KEIVAN MOKHTARPOUR, MSc

Montreal, QC♦ 438.722.5044 ♦ m.keivan73@gmail.com ♦ LinkedIn ♦ GitHub ♦ Homepage

SUMMARY

Dedicated and detail-oriented professional with hands on experience in creating, testing, and maintaining ETL pipelines and building machine learning models. Demonstrated expertise in identifying process improvement opportunities and executed projects within timely and budgetary constraints. Proved ability to collaborate well with team members and clients, to offer innovative solutions, and to rectify operational issues using problem-solving mindset to ensure timely deliverables.

TECHNICAL SKILLS

Programming
- Python
- SQL
- Linux
- C#
- .NET
- Git

Big Data
- Spark
- Databricks
- AWS EC2
- AWS Athena
- AWS S3
- AWS Kinesis

Machine Learning
- Scikit-learn
- Tensorflow
- Computer Vision
- NLP
- Classification
- Regression

Analytics
- Numpy
- Pandas
- Tableau
- Quick Sight
- Matplotlib
- Seaborn

EXPERIENCE

Software Developer **July.2021 – Present**

Polycontrols Inc., Montreal, Quebec

Collaborating closely with the engineering team in an agile environment to automate tasks through software solutions

- Building data pipelines to simulate process using Python, C# and SQL
- Using python multithreading, multiprocessing libraries to execute multiple tasks concurrently
- Collaborating with a team of scientists at the National Research Council of Canada (NRC) and performing exploratory data analysis to correlate manufactured products quality with relevant features
- Developing a machine learning strategy to optimize the process performance by 15% through feature engineering and hyperparameter tuning
- Presenting assigned tasks and outcomes to the manager on a weekly basis and documenting the work progress using Jira, Confluence and Monday.

Data Scientist **Jan.2018 – Jan.2021**

Concordia University, Montreal, Quebec

Worked with the engineering research team to design an image-based data pipeline to predict fluid flows behaviour

- Predicted the process performance using machine learning algorithms (**Regression**, **SVM**, **KNN**, **Random Forests**, **Neural Networks** etc.), for datasets with various levels of completeness
- Implemented production code for anomaly detection, as well as for the prediction of equipment performance
- Collaborated and coordinated works with other disciplines including electrical, chemical, and civil engineering.

DATA PROJECTS

Twitter Sentiment Analysis | Self-Guided Project

- Streamed data with Twitter API using Kinesis firehose (over 3 million rows of data) to analyze sentiment of tweets regarding Afghanistan situation
- Built an ETL pipeline using PySpark in Databricks for data preprocessing, modelling preparation and save clean data to S3 bucket
- Designed and trained a classification model using TF_IDF to predict the sentiment of tweets using the text blob library for labeling and dumped data with tags back into S3
- Implemented data analytics pipeline with Athena using SQL queries to connect database with QuickSight
- Converted data into actionable insights about the influence of the US government, prepared visualization and dashboard reports that interpret the volume of twitter posts plus sentiment analysis in different countries.

Web Scraping Booking.com & Data Analysis | Self-Guided Project

- Utilized web scraping techniques (Beautiful Soup and Selenium libraries) to extract Hotel information in Canada
- Implemented ETL pipeline injecting web data directly to MySQL database for data storage and wrangling
- Conducted data cleaning and data visualization with Matplotlib, Seaborn and Tableau. Compared the prices and services in different provinces and cities.

Instacart Dataset Analysis & Machine Learning | Self-Guided Project

- Trained, tested, and evaluated a model for anonymized data on customer orders over time to predict which previously purchased products will be in a user's next order
- Imported raw data from 6 sources into MySQL, and extracted, merged, preprocessed using Python, and Pandas
- Performed feature engineering, created machine learning pipelines using TensorFlow for predication
- Increased F1 score by 18% after engineering appropriate features using NumPy and SciPy.

EDUCATION

Applied Data Science and Big Data Diploma, 2021 | Toronto Institute of Data Science & Technology, Toronto
Deep Learning School, 2021 | Quebec AI Institute, Montreal
Master of Science Mechanical Engineering, 2020 | Concordia University, Montreal
Bachelor of Science Mechanical Engineering, 2017 | Polytechnic Tehran, Tehran, Iran

PUBLICATIONS

- K. Mokhtarpour, M. Jadidi, A. Dolatabadi, **Modal Analysis-Based Classification of Liquid Jets in Crossflow,** Atomization and Sprays Journal, 2021. (Linked)

- K. Mokhtarpour, M. Jadidi, A. Dolatabadi, **Dynamic Mode Decomposition of Elliptical Liquid Jets in Crossflow,** Transactions of the Canadian Society for Mechanical Engineering, 2021. (Linked)