Mustafa Aslamy

Data Scientist

📞 (862)207-9656   ✉ mustafa.aslamy96@gmail.com   📍 Lake Hiawatha, NJ   **in** www.linkedin.com/in/Mustafa-Aslamy

🌐 Website Link to Portfolio   ⬡ Github Link

EDUCATION

B.S. in Mechanical Engineering, minor in Applied Mathematics   Sep 2014 - May 2019

NJIT - Newark College of Engineering   Newark, NJ

- GPA: 3.1/4.0
- Relevant Coursework:
  - Linear Algebra | Calculus I, II & III | Engineer Probability & Statistics | Differential Equations | Partial Differential Equations | Vector Analysis

IBM Machine Learning Professional Certificate, Stanford Machine Learning Certificate, Duke Java Programming Professional Certificate, Harvard Intro to Computer Science, DevOps on AWS Certificate

EMPLOYMENT EXPERIENCE

Data Scientist   Apr 2022 - Present

Quality Homes LLC   Bloomfield, NJ

- Utilized **web scraping** techniques to gather real estate data from Zillow and Redfin public APIs, focusing on property prices in specific regions and square footage. Processed and analyzed this data using Python, employing TensorFlow neural networks for predictive modeling to enhance property cash flow by 6-8%
- Implemented **data cleaning** and maintenance procedures for financial data related to over 50 properties, encompassing various metrics such as monthly rental income, repair costs, maintenance expenses, and property taxes. Leveraged Pandas and Numpy for data processing and Scikit-Learn for developing linear/regression models to assess financial performance, resulting in a 12% increase in profit optimization
- Orchestrated accounting processes and generated comprehensive monthly financial reports for each property utilizing Tableau visualization tools, enabling clear insights into income, expenses, and overall profitability
- Integrated Microsoft SQL Server Integration Services (SSIS) to manage and query customer databases, streamlining data accessibility and ensuring compatibility for tax filing purposes. Leveraged SSIS for building high-performance data integration solutions, including extraction, transformation, and load (ETL) packages for efficient **data warehousing**.

Data Scientist   Apr 2020 - Apr 2022

Siemens Industry   Florham Park, NJ

- Utilized machine learning techniques and data science tools (Alteryx, Tableau Prep, SQL, Python) to analyze data from Siemens Industry's diverse portfolio of products and services, resulting in 6% increase in predictive modeling accuracy and operational efficiency across industrial automation solutions, digitalization tools, and software platforms.
- Collaborated with cross-functional teams to **acquire, merge, and mine data** from various sources such as Oracle DB, Tableau Data Source, and Salesforce, contributing to 5% reduction in data processing time across Siemens Industry's drive systems, motion control solutions, and industrial communication networks.
- Presented and explained performance KPIs derived from the analysis of Siemens Industry's products and services to stakeholders, facilitating informed decision-making and strategic planning, resulting in 10% increase in actionable insights and business value realization across energy management solutions, digital twin solutions, and comprehensive service offerings.

Data Scientist   Oct 2019 - Apr 2020

Advanced Solar Products   Flemington, NJ

- Led the extraction and transformation of solar energy production data using PVWatts, employing ETL (Extract, Transform, Load) tools such as Apache Spark and Apache NiFi, and imported it into Python utilizing Pandas and NumPy. Conducted in-depth analysis of solar energy production, consumption, and optimization, driving innovation and efficiency across ASP's product portfolio.
- Collaborated closely with **cross-functional teams** to acquire, process, and interpret data from various sources, including solar panel sensors, weather data, energy consumption patterns, and market trends. Leveraged Python (Pandas, NumPy), R, SQL, and ETL tools to inform strategic decision-making and product development initiatives.
- Utilized Excel's advanced functionalities, including **data manipulation** tools like PivotTables and VLOOKUP, to complement Python (Pandas, NumPy), SQL, and ETL analyses. This comprehensive approach facilitated seamless integration of diverse datasets from PVWatts, solar panel sensors, and market trends, enhancing data-driven decision-making processes and contributing to the continuous optimization of ASP's solar energy products.

KEY SKILLS

• Software Development Life Cycle • Natural Language Processing • Testing & Debugging • CI/CD • Performance Improvement • System Maintenance • Backup & Recovery • Client Servicing • Computer Architecture • Operating Systems • Networking Protocols • Data Capture • Data Mining • Predictive Modeling & Analytics • Data Warehouse • Data Pipeline • SAS • Programming • Process Improvement • Workflow Development • Model Deployment • Performance Validation • Data Science Tools • Data Sources • Statistical Analysis • KPIs • Model Evaluation • Iterative Data Exploration • Transform Data • Big Data • Hypothesis Testing • Resampling Methods • Database Architecture • Integrated Automation • Machine Learning Techniques • Big Data Technologies • Artificial Intelligence • Modeling Software • Data Analysis • Statistical Programming

TECHNICAL SKILLS

Languages: Python, Java, C, C#, C++, SQL, Javascript, HTML, R, Spark, Racket, Ruby, Matlab
Tools, Models, & Frameworks: Excel, Tableau, Scikit-Learn, Matplotlib, TensorFlow, Pytorch, NLP, Keras, Scipy, Apache Spark, PySpark, NoSQL, MongoDB, PowerBi, Visual Paradigm, PostgreSQL, Microsoft SQL, Transact-SQL, PHP, Oracle, CSS, React, Angular, Node.js, Express.js, Flask, Django, Visual Studio Code, Apache Hadoop, Scala, Git, Jenkins, Spring boot, Bootstrap
• **Cloud Computing Platforms:** AWS, Microsoft Azure, Databricks • **Container Software & Tools:** Docker, Kubernetes • **Operating Systems:** Linux, Unix •
Design Software 2D & 3D: Creo Parametric, SolidWorks, Ansys, AutoCAD

CERTIFICATIONS

• **Machine Learning with Python** | IBM | Coursera • **Deep Learning & Neural Networks Keras** | IBM | Coursera • **Computer Vision & Image Processing** | IBM | Coursera • **Deep Neural Networks with Pytorch** | IBM | Coursera • **Deep Learning Models with TensorFlow** | IBM | Coursera • **AI Capstone Project Deep Learning** | IBM | Coursera • **Introduction to Computer Science CS50** | Harvard University | Edx • **Java Programming: Arrays, Lists, Structured Data** | Duke University | Coursera • **Data Structures & Performance** | UC San Diego | Coursera • **Java Programming: Principles of Software Design** | Duke University | Coursera • **Java Programming: Build a Recommendation System** | Duke University | Coursera • **Programming Languages Part A,B,C** | University of Washington | Coursera • **Build a Modern Computer from First Principles: Nand to Tetris Part 1,2** | University of Jerusalem | Coursera • **Intro to Operating Systems** | Georgia Tech | Udacity • **Supervised Machine Learning: Regression and Classification** | Stanford | Coursera • **Database Management Essentials** | University of Colorado | Coursera • **Cryptography 1** | Stanford | Coursera • **Unix Workbench** | John Hopkins University | Coursera • **Computer Science: Algorithms, Theory, and Machines** | Princeton University | Coursera

ACADEMIC PROJECT

- **Real Estate Price Prediction** | Github Link
  - Created a real estate price prediction website using sklearn and linear regression for banglore home prices dataset from kaggle.com and python flask server. Worked on data load and cleaning, outlier detection and removal, dimensionality reduction, feature engineering, and k fold cross validation. Deployed machine learning model to production on amazon AWS ec2 instance
- **Chatbot Food Delivery System** | Github Link
  - Natural language processing project using Dialogflow to create the chatbot for a food delivery website. Analyzed text input and trained data to understand customer feedback/requests. The backend was developed with python and Fastapi, database created and managed with MySQL
- **Celebrity Image Classification** | Github Link
  - Created a website where you could upload images and see if they match any of the celebrity images in the existing database. Project implements Opencv library for face and eyes detection, wavelet transforms, model building(SVM, logistic regression, random forest), and Model fine tuning using gridsearchcv
- **Path Finding Visualizer** | Github Link | Web App Live
  - Built a React web applications for path finding visualizer. I used Dijkstra algorithm to find the shortest path from start to end node with any obstructions in the way. The algorithm works on the first in first out principle for queue data structures which is whatever node is visited first is dequeue first until all nodes are visited
- **Sorting Visualizer** | Github Link | Web App Live
  - Built a React application for sorting visualizer. An array is generated of random values each visualized as a bar with pixel height equal to array values. The array is sorted with merge sort algorithm which breaks array from middle index into sub arrays, sub array values compared and updated until final array shows smallest to largest values

ACTIVITIES AND PROFESSIONAL DEVELOPMENT

- **Research : Fuel Cells, Under Professor Dibakar Datta**
  - Compiled software with Linux terminal to run multi-core processes for input and data files and used LAMMPS and Ovito software to run simulations of atom association with extensive and intensive properties, and how they are impacted by force and movement
- **NJIT Drone Competition**
  - Worked with a team of four to build and solder drone with hardware, drone frame, and motors. Used Ardupilot software for GPS control of drone flight patterns
- **Theta Tau Fraternity - Service Chair**
  - Organized and participated in park cleanup at Weequahic Park with the help of 20 fraternity members. Assisted in Highlanderthon at NJIT to help raise money for the Children's Miracle Network Hospitals which is a non-profit organization that raises funds for children's hospitals