Omar Syed . Jarvis Consulting

I graduated from Ryerson University with a Bachelor's of Science (Honors) degree in Computer Science and a minor in Finance. During this time, I gathered project experience with a large varity of software tools and programming languages such as, Python, Java, PHP, C/C++, HTML/CSS, JavaScript, SQL, Prolog, MATLAB and much more. In university, I specialized in Data Engineering with courses such as Data Structures, Algorithms, Database Systems, Probability and Statistics and Data Mining. With a vast amount of experience in group projects, I leanred how to effectively collaborate and efficiently lead my team to desired goals. I joined Jarvis Consulting in July of 2022, where I worked on projects revolving around key concepts such as Computer Networks, Dockers and containers, Linux bash scripting, PostgreSQL, DDL, Python Data Analytics, Pandas/Numpy, Data Warehouse, Hadoop and more. Along with extensive technical experience, Jarvis also provided me with experience working under an Agile framework. The software and data industries excite me most for one simple reason, the extent of its reach and effect on our real and virtual world.

Skills

Proficient: Python/Pandas/NumPy/Scikit-Learn, Java, Spyder, Linux/Bash, RDBMS/SQL, Agile/Scrum, Git, MS Excel

Competent: Hadoop/Hive, MATLAB, HTML/CSS, Visual Studio, Docker, Oracle SQL, JavaScript, C++, Jupyter Notebook

Familiar: ZenHub, Zeppelin, PHP, yaml, Tableau, Google Cloud

Jarvis Projects

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

Cluster Monitor [GitHub]: Developed an application to collect and send real time resource usage data as well as hardware specifications on the Linux host to a PostgreSQL Database. Application system is built on Linux using bash scripts and PostgreSQL, while docker is used to manage the container the application runs on and the database. Crontab is used to automate the collection of the real time usage data.

Python Data Analytics [GitHub]: Analyzed a company's customer transaction histories in order to gain an insight to customer behavior for the purpose of better marketing stateigies. Setup a data warehouse using PostgreSQL on the Jarvis Docker Container and populated it using the provided CSV file of customer transaction histories. Implemented Python Scripts using the Pandas and NumPy libraries for preforming data analytics. Analyzed customer behavior and noted points of improvements for the marketing team. Recorded and documented the Python script using Jupyter Notebook

Hadoop [GitHub]: Prcoessed data using big data tools such as Apache Hadoop to evaluate core Hadoop components such as HDFS, YARN, and MapReduce. Hadoop clustering was provisioned using Google Cloud with the use of 1 master node and 2 worker nodes.Data was queried to answer business questions using Apache Hive and Zeppelin Notebook.

Highlighted Projects

Data Mining Association Analysis [GitHub]: Implemented association analysis using Python to a dataset, provided as a CSV file, of customer transaction histories in order to use machine learning for relational discovery in the dataset. Association analysis distinguished strong rules present and followed within the transaction history to be used by sales team for optimized marketing.

Machine Learning [GitHub]: Analyzed and processed a breast cancer tumor dataset, using Python scripts, holding over 17000 datapoint with 569 tumor sample specificiations. Implemented 5 different classification methods, Gaussian Naive Bayes Classification, SVM (Support Vector Machine), KNN (K-Nearest Neighbor Algorithm), Decision Tree Classification and ANN(Artifical Neural Network), by using the Scikit-Learn to render a model that can predict with high accuracy whether a breast cancer tumor is benign or malignant. Accuracy was measured with a portion of the dataset being used for testing in order to compare the preformance of the classicaition methods.

Professional Experiences

Data Engineer, Jarvis Consulting (2022-Present): Developed applications and managed Data using PostgreSQL, Python(Pandas, NumPy), Hadoop, Hive, MapReduce and Bash scripting. Implmented and used software tools such as Google Cloud Platform, VNC(Virtual Network Computing), Docker, Jupyter Notebook and Zeppelin to complete the

assigned Data Engineering projects. The position provided me with first hand experience using Agile/Scrum framework in an evolving work enviornment to host and run daily scrum meetings as well as bi-wekkly sprint retrospectives.

Financial Treasurer, SRCC (2015-Present): Managed Finances for the charity and projected optimal fund raising oppunities. Analyzed and calculated charity financial trends using MS Excel as well as developing plans to maximum attainable growth.

Education

Ryerson University (2018-2022), Bachelor of Sciences (Honors), Computer Science - Dean's List (2020-2021, 2021-22) - Minor in Finance

Miscellaneous

- [Certification] Azure Data Science (In-Process)
- [Volunteer] SRCC(Syrian Refugee Camp Charity)(2015-Present)
- [Sports] Volleyball, Cricket
- [Hobbies] Casual Gaming, Hiking, Board Games and eating out