

# Aditi Mulye

www.aditimulye.com | Email ID : aditi.mulye0@gmail.com | Mobile No.: +1-424-420-8338

## EDUCATION

INDIANA UNIVERSITY, Luddy School of Informatics

Bloomington, Indiana

Master of Science in Data Science | GPA: 3.8

Aug 2021 - Dec 2022

- **Coursework** : Applied Machine Learning, Introduction to Statistics, Network Science, Applied Database Technologies

MEDI-CAPS UNIVERSITY, Department of Computer Science

Indore, India

Bachelor of Technology in Computer Science & Engineering | GPA: 3.8

Aug 2016 - May 2020

- **Coursework** : Data Structures, Object Oriented Programming, Machine Learning, Design and Analysis of Algorithms

## WORK EXPERIENCE

Barclays | *Data Engineer*,

Feb 2023 - Present

- Wrote **reusable unit test** documents to assure code quality control and bug detection in the development phase, minimizing production issues by **68%**.
- Performed Continuous Integration (CI) and Continuous Deployment (CD) Jenkins pipeline creation for microservices to be deployed on Openshift v4 architecture.
- Collaborated with the Product Owner to understand business requirements, scrum planning, and backlog refinement.

Amazon | *Data Scientist Intern*

May 2022 - Aug 2022

- Performed customer segmentation and profiling of **125M+** customer records, enabling **Amazon's Choice** to accomplish their mission of 2022.
- Built an end-to-end Machine Learning project by developing a **Propensity Model** to identify the customers that have a higher likelihood to buy sustainable and environment-friendly products.
- Estimated the business impact of using the model by deriving a formula to calculate the increase in units sold of sustainable products.

Paytm | *Business Analyst*

Jun 2020 - Apr 2021

- Utilized Python and VBA to implement automation of orders of 15+ international clients and visualize previously inaccessible orders on a daily basis.
- Developed dashboards for individual clients using Tableau to keep track of changes in profitability and orders which **reduced the turnover time by 90%**.
- Designed a Python-based tool for better speed automation of daily activities carried out in Excel **thereby boosting the speed by 50%**.

University Of Cambridge, CJBS | *Data Science Intern*, Prof. Jaideep Prabhu

Jun 2019 - Aug 2019

- Trained massive datasets and performed **predictive data analysis** using Statistical and Machine Learning algorithms.
- Conducted text mining and ontological predictions on data for a behavioral study of over 200 college students.
- Used R to conduct Analysis of Variance (ANOVA) and analyze the creativity amongst college students.

Indian Institute of Management, Ahmedabad | *Analytics - Research Assistant*, Prof. Arvind Sahay

Jun 2018 - Aug 2018

- Carried out a **background study** to analyze the evolution of steel industry in India over the past 20 years (1995-2015).
- Conducted on-site interviews of 20+ customers and distributors of steel firms to understand their growth paths.
- Compiled the research as a **Management Case Study** emphasizing on future prospects and profitability of a firm.

## TECHNICAL SKILLS

- **Languages**: SQL, Python, R
- **Big Data & Machine Learning**: Python (eg. sci-kit learn, numpy, pandas, matplotlib, seaborn), Tableau, Google Data Studio, Jupyter, Hive, PySpark, Hadoop, AWS (EMR, S3, SageMaker).
- **Data Science & Miscellaneous Technologies**: A/B Testing, Data Science Pipeline (cleaning, wrangling, visualization, modeling, interpretation), Regression, Classification, Clustering, Statistics, Hypothesis Testing, OOP, CI/CD

## PROJECTS

Home Credit Default Risk | (Python)

- Combined 7 datasets with **20+ million rows and 200+ columns**, performed extensive feature engineering and exploratory data analysis to prepare data for modeling.
- Implemented Machine Learning pipeline to build a classification model for predicting the Credit Default Risk using ensemble learning methods

Adult Income Data-set | From Scratch (Python)

- Utilized Adult Census Income from UCI Machine Learning Repository to predict if a person makes over **\$50K a year**.
- Implemented Logistic Regression from scratch and compared it to sci-kit learn's inbuilt logistic function.

Mental Health at Workplace (Python)

- Designed the notebook to conduct Exploratory Data Analysis and derive logical conclusions from the data.
- Implemented logistic regression and decision tree models based on optimal variables and achieved an accuracy of **80%**.

Exploratory Data Analysis in R (R Markdown)

- Used the World Bank data to analyze patterns, identify anomalies interpret the global suicide rates between 1985-2015.