manual effort by 30% in resolving service requests.
- Created data visualization dashboards (Matplotlib & Seaborn) for real-time KPI monitoring by leadership teams.
- Contributed to Agile sprints by integrating ML solutions directly into enterprise-scale applications.

Environment: Python, Pandas, Docker, Pega, Isolation Forest, Autoencoders, Matplotlib, Seaborn, Agile

Python Developer (Intern)

Mandli Technologies Pvt. Ltd. | Bangalore, India | October 2020 – April 2021

Responsibilities:

- Enhanced ML pipelines for AI-driven e-learning modules, improving personalization by 30%.
- Trained ML classifiers on 10,000+ user interactions, raising module completion by 25%.
- Integrated ML model outputs into production-ready features within Agile sprints with cross-functional teams.
- Developed REST APIs in Flask to serve model predictions, enabling real-time feedback for learners.
- Performed data cleaning and feature engineering on 50,000+ learning activity logs, improving model accuracy by 15%.
- Implemented unit tests and CI/CD practices for ML workflows, ensuring robust and reliable deployments.
- Built interactive data visualizations (Matplotlib & Seaborn) to track learner performance and engagement trends.
- Collaborated with senior developers to optimize Python scripts, reducing training runtime by 20%.

Environment: Python, Scikit-learn, Pandas, Flask, Matplotlib, Seaborn, Git, Agile, CI/CD

Projects

Federated Learning IDS for Autonomous Vehicles · [*Github*] | *Graduate Project*

- Designed a federated learning Intrusion Detection System (IDS) using PyTorch & Vision Transformers (ViTs) for decentralized traffic sign recognition.
- Achieved 90% detection accuracy, reduced false positives by 15%, and improved overall system security by 40% without centralizing sensitive vehicle data.
- Implemented federated averaging (FedAvg) for distributed training across multiple simulated clients, ensuring data privacy and scalability.
- Optimized training with GPU acceleration (CUDA), reducing model convergence time by 25%.
- Applied data augmentation & imbalance handling (SMOTE) to improve model robustness across diverse driving environments.

Environment/Tools: Python, PyTorch, Vision Transformers, Federated Learning (FedAvg), CUDA, SMOTE, Weights & Biases

Marketing Campaign and Loan Status Predictor · [*Github*] | *Machine Learning & Soft Computing Project*

- Developed predictive models to forecast term deposit subscriptions and loan eligibility using ensemble techniques (Decision Tree, KNN, Naive Bayes).
- Reached 97% accuracy while integrating a differential privacy mechanism to protect sensitive PII during both training and inference.
- Conducted feature engineering on 50,000+ customer records, extracting behavioral and transactional attributes for improved predictions.
- Compared performance using cross-validation, confusion matrix, ROC-AUC, and hyperparameter tuning with GridSearchCV.
- Automated data preprocessing pipeline with Pandas & Scikit-learn, ensuring scalability for future datasets.

Environment/Tools: Python, Scikit-learn, Pandas, NumPy, Matplotlib, Differential Privacy, GridSearchCV

Education

Master of Science, Computer Science — **August 2022 – May 2024**
Southern Illinois University Carbondale | GPA: 3.879/4 — *Carbondale, IL*

Bachelor of Technology, Electronics and Communication Engineering — **July 2017 – April 2021**
Koneru Lakshmaiah Deemed University | GPA: 9.03/10 — *Vijayawada, India*