Saurabh Patel

Data Scientist/ Machine Learning - ERICSSON

Santa Clara, CA - Email me on Indeed: indeed.com/r/Saurabh-Patel/2c700d9cf35f9d1f

• Well versed in using statistical methods like z-test, p-test, Simpson's Paradox, Poisson distribution, Binomial Distribution.
• Familiar with various ingestion techniques to bring into R, Python and Azure ML environment from different big data platforms such as a HDFS, Spark, Azure Blob and Azure Data lake.
• Hands on with various Data cleansing process like handling missing values by using techniques such as a replacing by mean, forward or backward fill, removing entire rows or columns or values, removing outliers and normalizing, scaling data.
• Visualized data using different visualization tools R, Weka, Azure ML, Tableau and Power BI.
• Mentored by Sr. Data Scientist to create predictive model using supervised, unsupervised and ensemble machine learning algorithms.
• Familiar with predictive models using classification algorithms like KNN, Naive base, regression and decision trees.
• Familiar with predictive models using numeric and classification prediction algorithms like support vector machines and neural networks.
• Familiar with predictive models using ensemble methods like bagging, boosting and random forest to improve the efficiency of the predictive model.
• Conducted sentimental analysis by extracting unstructured streaming data from different social media platforms like twitter, Facebook and LinkedIn etc.
• Created a repository in GitHub for version control.
• Conducted a streaming analysis with IOT on Microsoft Azure using HDInsight, Blob Storage, Power BI.
• Familiar with predictive models using different cloud based Machine learning tools like Microsoft Azure ML, AWS ML, KNIME
• Worked on NLP, Text Mining and sentimental analysis for extracting the unstructured data from various social Media platforms like Facebook, Twitter and Reddit.
• Worked on KAGGLE Data sets and Microsoft azure Ml Predictive models as a part of Data science Boot camp.
• Familiar with Hive and pig scripting.
• Familiar with IBM SPSS predictive analytics Software.
• Strong understanding of CRM, Sales force and Net Suite.
• Strong Understanding of Forecasting and GAP analysis.
• Strong Understanding of SDLC, Agile Life cycles.
• Strong Experience for working in ambiguity to solve complex problems. Focus on strategy, development, operations and defining unique solutions.

WORK EXPERIENCE

Data Scientist/ Machine Learning

ERICSSON - Santa Clara, CA - January 2016 to Present

Network AI (Artificial intelligence)
Roles and Responsibilities
* Working on Network Data to find the Root cause in VOLTE call drops (Voice Over long term evolution).
* Performed Data preparation on a High dimensional (Big data with large volume and variety) Data sample collected from the live customer data.

* Data preparation Includes Mapping of unlined data from various formats, Identifying The missing data, Finding the correlations, scaling and removing the Junk data to further process the data for building a predictive model
* Closely working with Network engineers and Ericsson analytics Expert Team to find the rule sets to build a predictive model.
* Processed data using R programming and developed a Predictive model to predict KPI'S (Key performance indicators) Such as VOLTE accessibility, Connection success, Session Set Up success and Retain Ability.
* Worked on Microsoft Power Bi to gain the insights from the Alarm Data to find the Root Cause for Alarms that happened when there is Network Breakdown.
* Closely working with senior Artificial Intelligence Team to create Ontology and Build a Machine learning layer in the final Product.
* Trained Data with Different Classification Models such as Decision Trees, SVM and Random forest
* Random Forest was decided as Final Model Based on over all Statistics, Model Performance and Run Time.
* Creating the Desired Models in Spark Environment, Spark R.

Achievements: Presented Models and Results at Different Locations of Ericsson in Texas, New York and Los Angeles in Network AI workshops to different network operators for Business Attraction.

Technology Stack: R, spark R, Azure Ml, Python, Drools Rule Engine, SQL, Tableau, Data Mining, Machine Learning, Big Data, Microsoft Power Bi and HTM studio.

What was the spectrum for Prediction?
In depth analysis of data, Calculation of KPIS from the performance Metrics Data and based on the current KPI metrics predicting the KPI for Next Hour / Next Day / Quarter.
What was the action Taken?
Closely worked with senior data scientists for analyzing the predictive model this was built using Machine Learning Techniques Purely a Classification Problem

Junior Data Scientist / Data Analyst

JP Morgan - Chase, NY - August 2014 to December 2015

Roles and Responsibilities
• Analyze Data and Performed Data Preparation by applying historical model on the data set in AZURE ML
• Performed Data cleaning process applied Backward - Forward filling methods on dataset for handling missing values
• Under supervision of Sr. Data Scientist performed Data Transformation method for Rescaling and Normalizing Variables
• Developed a predictive model and validate Neural Network Classification model for predict the feature label
• Performed Boosting method on predicted model for the improve efficiency of the model
• Presented Dashboards to Higher Management for more Insights using Power BI
Technology Environment: R, social Media analytics and Microsoft Power BI.

What was the prediction?
To analyze and to predict the customer's credit history and past bill payments based on the Credit card offers, to create a predictive model, and send offers to customers on the base of Model and past data.

What was the Action Taken when a system was built?
A system was successfully created on the past data and Credit History and Payments Activity of customers and runs model against the historical data and get predicted label if customers eligible for credit card offer or no, on the base of that send them offer

What was the Measure Taken?
Customers was getting actual offer they like and accept those offers and that way Increase numbers of Credit Cards Users that leads to profit to bank

Data Analyst

DXCorr Design Inc, IND - August 2010 to May 2012

Roles and Responsibilities
• Analyze and mapping of data, create data dictionary by using SQL and SQL Server
• Created ER diagram and Read ER diagram and Created
• Data Gathering and Analyzing data, creates reports
• Created Database in Microsoft Access by using blank database and created tables and entered dataset manually and Data Types, performed ER Diagram and Basic SQL Queries on that database.
• All-time favorite Microsoft Excel used for formatting data as a table, visualization and analyzing data by using certain methods like Conditional Formatting, Remove Duplicates, Pivot and Unpivot tables, Created Charts and Sort and Filter Data Set
• Performed Statistical Analysis and Hypothesis Testing in Excel by using Data Analysis Tool.
• Used SQL and SQL Server for writing simple and complex queries like finding Distinct Values and Unique Values in Data set
• Created Entity Relationship Diagrams and Data mapping for better understanding of dataset
• Presented a Dashboard for better understanding of dataset to all Stock Holders

EDUCATION

Masters in Computer Science and Information Systems in Computer Science and Information Systems

California University of Management and Sciences
2014

SKILLS

Bi (2 years), BUSINESS INTELLIGENCE (2 years), Excel (1 year), Power Bi (2 years), SQL (3 years)

ADDITIONAL INFORMATION

Technical Skills
Data Sources Azure Blob, Data Lake, HDFS, SQL Server, MS Excel
Programming Languages R, SQL, Python
Data Visualization R, Python, Weka, Microsoft Power BI, Azure Machine Learning, Tableau
Data Exploration Azure ML, MS power BI, Azure IOT hub, HD Insight, Streaming Analytics, S3, Redshift
Cloud services Azure ML, AWS ML, MS power BI, Azure IOT hub, HD Insight, Streaming Analytics
Cloud platforms Azure ML, Aws ML
Repository GIT HUB
Machine learning Algorithms
Classification, KNN, Regression, Random Forest, Clustering(K-means), Neural Nets, SVM, Bayesian Algorithm, Social Media Analytics, Sentimental analysis, Market Base Analysis, Bagging, Boosting

Statistical Techniques Z-Test, T-test, Null Hypothesis, P-Values, Simpson's paradox.