Shayan Liaghat

shayanliaghat@gmail.com | Hamilton, ON, Canada | 289-527-5429

Personal Website: https://bit.ly/3oFh9Pn | **LinkedIn:** https://bit.ly/366s5ze | **GitHub:** https://bit.ly/3kQlfCl

Summary

· Data Scientist with +4 years' experience in engineering, data analytics, modeling, and building end-to-end machine learning web applications.

Technical Skills

· **Languages:** Python (Pandas, NumPy, SciPy, SciKit-Learn, Keras, TensorFlow, PySpark, PyTorch, SpaCy, NLTK, Seaborn, Plotly, Matplotlib), SQL (PostgreSQL, MySQL), Unix Shell Scripting.
· **Tools:** Git, API's, Flask, Django, AWS (EC2, S3, SageMaker, RDS, Elastic Beanstalk, Redshift, Lambda, Elasticsearch, CloudWatch), GCP BigQuery, Azure Synapse, Heroku, Docker, Kubernetes, Tableau, Jupyter Notebook, Google Colab, VS Code, pgAdmin, Jira.
· **Skills*:*** Exploratory Data Analysis, Data Wrangling, Data Cleaning, Data Visualization, SciPy Optimization, Feature Engineering, Feature Selection, Model Testing/Validation, Load Balancing, Time Series Analysis, Deep Learning, NLP, Transformers and BERT, Recommender Systems, Web Scraping, Version Control, Object Oriented Programming.

Professional Experience

Data Scientist - Branchy (contract) **-** Smart E-commerce Service Analyzer — August 2020 - Present
· Developing a full-stack e-commerce web application that can save the store owner's costs and increase their revenue. [https://branchy.ca/]
· Collaborating with a Senior Data Scientist in an Agile environment to push services into deployment using Django, Docker, and Elastic Beanstalk.
· Implementing Sentiment Analysis, NER using SpaCy, Topic Modeling using LDA, Keywords & Trend Identification, and Action Item Summarization.
· Implementing Ensemble and deep learning models for Text Classification.

Lead Design Engineer - National Steel Car Limited (NSC), Hamilton, ON. — April 2018 - Present
· Built an automation pipeline for inter-departmental data access/analysis that saved 10 hours of engineering time per week.
· Created an ARIMA model based on vehicle performance time-series data provided by our client for fuel consumption prediction.
· Generating technical and analytical reports to fully meet customer expectations.
· Work and communicate directly with clients as well as other departments such as Marketing, Statistical Engineering, and Manufacturing.

Research Assistant - Center for Advanced Micro Electro Fluidics (CAMEF), McMaster University, Hamilton, ON. — May 2015 - Feb 2018
· Developed a biomedical device for cell detection and performed A/B testing that ultimately increased the detection accuracy by 15%.

Projects

Tweet Detective
· An end-to-end web application using Flask, Docker, and Elastic Beanstalk by scraping tweets and providing in-depth insights about specific businesses that could save business owners hours of tweets reading and analyzing. [https://bit.ly/2XzqooB]

Movie Me
· An end-to-end Recommender System web application using PySpark, Flask, and Elastic Beanstalk using collaborative filtering with Spark's Alternating Least Squares (ALS) implementation. [https://bit.ly/38HW0it]

Telco Customer Churn Prediction
· Implemented data augmentation techniques such as Borderline-SMOTE and ADASYN to work with imbalanced target variable, and used Grid Search for hyperparameter tuning of the selected models that resulted in 11% increase in comparison to the base ROC-AUC score.

Professional Development

Python for Data Science and Machine Learning Bootcamp, on Udemy
· Applied Machine learning algorithms using SciKit-Learn library (Linear and Logistic Regression, SVM, K-Means Clustering, Random Forest, Decision Trees, XGBClassifier, Neural Networks). Implemented neural network models using Keras and TensorFlow using AWS EC2 instance. Familiar with Apache Spark, Hadoop for Big Data Analysis, and Optimization Techniques (Stochastic Gradient Descent).

NLP - Natural Language Processing with Python, on Udemy
· Worked with text and PDF files with Python, Regular Expressions, Text Classification, Analyzed Yelp restaurant review with Spacy for tokenization, parsing, NER and Lemmatization. Visualizing POS and NER with Spacy, Topic Modelling, Sentiment Analysis with NLTK's VADER module.

The Complete SQL Bootcamp, on Udemy
· Used complex SQL queries to manipulate and analyze various databases.

Deploying Scalable Machine Learning for Data Science, on LinkedIn
· Architecture design patterns for scalable systems, Machine learning models as services, Containerizing models, Kubernetes for container orchestration.

Education

Master of Applied Science - Mechanical Engineering - McMaster University, Hamilton, Ontario, Canada. — May 2015 - Dec 2017
Bachelor of Applied Science - Aerospace Engineering - Sharif University of Technology, Tehran, Iran. — Sep 2009 - June 2014

Honors & Awards

· Recipient of full scholarship for M.A.Sc. and B.Sc.
· Recipient of International Excellence Award (IEA) (2016 & 2017).