DEVANSHEE TANNA

Atlanta, GA | 612-800-2079 | dtanna1@student.gsu.edu | www.linkedin.com/in/devanshee-tanna08 | https://github.com/dtanna1

PROFILE

- Data Science and Analytics graduate student with a passion for insightful decision making, problem solving and driving business impact to enable data driven decision making and strategy.
- Four years of professional experience in creating analytics solutions across industries. Skilled in machine learning algorithms, quantitative analysis, unstructured data analysis, big data and visualization.
- Worked across global locations with clients to perform Data mining and Master Data Management and set up big data analytics and digital platform to help organizations generate actionable insights.

CORE COMPETENCIES

- Machine Learning: Regression, CART, Ensemble models, Clustering, Scikit learn, Time series modelling
- NLP: NLTK preprocessing techniques, topic modelling, sentiment analysis, information retrieval
- Data wrangling/analysis: Python(pandas, numpy), SQL, Apache Hive, R, Association rules
- Data management and Viz: Hadoop, Web scraping, Tableau, Google Analytics, OpenText Magellan Analytics suite

EDUCATION

- MS in Analytics - J. Mack Robinson School of Business, Georgia State University, Atlanta, USA - Dec 2020 – CGPA: 4.09
- B-Tech in Mechatronics - SRM Institute of Science & Technology, SRM University, India. – 2015 -CGPA: 8.65/10

PROFESSIONAL EXPERIENCE

Cognizant Technology Solutions, India Sep 2015 – June 2019

Associate – Analytics *Jan 2018 – June 2019*

- Developed a strategy to utilize the data captured by an Australian **Internet of Things** client for enhancing device performance. Studied the data processed by smart devices using statistical techniques such as **association rules** and analyzed correlations using Python packages apriori, pandas, matplotlib for improving device performance across the network.
- Delivered Dashboards on Retail analytics for an APAC Airport. Worked with product engineers to draw insights on **Customer behavior** and **Geographical trends** using data profiling, **data mining**, **market-basket analysis** and dashboards to increase trend visibility, customer satisfaction and optimize sales.
- Delivered an **On-Site** pilot implementation at Malaysia, to an Energy Company on **Master Data Management**. Analyzed Customer entity containing ~1,000,000 records and 200+ attributes. Eliminated inconsistencies, conducted **exploratory** data analysis and **visualization**. The pilot led to the signing of a $90,000 deal.

Data Analyst *Sep 2015 – Dec 2017*

- **BI Developer** for a Major Indian Bank to create intelligent reports for **Government Funding Schemes**. Created complex charts and crosstabs using Eclipse BIRT. Presented critical and actionable data to the client.
- **Big Data Engineer** for an American Bank to create a **Centralized Data Lake** on top of the bank's **Hadoop ecosystem** for data coming from different sources.
- Attended an official training on Big Data Analytics in Singapore, followed by certification by the client's product engineering team.

PROJECTS

TMDb Box Office Revenue Prediction:

- Utilized MultiLabel binarizer, OneHotEncoder, TF-IDF to encode features containing JSON lists of columns.
- Applied machine learning models like random forest and xgboost to **predict the revenue** of a movie with an RMSE value of 1.59. Utilized RFE feature selection technique to reduce computational power.

Analyzing the performance of Healthcare systems in Georgia:

- Performed **sentiment analysis** on patient reviews of a healthcare system extracted through Web Scraping (Python's scrapy and selenium packages). Preprocessed data using Python's NLTK package. Used **Bag of Words** for three key categories influencing the hospital reputation).
- Highlighted major factors affecting patient experience across branches of the hospital. Provided recommendations to improve 2 significant factors - the quality of care and infrastructure for hospitals with below average ratings.

Analysis of Marketing Strategies for Porsche's Electric vehicles:

- Developed a Machine learning model to target potential customers of Electric vehicles (EVs) through the information of EV registration counts per zip code and population demographics.
- Mined patterns and correlations between demographic features and EV sale count and further used factors that influenced the sales of EV to implement Lasso regression model with an R square value of 0.65.