

Varun Kumar Vats . Jarvis Consulting

I am a graduate student from University of Windsor specializing in IOT and Machine Learning implementations. During my course of studies, I published a couple of conference papers in IEEE focusing on implementation and use cases in the healthcare industry alongside working on my thesis. The thesis focuses on reducing the average time spent in queue by patients. The work of research instigated the mentality of effectively understanding the problem to debug it in correct way possible. Currently, I am working with Jarvis to gain and expand my skills by working on various projects in different programming languages such as Linux, SQL, Java, and Python in an agile environment.

Skills

Proficient: Java, Linux, RDBMS/SQL, Agile/Scrum, Git

Competent: Python, Data Analytics, Data Warehousing, Tableau, Power BI

Familiar: Apache Spark, Apache Kafka, Scala, Hadoop, C++

Jarvis Projects

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

Cluster Monitor [GitHub]: Linux Cluster Monitoring Agent(LCMA) is used to monitor and collect data of various hosts machines connected via network. The data collected consists of hardware specifications and system performance and is further stored in Postgres SQL Database which is an RDBMS using Docker Containers. Bash scripts are used in this project for intialising PSQL container using Docker, and monitoring the hardware specifications and system performance data. This project also uses crontab to collect and generate real-time data and provide statistics for resource utilisation. Finally, the monitoring agent sends data to a database for storage and further analysis.

Core Java Apps [GitHub]:

- **JDBC App:** The JDBC application developed primarily in Java deals with an RDBMS database and data manipulation is done using OOP concept such as Data Access Objects (DAO), Data Transfer Objects(DTO), and various other interfaces that joins both the java programming side and the relational database. At the core of this application, are the abilities to perform create, read, update, and delete operations from a RDBMS in a dockerized container of a PSQL instance. Apart from Java implementations in the project, MAVEN, Docker, PostgreSQL, and CLI are all interconnected to produce this application. The heart of the project lies in the concept of JDBC API which allows Java programs to access database management systems , and perform operations from a Java program interface.
- **Grep App:** The grep application in Java simulates the grep command used in Linux Bash to search for a particular regex pattern from an input file. As a result, the program goes through several checks in main program to search for a pattern and once the pattern is found , it returns the ouput in the form of .txt file. This application is tested for an example file stored in root directory. The very famous Romeo Juliet pattern was used an input pattern in the Shaksphere file and output returns the pattern in the lines and store the results in an output .txt file. This project is build using MAVEN and deployment is done using Docker container.

Highlighted Projects

Machine Learning for C-Section Analysis: This project focuses on performing predictive analytics for determining sucess or failure of C- Section based on current medical condition such as blood pressure, heart rate, and respiration rate. The data is further processed and modelled using Unsupervised Machine Learning algorithms.

Machine Learning for Liver Disease Prediction: This project focuses on performing a comparative analysis of Liver Disease Prediction based on content of enzymes found in the body. The data is further cleaned and organized using standard data mining algorithms using Python standard Libraries. Finally, data modelling is performed and further comparison is done to determine the appropriate method for a particular Liver Disease Dataset.

Professional Experiences

Data Engineer, Jarvis (October 2021-present): Collaborating with an agile team following Scrum and Sprint principles. Developing applications using wide range of technologies ranging from Bash Scripting, Java, Python, Spark, Scala, to Linux inside dockerized environment.

Graduate and Research Assistant, University of Windsor (September 2017- October 2019): This project focuses on performing a comparative analysis of Liver Disease Prediction based on content of enzymes found in the body. The data is further cleaned and organized using standard data mining algorithms using Python standard Libraries. Finally, data modelling is performed and further comparison is done to determine the appropriate method for a particular Liver Disease Dataset.

Education

University of Windsor (September 2017-October 2019), Masters of Applied Sciences, Electrical and Computer Engineering - Gold Medal - University of Windsor - GPA : 3.9/ 4.0

APJAKTU (August 2012-August 2016), Bachelor of Technology, Electronics and Communication Engineering

Miscellaneous

- Spark and Python for Big Data with Pyspark(Udemy)
- Chair, Young Professionals(IEEE)
- Chair, IEEE: Organized several technical and non technical events
- Judge, MasseyHacks:Coding Hackathon for school students