

Syed Waquas Hashmi

Machine Learning Engineer - Data Engineer & Cybersecurity Specialist

Email: waquashashmi@yahoo.com — Phone: +44 7771024056 — Location: Cambridge, United Kingdom — Website: waquashashmi.com — LinkedIn: linkedin.com/in/waquashashmi

Professional Summary

Experienced in building secure, high-performance data systems and machine learning models. Specialized in designing intrusion detection systems, ETL pipelines, and real-time analytics using Apache Spark, Hadoop, PySpark, and cloud platforms (AWS, Azure, GCP). Proficient in Python, SQL, and JavaScript, with expertise in TensorFlow, PyTorch, Scikit-learn, and Keras for developing NLP, computer vision, and deep learning solutions. Strong background in containerization (Docker, Kubernetes), CI/CD, and infrastructure-as-code. Skilled in data visualization (PowerBI, Tableau) and explainable AI techniques, improving investigation time by 60%. Experienced in applying technical innovation to both academic research and industry projects, including healthcare data analytics and security monitoring systems.

Professional Experience

Research Associate - AI Security (CyBOK Project)

January 2025 - Present

Anglia Ruskin University - Hybrid

Cambridge, UK

- Led development of educational resources for AI/ML security as part of funded CyBOK project
- Developed interactive lab environments showcasing 12+ AI security vulnerabilities using Python, TensorFlow and PyTorch
- Created 15+ educational modules with hands-on exercises in machine learning security
- Implemented 10+ deliberately vulnerable ML models across classification, NLP, and computer vision
- Built attack demonstrations for CNNs, RNNs, Transformers, GANs and Autoencoders with 98% reproducibility
- Integrated visualization tools for adversarial machine learning attacks with real-time model manipulation
- Designed secure coding curriculum for ML engineers, reducing production vulnerabilities by 65%
- Deployed lab environments using infrastructure-as-code across AWS, Azure and GCP platforms

Research Associate - Security Engineer (Sirocco Project)

January 2025 - Present

Anglia Ruskin University - On-site

Cambridge, UK

- Developed intrusion detection system with 95% accuracy for wind farm infrastructure security
- Designed distributed security network using Raspberry Pi devices integrated with Azure HDInsight
- Created high-performance PowerBI dashboard processing 100+ concurrent connections
- Implemented explainable AI visualizations for 35 security features, reducing investigation time 60%
- Reduced false positives by 75% using federated learning models deployed through Azure DevOps
- Built ETL pipelines with Azure Data Factory for transforming security logs from diverse sources
- Developed Apache Airflow workflows for security event processing handling 1000+ events/second
- Migrated security storage from Cassandra to Azure Synapse Analytics for improved performance
- Streamlined deployment with Kubernetes, reducing setup time from 2 hours to 10 minutes

Machine Learning Engineer

July 2024 - October 2024

Smallcap.ai - On-site

London, UK

- Developed, trained, and optimized ML models using Random Forest, XGBoost, and CNNs, improving accuracy by 15-25%
- Utilized Python (NumPy, Pandas, Matplotlib, SciPy, Seaborn) for EDA, uncovering insights and reducing data biases by 20%
- Collected, cleaned, and preprocessed large datasets (CSV, Excel, SQL), improving data quality by 30%
- Built ETL pipelines and complex SQL queries for efficient data extraction, transformation, and validation
- Designed advanced visualizations in Tableau and Excel (Heat Maps, Pareto Charts, Tree Maps)
- Deployed models using Docker, Kubernetes, and Flask, ensuring scalability on AWS, GCP, and Azure
- Implemented MLflow and DVC for version control and efficient model monitoring

Data Analyst

July 2022 - July 2023

SKYMe - Remote

Melbourne, Australia

- Analyzed 10,000+ MRI reports using NLP, achieving 92% accuracy in medical data extraction
- Implemented ML models with 90% average accuracy and deployed NLP solutions reducing diagnostic errors by 25%

- Designed ETL pipelines using Apache Spark/Hadoop and optimized SQL queries for data transformation
 - Developed scalable workflows on AWS (S3, Redshift, Lambda) and migrated pipelines to Azure services
 - Created real-time data streaming solutions with Kafka/Flink and automated visualizations with Tableau/Power BI
 - Optimized healthcare datasets, reducing redundancy by 40% while maintaining data integrity
- Software Engineer**
SMART HOME TUTORS - On-site

January 2016 - May 2019
Aurangabad, Maharashtra, India
- Built and managed data pipelines using Hadoop/MapReduce and implemented data storage solutions with HDFS/HBase
 - Created ETL workflows with Apache Pig/Hive and utilized Spark for distributed data processing
 - Developed data warehousing solutions with Amazon Redshift and optimized SQL queries to improve performance
 - Ensured data quality through validation practices and automated processing tasks using Python/Bash scripts

Education

- Anglia Ruskin University**
Master in Artificial Intelligence

September 2023 - January 2025
Cambridge, United Kingdom
- P.E.S. College of Engineering**
Bachelor of Technology in Computer Science & Engineering

June 2019 - September 2022
Dr. Babasaheb Ambedkar Technical University, India
- People’s Education Society (P.E.S.) Polytechnic**
Diploma in Mechanical Engineering

June 2012 - May 2015
Maharashtra State Board of Technical Education, India

Technical Expertise

Machine Learning/AI	TensorFlow, PyTorch, Scikit-learn, Keras, NLP, Computer Vision, Deep Learning, Federated Learning, Explainable AI, CNNs, RNNs, Transformers, GANs, Autoencoders
Programming	Python, JavaScript (ES6+), SQL, HTML5/CSS3, PySpark, Bash, T-SQL, SparkSQL
Cloud & DevOps	AWS (S3, Redshift, Lambda, IAM), GCP, Azure (DevOps, Synapse, Data Factory, Data Lake, Databricks, HDInsight), Docker, Kubernetes, CI/CD, Git, Containerization
Data Engineering	Apache (Spark, Airflow, Pig, Hive), ETL Pipelines, SQL Databases (MySQL, PostgreSQL), Hadoop, MapReduce, HDFS, HBase, Cassandra, Data Preprocessing
Security	Intrusion Detection Systems, Network Security, Real-time Monitoring, Threat Intelligence, Adversarial ML, Phishing Detection, PGD Attack Defenses
Visualization	PowerBI, Tableau, Matplotlib, Seaborn, Interactive Dashboards, Heat Maps
Development Tools	Raspberry Pi, OpenCV, Jupyter Notebooks, REST APIs, SOAP APIs, Flask

Key Projects

- Federated Learning Security System** | *TensorFlow, Scikit-learn, Python*

2024

 - Designed federated learning system with 10 clients and central server using TensorFlow and Scikit-learn
 - Implemented PGD attack defenses and enhanced model security through advanced mitigation strategies
- Healthcare Natural Language Processing System** | *Python, Flask, Apache Fuseki, ML*

2023

 - Developed semantic model achieving 98% accuracy in healthcare data processing

Certifications & Languages

- Key Certifications:**

- Machine Learning with Python
 - Deep Learning Specialization
 - Cyber Security Foundation
- Languages:**

- English (Professional)

- Hindi (Native)
- Urdu (Native)

- Marathi (Native)