ADITYA DIWANE

201-417-0047 • Harrison, New Jersey, 07029 • Email: aditya.diwane@rutgers.edu
github.com/adityadiwane • linkedin.com/in/adityadiwane • public.tableau.com/profile/aditya.sandeep.diwane

EDUCATION

RUTGERS UNIVERSITY | RUTGERS BUSINESS SCHOOL — Newark, NJ
Master of Information Technology and Analytics (GPA: 3.6/4) — Aug 2019 - Dec 2020
Coursework: *Machine Learning & Statistics, Data Visualization, Business Data Management*

UNIVERSITY OF MUMBAI — Mumbai, India
Bachelor of Engineering (Electronics and Telecommunications) (GPA: 7.04/10) — Aug 2015 - May 2019
Coursework: *Probability, Statistics, Object-Oriented Programming, Business Communication*

WORK EXPERIENCE

INCTURE TECHNOLOGIES — Pittsburgh, PA
Data Analyst Intern — May 2020 - Present
- Heading a team of 3 cross-functional members to develop a data-driven solution for repair task management module
- Fetched sensor data in JSON format with 13 million rows from multiple APIs using python scripts & Postman API tools
- Extracted information from multiple JSON files by parsing it into a single CSV file using pivot tables & VLOOKUP
- Designed a Tableau dashboard to keep track of incomplete tasks & associated reason, increasing operator efficiency by 30%
- Developed a classifier model to provide predictions of repair task type using Machine Learning libraries available for python
- Tuned hyperparameters of the model to optimize F-1 score using GridSearchCV to improve the quality of predictions by 5%
- Achieved 70% F-1 score on testing data, which will save $91,250/Mo by reducing the time required to create repair tasks
- Assembled digital architecture & deployed the Machine Learning model to cloud to provide predictions as an API

SKILLS

Programming Languages: Python, SQL, R, VBA
Databases: MySQL, Microsoft SQL Server, NoSQL(MongoDB)
Frameworks/Libraries: Pandas, Matplotlib, SciKit-learn, Tensorflow, dplyr, shiny, ggplot2, Extract Transform Load (ETL)
Cloud Services: Amazon Web Services(Elastic Compute(EC2), Lambda, DynamoDB, S3), SAP Cloud Platform, Cloud Foundry
Tools: Tableau, Advanced MS Excel, MS Office, Power BI, MS Business Intelligence (SSIS, SSAS, SSRS), ProjectLibre, Sharepoint

ACADEMIC PROJECTS

EXPLORATORY DATA ANALYSIS OF MOVIES DATA-SET | *R | GitHub pages* — Sep 2020 - Oct 2020
- Cleaned & Analyzed IMDB Movies data-set using dplyr & ggplot2 to find associations between IMDB rating & movie revenue
- Created a profitability plot that displays Avg. IMDB score V/s Profitability for every country in the data-set
- Observed positive correlation between revenue & IMDB score, also movies with rating 8.1 and above are top 5% in the world

INSTAGRAM DATABASE CLONE ANALYTICS & REPORTS | *MS Access | Microsoft SQL Server* — Aug 2020 - Oct 2020
- Gathered requirements, designed database & generated an ER Diagram to obtain a visual overview of the database
- Authored CRUD commands to insert data for 100 users containing multiple attributes into the MS SQL Server database
- Analyzed the database & identified 13 bots that liked all the images in the database for the admin to take required actions

ENGLISH PREMIER LEAGUE 2019-20 SEASON OVERVIEW | *Tableau | Excel* — Aug 2020 - Sep 2020
- Collected data of 20 teams & 584 players spread in multiple files and homogenized it using Excel and prepared it for analysis
- Shaped interactive Tableau dashboard that filters and exhibits top-performing teams & players from each team in the league
- Consolidated dashboards into a story, aiding the users to compare the performance of all teams based on multiple metrics

SUMMARIZING YOUTUBE VIDEO TRANSCRIPTS | *Python | NLP* — Jan 2020 - May 2020
- Acquired & cleaned the text data using Regex into a suitable format for the machine learning algorithm to be ingested
- Measured performance of text summarizing algorithms using Bleu & Rogue measures and selected TF-IDF algorithm
- Attained a 75% reduction in length of transcripts by using the TF-IDF algorithm implemented using multiple NLP libraries

YOUTUBE TRENDING VIDEO CLASSIFICATION | *Python | AWS* — Sep 2019 - Dec 2019
- Performed EDA on data-set of Trending & Non-Trending video statistics & visualized the same using Python Matplotlib library
- Trained Stochastic Gradient Decent classifier to sort YouTube videos into Trending & Non-Trending categories
- Achieved 97% model accuracy by tuning the hyperparameters of the model using a Grid-search method