

HIMALAY PATEL

himalay.patel@sjsu.edu | +1 (408) 338-7503 | <https://www.linkedin.com/in/patelhimalay/> | <https://github.com/himalaypatel24>

SUMMARY

Analytical and detail-oriented aspiring data Data Scientist or Data Analyst possessing strong knowledge in Statistics, ML Algorithms, Data Analytics, and Visualization, and Six-Sigma. Actively looking for full-time role.

EDUCATION

Master of Science (MS), Engineering Management | San Jose State University – CA, USA GPA: 3.66 | Jan 2020-Present

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 | GTU, India

GPA: 3.41 | May 2016-May 2019

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
- **Tools:** Spark, Anaconda, PyCharm, MS Office, R studio, Minitab

WORK EXPERIENCE

Digidott E Solutions- Ahmedabad, India

June 2019-Oct 2019

Data Analyst Intern

- 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-Sept 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.

KPMG Data Analytics Virtual Internship (Pandas, Numpy, Matplotlib, Seaborn)

June 2021-July 2021

- Successfully classified **high value customers** based on customer demographics and attributes using **RFM analysis** marketing technique for Sprocket Central Pty Ltd, a medium size bikes & cycling accessories organization.
- Transformed raw data into meaningful insights by applying **exploratory data analysis**. Recognized **top three** industry sector contributing the maximum profit.

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

May 2021-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:

911 Emergency Calls-EDA | Cricket ODI Analysis EDA | House Prices: Advanced Regression Techniques