| **Nabeel Ayyad**  **Senior Data Scientist**  nabeelayyad98@gmail.com  [LinkedIn](https://www.linkedin.com/in/nabeel-ayyad/)  [GitHub](http://github.com/nabeelbassam)  +962791078205  Amman, Jordan  **EDUCATION**  **Jordan University of Science and Technology (JUST)**  B.S. Computer Science  *Fall 2017 – Summer 2021*  **SKILLS**  **Programming Languages**  Python, Julia  **Data preprocessing & manipulation**  PyData stack (SciPy, Pandas, NumPy)  PySpark  **ML/DL Frameworks**  TensorFlow, Keras, Scikit-learn,  Xgboost, scikit-optimize, Turicreate  H2O.ai, Huggingface, LangChain, Llama-index  **Data Visualization & Dashboards**  Plotly, Matplotlib, Seaborn, PowerBI  **Cloud**  Azure, GCP, Databricks, CDSW  **Data Engineering and data pipeline**  Airflow  **Model Deployment**  FastAPI, AWS EC2, GitHub Actions  **Databases**  MySQL, Firebase  **Tools**  Streamlit, Docker, Git, MLflow  **CERTIFICATES**   * Azure Data Scientist Certified * Scalable Machine Learning using Spark * Dataiku ML practitioner Certified * DeepLearning.AI TensorFlow Developer Professional Certificate | **WORK EXPERIENCE**  **Deloitte - Senior Data Scientist**  *November 2023 - Present*   * Led a technical team to deliver an end-to-end GenAI project using RAG Architecture for Cognitive City in KSA. * Designed and implemented a GenAI use case for a Government entity in Dubai, enabling natural language interaction with tabular data and generative data visualizations. * Created end-to-end AI solutions for clients, including solution architectures for data science, data pipelines, and MLOps framework, with a focus on scalability, reliability, and performance. * Led a team to develop and deploy a GenAI assistant for demonstrations, enhancing our ability to showcase its innovative capabilities to clients.   **Medwrite.ai - Lead Data Scientist (Limited Contract)**  *August 2023 - November 2023*   * Designed and implemented an AI system using OpenAI LLMs to generate discharge letters for healthcare professionals, including a PII/PHI masking service for GDPR compliance and contributing to DPIA documentation.   **Multiple companies - Data Scientist (Part-time)**  *September 2022 - August 2023*  **CRDF Global**   * Collaborated with the CTO to shape and implement the company's AI strategy. * Led pilot AI initiatives, including invoice processing and budget forecasting projects. * Worked with finance and operations stakeholders to identify workflow issues and implement AI-driven automation for improved efficiency.   **Al Ghanem Group**   * Improved document organization and retrieval with advanced text clustering. * Automated document classification, reducing manual effort. * Utilized Large Language Models (LLMs) for efficient ERP system registration.     **Open Insights** - **Data Scientist**  *March 2021– August 2022*   * Collaborated closely with clients and product managers to understand operational and strategic requirements for data analysis. * Executed end-to-end projects using machine learning and data science to derive actionable insights from advanced analytics and big data solutions. * Developed and presented impactful proof of concepts (POCs) to guide decision-making and engage stakeholders effectively.   **E-marefa - Data science intern**  *January 2021- March 2021*   * Automated metadata extraction, indexing, web scraping, and data cleaning of Arab world research literature. |
| --- | --- |

**PROJECTS**

* **Cash Forecasting** | Open-Insights

Duration: two months

* Developed and executed a time series model for a financial company client to predict Cash Flow forecasts.
* Enhanced model performance through feature engineering and extraction using the Featuretools library.
* Employed MLflow for meticulous tracking, comparison, and evaluation of models, parameters, and outcomes.
* **Alerts classification** | Open-Insights

Duration: seven months

* Developed and implemented a classification-based system to detect false positive alerts for a major Middle Eastern bank, enhancing alert detection and operational efficiency.
* Developed an automated DataMart pipeline that consolidates data from diverse sources, enabling streamlined and efficient analysis.
* Scheduled the data pipeline using the CDSW scheduler and deployed multiple machine learning models, including XGBoost, neural networks, and Random Forest, as a REST API via CDSW.
* **KPI Insights via Big Data Analytics**| Open-Insights

Duration: Two months

* Developed analytical frameworks for feature extraction, effective KPI computation, confidence measurement, and lookalike modeling for a major telecommunications company.
* Collaborated with product owners to develop features that deliver impactful KPIs and actionable insights.
* Used Pyspark to build the framework to deal with large-scale data preprocessing.
* **Contract Renewal Prediction** | Open-Insights

Duration: Three months

* Assisted a leasing company in developing an ML framework integrated with their system, predicting contract renewal likelihood using CatBoost and Xgboost.
* Collaborated with a team of three data scientists to enhance a scalable framework, incorporating client input, domain research, and product manager feedback.
* Integrated a feature store and AutoML components into the framework for rapid scalability.
* **Automatic invitations registration |** Al Ghanem Group

Duration: Three months

* Streamlined Invitation Registration: Leveraged OpenAI's LLMs to automate the ERP invitation registration process, significantly reducing manual effort and time.
* Engineered Document Processing Framework: Developed a robust system to parse and optimize diverse documents for accurate LLM input.
* **MedWrite GO |** Medwrite.ai

Duration: Four months

* Led a small team in deploying our MVP and developed a Generative AI system utilizing Large Language Models (LLMs) to generate personalized discharge letters.
* Implemented the system as a REST API using Django, ensuring seamless integration and accessibility.
* Utilized an open-source model with PII masking on AWS to enhance security.
* Led the development of Data Protection Impact Assessment (DPIA) documentation, ensuring thorough compliance with data protection standards.