pdeshmuk@buffalo.edu
(716) 598- 4517
Buffalo, NY 14214

Parikshit S Deshmukh

Linkedin.com/in/parikshitd92
Github.com/parikshitdeshmukh
http://www.parikit.com

TECHNICAL SKILLS

Languages - **Python, R,** JAVA (Oracle Certified 2016), C++, SQL, Oracle PL/SQL, JavaScript, Unix
Technologies - AWS, Android, Mongo DB, REST API, Vue JS, jQuery, Linux, Git, Maven, Blockchain
Data Science - **Machine Learning, Deep Learning (**Keras, TensorFlow, pandas, numpy, scikit-learn, scipy**)**, Feature Engineering, Predictive Modelling, EDA, Statistical Data Mining, Hadoop, Spark, Scala, R Studio, D3js

EDUCATION

MS in Computer Science, University at Buffalo, State University of New York – GPA: 3.6 Sept 2018 - Feb 2019
Coursework*: Statistical Data Mining, Machine Learning, Pattern Recognition, Data Intensive Computing, Algorithm and Analysis*
BE in Electronics and Communication, Nagpur University, India Sept 2009 - Aug 2013

WORK EXPERIENCE

Research Assistant at DIDC Lab — University at Buffalo Jan 2018 – present

- Working on NSF funded One Data Share project which is a cloud-hosted data transfer tool that manages and optimizes the movement of data across arbitrary endpoints in a protocol-agnostic manner.
- Project is funded by NSF at The Data Intensive Distributed Computing Laboratory of University.

The Advisory Board Company, Chennai, India Oct 2016 – Jun 2017
Data Analyst

- Developed ETL products in Java, Groovy, Yaml and SQL to process and load patient's data into **PostgreSQL** database according to data mapping documents.
- Reduced manual efforts by around **30% by automating** QA report generation for business team with **JAVA, SQL** and **Maven.**
- Incorporated Log4j for logging, Ant/Maven for project deployment and JUnit for unit and integration testing.

Cognizant Technological Solutions, Chennai**,** India Jan 2014 – Oct 2016
Programmer Analyst

- Took an initiative to develop a junk filter for IT support team to cull out irrelevant emails using bag of words and tf-idf transformer of Scikit learn. Used **Support Vector Machine** and **Naive Bayes classifier** for prediction. Received appreciation from client.
- Performed Data mining techniques like **RFM Analysis** on Underwriting data to understand the insurer's behavior and how it impacts the value that can generate for the client with independent RFM scoring
- Designed, developed and tested core components of data management applications for Liberty Mutual Insurance (USA) in **JAVA, MVC, JavaScript, Ajax, jQuery, jUnit** and **MySQL** for integrating and managing customers' data in MySQL database
- Developed web application with Single Sign On (SSO) feature using **HTML5**, **Spring Boot**, **Hibernate** and **SQL** for vendors' registration
- Developed **REST** based web services secured with **oAuth2.0** to provide fraudulent transaction details to the security team.
- Implemented complex **PL/SQL** queries for various CRUD operations including functions and procedures to integrate insurance applications with MySQL and DB2 databases using **JDBC** and **Hibernate ORM**.

PROJECTS

EDA and Visualization of Influenza trends in USA (R, **dplyr, ggplot2, rCharts, TweetR**) - **GitHub** Feb 2018

- Gathered tweets data for influenza from **Twitter** API, and after processing and extracting meaningful tweets, used Geocode API to plot the tweets frequency on USA state wise map as **Heatmap**.
- Gathered data from cdc.gov/flu and performed EDA by plotting pie charts, bar plots, USA Statewise heatmap etc. to study different aspects of Influenza breakout in USA

Sentimental analysis and Visualization using Hadoop Mapreduce (Python, R, D3 js, Hadoop, **Mapreduce**) - **GitHub** Apr 2018

- Gathered articles' data from NYTimes API and tweets from Twitter API corresponding to current trends in USA. Applied MapReduce on both corpus, removed stopwords and published the result as a word cloud to a web page using D3 js for sentimental analysis

Predicting Airlines delay and the best airport to fly out from (Python, Logistic Regression, AWS S3) - **GitHub** Nov 2017

- Performed EDA, and Feature engineering and built predictive model on pickled dataset using Naive Bayes classifier and Logistic Regression on top of Gaussian Basis functions and validated using Early Stopping method.
- Tested the model to predict city's best airport with minimum flight delay.

Glasses Detection using Deep Convolutional Neural Network (Python, Deep CNN, TensorFlow): Nov 2017

- Implemented Deep CNN to determine if person in test image dataset is wearing glasses or not using TensorFlow. Used DeepCNN to train the model and SGD for hyper-parameters tuning with Dropout regularizer for improving performance. Accuracy achieved approximately 98%.

Predictive analysis on Boston dataset (**R, Logistic Regression, ggplot2**) - **GitHub** Oct 2017

- Performed feature engineering and implemented mentioned machine learning algorithms for selected feature subset.
- Obtained improved accuracy after dropping irrelevant features for KNN model. Analyzed LOOCV error for set of models.

Data Analysis using Neural Network and Support Vector Machines (R): Sept 2017

- Performed predictive Data analysis and spam email prediction using SPAM data (ElemStatLearn lib) using Neural Network (ANN).

Fitted Support Vector classifier, SVM with Radial Kernel and SVM with Polynomial kernel on Orange Juice dataset of 18 variables to estimate what kind of juice customers of Test dataset bought. Parameter tuning by 10-fold cross validation.

● **1st prize** winner at **Blockchain** Buildathon at Buffalo for building web application for supply chain traceability using Blockchain ● **BrickHack4** (Jan2018) hackathon **1st prize** winner at Rochester Institute of Technology, USA ● "Best Quarterly Performance" award for 2 continuous quarters in year 2015