

HIMALAY PATEL

patelhimalay1@gmail.com | +1 (408) 338-7503 | <https://www.linkedin.com/in/patelhimalay/> | <https://himalaypatel24.github.io/>

SUMMARY

Analytical and detail-oriented aspiring Data Scientist or Data Analyst possessing strong knowledge in Statistics, Machine Learning Algorithms, and Product Analytics. Actively looking for full-time role.

EDUCATION

Master of Science (MS) | Engineering Management, Specialization: Data Science **Jan 2020-Dec 2021**

San Jose State University, San Jose, CA, USA

GPA:3.66

Relevant Coursework: Big Data Management, Data Mining for Business Analytics, Engineering Probability and Statistics, Operations Planning and Control, Leading the Six Sigma Improvement Project.

Bachelor of Engineering | Mechanical Engineering

May 2016-May 2019

Gujarat Technological University, Gujarat, India

GPA: 3.41

SKILLS

- **Programming:** R, Python ,SQL, NoSQL
- **Business Intelligence :** Tableau, Power BI, MS Excel, Google Sheets
- **Machine Learning Algorithms:** Linear Regression, Logistic Regression, Decision Trees Classification, Random Forests, Time Series Forecasting, K-Mean Clustering, Ensemble Method, Naive Bayes, kNN,NLP
- **Databases:** Microsoft Access, MS SQL server, MySQL, PostgreSQL

WORK EXPERIENCE

Digidott E Solutions | Data Scientist

May 2019-Dec 2019

- Designed and developed ETL pipelines for data extraction in preparation for exploratory data analysis, modeling and reporting. Used Spark and Hadoop File System to make the process efficient, fast, and seamless.
- Identified issues in models and constructed model evaluation framework in python and generated analytical report with highly increased model accuracy and evaluated efficiency.
- Solved business problems with applying inferential statistical analysis, A/B experiments and machine learning solutions.

Coderedge | Data Analyst Intern

Dec 2018-April 2019

- Analyzed organization's sales data of revenue (\$2M) for different customers, products, and locations by executing data queries in SQL, Python, and developed Tableau dashboards and QBR report.
- Worked cross functionally across marketing and product business teams to build out the reporting structure in Tableau. Created interactive group analysis report with filters and parameters that assisted to improve **KPI's by 15%**.
- Identified procedural areas of improvement through customer data by performing business analysis and improve the **profitability of organization by 8%**.

PROJECTS

Sentiment Analyzer And Classifier (NLP,nltk,spacy,sklearn)

Aug 2021

- Performed Sentiment analysis on more than 100K rows of amazon fine food reviews with span a period of more than ten years. Refined the data with preprocessing methods such as stop words removal ,lemmatization, etc.
- Implemented different vectorization techniques like count vectorizer, bigram-ngram, and tf-idf to encode the text to numeric data. Developed the models that were up to **87% efficient** to determine the sentiment of a given review.

Air Fare Prediction (Random Forest Regressor,Label Encoder,Mutual Information)

June 2021

- Performed data preprocessing, feature engineering, feature selection over 100k rows of data with more than 30 features to identify the features that affect the target variable Price the most.
- Trained Machine Learning models to predict the prices of flight tickets for various airlines, also compared their performance and hyper-tuned models which were approximately give **85% R-square value**.

Yelp Dataset Challenge Big Data (SparkSQL, Pyspark, NoSQL)

Oct 2020 -Dec 2020

- Performed data wrangling over 8 million reviews running SQL queries on databricks with applying **CoNVOs** framework. Compared elite and non-elite reviews for voter count using Box-whisker Plot, Stacked Bar graphs in Tableau.
- Identified **Elite users with 35%** of total users.

Other Projects: (<https://github.com/himalaypatel24>)