**Shekh Firoz Alam**

Mobile: +966 580439713

Email:shekhalam@hotmail.com

**Professioal Summary:**

Senior Data Scientist with 5+ years of IT experience focused mainly on Data Engineering &Machine learning techniques, extensive data analysis, design, development, testing, and implementation of data analytics & solutions using Python ,BigData within a wide range of industries like Banking Telecom and Food ordering

* Strong hands on experience in programming languages like **Python,SparkSQL,**
* Working experience with open cv and computer vision
* Expertise in machine learning techniques like **Logistic Regression, Random Forest, Decision Trees, Support Vector Machines and Clustering.**
* Strong Technical knowledge on **Deep Learning, Neural Networks, Tensor flow**
* Good experience in creating Hierarchies, Custom Categories, Calculated Items, Aggregated Measures, New Data Source Filters and Report Linking in SAS Visual Analytics.
* Good experience in Data Collection, querying, modifying and analysis, coding, summarizing the findings, and presenting the results.
* Expertise in data processing techniques like merging, concatenating, interleaving, appending of datasets and data validation using **Spark ML(Action and Transformations)**
* Expertise in extracting data from various disparate data sources & (.xls, .csv, .txt) file formats, Oracle, SQL Server, MySQL Server, etc.
* Good expertise in building Dashboards using Microsoft BI tool
* A good team player with excellent interpersonal and communication skills to interact with Business analysts, Project managers, Developers, Customers and Stake holders.

**Technical Skills:**

Big Data Apache Hadoop,Hdfs,Spark(RDD,Sql)Hdfs,Hive,Metron,Strom,

MQ Kafka,Flume

ML(Algorithm) Linear, Logistic Regression, Random forest,Predective Analysis,open cv

Working Skills Python,Hive,Spark(Sql,Core,Streamining,ML)

Prog Language Python,Java

FrameWork Spring,Django

Design Pattern Factory,Singleton,Builder Design patterns

Databases: MySQL 5.7 ,Oracle 10g,NOSQL

Ide’s STS,Eclipse,R-studio,Pycharm,Python Notebook

Os: Macintosh,RedHat6.9 ,Windows

Cloud GoogleCloud (worked on 90 Node cluster )

VisualizationTool Arcadia(Open Source),Tableau

Tools GitHub,Jira,maven,sql developer

**Education:**

* Bachelors degree (B.Tech) in Information Technology from New Delhi in 2014.
* 12th from Doon Public School, New Delhi
* 10th from Air Force School, Uttar Pradesh

**Work Experience:**

* Worked as Senior Data Engineer for Innovative Software Solutions from Jan 2019 to till date
* Worked as Data Engineer for Micro Spark Software Ltd from Aug 2017 to 30 Dec 2018
* Worked as Data Engineer for Amdocs Development Center Dec 2016 to Aug 2017
* Worked as Software Developer in Appster Information Technology March 2015 to May 2016
* Worked as Software Trainee Kundoor Consultancy Services(Jan 2014 to Feb 2015)

**Current Company: Innovative Solutions, Riyadh(Jan 2019 till Date)**

Roles, Responsibilities and project Information During my tenure in Innovative Software Solutions

**Cyber Watch Big Data Application**

I am working on CyberWatch Product which will be responsible for Cyber Security, threat detectation ,enrichment, alerts,alarmingMy role will be to develop end to end product and bring it to production. I am using kafka,metron,strom,spring boot many open source technology to achieve my goal.

**Previous Company : Micro Spark Software Solutions(Sep 2017 till Dec 2019)**

**Project: Vodafone Germany**

**Description:** Vodafone is leading telecom company around the globe, Vodafone Germany is one of the best service provider in Germany.

**Segmentation Model:**

* Actively participated in the business requirements gathering workshops and meetings with the stake holders.
* Extracted the data from different data sources like SQL server, and Teradata etc.
* Used data imputation techniques for missing data.
* Applied various techniques using Python data structures to get the data in right format.
* Created various functions using Python which involves calculating various metrics such how many customers belongs to which bill range
* Used **K-means clustering techniques and Support Vector Machines (SVM)** in finding the groups of prospective customers which need to be targeted for marketing campaign using Python.
* Performed multiple iterations in arriving at the targeted customers to campaign.

**Data Analysis Using Machine Learning:**

Analyzed the historical data to calculate the profit made by different outlet stores across the world and provided insight to the business to help increase the profits.

* Organized and actively participated in the workshops and meetings with the stakeholders at various levels for gathering the business requirements.
* Data preparation is done in **Python** eliminating the null and missing values.
* Imported **pandas, numpy, sklearn, matplotlib, seaborn** packages to analyze, and calculate the sum of profits on each business day, in a month, and quarterly basis.
* Analyzed and used advanced Machine learning techniques to understand correlation between the variables (I.e. marketing spend) which helped the business to achieve the goals and gain profits.
* Used **Linear Regression, Decision trees & Random forests,** in comparing accuracies.
* Created various visualizations using matplotlib and seaborn as per the business requirements.
* Presented the modeling results and data insights to the managers and stakeholders.

**Project :Royal Bank of Scotland.**

**Role: Data Scientist**

**Description:**

Client is the one of the leading bank in UK, which issues loans on a daily basis which includes both personal and business loans including small scale to high scale businesses. They would like to come up with ways to increase their predictability of offering loans to personal due to an increase in the trend of number of defaulters during the last 2 years. RBS has hired a team of insight analysts in assisting this project.

* Worked on Oracle Database Server to extract the data, done extensive programming in SQL and Teradata databases based on the client requirement and has done joins between the tables to extract the variables which are required for the model.
* Used techniques such as univariate analysis for detecting the outliers
* Used bi-variate and multi variate analysis to understand the relationship between the variables, missing value treatment and imputation of missing values
* Used multi-collinearity analysis to understand the highly correlating variables
* Used Logistic Regression using **Python** to predict the defaulters in the customer who would not be able to pay to the bank.
* Used confusion matrix to interpret the output.
* Analyzed and used advanced analytical techniques such as **Logistic Regression, Decision Tree, Random forests, SVM, ROC, AUC.** Compared the accuracies various algorithms and used the one which has highest accuracy that help in decision making.

**Responsibilities:**

* Extracting data from oracle database through SQL Pass through query and Libname facility.
* Data formatting as per the analysis and reporting requirements
* Data sub-setting and summarizing as per the reporting needs
* Creating user define formats using PROC FORMAT.
* Developing MACROS for automatic running of all programs like MERGING, SORTING, PRINTING, and for changing the dates dynamically.
* Creating Stored process thru EG and Publishing Stored process reports into portal
* Creating reports like Sub-Broker Active AM report, Channel wise INWARDS and NCA report, NRI Conversation v/s Acquisition report, EAG report and ATS report etc...
* Used Output delivery system to generate reports in html using Proc Report

**Company: Amdocs Development, (December 2016 to Aug 2017)**

**3IL(3 Ireland)**

Analyzed cdr files using Machine Learning(Decision Tree,Random forest),Spark in Cloudera Distribution Analyzed the cdr files and suggested the mobile top-up plan to business In Three.ie(Three Ireland) my role is to write Shell/Python scripts and automate.

All the Billing jobs of AMC(job monitoring tool),At regular interval we have to update Amc tool using Python/Django framework for smooth work flow of jobs.

**BritishTelecom(EE UK)**

In British Telecom(Everything EveryWhere,Ensemble ,Abp,Amc),I had analysed the cdr files using BigData(Spark and Hive) deep learning(Svm,Random Forest,clustering ) most of the time we will go with clustering and random forest algorithm to get insights.

My role is to keep up and running the Amdocs Monitor Control Tool to up and running of the jobs Amc is heart of in production of billing.

**Previous Company: Appster Information Technology, (March 2015 to May 2016)**

Appster is an mobile application startup company with 500 employee in 3 country

Roles and Responsibilities:

Developed Restful WebServices,Cms Admin Pannel using Spring3.0 and JPA and other technology and frame work as per client requirement.few mobile application detail is as below

**2Stones:**

2Stone is of E-Commerce mobile application which is selling products around different states in Australia. My Role is to develop restfull webservices where customer can purchase articles,return or request money back.

Admin panel for Admin's of application using Spring-Mvc frame work As admin will verify the sellers products before listing it on application

**OGO:**

Ogo is an real time position tracking application of people of similar interest,hobbies,passion

My Role is to develop restfull services where can allow his/her geolocation on the basis of geo location similar interest people can meet and plan any event . Admin panel for Admin's of application using Spring mvc & Restfull services where admin can approve or blockapp users.

**Simply Sports:**

Simply Sports is an Application which notify about the sports you follow and need updates on regular time as admin put add rss feed from Admin Pannel all the user will get notified.

My Role is to develop Restfull Services(Jersy) .Admin panel for Admin's of application ,using mvc Design Pattern.

**Previous Company: Kundoor Consultancy Services(Jan 2014 to Feb 2015)**

Worked as Software Trainee, My role was to develop web component using java and spring frame work

**Doctor Connect:**

Doctor connect is about consulting doctors in your area via video call. My role is Developed backend i.e spring restful services there will be 3 tier Super Admin, AppAdmin,mobile user like doctor signup, upload doctor qualifications,admin will verify license of practice and approve doctor’s profile.

User can video call available doctor and discuss his health issue on the basis of discussion doctor will write his prescription and it will be automatically send to nearest medical store and medicine will be delivered to patient.

**Shekh Firoz Alam**

**Riyadh(Saudi Arabia)**