Faiyaz Ahmad

Dearborn, MI, USA | 313-632-1280 | faiyaza@umich.edu | linkedin.com/in/faiyaz106 | faiyaz106.github.io

Education

MS in Data Science, University of Michigan-Dearborn **Grade:** 4.0/4.0 **Expected Graduation April 2024**
Coursework**:** Database Systems, Natural Language Processing, Artificial Intelligence, Applied Regression Analysis, Introduction to Big Data, and Deep Learning

Bachelors in Mechanical Engg, Motilal Nehru National Institute of Technology Allahabad **Jul 2014-May 2018**
Coursework**:** C programming, Numerical and Statistical Methods, Quality Engineering, Mechatronics, Automatic Control Industrial Engineering

Technical Skills:

Big Data Tools: Apache Spark, Hadoop, Apache Airflow, NiFi, Kafka
Database: SQL(Postgres), NoSQL(MongoDB)
Data Visualization: Matplotlib**,** Seaborn**,**ggplot, plotly,Tableau,Power BI
Cloud Tools: AWS , Google Cloud
Microsoft Tools: Excel, Word and Power Point
Project Tools: Git
Programming: Python, R, SQL, Matlab
IDEs: Atom, VS Code, Anaconda
Data Analysis: Scikit Learn, Nltk, Tensorflow, Keras, Pandas, Numpy, BeautifulSoup
Other Tools: Hypothesis testing, Statistical Analysis, Machine Learning, Deep Learning, Data Modeling, NLP

Professional Experience

Hero MotoCorp Limited, Gujarat, India **Jul'18- Mar'22**

Data Analyst

- Building data analytics solution, digitization, project management, drive sustainability, and process improvement project.
- Experienced in Data Analysis, Data Modeling, Data Normalization, Data Visualization, SQL, Python, Excel, and Tableau.
- Implemented a data acquisition and analysis system for vehicle acceleration testing, ensuring a 100% test pass rate before Shipment. Performed data modeling, and loading into local SQL database.
- Designed and executed end-to-end data pipeline for fuel tank leak testing comprising of data extraction, transformation, loading into PostgreSQL database, and built Tableau dashboard for real-time productivity and quality data visualization.
- Extracted and analyzed materials data from SAP for over 500 export model vehicle components using python for warehouse floor occupancy and infrastructure planning. Reduced the planning and execution time by 50%.
- Built the computer vision based system to reduce the rework of crankshaft assembly by 96% due to wrong assembly of component in engine assembly. Annotated the images and trained the convolution neural network with 93% accuracy.
- Developed and maintained an Excel-based report and dashboard to track the plant-level KPIs that is productivity, capex, indirect tooling costs, major project progress, and overall equipment efficiency on a monthly basis.
- Project coordinator for setting up vehicle assembly for 1000 vehicles per day. Involved in project planning and execution.
- Collaborated with cross functional team for successful installation and commissioning of 20+ production related Machines.
- Lead the implementation of data driven 10+ sustainability, and process improvement projects, saved cost by $20,000/year.

Academic Projects **Aug'22- Feb'23**

ETL Pipeline for Football Leagues Analysis and Prediction

- Built python script for data extraction through football API and stored in AWS S3 bucket. Orchestrated the data using Airflow.
- Performed the data modeling, data cleaning, data transformation and data loading into PostgreSQL database by using Spark.
- Built Tableau dashboard and trained the deep learning model for predicting upcoming matches with 60% Accuracy.

Real-time Streaming of Twitter data Sentiment Analysis

- Created streaming data pipeline by using the Twitter API, NiFi, Kafka, Spark Streaming, and MongoDB in AWS EC2 instance.
- Trained decision tree, random forest, and SVM classifiers for sentiment prediction using 1.6 million annotated tweets.
- Developed a Python script for a visual dashboard that displays real-time insights on a specific topic.

SQL Queries Performance Evaluation

- Extracted the IMDB data from SQL server. Built python script for data cleaning, and transformation.
- Performed data modeling for data ingestion into PostgreSQL database.
- Evaluated the performance of lookup queries with and without indexing. Indexing improved the lookup query time by 75%.

Text Summarization

- Developed the text summarization by using the concepts TF-IDF, Centroid, and Graph (text rank) on CNN News Dataset.
- Graph based approach is providing better central idea for given text and achieved the average 0.42 ROUGE-1 score.

Leadership and Achievements

- Raised 300$ and provided the 3 hours community services for Buildon UM-Dearborn club.
- Public Relation Coordinator in culrav-2015 fest at Mnnit Allahabad, increased the fest exposure through different media.
- Won Bronze Medal for completing 21 Kilometer half marathon in Vadodara in 2 hr 38 Minutes.
- Won Shining Project award in Hero Motocorp for developing the acoustic chamber for 10dB Noise reduction