



<https://github.com/kshitijyad>

<https://www.linkedin.com/in/kshitijyad>

Kshitij Yadav

+1 916.793.9923

Yadav.kshitij23@gmail.com

Naveen Jindal School of Management, The University of Texas at Dallas

M.S. Business Analytics concentration in Data Science

Coursework: Machine Learning Algorithms, Statistics, Linear Algebra, Probability, Advance data science concepts

May 2018

GPA 3.7 / 4

Army Institute of Technology

BE, Electronics and Telecommunication

July 2014

GPA 3.9 / 4

PROFESSIONAL EXPERIENCE

Colaberry Inc

Boston, MA

Data Science Intern (Python, NLTK, Scikit-Learn, Pandas, Scrapy)

February 2017 – Present

- Worked on machine learning projects related to text classification to develop category/tags of course content
- Used Natural Language Processing techniques to understand interests of users, identifying top keywords using LDA to plot word clusters
- 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
- Increased company's revenue by 9% by deriving insights on company's internal data, applying machine learning techniques and advance statistical models
- Worked as business development manager and expanded companies existing business in North America and India
- Curriculum development and testing of Data Scientist and Analytics for [Refactored.ai](#)

ZS Associates

Gurgaon, India

Business Associate (SAS, R, SQL, e1071, nnet, caret)

July 2014 – May 2016

- Designed, built and fine-tuned predictive models to determine sales-rep goals, helping the client with the decision making, cutting their losses substantially and making them 7% more profitable
- Implemented various machine learning algorithms for customer segmentation, using it for targeted marketing
- Conducted market research for client; analyzed the survey data for pricing and availability of clients' products
- Secured \$12 million-dollar funding of client renewal business for the company for which was awarded **Operational Excellence award**
- Worked on SQL databases including Normalization, Stored Procedures, Constraints, Querying, Joins, Keys, Indexes, Data Import/Export and Triggers

Graduate Student - University of Texas at Dallas

August 2016 – May 2018

- **World Development Indicator Analysis using R Programming (e1071, caret, gbm)**

Analyzed data using various linear and non-linear models to implement predictive and descriptive algorithms on the dataset.

- **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.

- **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

- **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.

TECHNICAL SKILLS

Languages & Technologies: C/C++, SQL, MATLAB, SAS, Python, R, VISIO, PIG, Julia, Hadoop, HIVE, Excel(advanced), Word, PowerPoint, Tableau, Adobe Analytics

Techniques and Processes: Feature Engineering, Linear/Non-Linear Models, Neural Networks, Machine Learning and Data science algorithms, Text mining, Natural Language Processing, Pandas, Numpy, Time Series

ADDITIONAL INFORMATION

- Deans Scholarship Awardee Fall 2016
- **Operation Excellence Awardee**
- College level golf player
- **Associate at Deans Council University of Texas at Dallas**
- State level Badminton and Cricket player
- Got sponsored by my employer Colaberry for my Tuition Fee for Fall 2017 and Spring 2018 Semester