

DANA DIAB

+96181212095 • diab.danawork@gmail.com • LB • Open to Remote • [LinkedIn](#) • [Credly](#) • [GitHub](#)

AI Engineer | Data Scientist

As a data storyteller with expertise in artificial intelligence, I am dedicated to revealing the stories embedded in data. Utilizing cutting-edge AI technologies, I convert raw data into meaningful insights that facilitate strategic decision-making and deliver significant results.

Professional Experience

Coding_C_Lebanon

Sep 2023 – Jun 2024

AI Engineer Trainee, Remote

- Implement data augmentation and feature engineering techniques to enhance model performance.
- Assist in the design and development of machine learning models using frameworks such as TensorFlow, PyTorch, and Scikit-Learn.
- Implement and optimize machine learning algorithms for various AI applications, such as natural language processing (NLP), computer vision, and predictive analytics.

Google Developers Student Club

Oct 2023 – Jul 2024

Vice President and Former Core Team Member, Tripoli

- Cultivated and maintained strong relationships with key stakeholders, effectively managing communication strategies that enhanced brand reputation.
 - Led and motivated a diverse team, fostering collaboration and guiding projects to successful completion with a focus on achieving organizational goals.
 - Collaborated effectively with cross-functional teams, contributing to a cohesive work environment and the achievement of project milestones.
 - Designed compelling visual content and branding materials, enhancing marketing campaigns and increasing audience engagement.
 - Developed and executed targeted social media strategies, resulting in increased online presence and follower growth across multiple platforms.
-

Projects

Analyze Common Crawl Data with PySpark, [Link](#)

- Developed a data analysis pipeline using PySpark to process and transform large-scale domain graph datasets from Common Crawl.
- Implemented custom data transformation functions and applied RDD operations to extract, format, and analyze domain and subdomain information.
- Configured and utilized Apache Spark for distributed data processing, optimizing performance for handling high-volume datasets.
- Demonstrated expertise in data engineering by building scalable solutions to read, process, and manipulate data from CSV files into structured formats.

StyleVision: Automated Fashion Image Classification, [Link](#)

- Processed and pre-processed the dataset containing 60,000 training images and 10,000 test images across 10 different clothing categories.
- Designed and implemented a convolutional neural network (CNN) to achieve high accuracy in classifying clothing items.
- Conducted hyperparameter tuning to optimize model accuracy, achieving 91% accuracy on the test set.
- Visualized and analyzed model performance through confusion matrices, accuracy plots, and loss curves.

Codecademy: R projects, [Link](#)

- Applied R programming skills to real-world data analysis tasks.
- Conducted data manipulation, visualization, and statistical analysis using R.
- Worked with datasets to create plots and draw insights.

InstaInsights: Analyzing Top Instagram Influencers, [Link](#)

- Imported and managed large datasets for influencer analysis.
- Cleaned and transformed influencer data, including conversion of key metrics (followers, likes) for accurate analysis.
- Analyzed top influencers and countries, and calculated engagement metrics such as the ratio of average likes to followers.
- Created and interpreted visualizations (e.g., bar charts, pie charts, scatter plots) to uncover and present key trends and relationships.

Certifications	
Applied Artificial Intelligence, Beirut Arab University	Nov 2024
Introduction to Big Data with PySpark , Codecademy	Oct 2024
Learn R , Codecademy	Oct 2024
Microsoft Certified: Azure AI Fundamentals , Microsoft	Jul 2024
IT Specialist – Python , Certiport – A Pearson VUE Buisness	Jul 2024
Google Cloud AI Track , Google Cloud	Jun 2024
Generative AI Overview for Project Managers , Project Management Institute	Mar 2024
IC3 Digital Literacy Certification, Pearson VUE	Dec 2023
Cisco IT Essentials, Cisco	Sep 2022

Core Competencies	
Programming Languages	Python, R, SQL, JavaScript, Bash, Java
Big Data Technologies	Spark, Databricks
Machine Learning	PyTorch, TensorFlow, Scikit-Learn
Data Analysis and Manipulation	Data Visualization, Pandas, NumPy
Cloud Platforms	Google Cloud Platform, Azure Cloud Platform
Natural Language Processing	NLTK, SpaCy, GPT
Computer Vision	OpenCV, YOLO
BI Tools	Tableau
Version Control	Git, GitHub, Git LFS
Interests	Writing, Reading, Photography, Design, Sport, Psychology, Internet Of Things, Android
Languages	French, English, Turkish, Arabic

Education	
BEIRUT ARAB UNIVERSITY, Tripoli, BACHELOR OF SCIENCE in COMPUTER SCIENCE, Sep 2020 – May 2024	

Achievements	
--------------	--

Perfect Score on IC3 GS5 Key Applications Exam

Scored 1000/1000

- Achieved a flawless score on the IC3 GS5 Key Applications exam, reflecting advanced proficiency in key software applications critical for data analysis. This accomplishment demonstrates exceptional skills in data management, analysis, and presentation, vital for data-driven decision-making and reporting.

Extracurricular Activities

Accenture North America

Data Analytics and Visualization Job Simulation on Forage - August 2024

- Completed a simulation focused on advising a hypothetical social media client as a Data Analyst at Accenture
- Cleaned, modelled and analyzed 7 datasets to uncover insights into content trends to inform strategic decisions
- Prepared a PowerPoint deck and video presentation to communicate key insights for the client and internal stakeholders