Saad Khan

Data Science | Machine Learning

saadkhan321@gmail.com | 571.420.1920 | saadkhan321 | saadakhan1 | saadakhan321

4251 Parklawn Ave, Apt 204, Edina, MN 55435

ABOUT ME

Seasoned Professional with more than 11 years of consulting experience in various wireless technologies. Extensive experience using stats and machine-learning to find and present useful insights in data with the aid of dashboards and visualizations. In addition to that, considerable knowhow about the broader field of artificial intelligence. Self-motivated, willing to take initiative to learn quickly and also good at public speaking.

TEHCNICAL SKILLS

PROGRAMMING LANGUAGES:	Python • R • Java • MySQL • HTML, CSS & Javascript (d3.js)
SOFTWARES:	MS Office • MS Visio • R-Studio • Jupyter Notebook • IntelliJ IDEA • Android Studio • Weka
PYTHON PACKAGES:	numpy • scikit-learn • pandas • ggplot • matplotlib • scipy • pymongo • PIL • pyretic
R PACKAGES:	ggplot2 • ggmap • dplyr • ggally • ggpairs • reshape2 • mclust • nbclust • clvalid • caret • tm
DATA ANALYSIS TOOLS:	Tableau • Alteryx • FME • SQL Server
DATABASE FORMATS:	XML • JSON • SQL
STATISTICAL SKILLS:	std. dev./variance • confidence level • hypothesis testing • significance tests (z-test, t-test, χ^2)
SUPERVISED LEARNING:	Decision trees • Logistic regression • SVM • K-NN • NLP • Neural networks • Linear regression
UNSUPERVISED LEARNING:	K-means Clustering • PCA • ICA • natural language processing • feature selection & transformation
REINFORCEMENT LEARNING:	Q-Learning • Value iteration • Policy iteration • TD-Learning • performance analysis

WORK EXPERIENCE

TELECOM TECHNOLOGY SERVICES INC. | Data Science & LTE Consultant (Under Contract to AT&T)

July 2012 – Present | Minneapolis, MN

- As part of the AT&T engineering team learned key requirements for wireless network design and spatiotemporal analysis of mobile data in the Upper Midwest Region.
- Main responsibilities include cellular design using Forsk ATOLL and AT&T internal databases and data manipulation and analysis using R/Python.
- Considerable troubleshooting skills in maintaining network design KPIs along with good understanding of design fundamentals.
- Ownership of generating 4G/3G Technical Approval projects for new sites & site modifications.
- Experience of creating python scripts to auto-populate network template and minimize down time for site data bulk creation.
- Leveraging data visualization features of R (ggplot2, ggmap, etc) to assist with analysis of crowd-sourced mobile data for intelligent business decision making.
- Considerable experience of performing basic statistical computation such as computing mean, median, variance, etc for mobile user data based signal level and throughput.
- Well versed with applying unsupervised learning techniques such as k-means clustering in R to geographically segment crowd-sourced mobile data based on network coverage & quality in order to identify potential areas for network improvement.
- Experience of processing and analyzing complex 4G network related data sets using advanced querying and visualizations with the help of data analysis tools such as Alteryx and FME.
- Ability to create presentations for AT&T Management based on 3G and 4G Optimized KPIs, Coverage Maps and visualization based on analysis performed using various data analysis tools.

T-FORCE INC. | LTE Design Consultant (Under Contract to Ericsson)

December 2010 – November 2011 | Plano, TX

- As part of the Ericsson's RF design team learned key design requirements for the AT&T LTE nationwide RF design.
- Main responsibilities included providing remote LTE RF design support to various AT&T Markets utilizing Forsk ATOLL.
- Performed audit & validation of the UMTS RF design project put forward by A&T as the baseline and determined suitable conversion to LTE for 10/5 MHz bandwidths as per spectrum availability.
- Proficient with ATOLL parameters, project set-up and design aid tools as Path Profile, Point Analysis, etc.
- Considerable troubleshooting skills in maintaining KPIs & good understanding of LTE fundamentals - link budget, PCI, KPI's, etc.
- Generated UMTS monte carlo simulations and KPI reports for specific AT&T markets to be used as reference for the optimized LTE RF design.
- Created KPI reports for specific AT&T markets to be used as reference for the optimized LTE RF design.
- Performed iterative optimized RF design to validate improvement in the KPIs such as RSRP, No. of best servers, RSRQ, SINR, throughput, etc.
- Generated KPIs and ran monte carlo simulations to show progress in the design work every week.

LCC INTERNATIONAL INC. | WiMAX Network Consultant (Under Contract to Clearwire)
December 2009 – December 2010 | Herndon, VA

- Main responsibilities included handling access performance issues through centralized support for various commercial Clearwire markets.
- Conducted FDD-LTE tests using 2.5 GHz spectrum, monitoring KPIs such RSSI, RSRP, CINR, MCS & throughput, using XCAL.
- Performed stationary & mobility baseline tests for performance critical indicators such has signal strength (RSSI), reference signal received power (RSRP), interference (SINR/CINR), MCS Usage, peak/average throughput, using XCAL (LTE + WiMAX).
- Developed good understanding of TCP/IP and UDP layer 3 messages for troubleshooting critical performance issues.
- Hands on experience of generating, analyzing & processing tabular/graphical data pertaining to RSSI, RSRP, SINR, throughput.

ABACUS CONSULTING | WiMAX Planning Consultant (Under Contract to Motorola)
February 2008 – December 2009 | Lahore, Pakistan

- Learned key basics of WiMAX RF design and optimization services provided by Motorola to Wateen Telecom.
- Worked on design/analysis of the WiMAX RF network based on coverage, interference & modulation using ATDI RF design tool.
- Contributed in detailed RF design pre-sales proposals/response to RFPs for different customers in Pakistan which included capacity dimensioning and detailed coverage analysis to meet RF requirements.
- Monitoring critical KPIs such as UL/DL utilization, dropped packet rate, total AP sector usage, DL/UL data and control channel utilization, total AP Sector Throughput, initial ranging failures, etc using Maxzilla (Motorola Internal Network monitoring tool).

ERICSSON PAKISTAN (PVT) LTD. | Services Engineer
March 2007 – February 2008 | Lahore, Pakistan

- Learned key basics of GSM RF and transport network design and optimization services provided by Ericsson to Warid Telecom.
- Performed drive tests at customer premises and suggesting nominal coordinates for new sites based upon signal strength (SS) and C/I plots extracted in MapInfo using TEMS Investigation.
- Maintenance of important documents such as microwave scope of work, MUX plan (Microsoft Visio), detailed RBLT assignment (BSC termination points) diagrams, etc.

PROJECT EXPERIENCE | DATA SCIENCE / MACHINE LEARNING

Linear Regression to Predict Gross Revenue for Movies
April 2017 | GEORGIA TECH - Masters in Computer Science — Language: R

- Linear regression in R was applied to selected, transformed features of the movies database for movies released after the year 2000.
- Both Numeric and Non-numeric features along with there interactions were tried out to improve the RMSE for gross revenue.

Logistic Regression to Predict MNIST digits
March 2017 | GEORGIA TECH - Masters in Computer Science — Language: R

- R was used to implement logistic regression from scratch to predict MNIST digits by comparing performance for training and test sets.
- Further performance improvement was conducted by tuning parameters such as learning rate, initial θ's, convergence criteria, etc.

Rule-Based and Machine Learning based Stock Market Trader
November 2016 | GEORGIA TECH - Masters in Computer Science — Language: Python

- Objective was to implement rule-based and ML based trading strategies in python and compare performance for a training and test periods.
- Strategies were built using technical market indicators such as relative strength index, momentum, moving average, bollinger bands, etc.

Analyzing Random Tree Learner with and without Bagging
October 2016 | GEORGIA TECH - Masters in Computer Science — Language: Python

- Decision Tree classifier and bagging API were implemented in python with tree and node decision based on randomness.
- Extent of overfitting was analyzed with and without bagging in terms of RMSE and correlation coefficient for the wine dataset.

Correlated-Q Learning and General-Sum Markov Games
July 2016 | GEORGIA TECH - Masters in Computer Science — Language: Python

- Python was used to create experiments for Friend & Foe-Q, Correlated-Q and Q-Learning when applied to 2-player zero-sum games.
- This was bsed on the paper published in 2003 by Greenwald & Hall and results replicated were for the soccer game experiment.

Temporal Difference Learning and Random Walk Experiment
June 2016 | GEORGIA TECH - Masters in Computer Science — Language: Python

- Random Walk experiment from Richard Sutton's 1988 paper on Temporal Difference methods was implemented in python.
- Two different settings for the experiment were setup and the best performing ranges of λ values were highlighted.

Markov Decision Processes
November 2015 | GEORGIA TECH - Masters in Computer Science — Language: Java

- Objective was to run value iteration, policy iteration and Q learning on 2 different MDPs, (a) with less no. of states, (b) with larger no. of states.
- Used Brown-UMBC Reinforcement Learning and Planning (BURLAP) based on Java to create and run all the experiments.

Unsupervised Learning & Dimensionality Reduction

November 2015 | GEORGIA TECH - Masters in Computer Science — Language: R — Tool: Weka

- 6 algos were implemented; 2 clustering: k-means, Exp. Maximization & 4 dim. red. algorithms: PCA, ICA, Rand. Projections and InfoGain.
- R and Weka were utilized for applying clustering and dimensionality reduction 2 interesting data sets from the UCI repository.

Analyzing Supervised Learning Algorithms

September 2015 | GEORGIA TECH - Masters in Computer Science — Tool: Weka

- 5 algorithms; decision trees, neural networks, boosting, support vector machines and k-nearest neighbors were evaluated against each other.
- Weka Explorer and CLI were utilized for applying classification algorithms to 2 interesting data sets from the UCI repository.

Data Visualization using D3

August 2015 | Udacity - Nanodegree — github visualization — Language: Javascript (d3.js)

- Generated data visualization using javascript libraries such as D3.js and dimple.js that are used for manipulating documents based on data.
- Visualization included elements of interaction and animation for representation for user to have a better understanding of the data.

Predicitions with scikit-learn

June 2015 | Udacity - Nanodegree — github report — Language: Python

- Created a person of interest identifier in python using resources from the Scikit learn library.
- Utilized machine learning to build an algorithm to identify Enron employees who committed fraud based on public Enron financial/email data.

Data Analysis with R

May 2015 | Udacity - Nanodegree — github markdown — Language: R

- Using R, applied exploratory data analysis to find relations between variables of white/red wine data taken from samples of Vinho Verde wines.
- Analysis was carried out with relations in one variable to multiple variables along with exploration of selected data set outliers and anomalies.

Wrangling using MongoDB

March 2015 | Udacity - Nanodegree — github report — Language: Python

- Scripted in python to analyze/clean open street maps data in xml format and converted it to json for better data storage in MongoDB.
- Applied data munging techniques such as assessing the quality of the data for validity, accuracy, completeness, consistency and uniformity.

VOLUNTEER EXPERIENCE | DATA SCIENCE

SCIENCE MUSEUM OF MINNESOTA | Predict annual Membership renewal rates

May 2016 – August 2016 | Minneapolis, MN — Languages: Python, R — Tool: Tableau

- A 3-month Kaggle-like community data science competition based in Minneapolis, MN.
- Worked in a multidisciplinary team to help the science museum predict and improve its annual membership renewal rates.
- Helped with data preprocessing on museum customer dataset and training machine learning models to make predictions using Python libraries such as Pandas, Numpy, and Sklearn.
- Performed feature engineering based on various events that take place at the museum throughout the year.
- Was member of the team that won the pre-qualification round, had highest score in the rubric measuring categories - prediction accuracy, action-ability, and storytelling. Team further won first runner-up in the grand final at the Financial and Retail Conference on Analytics (FARCON) 2016 – MinneAnalytics.
- For more information, please visit: https://www.meetup.com/AnalyzeThis

EDUCATION

GEORGIA INSTITUTE OF TECHNOLOGY | Atlanta, GA

MS in Computer Science | Expected Graduation - December 2018
Specialization: Machine Learning
Cumulative GPA: 4.00
Machine Learning • Knowledge Based Artificial Intelligence • Reinforcement Learning & Decision Making
Machine Learning for Trading • Data & Visual Analytics • Computer Networks • Software Development Process

UDACITY | Mountain View, CA

Nanodegree | Graduated - August 2015
Specialization: Data Analyst

UNIVERSITY OF ENGINEERING & TECHNOLOGY | Lahore, Pakistan

BS in Electrical Engineering | Graduated - March 2007
Specialization: Computer Engineering
Cumulative Percentage: 78%
Honor's List