Dhieddine BARHOUMI

AI & Data Engineer

📍 Offenburg, Germany 　📞 +49 17626630599 　✉️ dhieddine.barhoumi@gmail.com 　in barhoumi-dhieddine 　○ dhiabarhoumi
🌐 dhieddinebarhoumi.me

PROFESSIONAL SUMMARY

AI & Data Engineering Specialist with expertise in developing scalable solutions for unstructured data processing and analysis. Proven track record in implementing AI-powered systems, building data pipelines, and leveraging NLP techniques for text analysis. Experienced in cloud platforms and big data technologies. Strong programming skills in Python with hands-on experience in data engineering tools and frameworks. Demonstrated success in cross-functional collaboration and implementing solutions that drive business value.

EDUCATION

National Engineer's Degree in Industrial IT and Automation　September. 2020 – September. 2025
National Institute of Applied Sciences and Technology (INSAT)　*Tunis, TN*

　Relevant Courses: *Industrial Automation, Control Systems, Database Management, Algorithms and Data Structures, Deep Learning, NLP, Cloud Computing, Big Data, Optimization Methods, Reinforcement Learning, Industrial IoT*

Scientific Baccalaureate　September. 2017 – June 2020
Pioneer High School Hammam-Lif　*Ben Arous, TN*

　Top 5% of bachelors with an excellent average grade.

EXPERIENCE

AI Research Intern　March. 2025 – August. 2025
Institute For Machine Learning And Analytics (Hochschule Offenburg)　*Offenburg, DE*

- Developed and maintained data pipelines for processing multi-modal sensor data, implementing ETL processes for real-time analytics.
- Engineered scalable data processing systems using Apache Spark, improving data throughput by 40% and reducing processing time.
- Implemented NLP techniques for scene understanding and text analysis, achieving 25% better accuracy in complex scenarios.
- Deployed AI models on cloud platforms (AWS, Azure) with containerized solutions using Docker and Kubernetes.
- Designed and implemented data validation and quality checks, ensuring 99.9% data integrity across all pipelines.
- Collaborated with cross-functional teams to integrate AI models into existing systems, ensuring seamless deployment and monitoring.
- **Skills:** Data Engineering, ETL Processes, Apache Spark, NLP, Cloud Platforms, Docker, Kubernetes

AI in Security Systems Intern　June. 2024 – August. 2024
All Points Smart Solutions　*Amman, JO*

- Designed and implemented data pipelines for processing unstructured video and image data, optimizing storage and retrieval efficiency.
- Developed NLP-based text analysis systems for processing security logs and reports, improving information extraction accuracy by 30%.
- Implemented real-time data processing solutions using Apache Kafka and Spark Streaming for instant threat detection.
- Utilized MongoDB for efficient storage and querying of unstructured security data, reducing query time by 50%.
- Integrated AI Models with existing systems using RESTful APIs and microservices architecture.
- **Skills:** Data Pipelines, Apache Kafka, MongoDB, NLP, RESTful APIs, Microservices

Computer Vision Intern　June. 2023 – July. 2023
DidaMind　*Ben Arous, TN*

- Conducted research on multiple computer vision algorithms, choosing YOLOv8 for optimal speed and accuracy.
- Fine-tuned YOLOv8 for detecting angle and gusset objects in boxes, achieving 92.5% precision and 60% mAP50-95.
- Deployed the model as a real-time API on Microsoft Azure, integrating it with a dashboard for defect detection.
- Automated quality control, triggering alerts for misaligned or missing objects as boxes moved on a conveyor.
- Collaborated with a multidisciplinary team to replace manual defect detection for an industry client, boosting efficiency by 30%.
- Completed Microsoft Azure AI Fundamentals training and certification during the internship.
- **Skills:** YOLOv8, Computer Vision, Microsoft Azure, API Development, Real-Time Systems, TensorFlow, Python

PROJECTS

Unstructured Data Processing Platform　October. 2024 – November. 2024

- Developed a scalable platform for processing and analyzing unstructured data using Apache Spark and NLP techniques.
- Implemented ETL pipelines for processing text, images, and social media content, improving data processing speed by 60%.
- Utilized BERT and SpaCy for advanced text analysis and information extraction from various data sources.
- Deployed the platform on AWS using containerized microservices architecture for improved scalability.
- Implemented data quality checks and validation processes, ensuring 99.9% data integrity.
- **Skills:** Data Engineering, Apache Spark, NLP, AWS, Docker, Microservices

Real-time Data Analytics System　September. 2023 – December. 2023

- Designed and implemented a real-time data processing system using Apache Kafka and Spark Streaming.
- Developed data pipelines for processing streaming data from multiple sources with low latency requirements.
- Implemented data validation and quality monitoring systems using Python and Apache Spark.
- Created interactive dashboards using Tableau for real-time data visualization and monitoring.
- Integrated with MongoDB for efficient storage and retrieval of processed data.
- **Skills:** Real-time Processing, Apache Kafka, Spark Streaming, MongoDB, Tableau, Python

TECHNICAL SKILLS

Programming Languages: Python, SQL, Java, C++
Data Engineering: Apache Spark, Apache Kafka, Hadoop, ETL Processes, Data Pipelines
Cloud & Infrastructure: AWS, Azure, Google Cloud, Docker, Kubernetes
AI/ML Frameworks: PyTorch, TensorFlow, Scikit-learn, NLTK, SpaCy, BERT
Databases: MongoDB, PostgreSQL, MySQL, Redis
Development Tools: Git, CI/CD, RESTful APIs, Microservices, Tableau

CERTIFICATIONS

Professional Machine Learning Engineer (Google Cloud)	October. 2024
TensorFlow Professional Developer (DeepLearning.AI)	February. 2024
Deep Learning Specialization (DeepLearning.AI)	October. 2023
AI Engineering Professional Certificate (IBM)	July. 2023
AI Fundamentals (Microsoft Azure)	June. 2023

LANGUAGES

English: Professional Working Proficiency (TOEIC C1: 965/990)
French: Professional Working Proficiency
German: Pre-Intermediate
Arabic: Native Proficiency