Alaa Alghwiri

Pittsburgh, PA 15239 | Mobile: 001-330-389-2904 | alaa.alghwiri84@gmail.com | LinkedIn Account

Professional Summary

Highly skilled Data Scientist with a Ph.D. in Industrial Engineering and 7+ years of impactful experience. Proven expertise across the entire data science life cycle, from data acquisition to solutions deployment using both R and Python. Strong background in cutting-edge machine learning algorithms and statistical modeling at scale. Applied knowledge in Healthcare, Energy, Higher Education, and Transportation industries. Entrepreneur, teacher, researcher, effective communicator, and collaborative team player

Technical Skills

- Proficient in R and Python for robust data analysis, manipulation, and advanced modeling
- Strong skills in data cleaning, preprocessing, and feature engineering to ensure high-quality datasets
- Ability to tell compelling stories through data visualization and presentation using tools like Tableau and Power BI
- Demonstrated ability to apply machine learning algorithms for predictive modeling
- Sound knowledge of statistical methods and their application to draw meaningful conclusions
- Experience with both SQL and No-SQL databases for efficient data storage and retrieval
- Experience with big data analytics using Spark for scalable and distributed processing
- Proficiency in using Git and GitHub to manage codebase and collaborate effectively
- Experience working with cloud platforms such as AWS and DigitalOcean
- Skills in deploying machine learning models using tools like Docker

Education

PHD | August 2017 | University of Akron

- Major: Industrial and Systems Engineering | GPA: 3.83/4.00

Master | August 2014 | University of Akron

- Major: Industrial and Systems Engineering | GPA: 3.90/4.00

Bachelor | January 2007 | Jordan University of Science and Technology

- Major: Electromechanical Engineering | GPA: Good

Professional Experience

Senior Data scientist | University of Pittsburgh | School of Medicine | Full Time | July 2022 - Current

- Spearheaded the development of precise predictive models aimed at identifying patients susceptible to adverse events through the implementation of sophisticated machine learning methodologies
- Tested the effectiveness of a multifaceted Electronic Health Records (EHR)-based Population Health Management (PHM) intervention to improve evidence-based Chronic Kidney Disease care in high-risk patients (Utilized A/B testing, Mixed-effect modeling and Survival Analysis)
- Utilized mixed-effect models to analyze the impact of various factors on pain, fatigue, and depression, providing valuable insights into effective intervention strategies
- Designed and implemented a Tkinter application in Python to map addresses into the Area Deprivation Index (ADI), facilitating targeted resource allocation for underserved communities

Data Science Instructor | Simplilearn | Part Time | April 2023- Current

- Instructed comprehensive courses on Data Science, covering a spectrum from foundational mathematical and statistical concepts to advanced topics

- Designed and delivered engaging modules on data exploration, visualization techniques, and best practices, fostering a deep understanding among students
- Equipped learners with practical skills in applied Data Science and Machine Learning, utilizing both R and Python programming languages
- Received consistently positive feedback for adeptly balancing theoretical knowledge with hands-on learning experiences, ensuring students grasp concepts effectively and confidently apply them in real-world scenarios

C0-Founder and CTO | Algebra Intelligence | Part Time | January 2020 - June 2022

- Led a pioneering team in the creation and deployment of Taqtak, a cutting-edge mobile/web application revolutionizing real-time monitoring within energy plants, successfully secured $700,000 in funding in the pre-seed stage
    - Spearheaded the development lifecycle from conceptualization to execution, overseeing database architecture, backend infrastructure, and front-end design to ensure seamless integration and optimal performance
- Led the data science team in the conceptualization and development of an energy generation forecasting tool using advanced statistical analysis and machine learning algorithms

Data scientist | University of Pittsburgh | Office of the Provost | Full Time | October 2017- June 2022

- Utilized data analytics tools to automate generation of yearly/ad-hoc insights in static and dynamic formats, integrating data from multiple resources for students, faculty, and University data management
- Improved University of Pittsburgh's US News Ranking by 4 steps (between 2020 and 2021) using cutting-edge Machine Learning techniques including SVM, Random Forest, and Gradient Boosting
- Developed highly accurate (96%) predictive models for students' success indicators (retention, dropout, and graduation) across all 3 regional campuses, enabling proactive advising for students in need
- Collaborated with schools' deans to make data-informed investment decisions, optimizing the school's US News and World Report ranking
- Automated recurring reports on gender and ethnic group equity, as well as admission and financial aid reports for the office of the provost and deans

TRAINING & CERTIFICATES

- Computational thinking Using Python XSeries (**MITx Certificate**), March 2022
- Git Branches, Merges, and Remotes (**LinkedIn Learning Certificate**), October 2019
- Extending Hadoop for Data Science (**LinkedIn Learning Certificate**), November 2018
- Machine Learning A-Z: Hands on Using R and Python (**Udemy Certificate**) October 2018
- Mastering Tableau for Visualization (**LinkedIn Learning Certificate**), March 2018
- Advanced SQL for Data Scientists (**LinkedIn Learning Certificate**), July 2017
- Apache Spark for Big Data Applications (**LinkedIn Learning Certificate**), July 2017
- R Programming A-Z (**Udemy Certificate**), August 2017
- AWS Big Data Training (**LinkedIn Learning Certificate),** September 2017
- Certified Six Sigma Green Belt, **American Society for Quality** (ASQ), April 2017

Selected Publications

- Electronic Health Record Population Health Management for Chronic Kidney Disease Care: A Cluster Randomized Clinical Trial
- A clinician's guide to understanding and critically appraising machine learning studies: a checklist for Ruling Out Bias Using Standard Tools in Machine Learning (ROBUST-ML)
- Predictors of sleep quality among university students: the use of advanced machine learning techniques