Mitchell Acton . Jarvis Consulting

I graduated with a Master of Engineering degree from the University of Western Ontario and a Bachelor of Applied Science degree from the University of Waterloo. I am very interested in the field of Data Science and Engineering due to its wide range of applications, and I have professional experience in real-time data analytics, machine learning algorithms in Python, and object oriented programming in Java. Since starting with Jarvis, I have become comfortable working in an Agile/Scrum development cycle, have gained valuable experience with Linux, PostgreSQL database management, and Java. I'm looking forward to getting the opportunity to put these tools to use along with my existing data analysis and machine learning skills to make high quality software products with significant business impacts.

Skills

Proficient: Java, Python, JavaScript, RDBMS/SQL, Agile/Scrum, Git

Competent: Linux/Bash, Front End (HTML/CSS), Docker, Machine Learning (Python scikit-learn, Pandas, NumPy), Cloud Computing (MS Azure/IBM Watson), Computer Hardware/Assembly

Familiar: C++, C#, Neural Networks, Networking (TCP/IP), Computer Vision (Python OpenCV, PyTorch)

Jarvis Projects

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

Linux Cluster Monitor [GitHub]: Developed a Linux cluster hardware resource management tool in CentOS 7 intended to be used by Linux system administrators. The project includes a bash script to create and start a docker PostgreSQL container, and a SQL script to initialize the database and tables used to store system hardware and usage data. I wrote additional bash scripts to automatically collect the hardware and usage data, and persist the data into the PostgreSQL database provisioned with Docker. Finally, a crontab file was set up to execute the resource usage script every minute, and sample SQL commands were included to perform useful analytics on the data.

Core Java Apps [GitHub]:

- Twitter App: Under Development
- JDBC App: In this project I built an application using the Java Database Connectivity (JDBC) API to connect to a database and run queries to access the data. The app provides a tool for performing CRUD (Create, Read, Update, Delete) operations on the data within the database. The project build was managed using Maven, and the Postgres database runs within a Docker container.
- Grep App: I built this app to mimic the Linux command line grep function. The program first recursively searches through the provided directory to find files containing a user-defined keyword. The application reads through the files and outputs lines containing the search term to the provided output file. The project was completed using the IntelliJ Ultimate IDE, and various methods were written using the Java 8 Lambda Stream API where applicable. The project was packaged using Maven and the Maven Shade plugin, and uploaded to DockerHub as a Docker Image.

Highlighted Projects

Algorithm and Simulation Visualizer Website [GitHub]: Created an interactive website to host various sorting and pathfinding algorithm visualizers, as well as some creative simulations written in JavaScript. The algorithm visualizers include sorting algorithms such as bubble sort, insertion sort, selection sort, quick sort, and merge sort, pathfinding algorithms such as Djikstra's Algorithm and A* Search, with an included recursive maze generator. The simulations include an ecosystem and evolution simulator, a gravity physics engine, and a boids flocking algorithm with quadtree optimization.

City Neighbourhood Clustering Algorithm: Built an unsupervised machine learning tool in Python with Jupyter Notebook to cluster areas (postal/zip codes) in a given city based on the area's facilities (restaurants, green space, schools, etc.). The algorithm resulted in 5 general clusters of postal codes (suburban, urban, industrial, downtown, airport). Analyzing the data revealed the most popular facilities in these clusters (e.g. Suburban areas were high in schools, green space, and yoga studios). The tool's intended use was to pinpoint the best location for a potential business owner to open their business (A suburban area with no yoga studios would have a demand for a new yoga studio). The tool was built using scikit-learn, Pandas, NumPy, and matplotlib. The model was trained using data from Toronto, and tested on similar cities including Chicago, Boston, and London England, with yoga studios as a test case. All cities were found to have at least one neighbourhood with a demand for a new yoga studio, proving the feasibility of the tool.

Oil Refinery Energy Cost Prediction Model: Developed a large tool in MS Excel for compiling and analyzing process data for a large oil upgrader/refinery facility. Used various regression models on historic data to predict future energy costs. The tool was used to plan the annual energy budget for one of the largest oil and gas facilities in Canada. The tool consisted of 19 regression models, one for each process unit in the facility. When used to predict test data, the models performed with a minimum R-squared value of 0.88, which was a significant improvement over the previous model.

Professional Experiences

Software Developer, Jarvis (Sept. 2021 - present): Responsible for developing various data engineering and analytics tools using an Agile/Scrum methodology. Completed software development projects using tools such as Java, Linux Bash, RDBMS, SQL, Docker, and Maven, to create robust solutions to real-world business problems. Some examples of these projects include a Linux cluster monitoring agent, a grep-like application to recursively search directories for files containing search terms, and a JDBC app to perform CRUD operations against data in a PostgreSQL database running in a Docker container.

Product Development Engineer, Pharmasees Inc. (Apr 2021 - Sept 2021): Designed the hardware and software for a medical device used to image and count white blood cells non-invasively. The hardware aspect involved optical system design and simulation, and 3D modelling of components, using AutoCAD and Zemax OpticStudio. The software aspect involved developing a machine learning computer vision algorithm (Python: PyTorch, OpenCV, scikit-learn) to detect and count white blood cells in videos of blood vessels in real-time.

Process Engineer Intern, Shell (Jan 2018 - Aug 2018): Performed data analysis on process data from a large oil refinery. Supported projects team in quantifying results from energy saving projects, and built a tool to predict annual energy cost for the following year's budget. Worked on various multidisciplinary teams, gained experience in project management, giving presentations, and leading meetings.

Education

University of Waterloo (2014-2019), Bachelor of Applied Sciences, Nanotechnology Engineering - Graduated with Distinction - Winner of First Year Design Competition - Norman Esch Entrepreneurial Competition Finalist

University of Western Ontario (2019-2020), Master of Engineering, Mechanical and Materials Engineering

Miscellaneous

- Coursera IBM Data Science Professional Certificate (2021)
- Cooking
- Reading
- Guitar
- 3D Art
- Running
- SCUBA Diving
- Gaming