Data scientist with three years of experience in predictive modelling, data processing, and data mining algorithms to
solve challenging business problems. Strong background in computer programming language, and knowledge of
various types of machine learning and natural language processing techniques.

Willing to relocate to: Pune, Maharashtra - Hyderabad, Telangana - Bangalore Urban, Karnataka

WORK EXPERIENCE

DECISION SCIENTIST

MU SIGMA - Bengaluru, Karnataka -

2015 to Present

Projects
TEXT ANALYTICS | MUFUSION - FOR MU SIGMA
‣ Developed a text analysis engine used by fortune 500 clients. The text analysis engine developed in Python supporting
English and Japanese text leverages Sentiment analysis, Theme identification, Categorisation, Named Entity Recognition,
Clustering and Word co-occurrence modules to generate insights for decision making.

DATA QUALITY AND VALUE ASSESSMENT - FOR A US BASED ASSET MANAGEMENT COMPANY
‣ Designed and developed a comprehensive and versatile Data quality and Value assessment module for a US based asset
management company. The module developed on Jupyter notebook with a Python kernel brought down the time of data
quality assessment of more than 500 million rows of data from 60hours to 50mins, was compatible with different types of datasets (transaction, health-care, geolocation etc.) and generated 2 reports. It allowed flexibility to the user to tailor the
conditions and checks according to the business and assess the quality of data.

THE ENERGY INSIGHTS APP - FOR A TOP ENERGY FIRM
‣ Developed a one stop application hosted on Predix platform for a fortune 100 energy firm which helps the business
development managers throughout customer engagement journey by leveraging the data available from multiple sources to have richer conversations with customers and help them in taking better decisions to achieve their energy efficiency goals
(by generating realistic project roadmap using mixed integer optimisation and linear programming) . Handled the entire
backend of the application using python.

FINDING DONORS - FOR UDACITY ML BASIC NANO-DEGREE
‣ In this project I applied and evaluated supervised learning techniques like Gaussian Naive Bayes, SVM and Random Forest on data collected for the U.S. census to help CharityML (a fictitious charity organisation) identify people most likely to donate to
their cause. Afterwards, I optimised the Random Forest model and presented it as my solution to CharityML. (GitHub)

CUSTOMER SEGMENTS - FOR UDACITY ML BASIC NANO-DEGREE
‣ In this project, I analysed a dataset containing data on various customers' annual spending amounts of diverse product
categories for internal structure using PCA and Gaussian Mixture Model clustering algorithm to best describe the variation in the different types of customers that a wholesale distributor interacts with to provide insight into how to best structure their
delivery service to meet the needs of each customer. (GitHub)

Data scientist

Mu Sigma -

July 2015 to June 2018

EDUCATION

BE in Information Technology

University of Pune - Pune, Maharashtra

2015

SKILLS

NLP, ML, deep learning, predictive modelling

ADDITIONAL INFORMATION

Technical Skills Certifications & Awards

Python NLP Machine Learning
Decision Scientist

Data Mining Linux Scala Mu Sigma University - JUN, 2017

Spark GIT IBM doCloud Machine Leaning Nano-degree

Udacity - MAY, 2018

Libraries - Data Mining

NPTEL - MAY, 2018
NLTK Sk-learn Numpy

Impact Award, Mu Sigma - APR, 2018
Mllib Pandas Gensim
Spot Award, Mu Sigma - JAN, 2018
spaCy Pattern Pulp