Sowon Ham . Jarvis Consulting

I am a recent graduate from the University of Toronto - St. George Campus that studied statistics and mathematics for my Bachelor's degree. During my studies, it focused mostly on regression models, A/B testing, hypothesis testing, design and analysis of experiments, and a few machine learning methods. As I have been performing elementary to advanced data and business analysis on smaller data sets in my degree, to get further involved in this industry on an organizational level scale I have joined Jarvis as a data engineer to pursue my passion. Jarvis uses the agile framework implemented using scrum for the project and team management. My team and I have been holding daily scrums and biweekly scrum retros to discuss daily tickets to complete and sprint backlog for the next sprint respectively. Additionally, I deployed Minimum Viable Product (MVP) and a Proof of Concept (PoC) for the Jarvis team (Linux Cluster Administration team) and their clients (London Gift Shop) respectively by using the most recent and growing technologies in the data and business field such as Docker, SQL, Pandas, NumPy, Seaborn, Jupyter Notebook, Linux, etc. I strive to become a professional data/business analyst or any related profession to deploy products and reports of the utmost quality that can satisfy the client and easily be managed by other engineers.

Skills

Proficient: RDBMS/SQL, Pandas, NumPy, Matplotlib/Seaborn, Jupyter Notebook, R/RMarkdown, Microsoft Excel, Statistics (A/B Testing, Hypothesis Testing, Time Series Analysis), Statistical Modelling (Regression and Classification Models), Linux/Bash, Agile/Scrum, Git/GitFlow, Docker, LaTeX, AutoIt, Experimental Design

Competent: Python, Tidyverse, Julia, Java, Stochastic Processes, Predictive Forecasting

Familiar: Machine Learning (Neural Network, Cluster Algorithms), Cloud Networking, Virtual Network Computing (VNC), Verilog, C#

Jarvis Projects

Project source code: https://github.com/jarviscanada/jarvis_data_eng_demo

Linux Cluster Resource Monitoring App [GitHub]: Developed a minimum viable product for the Jarvis Linux Cluster Administration team for monitoring the resource usage and hardware specifications for all nodes. The application utilizes bash scripting to automatically collect hardware specifications from each cluster node. Additionally, the recording of the resource usage is automatically collected every minute into a relational database management system, PostgreSQL which Docker was used to setting up this container, using CRONTAB on Linux. Lastly, the project consisted of querying a few elementary data/business questions using SQL with the resource usage collected in our database.

Python Data Analytics [GitHub]: Performed data analysis on an e-commerce company using the data they provided through ETL processing which then was loaded into our PostgreSQL data warehouse. To grasp a basic understanding of the data set we used SQL commands on IntelliJ as a PostgreSQL CLI tool to perform elementary data analysis. In addition, to provide the e-commerce company with the reports and diagrams the Jarvis data analytics team used Jupyter Notebook as a primary platform. The connection between Jupyter Notebook and the data warehouse was established using Docker network. Furthermore, we used several packages in Python such as Pandas, NumPy, Matplotlib, etc. to give the client a meaningful analysis that can help them improve their revenue. Overall, this project identified different groups of customers based on their transactional data to increase this e-commerce company's profitability and efficiency.

Highlighted Projects

Computer Automation: - Implemented a complex and lengthy set of instruction for the computer to perform using Windows API as a base to diminish labour. - Researched and incorporated image and pixel searching tools to precisely determine which stage of instructions the computer is performing to reduce automation time by 30%. - Constructed separate executable for different stages of the automation to cut down the testing time and cost by a significant amount.

Data Analysis on Computer Automation: This project based on the computer automation that I performed on the above project. - Designed and performed a replicated exponential factorial experiment to abolish the ambiguity and precisely determine the efficiency of two contrasting scripts. - Enhanced the computer automation efficiency by 10% through utilizing the statistical conclusions written from the experiment.

Non-linear Regression Model with a Neural Network: - Implemented a fully-connected neural network (multi-layer perceptron) with 10-dimensional hidden layer with a trigonometric non-linearity. - Performed gradient-descent using

automatic differentiation to estimate parameters given the batch-size, model variance and learning rate. - Adjusted and improved the neural network to learn both the mean and variance of the Gaussian model.

Professional Experiences

Data Analyst Consultant, Jarvis Consulting Group (2021.12-Present): - Provided the Jarvis LCA team with a MVP that collects hardware specification and resource usage on nodes using Bash scripting and store the data into a RDBMS, PostgreSQL. - Query the database several times to answer business related questions and observe the data. - Collaborated with teams to solve problems that were rather not trivial.

Volunteer Work, St. Paschal Baylon Church (2016.05-2016.09): - Enthusiastic to learn new methods to quickly incorporate it into the working environment. - Immensely communicated with other crew members to work efficiently and not overlap cleaning the same areas. - Flexible with abrupt changes in work schedule to accommodate and prepare for any uprising matters.

Education

University of Toronto - St. George Campus (2017-2021), Honours Bachelor of Science, with Distinction, Statistics and Mathematics

Miscellaneous

- Badminton Player
- PC Gaming
- Build automation scripts for convenience