Mohamad Dolatshah

Location: Burnaby, Canada
Website: sfu.ca/ mdolatsh/
Email: mdolatsh@sfu.ca Phone: +1-604-710-1261
LinkedIn: linkedin.com/in/mohamad-dolatshah/

Skills

Programming Languages	Python, R, C, C++, CUDA, SQL, CQL, Java (basic)
Data Science	Spark, Hadoop, Hive, SparkSQL, Cassandra, Tableau, Trifacta, Keras, TensorFlow, Microsoft Azure, IBM Watson, Amazon Web services, Jupyter, Pandas , WEKA, MATLAB
Database Technologies	MySQL, Oracle, PostgreSQL, Microsoft SQL Server, SQLite, MongoDB, OrientDB, Neo4j

Work Experience

Data Scientist, Part-time *Mar 2018 - Sep 2018*
Traction on Demand, Vancouver, Canada

- Built data pipeline to extract and map different sources of information related to customers and employees resulted in 800% speedup in data collection.
- Designed machine learning models for customer lifetime value (LTV) analysis, Sales lead scoring, customer churn prediction, project assignment and etc, using **Pandas, Scikit, Tensorflow, Keras** in order to increase profitability and throughput.

Research Assistant, Full-time *Sep 2016 - Sep 2018*
SFU Data Science Laboratory, Vancouver, Canada

- Explored a variation of the typical active learning setting where a learning algorithm is able to interactively query the user to obtain the desired label for training data.
- Developed an intelligent system that would reduce the cost of manually obtaining labels with the estimation error up to 3× smaller than the existing methods resulted in publishing a paper titled Cleaning Crowdsourced Labels Using Oracles For Supervised Learning in **VLDB 2019** conference.

CEO & Founder, Part-time *Jun 2013 - Present*
Zirend.com

- **M**anage an online outsourcing marketplace which allows employers to post projects for freelancers. Created by **Django**(**Python** MVC web framework), Zirend lets anyone to post works to get done by thousands of active skilled users.

Research Assistant, Full-time *Jul 2015 - Jul 2016*
IUST Data Mining Laboratory, Tehran, Iran

- Proposed a new indexing method for data mining which decreases search time by more than 50% resulted in publishing a paper titled Ball*-tree: Efficient Spatial Indexing for Constrained Nearest-Neighbor in **ICMSCE** 2016 conference.

Founder & Software Manager, Part-time *Apr 2015 - Jul 2015*
Elexir

- Used **Neo4j** for handling user relations and **MongoDB** for document-based objects in a social network mobile app, resulted in cutting the access time in half compared to traditional RDBMS.

Database Manager, Full-time *Apr 2014 - Jul 2014*
Emersun Industries Co, Tehran, Iran

- Refined the pipeline for data transmission between two **Microsoft SQL Server** databases in one of the biggest manufacturer of home appliances in middle-east led to reducing the human effort.

Projects

- **Pattern Recognition in the 2016 Presidential Election:** Leveraged natural language processing (NLP) and pattern recognition in top candidates tweets along with statistical test of the **null hypothesis** in **Pandas**.
- **Re-indexing Model for Webpage Crawler:** Performed a **Monte Carlo** study for the Update Rate of webpages using **Poisson** distribution and maximum-likelihood estimation (MLE).
- **Spark Image Clustering:** Applied exploratory Data Analysis (EDA) and visualization in **matplotlib** along with implementation of the **k-means** algorithm on a real dataset using **Spark**(PySpark) on **Databricks** Cloud Computing framework.
- **Analyzing NASA Server Logs:** Examined the NASA web server logs using **HDFS** clusters in order to analyze the **Pearson** correlation between features in **SparkSQL**.
- **Deep Feature Selection and Classification:** Used ensemble **feature engineering** and cross validation in **Matlab** for a Deep Belief Network(DBN) resulted in %71 accuracy.
- **Customer Lifetime Value (LTV) Analysis:** Calculated LTV values along with training supervised learning models to predict important customers based on behavioral features using **KaplanMeier** estimator resulted in increasing profitability by 77%.
- **Churn Rate Prediction:** Conducted **Data Wrangling** on business order information in order to predict if a sales lead is going to be converted, by training a LSTM(Long short-term memory) neural network model based on **Weibull distribution**.
- **Customer Activity Forecast:** Predicted if a customer is going to be inactive for a window of time by feeding CRM cycle history to a **Recurrent Neural Network(RNN)** by 96% accuracy.

Education

Sep 2016 - Sep 2018 — Master's (M.Sc.) in **Computing Science**
Simon Fraser University **(SFU)**, Vancouver, Canada
Courses: Data Science, Machine Learning, Big Data Programming
Database Systems, Data/Knowledge-Base, Algorithm Design
GPA: **3.8**

Sep 2011 - Aug 2016 — Bachelor of Science (B.Sc.) in **Computing Engineering**
Iran University of Science and Technology **(IUST)**, Tehran, Iran
Awards: Golden Medalist in ACM contest
Outstanding BSc student for 2 semesters

Activities

November 2014 — **Conference Committee, Lecture Presenter**
PyCon 2014, Tehran, Iran
Organized the largest gathering for the Python community in country along with holding a workshop about **Django**.