Muhammad Arslan
Senior Data Scientist

✉ muhammadarslan.se@gmail.com
📍 Karachi, Pakistan.
📞 +92331 3113789

Profile

📅 Date of birth 25/04/1993

👥 Single

Highly analytical and a passionate Data Scientist with demonstrated history of 4+ years of working in financial industry. Primary strength lies in analyzing problems to propose robust solutions and putting complex scenarios into perspective mathematically

Work experience

Senior Data Scientist HBL - Habib Bank Limited, (Head Office). Karachi, Pakistan

Since January 2019

HBL is a Karachi based multinational bank. Working as Senior Data Scientist, my responsibilities include:
1. Manage and lead data science team on critical projects (technical)
2. Formulate analytics opportunities by engaging with other bank units
3. Work as a liaison between technical team and management (convert ML jargons into explainable insights)

Achievements:
1. Developed a data science team
2. Setup data science Infrastructure & workflows
3. Operationalize data science use-cases
4. Developed Non-traditional credit scoring model

Data Scientist HBL - Habib Bank Limited, (Head Office) Karachi, Pakistan

From February 2018 to December 2018

As Data Scientist / Engineer, my responsibilities include:
1. Data Exploration, ETL, Feature Engineering, model building & optimizing using machine learning techniques
2. Collaboration with other units to develop effective analytics, processing, cleansing, and verifying the integrity of data.
3. Presenting the insights / analytics in a meaningful manner.

Software Engineer & BI Expert Central Depository Company (CDC) Pakistan Limited, (Karachi Office) Karachi, Pakistan

From April 2016 to February 2018

At CDC, my responsibility was to write well designed, testable and efficient code in java using object oriented techniques and to model data by writing scripts and design dashboards on Qlikview reflecting business needs. Working as a team member, I contribute in all phases of the development life-cycle.

Intern Microsoft Innovation Center Lahore, Pakistan

From June 2014 to July 2014

LinkedIn Profile

in @muhammadarslansiddiqui

Technical Skills

Programming Languages:
Python, R, Java, Visual C#, SQL

Database:
Oracle, Microsoft SQL Server & MySQL

Data Science Tools:
Sci-kit Learn, Tensorflow, Qlikview, Tableau, Power BI

Big Data Tech:
Spark, Hadoop (MR), Impala, Hive, YARN

Other Tools:
Rapid Miner, Jupyter hub, Cloudera, Docker, Linux, Putty, SVN / GIT

Key Skills:
- Machine learning & Deep learning
- Data cleansing, modeling, & visualization
- Proficiency in programming
- Database design & ETL pipelines
- Designing dashboards

AWARDS

Awarded Gold Medal in BS Degree (2016) @Sukkur IBA, Sukkur, Pakistan

Awarded National ICT R&D Scholarship (2012-2015) @Sukkur IBA, Sukkur, Pakistan

Awarded Institutional Merit Based Scholarship (In Deans List) for outstanding performance in academics @Sukkur IBA, Sukkur, Pakistan

Awarded Bait-Ul-Hikma Award (2015) @ Hamdard University Karachi for conducting one day workshop.

Education

Masters in Data Science FAST - National University of Computer and Emerging Sciences Karachi, Pakistan

From June 2017 to December 2019

CGPA: 3.51

BS (Software Engineering) Sukkur Institute of Business Administration (SIBA) Sukkur, Pakistan

From January 2012 to December 2015

Gold Medalist - CGPA: 3.54 - **Percentage**: 88.79%

Data Science Projects

Non-traditional Credit Scoring Model	Developed a credit scorecard using non-traditional behavior to identify potential customers for cross sell @HBL Consumer Banking Dept. *(Python, MSSQL, MS Excel, Tableau,)*
Segmentation - Anti Money Laundering	Applied clustering to segment customers based on their transaction behavior @HBL Compliance Dept. *(Python, Oracle, MS Excel, Tableau)*
Threshold Optimization - Anti Money Laundering	Re-calibrated AML scenario thresholds using statistical techniques @HBL Compliance Dept. *(Python, Oracle, MS Excel)*
Alert Risk Scoring - Anti Money Laundering	Developed scoring model for prioritization of alerts based on risk level @HBL Compliance Dept. *(R Studio, Oracle, MS Excel, Power BI)*
Cross Sell (Personal Loan & Credit Card)	Identifying potential customers likely to buy personal loan using machine learning and deep learning techniques *(DTrees & CNNs)* @HBL *(R Studio, Python, SQL Server, MS Excel)*
Fuzzy Matching (Screening Customers' Portfolio)	Assessment of Terrorist Financing Risk by screening customers' portfolio against FATF statement entities using fuzzy matching algorithms @HBL *(Python, SQL)*
Anomaly Detection in Agent Transacting Pattern	Identify the fraudulent agents of Branchless banking using *Unsupervised machine learning anomaly detection techniques* @HBL *(Python, SQL, R, Excel)*
ATM Cash Optimization	To forecast cash demand at an ATM for a given period using historic data through time series analysis and deep learning models @HBL *(LSTM, RNN, ARIMA, Python, SQL)*
Addressing anomaly detection in high volume setting	**Thesis project: (Research)** An attempt to a "*Data Driven approach*" in the place of existing '*Rule-based approach*' for detecting money laundering transactions - The goal to reduce false positives (Under Class Imbalance Constraints) and information retention @HBL/FAST *(GANs, VAEs, RBMs, IHT, Cramer Rao Lower Bound Optimization)*

Certificates Earned (MOOCs)

Coursera:
- End to End Machine Learning with Tensorflow on GCP

DataCamp:
- Cleaning Data in Python
- Intro to Python for Data Science
- Intro to Data Visualization with Python
- NLP Fundamentals in Python
- Introduction to R

Other Courses:
Big Data Analytics | Machine Learning | Deep Learning | Cloud Computing | Stat. and Math. Methods in Data Science | Advance Data Warehousing and Data Mining | Block Chain

Software Development Projects

Central Depository System (CDS): Extending core system with the new functionalities and reports @CDC *(EJB , Web Logic Server, Oracle 12c)*

Evoting: Incorporating business needs for electronic voting system @CDC *(JSF2.2, PrimeFaces6, Web Logic Server, Oracle 12c)*

Registrar & Transfer Agent (RTA): Developed a module in a system providing facilities of RTA @CDC. *(JSF2.2, Web Logic Server, Oracle 12c)*

Reporting Automation: Reporting requirements were automated using BI Tool. Data modeling, transformation and designing interactive dashboards fulfilling business needs @CDC *(Qlikview)*

Publications

https://www.researchgate.net/publication/334720715_Deep_generative_models_to_counter_class_imbalance_a_guided_model_selection_strategy

Languages

English — Fluent

Urdu — Native