Razzak Lebbai

Sr. Data Scientist - 8x8 Inc

Santa Clara, CA - Email me on Indeed: indeed.com/r/Razzak-Lebbai/6f07c4464ed03719

A Particle physics doctorate holder and a Data scientist/Research Scientist with overall Thirteen years of experience in data analysis and Machine Learning using numerical and statistical approaches in both business and scientific industries.

* Strong background in statistical data analysis and numerical methods
* Strong background in manipulating and analyzing complex big data (Terabyte/Petabyte scale)
* Strong background in data simulation, data visualization and data modeling
* Hands on experience in developing physics analysis software tools using C++/C, Python
* Proficiency in statistical data analysis packages Python, R and ROOT (comparable to R)
* Hands on experience in Distributed File System (DFS) using GRID
* Hands on experience in text mining, and NLP
* Hands on experience and depth knowledge on Supervised and Unsupervised Machine learning techniques
* Hands on experience in MySql (SQL), postgresql and Hive/Hbase

WORK EXPERIENCE

Sr. Data Scientist

8x8 Inc   -   San Jose, CA - March 2016 to Present

* Data mining and Data visualization using business and call logs
* Upsell customer prediction using Machine Learning methods
* User retention studies
* Customer Churn prediction using Machine Learning methods
* Call quality recommendation using Machine Learning
* Translate data into business-friendly visualization view using visualization tool Looker
* Setting A/B testing on predictive recommendation

Data Scientist (contract)

GE Digital - September 2015 to March 2016

* Data mining and Data visualization using business and sensor/GPS data
* Predict business decision models using Machine Learning method (random forest, svm, bayesian network, etc..) using both numerical and textual data (unstructured data)
* Use clusterization techniques for unsupervised text data
* Predict faulty reasons on GE locomotives (GE transportation) using Machine Learning techniques (random forest, svm, neural network, etc.) from both sensor and GPS data.
* Python (pandas, numpy, scipy, matplotlib, scikit, bayespy, fuzzywuzzy, ntlk) and R, GreenPlum database, spark/hadooop are being used in this project

Data Scientist

Cisco System   -   San Jose, CA - June 2015 to September 2015

* Multivariate performance prediction of switches and security hardware using Machine Leaning techniques from historical data using python (pandas, numpy, scipy, matplotlib, and scikit) and R
* Predict future hardware performance using Machine Learning techniques from historical and customer feedback data using python (pandas, numpy, scipy, matplotlib, and scikit) and R

Data Analyst/Scientist

DELL Inc   -   Santa Clara, CA   -   January 2014 to June 2015

* VMware and netapp storage performance analysis and prediction using Python (numpy, Matplotlib, scipy) and R
* Log analysis of storages (Netapp, Equalogic, compellent) and Linux/Unix servers
* Anomaly detection in Network traffic using k-mean clustering

Data Analyst/Scientist

IBM   -   San Jose, CA   -   August 2012 to January 2014

* VMware and IBM storage performance analysis and prediction using Python (numpy, Matplotlib, scipy) and R
* Log analysis of storages (Netapp, IBM) and Linux/Unix servers

Research Scientist/Assistant

University of Oklahoma   -   Norman, OK   -   July 2010 to June 2012

* Data analysis and data mining using ROOT (C++), Python and R
* Upgrade and maintain physics simulating packages for ATLAS software using C/C++ and Python
* Worked on data mining, modeling and analysis of complex, high-volume, high-dimensionality data with ROOT (C++) and Python
* Developed algorithms and numerical methods for investigating and measure of cut flow for SUSY, top cross section, and minimum bias analysis in ATLAS detector using ROOT (C++)
* Developed computer algorithms and strategies for ATLAS experiment using C/C++ and Python
* Used Multivariate Machine learning tools TMVA, scikit, and R for signal and background studies

Research Assistant/Scientist

CERN   -   Geneva, CH   -   June 2006 to July 2010

* Worked on data mining, modeling and analysis of complex, high-volume, high-dimensionality data with ROOT (C++) and Python
* Fully participated in building ATLAS pixel detector such as pixel sensor installation, opto-board installation, network connection, fiber connection, and installing back-end severs
* Data management (TB and PB scale) for data collected from the ATLAS detector for Minimum Bias and SUSY physics analysis
* Developed computer algorithms and strategies for ATLAS experiment using C/C++ and Python
* Developed Physics Analysis packages for data analysis using ROOT (based on C/C++ and Python)
* Used Multivariate machine learning tools TMVA, scikit, and R for signal and background
* Carried out pixel detector data analysis for detector tuning using ROOT and Python

EDUCATION

PhD in Experimental High-Energy Physics
University of Oklahoma
2005 to 2011