Tarek Djeddi . Jarvis Consulting

I am a Data Engineer at Jarvis Consulting Group, I am working in an Agile enviroment and I have worked on several projects where I used defferent tools like Python, SQL, Docker, Data Analysis, Bach Scripts, etc. I completed a Diploma in Data Science at Concordia University which I took due to my strong interest in data analysis and modeling such as":" visulazing data, cleaning and preparing data, anaylyizing data to make useful insigts, and creating various machine learning modele. I am a holder of a Master of Applied Science in Electrical and Computer Engineering whose thesis focused on modeling and simulating Digitally Controlled Microwave Components for High-Power RADAR Systems. I aspire to build a successful career as a data engineer or as a data analyst.

Skills

Proficient: Python, SQL/HiveQL, Data Analysis, Machine Learning, Juypter Notebook, Agile/Scrum, Git

Competent: Linux/Bash, PostgreSQL Database, Predictive Modeling, Hadoop, Zeppelin Notebook

Familiar: HTML, CSS, C++, Java, MATLAB, Google Cloud Platform

Jarvis Projects

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

Cluster Monitor [GitHub]: Developed a Linux Cluster Monitoring Agent which helps Jarvis Linux Cluster Administration (LCA) team to track hardware information and resource usage of each node in real-time, which are running on CentOS. The resource usage data were collected and saved in a PostgreSQL database automatically every minute, while the hardware specifications were assumed to be static, which means that they will be collected once only. Furthurmore, two tables were created using Data Definition Language (DDL) for hardware specifications and resource usage, SQL were used to solve some businies problems. Technolgies that been used to successfully develop the project are IntelliJ IDEA, Bash, Docker, PostgreSQL, SQL, Git, and GitHub.

Python Data Analytics [GitHub]: Developed a proof-of-concept (POC) project to help an online store that sells giftware called London Gift Shop (LGS) by visualizing and analyzing their data to understand their customer's behavior, which will help the LGS marketing team to make decisions in the hope of increasing their revenue since it did not grow in recent years. LGS retail data was stored in a PostgreSQL database and provisioned using a docker container, and then it was connected to the Jupyter Notebook to answer all the business questions. Recency, Frequency, Monetary Value (RFM) analysis were used to quantitatively rank and group their customers in order to identify the best customers and perform targeted marketing campaigns.

Hadoop [GitHub]: Apache Hadoop is used in this project to process and analyze the World Development Indicators (WDI) dataset, which contains around 21 million data points. This project helped us in evaluating the Hadoop ecosystem and how it interacts with a distributed system for data storage and processing. I provisioned the Hadoop cluster using Google Cloud Platform (GCP) DataProc service with 1 master node and 2 worker nodes. Apache Hive and Zeppelin Notebook were used to solve some business problems and analyze the data using HiveQL queries. Optimization techniques like Hive Partition and Columnar File Optimization have been used to speed up the execution time of queries and compare their performance.

Highlighted Projects

Deep Learning [GitHub]: I created two Stock Market Forecasting Models using Recurrent neural network (RNN) and AutoRegressive Integrated Moving Average (ARIMA). The NASDAQ data were gathered using Yahoo Finance API. The data were splitted into train and test data, and the test data were compared with the predictated data to specify the performance of the two models.

Predictive Modelling [GitHub]: We created an Image Classication Model on 400-bird speccies and tested the performance of our model, the model was created using Convolutional Neural Network (CNN). The Bird Species dataset was downloaded from Kaggle and Google Colab has been used to boost the model's training time

Unsupervised Machine Learning [GitHub]: We devoloped a model to cluster COVID-19 tweets using Natural Language Processing (NLP). The dataset were generated Twitter API, dimensionality reduction were used to visulize the clustered data in Two-dimensional space.

Regression [GitHub]: We used a Regression discontinuity design (RDD) method to visiliaze and identify the efect of COVID-19 In Quebec over three different events“:” The 20/3/2020 lockdown, The Reopening of Schools 31/08/2020, and The 25/12/2020 lockdown.

Algorithms [GitHub]: Nearest Neighbor Algorithm is used on a given list of bank locations, how much money each one holds and the time it would take to rob each one. The algorithm is separated into several functions which used together to return a list of bank ID's that could be robbed in 24 hour to make as much money as possible and then escape.

Professional Experiences

Data Engineer/Data Analyst, Jarvis (2022-present): Working with a team in an Agile environment. Developed a Linux Cluster Monitoring Agent to help the Linux Cluster Administrator team monitor the hardware specifications and resource usage. Visualized and analyzed London Gift Shop data to help them make decisions to increase their revenue. Used Hadoop to analyze big data, optimize query execution time, and gain an in-depth understanding of Hadoop key components such as HDFS, MapReduce, and YARN.

Fire Alarm Technician, Amike Fire Alarm Security Service (2020-2021): Assessing the performance of different systems, writing inspection and verification reports, and communicating with clients to understand their problems and solve them.

Web Desinger, Services VortexApp Inc. (2019-2019): Creating and designing Websites for clients, in English and French using HTML and CSS. Creating several one-page and multi-page templates. Set up VortexApp phones by enabling the necessary applications for the encrypted phones, and activating SIM Cards.

Education

Concordia University (2022-2022 Intensive mode), Diploma in Data Science,

Concordia University (2016-2018), Master of Applied Science in Electrical and Computer Engineering,
- GPA: 3.68/4.3

Ajman Univerity, United Arab Emirates (2012-2016), Bachelor of Science in Electrical Engineering/Telecommunication,
- Merit: 1st in cohort - GPA: 3.8/4

Miscellaneous

- Published a paper in IEEE and presented it in the 18th International Symposium on Antenna Technology and Applied Electromagnetics (ANTEM), Waterloo, ON,
- Playing soccer on the weekend (Barcelona fan)
- Mobile Gamming