KEVIN DOSHI

Austin, Texas 78704 doshikevin95@gmail.com
https://www.linkedin.com/in/kevin-doshi +1 (732) 983 1042
https://github.com/kevin21895 https://kevin21895.github.io

PROFESSIONAL EXPERIENCE

Hewlett Packard Enterprise Company **January 2022 – Present**
Data Scientist IV **Austin, Texas**

- Leading the development of machine learning model to determine the misclassified purchase orders based on the key components using Python that will save 30 hours of human efforts monthly.
- Developing a time-series model to forecast the sell out for next quarter using historical data & other indicators using R, SQL which will result in $100,000 savings annually.
- Designing a classification algorithm to identify fraudulent credit-card transactions based on the user's spending behavior using Python, SQL which will help reducing human efforts by 30%.
- Developing automated solution to load the data from data warehouse for different use cases and solution using Python, Hive.
- Developing a web-based solution to automatically ingest data from end users using Django, SQL to manage organization budget efficiently and reduce human efforts by 50%.

Hewlett Packard Enterprise Company **January 2020 – December 2021**
Data Scientist III **Roseville, California**

- Developed a production ready machine-learning model to detect anomalies and potential frauds in the purchase orders resulting in reducing review of ambiguous POs for attorneys by 50% using Python, SQL.
- Designed a machine learning model using K-Means algorithm to cluster similar vendors based on cosine similarity index from credit-card transactions to get a consolidated spend across HPE.
- Performed topic analysis using Latent Dirichlet Allocation algorithm and sentiment analysis to extract meaningful insights from executive survey data in-order to identify improvement opportunities for internal teams using R.
- Developed a data pipeline to integrate more than 300 million rows from 2 different data sources efficiently and performed calculations to detect grey marketing activities of HPE products.
- Built a framework for the team to automate data loads from SharePoint sites which results in saving of 20 hours of human efforts per week using Python and REST APIs.
- Spearheaded development of a web-based application using Django to get user's feedback on the model results that reduced human efforts by 5%.
- Designed data pipeline to automate loads from 7 different data sources and calculated carbon emissions to understand HPE's carbon footprint using web scraping technologies, SQL, Python which saves company $50K annually.

Hewlett Packard Enterprise Company **September 2018 – January 2020**
IT Engineer – Analytics **Santa Clara, California**

- Automated data loads and developed visualization of various Networking and Campus use-cases using Python, SQL
- Defined key performance indicators (KPIs) for various networking, campus, and security use-cases.
- Designed interactive real-time dashboards and executive reports for the senior management & marketing team.
- Developed an advanced analytical solution for the networking team to identify network threats and alerts using Python, PySpark, Natural Language Processing (NLP) Toolkit
- Built dashboards to monitor Conference room usage, Guest Registration, Service Desk Metrics, Network Usage using Microsoft Power BI.

Hewlett Packard Enterprise Company **May 2018 – August 2018**
Data Engineering Intern **Santa Clara, California**

- Gathered requirements and designed a database architecture for the Radio Frequency test data.
- Remodeled legacy analysis tool using SQL, Python, VBA which resulted in 30% reduction of human efforts.
- Performed Regression Analysis to reduce the number of test flows which reduced the testing time by 20% using Python, SQL.

EDUCATION

Rutgers University, Rutgers Business School – Newark & New Brunswick, New Jersey **January 2019**
Master of Information Technology and Analytics **GPA:** 3.80/4

University of Mumbai - Mumbai, India **May 2017**
Bachelor of Engineering, Information Technology **GPA:** 3.73/4

TECHNICAL SKILLS

- **Programming Languages:** Python, SQL, R, Java, C, VBA, PySpark, Hive
- **Machine Learning & Statistical Techniques:** Time Series Analysis, Regression, Classification, Clustering, Tree-based Algorithms
- **Data Visualization & Analytics Tools:** Microsoft Power BI, Tableau, Advanced Excel, Qlik Sense
- **Web Development Technologies:** Django, JavaScript, PHP, HTML, CSS