

PROFESSIONAL SUMMARY

Data Science Analyst and a certified **Six Sigma Green Belt** professional with a master's degree in **Industrial Engineering** having hands-on experience in leveraging **statistics, data mining** and **machine learning models** to uncover **trends, develop forecasting models**, create **dashboards & visualizations** to support executive decisioning. Ability to initiate and drive end to end project to completion with minimal guidance. **Proficient in Excel, VBA, Python, SQL, R** technologies.

EDUCATION

Master of Science in Industrial Engineering

CGPA: 3.76/4.0

University of South Florida, Tampa, FL

May 2019

Relevant Courses: Tech and Finance, Advanced Analytics, Engineering Analytics, Statistical Design Modeling, Operational Research.

Bachelor of Technology in Metallurgical Engineering

CGPA: 7.0/10

College of Engineering Pune, Pune, MH, India

May 2016

SKILLS

- Data Science: **Excel, NumPy, Pandas and Matplotlib, Django, Pyspark, Tableau, Git.**
- Programming Technologies: **Python, SQL, VBA, Shell Script, Deluge.**
- Machine Learning: **Regression, Decision Tree, Random Forest, Neural Networks, Ensembles, Gradient Boosting, Regularization, KNN, GAN, Statistical Analysis.**

WORK EXPERIENCE

Business Analyst- Senior, Citi - Capital Planning (Tampa, FL) (Contract)

February 2020- Present

Tools: R programming, Advance Excel, Power point, Outlook

- Perform advanced quantitative analysis supporting Outlook and CCAR reporting.
- Assist in performing attribution analysis to forecast scenarios for Citi retail banking.
- Developing and maintaining R scripts for annual forecasting, model variance analysis.
- Develop and modify pivot reports and power point presentations.

Data Science Analyst, Redperit LLC. (Defynance) (Tampa, FL)

July 2019- February 2020

Tools: Python, Tableau, Excel (VBA, Pivot Table, VLOOKUP), SQL, Zoho CRM, HTML, CSS, JS, Git, AWS.

- Collaborated with the senior management for new business process engineering efforts to implement analytics and automate internal processes using technologies like Python, Excel, VBA, SQL, Rest API's, Zoho CRM, AWS.
- Developed causal forecasting algorithm for pricing income share agreement.
- Implemented REST API's on excel based tool via python win32 client for **ETL data pipeline** for CRM data.
- Designed and developed analytical **visualization dashboard** for tracking monthly KPI's of Defynance ISA business using BI tool.
- Performed **data analysis** to provide **strategic recommendations** to improve operational rules of an income share agreement based refinancing product.
- Designed and developed an ISA calculator using Django web development framework and deployed it on AWS EC2 instance. Main purpose of the calculator was to gather data and improve customer interaction.

Data Science Analyst Intern, Redperit LLC. (Defynance) (Tampa, FL)

July 2018 - May 2019

- Created **forecasting models** to predict **customer lifetime value (CLV)**.
- Wrote **web scrapping and automation** algorithm using **python** libraries BeautifulSoup, Selenium, and Pandas to gather 220,000+ records by crawling **221,400+** pages.
- Performed data mining and econometric research for ISA investment fund and an annuity-based product.
- Developed an analytical dashboard tool using **V/H LOOKUP, VBA Macros, and custom functions in Excel** to calculate financial metrics to help analyze candidate risk profile based on the underwriting criteria and measure **KPI's**. Reduced entire process time by 30%.

Data Analyst, TATA Consultancy Services (Pune, MH)

January 2017-July 2017

Tools: SQL, C#, HTML/CSS, Microsoft Visual studio.

- Developed database schema for capturing the status of each inter-bank transfer transaction.
- Extracted and analyzed data from three large DB schema, provided **SQL** scripts to rectify records.
- Performed database **integration** with web interface using C#.

INTERSTING PROJECTS

Future transaction classification Kaggle challenge.

- Develop a **LGBM model** for predicting customer transaction probability using 200 anonymous features provided by Santander bank.
- I built this model with different types of **regularizations** (L1 and L2), utilizing Python libraries such as NumPy, LGBM and matplotlib. I ranked in top **10%** teams in this Kaggle competition.

Toxic comment classification

- Built a multi class classification logistic and NN model that can classify comments based on different types of toxicity (like threats, obscenity, insult and identity-based hate) in a comment with an accuracy of 97.90%.
- Performed text analytics (Word cloud, TFIDF, feature engineering) to understand the nature of dataset.

Personal Website and Data Science Portfolio

- Developed a personal website using Jekyll web framework and deployed it on Github site.
- Here I write blogs on topics like data science, predictive analytics, finance, stock investment, and some random topics etc.

Data Analytics for comparing 2 investment opportunities. ([Link](#))

- Performed statistical analysis of both investments using Sharpe Ratio in python and interpreted results of analysis.

Python google maps API to extract location details.

- Built a python module to gather location data related to any entity listed on google maps utilizing web scraping.
- This module currently generates csv file with attributes like address, latitude & longitude values, and type of property. Very useful to perform real estate valuation analysis.