RISHI DHAKA

Austin 78701, Texas, USA | rishi.dhaka@gmail.com | +1 737-217-9097

EDUCATION

University of Texas at Austin, Austin, Texas, USA **Aug 2019 - May 2021**
Masters of Science in Information Systems *GPA : 3.93*

Manipal Institute of Technology, Manipal University, Karnataka, India **Aug 2014 - May 2018**
Bachelor of Technology in Electronics and Communications Engineering *GPA : 3.4*

SKILLS

Languages and Suites: Python, SQL, Tableau, Excel, PHP, HTML
Popular Libraries: numpy, pandas, TensorFlow, Scikit-Learn, nltk, matplotlib, BeautifulSoup
Statistical Skills: Regression, Classification, Hypothesis Testing, Clustering
Agile Development

EXPERIENCE

Data Science Researcher - *University of Texas Development and Analytics, Austin, TX* *Jan 2021 - present*

- Scraped 2250 IRS 990 forms of foundations as XML files from an AWS server. Improved data collection efficiency by 53.5% and reduced resource engagement time by 33%. **(fuzzy data matching, clustering, NLP)**
- Raised the university's fundraising revenue by $1.4M per year by implementing predictive models to gauge the giving capacity of foundations to facilitate the university's fundraising. **(k-Nearest Neighbours)**

Data Science Research Assistant - *University of Texas at Austin, TX* *Jan 2020 - Dec 2020*

- Web scraped a dataset of 4000 questions on user data privacy and their metrics (titles, summaries, tags, accepted answers, most voted answers) from StackOverflow. **(Python, BeautifulSoup)**
- Created a bag of words corpus using the downloaded text. **(gensim, pandas, nltk)**
- Used Natural Language Processing for analysing the text data to find trends in developers' questions. **(Latent Dirichlet Allocation for topic modeling, n-gram algorithms, TF-IDF and lemmatization)**

Associate Product Specialist - *Sabre Travel Tech., Bangalore, India* *Oct 2018 - Jul 2019*

- Analysed and resolved complex problems at multiple layers (application, environment, database, networking) using proprietary software and diagnostics tools. **(Postman, SQL, API Testing)**
- Saved Sabre $300,000 in fines to the client by resolving 25 high severity issues preventing system failure.
- Trained on product installation and data migration on HP servers.

Associate Product Specialist - *Sabre Travel Tech., Singapore* *Aug 2018 - Sep 2018*

- Developed an employee leave module for SilkAir crew members. Defined criteria for leave allocation. Synched the module with the customer database to ensure crew rosters were published on time. Processed leave requests, and used web services to transfer the data as XML files. **(SQL, MS Excel, Postman)**
- Saved the client $32,000 by publishing the final rosters a month in advance.

ACADEMIC PROJECTS

Time Series Analysis of Superstore Sales Data *(Python, pandas, numpy, Forecasting)*

- Implemented ARIMA time series forecasting on Superstore furniture and office supply sales.
- Used Prophet forecasting tool to compare the trends and patterns of the Superstore sales.
- Obtained a root mean square error of 151.46 using this model.

Premier League Analysis Dashboard *(Tableau, MS Excel, Data Visualisation)*

- Created a dashboard on Tableau presenting Premier League team performances over the past 10 years.
- Analysed patterns, relationships, and trends in the teams' performances over the years.

Automated Database for Sentiment Analysis of Song Lyrics *(Python, SQL, nltk, pandas, NLP)*

- Created a web page using PHP/HTML that enabled users to web scrape song lyrics of various artists.
- Created a bag of words corpus using the lyrics from various artists' songs.
- Applied Natural Language Processing for Sentiment Analysis and Lexical Analysis on the lyrics.

Personal Protective Equipment Detection *(Python, YOLOv3, CNN, Computer Vision, Object Detection)*

- Implemented a safety system to detect PPE compliance and fires in construction sites and industries.
- Employed Computer Vision (Convolutional Neural Network-YOLO) for object detection in videos.
- Obtained a mean average precision of object detection of 77.58%.