**SAMIA SAAD**  
[samia.fazal@hotmail.com] | [3477033009] | [[(19) Samia Saad, PMP, CSM | LinkedIn](https://www.linkedin.com/in/samia-saad-pmp-csm-7090b449/)] | [https://github.com/Samiafazal]

**Professional Summary**

Data professional with **6+ years of experience** in the data space and a **Master’s in Computer Science**. Passionate about **Machine Learning, Predictive Analytics, and Data Science**, with a strong background in **data wrangling, exploratory analysis, statistical modeling, and machine learning algorithms**. Experienced in deploying data-driven solutions to optimize operations and drive business growth. Adept at presenting insights to both technical and non-technical stakeholders. Seeking to leverage expertise in **Python, R, SQL, and cloud technologies** to solve complex business challenges.

**Skills & Tools**

* **Programming Languages:** Python, R, SQL
* **Machine Learning & AI:** Scikit-learn, TensorFlow, PyTorch
* **Data Visualization:** Tableau, Power BI, Matplotlib, Seaborn
* **Big Data & Cloud:** AWS, Google Cloud, Hadoop, Spark
* **Databases:** MySQL, PostgreSQL, MongoDB
* **Other:** Feature Engineering, Data Wrangling, Statistical Analysis, Predictive Modeling

**Projects**

**Banking Marketing Analysis**

**Goal:** This project focuses on predicting whether a client will subscribe to a term deposit based on various demographic and marketing-related features. The dataset comes from a bank's marketing campaign, where the target variable is whether or not the client subscribed to a term deposit (y). The goal is to build a machine learning model to predict this outcome using a variety of features such as the client's age, job, education, and past marketing interactions.

* **Skills & Tools Used:** **Python, Pandas, NumPy, Scikit-learn, Matplotlib/Seaborn**
* **Outcome:**

**Improved Targeting:** The predictive model helped identify potential customers likely to subscribe to a term deposit, allowing the bank to focus its marketing efforts more efficiently.

**High Model Accuracy:** The **Random Forest Classifier** achieved an **accuracy of 91.2%**, effectively distinguishing between customers who would subscribe and those who wouldn't.

**Enhanced Marketing Strategies:** The insights from the model provided **data-driven recommendations** for personalized marketing campaigns, improving customer conversion rates.

**Optimized Resource Allocation:** By targeting only the most promising leads, the bank could **reduce marketing costs** and improve its return on investment.

**Future Scalability:** The project laid the foundation for further improvements, including **deploying the model for real-time predictions** and integrating additional customer data for better accuracy.

**Wine Quality Prediction**

* **Goal:** Predict wine quality based on physicochemical properties of Portuguese *Vinho Verde* wines.
* **Skills & Tools Used:** **Python, Scikit-learn, Feature Selection, Classification Models**
* **Outcome:** Built a classification model with 85% accuracy, identifying key factors affecting wine quality.

**Breast Cancer Diagnosis with Machine Learning**

* **Goal:** Classify breast cancer tumors as benign or malignant using digitized fine needle aspirate (FNA) images.
* **Skills & Tools Used:** **Python, SVM, Random Forest, Feature Engineering**
* **Outcome:** Developed a diagnostic model with 95% accuracy, improving early detection of breast cancer.

**Income Prediction Using U.S. Census Data**

* **Goal:** Predict whether an individual’s income exceeds $50,000 based on demographic and employment attributes.
* **Skills & Tools Used:** **Python, Pandas, Scikit-learn, Data Imputation, Classification Models**
* **Outcome:** Built an income classification model, achieving an F1-score of 0.82 after handling missing values and balancing the dataset.

**Professional Experience**

**Technical Project Management at Internal Revenue Services July 2023 – Present**

* Managed large volumes of client ordering data across multiple systems, developed 10+ custom reports for leadership using Power BI and Service Now
* Coordinated delivery of key client deliverable, Monthly Status Report (MSR) and Weekly Status Report (WSR). The report summarized previous project team activities and showcased upcoming months’ activities.
* Prepared Integrated Monthly Schedule as part of the deliverable using MS Project.
* Data Analysis of 50+ large data sets using Advanced Excel functions, VLOOKUP and Pivot tables and created reports.
* Provided PMO support, prepared Deck, presented and delivered to the client on Ad hoc basis.
* Coordinated delivery of key client deliverable, Monthly Status Report (MSR). The report summarized previous months’ project team activities and showcased upcoming months’ activities.
* Driving SAP implementation projects or supporting enterprise scale SAP.
* Prepared Integrated Monthly Schedule as part of the deliverable using MS Project.
* Understanding and making several processes flows of the system.
* Updated the Playbook.
* Ordering product on KISAM and then transitioning to ServiceNow.
* Worked on Spend Plan and delivered CIO Brief Deck.

**Advanced Skills:** Data Reporting, Project Management, SAP, Data Analysis

**Data Scientist at Deloitte Consulting March 2022 – June 2023**

* Collaborated with stakeholders to identify business problems and opportunities for data-driven solutions.
* Developed statistical models to predict customer behavior and optimize marketing campaigns.
* Analyzed large datasets to identify key patterns and trends, informing product roadmap decisions.
* Designed and built data visualization dashboards to communicate insights effectively.
* Led cross-functional teams in agile development of data products.

**Advanced Skills:** Data Science (Modelling, Pattern Analysis), Data Engineering (ETL, Infrastructure), Python, Machine Learning, Data Visualization (Power BI, Tableau), Web Application Development, Business Stakeholder Collaboration, SQL, NoSQL, Agile Methodologies

**Data Engineer at Drew Marine,** *Whippany, NJ* **January 2018- March 2021**

* Designed and implemented ETL pipelines to integrate disparate data sources.
* Built scalable data warehousing infrastructure on cloud platforms.
* Developed data quality monitoring tools to ensure accurate and timely data delivery.
* Collaborated with data scientists and stakeholders to ensure data availability for analysis.

**Advanced Skills:** Data Science (Modelling, Pattern Analysis), Data Engineering (ETL, Infrastructure), Python, Machine Learning, Data Visualization (Power BI, Tableau), Web Application Development, Business Stakeholder Collaboration, SQL, NoSQL, Agile Methodologies

**Data Quality Intern at Critical Mention,** *Newark, NJ* **July 2017 – December 2017**

* Conducted 100+ Data Quality tests to validate Data Quality issues using Data Quality Scorecards, best practices, and tools.
* Managed 15 large databases using MS SQL Server, advanced queries and performed completeness checks.
* Improved sever operations by 15% using Nagios and Sensu to monitor and alert server issues.
* Resolved Data Quality issues and prepared 20+ DQ error reports by applying Data Quality rules to source systems.

**Advanced Skills:** Big Data, Databases, Data Analytics, Reporting, Statistical Analysis, Data Quality, Automated Testing

**Research Assistant at Montclair State University,** *Montclair, NJ*  **January 2017 – June 2017**

* Conducted research on Cloud Computing to improve Data Quality by 50% and resolve associated big data issues.
* Published a research paper on “how to preserve privacy in a federated cloud while performing different computations.”

**Business Data Analyst at SILKBANK,**  **April 2010 – October 2013**

* Defined 10 Data Quality business rules and developed 30+ scorecards to track Data Quality and reconciliation issues.
* Converted data from 3 legacy systems to the new system using ETL pipelines.
* Provided daily/weekly status updates to Project Team Stakeholders

**Advanced Skills:** Data Analytics, Reporting, ETL pipelines, Data Quality, Data Visualization, Dashboards, Data Mining

**Data Analyst at KALSOFT,**  **August 2008 – March 2010**

* Implemented data collection systems to optimize statistical efficiency and quality by 33%.
* Maintained 50+ datasets from primary or secondary data sources using SQL and T-SQL for data extraction.
* Analyzed 20+ datasets using advanced statistics and reported findings to the executive team.

**Education**

**Harvard Business School**- Executive Education in Business Analytics **2021**

**Master’s in science (MS), Computer Science,** *Montclair, NJ* **2019**

Montclair State University (MSU), Overall CGPA: 3.8/4.0

**Bachelor’s in science (BS), Software Engineering** **2007**

Bahria University

**Certifications & Additional Information**

* Project Management Professional (PMP) 2022
* Google Cloud Professional Certified (2022)
* AWS Solution Architect-Associate (Amazon AWS, 2019)
* ITIL V3 Foundation (ITIL, 2013)
* Google Analytics Individual Qualification (GAIQ) 2021
* Scrum Master
* Google Professional Machine Learning Engineer (In Progress)