

# Arpit Chaukiyal

(469) 929-4740 | [chaukiyal.arpit@gmail.com](mailto:chaukiyal.arpit@gmail.com) | <https://www.linkedin.com/in/chaukiyalarpit> | <https://github.com/arpitchaukiyal/>

## EDUCATION

Master's in Business Analytics, [The University of Texas at Dallas](#), USA

May 2019 (Exp.)

Bachelor's in Information Technology, [Uttar Pradesh Technical University](#), India

June 2012

## SKILLS

**Certification:** *AWS – Solution Architect Assoc., EMC – Data Science Associate*, Big Data & Hadoop, CCA175 Spark & Hadoop (in progress), AWS-Big Data Specialty (in progress)

**Statistical Techniques:** Supervised & Unsupervised learning, Social Network Analysis, Naïve base, PCA, Decision trees, NLP, Clustering, SVM, Decision Trees, KNN, Markov chains, Ensemble methods, Boosting – XGBoost, ADABoost etc.

**Programming:** R, Python, R-shiny, Java, Hadoop, MapReduce, HIVE, Scala, Spark, Pig, Flume, Sqoop.

**Software & Databases:** SAS, R Studio, Tableau, Advanced Microsoft Excel, Eclipse, Git, Python Notebook, MS Office.

## PROJECTS

**Machine Learning- Classification**, using Python, Numpy, Pandas, Sklearn

Sept 2018 – Oct 2018

- Implemented SVM, Decision Trees and ADA-boosting models on Telecom Churn & Student Performance datasets.
- Experimented with different kernels of SVM & performed hyperparameter tuning to maximize the accuracy of 97%.
- Implemented Decision Trees and ADA-boosting models & experimented with Gini indexes/InfoGain, pruning etc.
- **Technique Used:** Support Vector Machines, Decision Tree, Ensemble methods, ADA boosting.

**Supervised Learning**, using Python, Numpy, Pandas

Aug 2018 – Sept 2018

- Predicted final grades for Student Performance dataset.
- Implemented supervised learning regression & logistic regression model using gradient descent algorithm.
- Compared train/test error metrics and plotted the results against learning rate for gradient descent & convergence threshold.
- **Technique Used:** learning regression, logistic regression, gradient descent, learning rate curves.

**Natural Language Processing**, using R and R Shiny

March 2018 – May 2018

- The reviews of the users have been scraped from the website by R code-www.cars.com along with the star rating.
- Performed normalization measure like stemming & lemmatization along with document tagging.
- Constructed logit model that predicts user rating using sentiment scores of user's reviews with an accuracy of 80%.
- **Technique Used:** Web Scraping, NLP, stemming & lemmatization, Regex, Logistic Regression, TF-IDF.

## EXPERIENCE

**Mphasis Limited**, Data Analyst

June 2016– June 2017

- Developed data solution based on predictive and behavioral models via statistical analysis like regression and classification techniques like decision trees and K-nearest neighbor for life insurance client which optimized their new business inflow by 17%.
- Developed model that predicts the insurer's fraudulent behavior in paying the premium as well as estimating the premium amount depending on undertaking data resulting in 11% reduction in defaulter volume.
- Created interactive Tableau dashboard that helped the stakeholders and VPs to understand data-driven actionable insights and take business decisions.

**Birlasoft India Limited**, Data Analyst

October 2013– June 2016

- Liaised with the client team to develop a model that predict the work-hours required to complete the task depending upon resources involved and the complexity of the task to align the project with CMMI-5 standards resulting in a 15% increase in team's efficiency.
- Incorporated linear regression, logistic regression and random forest techniques to develop statistical models for demand forecasting and procurement analytics which had in decreased the procurement cost by 7% annually.
- Trained 30 plus employees over the period of 8 months in technical and process-oriented modules which reduced on-boarding time by up to 20%.

## ORGANIZATIONS

**Enactus UTD Chapter, University of Texas at Dallas** – Co-Project Leader

October 2017–Present

**Data Science Club, University of Texas at Dallas** – Officer

January 2018–Present

**The Research Club, University of Texas at Dallas** – Officer

August 2017–Present

**Orators at UTD Toastmasters Club, University of Texas at Dallas**

August 2017–Present

**Eligibility:** Eligible to work in the U.S. for internships and for full-time employment for up to 36 months without sponsorship.