Ifraz Ahmed

Houston, TX

(678)-353-8995 | iahmed10@uh.edu | https://www.linkedin.com/in/ifraz-ahmed/ | https://ifraz93.github.io/

EDUCATION

UNIVERSITY OF HOUSTON, Houston, TX **Fall 2021**
C.T. Bauer College of Business
Master of Science in Business Analytics, GPA : 3.8/4.0
Courses taken: Quantitative Analysis, Basic Programming, Database Management Architecture, Research and Design, Business Analytics Frameworks
Awards: Dean's Award for Academic Excellence

GEORGIA STATE UNIVERSITY, Atlanta, GA **Spring 2018**
J Mack Robinson College of Business
Bachelor of Business Administration in Computer Information Systems
Courses taken: Game Theory, Corporate Finance, Risk Management, Intermediate Accounting, Marketing
Study Abroad: Computer Information Systems Study Abroad Program at Université de Nantes, France

EXPERIENCE

Internal Revenue Service, Houston, TX **June – August 2021**
Data Science Fellow
- Improved existing processes and methods for aggregating and using operational data to conduct research and support high priority IRS operations.
- Engineered ETL workflow by creating a scp connection between Mercury Server and EdgeNode on the Hadoop Cluster.
- Designed documentation using white paper methodology for easy replicability of SparkR environment.
- Improved processing times and increased stability of R code through the implementation of SparkR.
- Performed transformations and actions on SparkR DataFrames for statistical analysis, validation, program stability, and performance improvements.

University of Houston, Houston, TX **December – May 2021**
Instructional Assistant
- Provided supplemental instruction for MBA and MS students in Quantitative Analysis.
- Assisted students in understanding and internalizing fundamental concepts in Quantitative Analysis through one-on-one and group tutoring sessions.
- Collaborated with professors to highlight key topics discussed during class for supplemental instruction.

Harris County Public Health, Houston, TX **March – August 2019**
Business Intelligence Data Warehouse Intern
- Developed Microsoft Power BI reports visualizing at risk populations within Harris County based on demographic factors such as race, household income, and educational attainment.
- Created scalar functions and stored procedures in SQL Server to validate data contained within Power BI.
- Developed a Python script to extract veterinary public health data into Python, apply transformations for data clean up, and upload back for use within data warehouse environment.
- Leveraged four fuzzy logic algorithms to give a confidence rating on how close an input matches a record.
- Presented insights and conclusions to key stakeholders and executive directors within Harris County.

PROJECTS

Frameworks and Methods, University of Houston **December 2020**

Predicting the Likelihood of Success in Cross Selling Insurance Through Artificial Neural Networks

- Imported, preprocessed, and split original dataset into training and test sets using Python.
- Performed exploratory data analysis through descriptive statistics, transformation, and visualization.
- Built a Multi-Layer Perceptron model with backward propagation, L2 regularization, and sigmoid activation.
- Compared model results to actual results with 93% accuracy.

Database Management Tools, University of Houston **October 2020**

Asia and North America COVID-19 Analysis using HBase and Hive

- Collaborated with three other students to perform analysis on countries in Asia and North America exceeding 50,000 cases after March 15 to determine leading factors in cases and deaths related to COVID-19.
- Ingested data into HBase, created an external table using Hive, and connected data contained within HBase to Hive using Hive API.
- Determined the strongest indicators of success for countries are timely government response, strong healthcare infrastructure, widespread testing, and strategic communication in contrast to population density, average age and pre-existing conditions.

Quantitative Analysis, University of Houston **May 2020**

Data Mining Project for Restaurant Profitability using R

- Ingested and preprocessed data for performing analysis using R.
- Created simple and multiple linear regression models using R to identify key metrics client must focus on to increase customer volume (table turns).
- Discovered advertising, ads, price, and rating have a high impact on predicting table turns in the regression model while year, days in business, parking, and cuisine have a low impact.
- Validated the regression model by predicting seven out of ten restaurants with the highest table turn metric correctly.

Digi Safari (Big Data Technology Learning Center), Alpharetta, GA **January – July 2018**

Cloudera Hadoop and Spark Developer Trainee

- Transformed data in Spark leveraging SQL Context and Hive Context APIs to discover profitable variables.
- Performed import and export operations between HDFS and MySQL using Sqoop.
- Ingested data into Spark in Avro, Parquet, JSON, CSV, TSV, and text file formats, transformed into DataFrame, performed SparkSQL queries, and saved back on to HDFS while applying Gzip compression.
- Designed external tables in Hive, ingested JSON formatted files from HDFS, and ran queries using HQL.
- Leveraged RDD APIs to calculate aggregations of orders inside spark-shell environment.
- Installed, configured, and tested Hadoop Ecosystem inside VirtualBox and VMware.

TECHNICAL SKILLS

- Scala
- Python
- R
- Unix/Linux
- HDFS
- Sqoop
- Spark
- Hive
- HBase
- SparkR
- SQL
- Quantitative Analysis

ADDITIONAL

- President of Bauer MBA Society
- President of Consulting and Technology Club
- Judge for 2020 Science and Engineering Fair
- Notary Public for the State of Texas
- United States citizen