



Kshitij Yadav

916-793-9923 | yadav.kshitij23@gmail.com | <https://www.linkedin.com/in/kshitijyad> | <https://github.com/kshitijyad>

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

SKILLS:

Machine Learning | Web Analytics | Statistical Methods

Web analytics, logistic regression, linear/non-linear models, hypothesis testing, natural language processing (NLP), time series analysis

CERTIFICATION:

Google Analytics, Customer Insights (Wharton School), R Programming, Python for Data Science

TECHNICAL SKILLS:

Machine Learning: Classification, Regression, Clustering, Probability Theory, Feature Engineering

Statistical tools and programming languages: SAS, R, Apache-Spark, Scala, Python (scikit-learn, numpy, scipy, pandas, seaborn, matplotlib), HIVE-SQL, MySQL, SAS-Enterprise Miner, LaTeX, GitHub version control, Microsoft Office, AWS, MicroStrategy, Tableau, STATA

Selected coursework: Applied machine learning, Big Data architecture, Data Visualization, Web Analytics, Marketing Management, Marketing Customer Insights

EXPERIENCE:

Colaberry, Boston, MA

Feb 2017 – Dec 2017

Data Scientist Intern

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

Responsibilities:

- Participated in all phases of research including data collection, data cleaning, data mining, developing models and visualizations
- Increased company's revenue by 12% by deriving insights on company's internal data, applying data science modelling techniques and advance statistical models
- Engineered various machine learning models for both descriptive and predictive analysis from scratch, utilizing business requirement to propose various business and strategic recommendations
- Developed a web-scrapper from scratch using python, to collect data from relevant blogs/articles/news on internet and performed various analysis on the data using NLP techniques to get insights
- Leveraged data to provide strategic business decisions to increase sales and effectiveness of marketing efforts of the team
- Created dashboards to provide information on current marketing KPIs and actionable data insights using tableau

ZS Associates, Gurgaon, IN
Business Operation Associate

July 2014 – May 2016

Overview: *Consultant | Analytics | Data Strategy*

- Led teams in various client engagements, managed client relationships and participated in technical discussion
- Designed, built and fine-tuned predictive machine learning 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
- Secured \$12 million-dollar funding of client renewal business for the company for which was awarded Operational Excellence award
- Managed to retain my project team with zero attrition and 100% client satisfaction

Academic Projects:

Risk Profiling and Severity prediction (Scikit-Learn, Seaborn, Pandas, Statsmodels, Stratified K Fold)

- Used 2017 insurance data to predict the probability of claim per customer for 2018, used gradient boosting to predict USD amount of claim in 2018 for those customers and sort customers on the basis of their Risk Profile.

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.

IMDB Rating Prediction (Keras, Tensor-Flow, Scikit-Learn, Seaborn, Pandas)

- Engineered script to pull data from IMDB sites, cleaned, pre-processed it and used various regression models such as Linear Regression, Random Forest and Gradient Boosting to Predict IMDB rating. Compared the models with Neural Network model with respect to MSE.

Time Series Analysis using SARIMAX and RNN (Keras, Scipy, Seaborn, Pandas, PyMix)

- Built Seasonal ARIMA model to forecast future sales of fortune 500 products and compared the model with Recurrent Neural Network using Train-Test Data accuracy.

EDUCATION:

University of Texas at Dallas, TX

Aug2016–May2018

M.S. Business Analytics concentration in Data Science | GPA: 3.75/4.0

Coursework: Machine Learning Algorithms, Statistics, Web Analytics, Probability, Advance data science concepts

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.

References: Provided on request

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

Portfolio: <https://kshitijad.github.io/>