

Deepansh Parab

E: dparab@stevens.edu • M :201-736-9202 • Jersey City, NJ

[LinkedIn](#) • [GitHub](#) • [My Website](#)

DATA ANALYTICS PROFESSIONAL

Creative and self-starter data analytics professional with 1.5 years of academic experience in technology space, and the ability to mine business insights from large structured, and unstructured datasets. Proven ability to combine vision, creativity and business acumen with well-developed communication skills to drive data driven decision making. Core competencies include

Data Mining • Machine Learning • Text Mining • Data Visualization • Statistical Analysis

PROFESSIONAL EXPERIENCE

STEVENS INSTITUTE OF TECHNOLOGY • HOBOKEN, NJ

Graduate Teaching Assistant – Intro to Data Structures and Algorithm

- Assisted the professor by generating labs and assignments for students
- Generated automated testing modules to grade to assignments
- Held review session to assist the students in the core competences of the subject

STEVENS INSTITUTE OF TECHNOLOGY • HOBOKEN, NJ

Graduate Teaching Assistant – Knowledge Discovery and Data Mining

- Evaluated homework, test, and held office hours to ensure students understood course concepts
- Organized and Facilitated classroom lessons and presentations; effectively managed and assisted 40 students with programming in **R** and **Python**

ACADEMIC PROJECTS

Segmentation of Twitter Timelines using Topic Modelling

- Collected 50000 Twitter Timelines using R's Tweepy package and Twitter API and stored it into a csv file
- Constructed a corpus of cleaned and tokenized twitter timelines using R's text mining package
- Performed LSA dimension reduction and Kmeans Clustering on the corpus
- Compared the results with LDA (Gensim version) and visualized the outcome using various plots in R

Data Analysis on Image DataSet (MNIST)

- Extracted the mnist dataset file and converted the ubyte.gz file into csv containing 60000 train data and 10000 test data, using python's pandas and numpy modules
- Analyzed the dataset using four supervised classification algorithm Random Forest, SVM, KNN and ANN using various R packages
- Compared the accuracy of all the models and achieved an accuracy of 98.87 using the deeplearning function in R's H2O package

Sentiment Analyzer

- Extracted companies related data from indeed.com and dice.com using Python Web Scrapping
- Interpreted the Sentiment Analysis for the company reviews using python's NLTK module
- Displayed the overall insights of a company based on current job opening using RStudio's Shiny App

Text Analysis and Entity Resolution

- Performed entity resolution across two datasets of commercial products using Apache Spark with pySpark to apply powerful and scalable text analysis
- Performed text similarity – Bag-of Words using TF-IDF, and cosine similarity to compare the two datasets

EDUCATION

Masters of Science in Computer Science GPA: 3.7 • Stevens Institute of Technology • Hoboken, NJ • May-2018

Bachelor of Engineering in Information Technology CGPA: 7.19/10 • University of Mumbai, Mumbai • INDIA • 2016

TECHNICAL SKILLS & DATA ANALYTICS SKILLS

Programming Skills: **Python** (numpy, pandas, scikit-learn), **SQL**, **R**, **SAS**

Big Data/Distributed Tools: Hadoop, Apache Spark, Pig, Hive, pySpark, sparkML

Data Visualization: matplotlib, ggplot2, Tableau, Shiny App

Databases: MySQL, Postgre, Oracle 10g