

OBJECTIVE: To use my talents as a mission-centered data-driven leader, among a team of results-driven Data Scientists & Engineers, dedicated to achieving mutual goals, while providing opportunities for personal growth and advancement.

SUMMARY:

Qualified data scientist with strong math background and expertise in Regression models, Predictive analysis, Decision trees, Machine learning and NLP. Self-motivated and a quick learner, passionate about data science and machine learning. Excellent interpersonal and communication skills.

3+ years of experience in IT business consulting with expertise in Project Management, Analysis with a unique mix of business development, functional, domain, technical and client handling skills.

SKILLS:

Machine Learning | Data Science | Decision Analytics | Text Analytics | Statistical Methods

Logistic regression, linear/non-linear models, decision trees, random forest, hypothesis testing, Feature Engineering, neural networks, natural language processing (NLP), time series analysis.

TECHNICAL SKILLS:

Machine Learning: Classification, Regression, Clustering, Probability Theory.

Statistical tools and programming languages: SAS, R, Spark, Scala, Python (scikit-learn, numpy, scipy, pandas, seaborn, matplotlib), SQL, SAS-Enterprise Miner, LaTeX, GitHub version control, Microsoft Office.

Selected coursework: Applied machine learning, Big Data architecture, Applied Bayesian methods, Applied Analytics, Probability and Statistics.

EXPERIENCE:

Colaberry, Boston, MA

Feb 2017 – Dec 2017

Data Scientist

Overview: *Colaberry* <http://colaberry.com/> | *RefactorEd* <https://www.refactored.ai/>

Colaberry provides cutting-edge Artificial Intelligence and Machine Learning solutions. The *RefactorEd Lab* provides companies, universities, and students with an interactive visual platform & tool. The platform enhances the learning of bleeding-edge curriculum, through applying abstract Data Science and Machine Learning concepts in Python.

Responsibilities:

- Worked on machine learning projects related to text classification to develop category/tags of course content
- Responsible for creating course curriculum in data science and machine learning tracks, using visualizations to explain statistical and machine learning concepts.
- Analyzed complex datasets to create labs for users to realize the process involved in data modeling.
- Worked on large datasets, transformed information from raw data into meaningful analysis that identifies trends and predicts outcomes.
- Developed a web-scraper from scratch using python, to collect data (information from LinkedIn, google scholar, Facebook etc.) from student and create its profile, to be used for recruitment by company's clients
- Created word cloud from twitter feeds using text feature extraction and NLP techniques in Python.
- Helped colleagues to apply optimum modeling techniques in their data analysis.

Data Science | ODSC East: <http://tinyurl.com/k3qlkw5>

Machine Learning | SUNY Buffalo: <http://www.cedar.buffalo.edu/~srihari/CSE574/index.html>

ZS Associates, Gurgaon, IN

July 2014 – May 2016

Business Operation Associate

Overview: Consultant | Analytics | Data Strategy

- Led teams in various client engagements, managed client relationships.
- Designed, built and fine-tuned predictive models to determine sales-rep goals, helped the client with the decision making, cutting their losses substantially and making them 7% more profitable.
- Conducted market research for client; analyzed the survey data to strategize client's product launch using data science
- Engineered complex incentive compensation model using SQL based tool for a client's business unit having 1000+ sales representatives.
- Managed to retain my project team with zero attrition and 100% client satisfaction.

Data Science and Machine Learning Projects:

- **Credit Data Analysis (Scikit-Learn, Seaborn, Pandas, Statsmodels)**

Implemented and compared various machine learning algorithms such as - KNN, Decision tree, Naïve Bayes, Support Vector machines etc. with respect to their accuracy to classify customers having Good/Bad credit.

- **Churn Analysis Using R programming (e1071, randomForest, tree)**

Implemented various tree-based algorithm such as Decision tree, Random forest, XG-Boost and Bagging to find out the various parameters leading to churn of customers.

- **Natural Language Processing and Text Learning (Beautiful soup, NLTK, TextBlob, Tweepy)**

Web scrapped 1million tweets from twitter on presidential elections of 2016, 100 days before the election to try to find sentiments and predict the results for the elections of 2016.

- **Clustering for Machine Learning Engineer (Scipy, Seaborn, Pandas, PyMix)**

Researched and implemented various clustering algorithms such as K-Means, Spectral, DBSCAN and Gaussian Mixtures to study their properties and know when to implement

EDUCATION:

University of Texas at Dallas, TX

Aug 2016 – May 2018

M.S. Business Analytics concentration in Data Science | GPA: 3.7/4.0

Army Institute of Technology, India

June 2010 – July 2014

Bachelor of Engineering in Electrical and Telecommunication Engineering | CGPA: 3.9/4.0

AWARDS | HONORS:

- Operational Excellence award at ZS Associates for exceptional performance for the year 2015
- Council Member at Dean's Council and Data Science Club member at University of Texas at Dallas.
- Top 5% in a class of Master's students within the department of analytics at University of Texas at Dallas.
- Held many position of responsibility in High school (Junior Captain) and undergraduate college (General Secretary)
- Ranked top 0.1% in competitive examination for engineering.
- National top 0.1%, Certificate of Merit from President of India, in CBSE class X mathematics examination.

References: Provided on request

Employment eligibility: Eligible to work in U.S. on OPT for 36 months