Modou K Touray

Phone number: (+1) 5062927713 (Mobile) | **Email address:** modou.touray@unb.ca | **LinkedIn:** Modou K Touray |

Address: 523 Canterbury Drive, E3B 4M4, Fredericton, Canada (Home)

ABOUT ME

Results-driven AI Engineer with over 3 years of hands-on experience building, training, and deploying machine learning, deep learning, and generative AI systems. Strong background in end-to-end ML lifecycle, including data preprocessing, feature engineering, model training, evaluations and deployment. Skilled in PyTorch, TensorFlow, Keras, Scikit-learn, Hugging Face, and cloud platforms such as AWS SageMaker AI, Lambda, API Gateway, and Bedrock. I bring a unique blend of software engineering, cloud engineering, and ML skills to ship AI solutions that are scalable, production-ready, and business-impactful.

SKILLS

Sagemaker AI | Machine Learning and Generative AI | Amazon Bedrock | LLM Fine-Tuning | Langchain | CI/CD & Automation | Monitoring and Observability | Kubernetes | Docker | Python | S3 | EKS | AWS API Gateway | Prompt Engineering | OpenCV | Retrieval Augmented Generation (RAG) | Classical ML | Pandas | Numpy | Keras | Matplotlib | Seaborn

WORK EXPERIENCE

AI ENGINEER – DUNDAL AI – 05/2024 – Current

- Led client-facing design and deployment of a real-time generative AI proof of concept on AWS, architecting a scalable inference endpoint with SageMaker, API Gateway, Lambda, and S3.
- Developed and deployed a FastAPI-based AI endpoint for deepfake detection across images, audio, and video, using Reality Defender API to deliver scalable, production-ready content verification solutions.

SOFTWARE DEVELOPER – GAMBIA NATIONAL PETROLEUM CORPORATION – 12/2021 – 08/2023 – BRUSUBI, THE GAMBIA

- Built the corporation's first official website using Laravel Livewire and Tailwind CSS, delivering a modern website.
- Implemented a dynamic, component-driven architecture that streamlined content updates for non-technical staff, which enabled easy maintenance for the website.
- Enhanced in-house web applications by integrating interactive analytics dashboards using JavaScript libraries and Laravel.
- Improved operational efficiency by delivering features that provided live insights to stakeholders, enabling faster, data-driven actions.
- Developed and deployed a nationwide coupon tracking system using VueJS and MySQL, automating coupon validation across petrol stations.

AI SOFTWARE DEVELOPER – CYPEFINLAND – 12/2020 – 11/2021 – BAKAU, THE GAMBIA

- Researched, trained, and evaluated an XGBoost model for space-availability prediction using sensor time-series data, achieving 95% accuracy.
- Handled end-to-end model development in Python, including data processing, training, and evaluation.
- Built Flask applications to serve and reference AI models, enabling real-time predictions and testing.
- Developed an automated invoice extraction tool using PyTesseract and advanced data parsing techniques to capture key financial data and exported structured data to CSV, reducing manual processing time and significantly improving data accuracy.

EDUCATION AND TRAINING

09/2023 – 05/2025 Fredericton, Canada

MASTERS IN COMPUTER SCIENCE University of New Brunswick

Key Courses

- Cloud and Information Management (AWS)
- Machine Learning and Data Mining

- Natural Language Processing
- Software Architecture
- Big Data Systems

09/2016 – 10/2020 Kanifing, The Gambia
BACHELORS DEGREE IN COMPUTER SCIENCE University of The Gambia

PROJECTS

Digital Twin

- Designed and deployed a production-grade serverless AI Digital Twin of me on AWS, leveraging FastAPI, AWS Lambda, API Gateway, Bedrock (Nova lite model), S3, and CloudFront to deliver a scalable, low-latency conversational AI platform.
- Provisioned and managed multi-environment cloud infrastructure using Terraform (IaC), enabling dev/test/prod isolation with automated setup of IAM, Lambda, API Gateway, S3, and CloudFront backed by remote state and locking.
- Built an end-to-end CI/CD pipeline with GitHub Actions and OIDC-based AWS authentication, automating Lambda packaging, infrastructure deployment, frontend builds, S3 publishing, and CloudFront cache invalidation.

Dialogue Summarization (AWS SageMaker + Flan-T5)

- Fine-tuned the Flan-T5 model using PEFT/LoRA in SageMaker Notebooks.
- Evaluated summarization quality using ROUGE metrics.
- Produced a lightweight, cost-efficient model for conversation summarization

Farm Intrusion Detection System

- Led full-cycle development of a real-time intrusion detection system for farmers.
- Conducted data pre-processing and applied augmentation using OpenCV to simulate nighttime conditions across 75% of the dataset.
- Trained a deep learning model using Transfer Learning (InceptionV3) with hyperparameter tuning and iterative experimentation, achieving 98% accuracy on the final model.
- Deployed a lightweight detection pipeline with Streamlit and integrated Twilio for instant SMS alerts to farmers.

AI-Driven E-Commerce Customer Support Agent

- Architected and deployed a Node.js e-commerce platform using AWS, GCP, and Azure for scalable, intelligent operations.
- Automated secure AWS infrastructure with Terraform (VPC, Lambda, DynamoDB) for repeatable, production-ready deployments.
- Containerized and orchestrated workloads using Docker and Kubernetes (Amazon EKS) for high availability and portability.
- Built a full CI/CD pipeline with AWS CodePipeline, enabling automated build, test, and zero-downtime releases.
- Designed and optimized an AI-powered customer support agent using AWS Bedrock (Claude Sonnet 3 and GPT-4) to securely query DynamoDB, generate accurate insights, and deliver personalized customer responses.
- Implemented a real-time, event-driven data pipeline by streaming DynamoDB updates through Lambda into Google BigQuery, enabling analytics and reporting on order data.
- Integrated Azure AI Language for sentiment analysis to enhance customer experience by detecting user intent and emotional context during agent interactions.

Spam Classification System

- Led and built a complete NLP pipeline for spam detection.
- Developed a script to automate the data extraction pipeline to parse raw emails from the Enron corpus's multi-folder directory structure and transform them into a structured CSV dataset.
- Engineered an end-to-end preprocessing workflow, including text cleaning, tokenization, and stop-word removal.
- Applied TF-IDF vectorization to transform text into numerical features optimised for Logistic Regression.
- Achieved 98% accuracy, 97% precision, and 99% recall, representing a 43% improvement over baseline.

CERTIFICATIONS

Amazon Web Services
AWS Certified AI Practitioner

Amazon Web Services
AWS Certified Cloud Practitioner

Standford Online
Machine Learning

DeepLearning.AI
Generative AI with Large Language Model