Mohammed-Taqi Jalil

512-696-1614 | mohammed.jalil.email@gmail.com
Austin, TX 78660 | https://www.linkedin.com/in/mohammedtaqijalil/
https://github.com/MohammedJalil/Projects

SUMMARY

Data-driven Computational Biology graduate with hands-on experience in data analysis, machine learning, and visualization. Proficient in Python, SQL, and Tableau, with a track record of developing predictive models and interactive dashboards that drive decision-making. Passionate about applying statistical insights to solve real-world problems across healthcare, education, and environmental data domains.

SKILLS

Languages & Tools: Python (Pandas, NumPy, Matplotlib, Scikit-learn), R (ggplot2, tidyverse), SQL (MySQL, MS SQL Server), Tableau, Power BI, Jupyter Notebook, VSCode, Azure DevOps

Data Expertise: Data Wrangling, Regression Analysis, Predictive Modeling, Business Intelligence, Exploratory Data Analysis (EDA), ETL, Dashboard Development, Statistical Inference

WORK EXPERIENCE

Data Engineer Intern

SNH AI, Austin, Texas June 2025 - Present
- Developed and configured data labeling workflows in Label Studio, enabling streamlined annotation of key features for upcoming AI agent models.
- Collaborated on the design and implementation of data transformation pipelines to standardize heterogeneous XML data into a unified "golden dataset" for training and evaluation.
- Created and validated XSD (XML Schema Definition) schemas to ensure consistency and integrity of incoming XML data formats across diverse data sources.
- Contributed to cross-functional discussions with ML engineers and product stakeholders to align data preparation efforts with model training needs.

Data Analyst Intern

JSoftUSA, Austin, Texas, **February 2024 – August 2024**
- Analyzed student attendance and performance data, uncovering trends that led to a 15% increase in platform adoption.
- Automated ETL workflows using SQL scripts, decreasing manual report time by 30%.
- Built Tableau dashboards to track performance and resource use, improving educator efficiency.
- Conducted regression analysis to identify student success drivers, enabling targeted academic interventions.
- Partnered with product teams to improve reporting tools in alignment with regulatory standards.

PROJECTS

Ambient Air Pollution Prediction Model
- Built machine learning models (Random Forest, XGBoost, k-NN, Lasso) to predict PM2.5 levels across the U.S.
- Performed EDA and feature selection; optimized models using cross-validation and RMSE metrics.
- Tools: Python, Pandas, Scikit-learn, Jupyter

RNA-Seq Gene Expression Analysis
- Performed RNA sequencing differential expression analysis using R and Bioconductor packages (DESeq2, tximport).
- Conducted gene-level normalization, quality control, and log fold-change analysis on high-throughput sequencing data.
- Identified statistically significant gene expression patterns across experimental conditions, enabling biological interpretation of genomic results.
- Visualized key findings using PCA plots, volcano plots, and hierarchical clustering heatmaps.

EDUCATION

The University of Texas at Austin, Austin, TX   May 2024
Bachelor of Science, Biology, Computational

Relevant Coursework: Data Science, Regression Analysis, Biostatistics, Statistical Modeling, Genomics, Data Mining, Mathematical Statistics