

# Saif Ali Kheraj

Data Scientist

Contact: +92-306-5830666

Email: [saifalikheraj@hotmail.com](mailto:saifalikheraj@hotmail.com)

Profile: <https://saifkheraj.github.io/>

Medium: <https://medium.com/@saifalikheraj>

---

I have spent my time working on structured and unstructured data to make useful decisions. My area of expertise includes Machine Learning, Deep Learning and Business Analytics.

In these areas, I have worked on lot of different use cases for the business including Churn, Segmentation, Credit Scoring for Microlending, HR Analytics, real time processing for campaigns and data propensity models. In addition to this, I have also added value for the business by performing analysis on the customer base and other data, and of course working for the client facing company has enabled me to understand business processes and customer lifecycle making me able to identify business pain points.

## EXPERIENCE

---

DATA SCIENTIST • JAZZ (SUBSIDIARY OF VEON) • FEBRUARY 2018 - CURRENT

- Working closely with the Customer Value Management, Marketing, Pricing, Digital Financial Services, Human Resource and other teams to perform analytics that can generate business value.
- Generated Revenue for the company by converting non data users to data users using data propensity model. Performed Value and behavioural segmentation on vast customer base to identify rotational customers and the segment that needs to target which converted customers into using 4G.
- Identified huge opportunity for the company to work on in the area of Credit Scoring using non-financial data (Psychology Traits) to target thin file customers. Generated great results for the company.
- Reduced 45 Working days of Human Resource to just 1 working day by designing Resume Shortlisting system.
- Generated Revenue using real time campaigns.
- Translating Data into analytical outputs that enables the business analyst to answer questions that are central to their business success.

TRAINEE CONSULTANT/APPRENTICE • TERADATA • AUGUST 2017 - JANUARY 2018

- Worked with the international team in the areas of Big Data Analytics for the telecommunication client.
- Worked on Real Time processing for live campaigns.
- Worked for clients in Pakistan, Algeria and Bangladesh.

INSTRUCTOR • UDEMY • AUGUST 2017

- I teach online courses.

INTERNEE • TERADATA • JUNE 2016-JULY 2016

- Worked on building machine learning model to cluster text.

## EDUCATION

---

BS COMPUTER SCIENCE • 2013-2017 • NATIONAL UNIVERSITY OF COMPUTER AND EMERGING SCIENCES

- Majors: Data Mining, Deep Learning
- Dean's Honour List more than 3 times
- CGPA: 3.48

## SKILLSET

---

Data Science and Machine Learning/Deep Learning Stack: Python (NumPy, Pandas, Python Scientific Toolkit, scikit-learn, TensorFlow, Keras), Machine Learning with Spark, Spacy

Big Data Stack: Apache Ignite, Apache Kafka, Apache Sqoop, Apache Hive, Apache Spark, NoSQL, Java and python.

Business Analytics: Microsoft Excel (Reporting and KPI Analysis), Microsoft PowerPoint

Soft Skills: Communication, Collaboration, Leading, Management, Public Speaking

Visualization: Chart.js, [plotly](#), matplotlib, ggplot, Tableau

Relational Databases: MySQL, Teradata, Oracle, etc

Web Development Stack: Java Spring, Flask

Agile: JIRA, Confluence, Sprint Planning, Kanban

## PROJECTS

---

GSM Based Credit Scoring for Microlending

Developed proposals and aligned key stakeholders on this project. Developed Credit Scoring Model using analytical and machine learning approach and presented key findings to the stakeholders and leadership. Developed pipeline and dashboard for run time metrics for the key stakeholders.

Value and behavioural segmentation on Data Subscribers

Using the analytical techniques along with the machine learning, we identified segments of customers clearly indicating rotational behaviour. These customers were then further profiled to see behaviour in order to send right campaign.

Resume Shortlisting for Talent Screening

Worked with a team of recruiters and talent to develop business case and proposals. Automated their entire talent screening system by connecting custom tool with their job portal and automatically shortlisting resumes by job title. Using Text Mining approach, this achieved great results. Presented this to Executive Leadership Team.

Autoregressive Modelling to Predict Real Time Hourly Recharge of Subscriber

Company needed to know hourly and daily projection amount for the Recharge so that they can set overall target for the day and track each hour's prediction. Created Autoregressive Model which runs every hour and makes projection for the entire day hour by hour with the total projection amount.

Gender Prediction Model for Digital Financial Services

Business Goal was to target female customers for targeted campaigns to provide special offers to women. Achieved around 80% of precision required by the business.

HR Analytics/Employee Churn using Machine Learning

Business Goal was to identify why top performer churn and what strategy company needs to work on to prevent top performer from churning. Achieved around 82% precision with Random Forest as the best model. Presented this to Chief People Officer.

Real Time Processing for Campaigns

Worked on real time processing system that generates notification to IBM Campaign Management on User's event. This is an important use case for the business as this is currently being used to send different campaigns to the user in near real time.

A Smartphone User Activity Prediction (Academic Project)

Predicted human activities based on smartphone usage pattern. Achieved great results with top 10 on final leader board.

## CERTIFICATIONS

---

- MicroMasters Marketing Analytics • University of California, Berkeley • <https://credentials.edx.org/credentials/3b28901a7a2e4bae9e2caaf7c45d003a/>
- Data Science with Scala • IBM • <https://www.youracclaim.com/badges/db861903-e5b5-4706-8f92-17b77924ab26>
- Enterprise Design Thinking - Team Essentials for AI • IBM • <https://www.youracclaim.com/badges/8a40a3ae-01a0-496e-8cf2-382764b91860>
- IBM Watson Machine Learning Essentials • IBM • <https://www.youracclaim.com/badges/8d12b4d2-0f5e-46ee-afe6-9a8d6e1c0b30>
- Watson OpenScale Foundations • IBM • <https://www.youracclaim.com/badges/c4e0905e-b64e-4525-9929-38d4bc8d8d2c>
- Watson Studio Essentials • IBM • <https://www.youracclaim.com/badges/f0315790-744b-474d-b7ef-5a4fa1c6b26f>
- Enterprise Design Thinking Practitioner • IBM • <https://www.youracclaim.com/badges/3b2815d1-932a-444b-8c8f-20d45699ac09>
- IBM New Collar Professional Skills • IBM • <https://www.youracclaim.com/badges/9ada766f-fb8f-4e58-8d64-fd94a03af99b>
- Deep Learning with Honors Track • National Research University Higher School of Economics • <https://www.coursera.org/account/accomplishments/certificate/B9R7AA3L3G5C>
- Business Metrics for Data Driven Companies • Duke University • <https://www.coursera.org/account/accomplishments/verify/SKEAK46V99BA>
- Technology Entrepreneurship • Stanford University • <https://sucourses.novoed.com/#!/courses/venture/statements/76232>