Sarah Mohammed Aljudaibi

Data Scientist

LinkedIn: [Sarah Aljudaibi](#) **Email:** sarahal.jodaiby@gmail.com **Phone:** 0567525888

Summary

Data Scientist with a strong background in developing machine learning solutions, specializing in large language models, computer vision, time series forecasting, and predictive analytics. Proven experience in building scalable data pipelines, conducting exploratory data analysis, and applying advanced analytics to solve real-world problems. I have experience in working in environment, real estate, construction, and Transportation. Currently working on AI-driven projects at Artificial Intelligence Global Company, implementing techniques such as object tracking, classification, sentiment Analysis, and translation. Proficient in Python, R, HTML, CSS, JavaScript, PHP and PostgreSQL. For data visualization tools Power BI, Plotly Studio, Tableau, GCP Looker, and other Python-based analytics libraries.

Experience

Artificial Intelligence Global Company, Data Scientist – (Jan 2025- Present)
- Extract and transform manual KPIs into automated data pipelines for dashboard or reporting.
- Build Dashboard using Power BI with database to show KPI's of Client
- Support and debugging in predictive analytics use case related industries field
- Developed a translator system using GenAI to support multilingual communication.
- Designed and implemented a GenAI agent system to analyze customer reviews for a transportation company and a RAG solution to provide accurate, database-driven responses to user queries.
- Developed & enhanced algorithms for CV systems used in safety-focused CCTV on construction & substation sites.

KFUPM Center for Environment and Marine Studies, Data Scientist – (Feb 2023- Jan 2025)
- Leading machine learning projects focused on environmental data and image recognition.
- Developed CV models for Red Sea fish classification using ARCEMS datasets.
- Design and develop a system by collaborating with a cross-functional team to streamline data and automate workflows, contributing to increased efficiency in the center
- Design & develop an interactive dashboard for data analysts using python tools to improve data visualization and decision-making efficiency.
- Key organizer and member of the Scientific Committee for the 2024 center conference (GiKS)

Machine Learning Projects

"Mutarjim" using LLM: Translate document from Arabic to English and vice versa.
- Designed and built an end-to-end pipeline that ingests PDF/Word files, extracts content and translates using LLM
- Fine-tuned the GPT-4o model on domain-specific corpora, raising translation quality, and style consistency.

Customer Journey using generative AI: Optimize customer experience with GenAI.
- Developed a sentiment model that can classify survey responses and social-media posts to measure traveler satisfaction at every journey touchpoint and be able to optimize the given services.
- Employed prompt-engineering and a multi-agent architecture to drive answer accuracy and reduce response latency.

Time Series Forecasting for Real Estate Valuation: Developed and implemented time series models to forecast real estate market trends, enabling data-driven valuation.
- Collected real estate sales data from 47 Japanese regions (2006–2019).

- Cleaned data, handled missing values, and converted categorical data to numerical form.
- Implemented ARIMAX and SARIMAX as baseline models.
- Built an LSTM to forecast prices with long-term dependencies.

Education

Master of Computational Analytics - *King Fahd University of Petroleum and Minerals (KFUPM)*
Graduate: 2022, GPA: 3.62
- Thesis: Application of Time Series Forecasting in Real Estate Valuation
- Key Projects: Bank Financial Model Classification

Bachelor of Computer Science - *Taibah University*
Graduate: 2019, GPA: 4.58
- Member of Students Council
- Undergraduate Project: Date-Fruit Quality Classification using Computer Vision

Associate Degree in Computer Programming & Information - *Taibah University*
Graduate: 2014, GPA: 4.83

Technical Skills

Programming Language: Python, R, SQL, HTML, CSS, Bootstrap, JavaScript
Machine learning Problems: Analytics, Computer Vision, Large Language Model, Generative AI, & Time Series.
Data Processing: Data collecting, data cleaning, image processing, Feature selection and extraction.
Model Development Libraries: TensorFlow, Keras, Scikit-learn, OpenAI, Ollama, Huggingface.
Visualization Libraries: Matplotlib, Seaborn, Power BI, Tableau, Looker
Tools and Platforms: AWS, Visual Studios, PyCharm, Notebooks, Google Products
Database: PostgreSQL, PHP, SQLite
Deployment: Streamlit, FastAPI

Soft Skills

- Teamwork
- Clear Accountability
- Continues Improvement
- Achievement Drive
- Time Management, Planning and Organization
- Empowerment
- Mentoring and Coaching
- Being Resource for others

Certificates

- **Aveva Tool Predictive Analytics for Clients,** Aveva 2025
- **Practical Application of Generative AI for Project Managers**, PMI 2025
- **Generative AI Overview for Project Managers**, PMI 2025
- **Generative AI with Large Language Models**, Coursera (DeepLearning.AI) 2024
- **Front-End Web development (HTML, CSS, Bootstrap)**, Udemy 2022
- **Data Scientist Immersive course**, General Assembly 2020
- **Digital Marketing Nanodegree**, Udacity 2020