PRATHEEP NADAR

Mumbai, India | +91 8976504763 | pratheepknadar@gmail.com

Kaggle Profile - https://www.kaggle.com/pratheepknadar

Professional Summary

- Around 11 months of experience as a Data Analyst with a good understanding of evaluating data sources, data warehouse, SQL, Tableau, Power BI, Python, AWS, and SQL Server.
- Proficiency in all phases of the Software Development Life Cycle (SDLC) using methodologies like Agile and Waterfall.
- Great knowledge of AWS Cloud platform, Google Cloud Platform (GCP), and its features.
- Good knowledge of Python-based environment, along with data analytics, data wrangling and MS Excel data extracts and Familiar with Pandas, NumPy, Matplotlib and SciPy.
- Proficient in developing different reports, analytics, data mining solutions to various businesses generating data visualizations using Python, Tableau, Power BI, QlikView and Google Sheets.
- Capable in the development of Big data projects using Hadoop, Hive, Spark, HBase, Pig and open-source tools/technologies.
- Strong experience with Database performance tuning and optimization, query optimization, index tuning, caching and buffer tuning.
- Working knowledge of relational databases like MySQL, SQL Server and MS Access.

Skills

Languages: Python, R

Database Systems: MySQL, SQL Server, MS SQL, T- SQL, PL/ SQL

Visualization Tools: Tableau, Power BI, MS Excel (Pivot tables and charts, Power Query)

Python Packages: NumPy, Pandas, Matplotlib, Seaborn, SciPy

Data Analytics: ETL Concepts, Google/ Adobe Analytics, SSIS, Pentaho

Certification: Tableau Desktop Specialist

Education

Master of Science, Information Technology and Management

The University of Texas at Dallas - Richardson, Texas May 2021

Relevant Coursework: Applied Machine Learning, Database Management, Object Oriented Programming (OOPS) with Python, NLP, Business Analytics using R, Statistical Data Analysis

Bachelor of Science, Electronics and Telecommunication Engineering

University of Mumbai - Mumbai, India June 2018

Relevant Coursework: Calculus and Probability, Data Structures and Algorithms, Introductions to Neural Networks.

Work History

Data Analyst Sep 2021 – July 2022

Pizza Hut LLC Plano, TX

- Tracked the revenue of franchise to ensure smooth transition of ownership from one franchise owner to another.
- Removed redundancy in survey questions from customer feedback platform based on correlation matrix values.
- Built intuitive dashboards in tableau to visualize the YoY sales growth, average guest check size for franchise reporting
- Generated automatic weekly performance reports using Python, sent out the franchise and reduced time by 50%
- Cleaned, mapped, and organized two years of instore data using Python scripts using NumPy, pandas libraries, etc.
- Used Alteryx and Python to automate ETL jobs for update tables inside snowflake cloud warehouse daily.
- Identified integration impact, data flows and data stewardship.
- Generated, wrote and run PL/SQL script to implement the DB changes including table update, addition or update of indexes, creation of views and store procedures.
- Consolidated and updated various data models through reverse and forward engineering.

Data Analyst Intern July 2020 – Aug 2020

APILATION.AI INC Allen, TX

- Built an automated ETL process to integrate data from multiple sources in AWS S3 and load using data integration tools for further analytics and saved time on implementation by 100 hours
- Extracted data from 20+ tables in database using PL/SQL queries for creating powerful dashboards in Tableau for reporting to Stakeholders and management
- Interacted with stakeholders and data engineers to put into production and to implement the model within AWS EC2 for live forecasting in client's monthly and weekly revenue with over 82% accuracy

PROJECTS | (https://github.com/pratheepknadar)

Online Payments Fraud Detection with Machine Learning Jan 2020 – Feb 2020

- Developed a machine learning-based online payment fraud detection system using Python, aimed at identifying fraudulent transactions and ensuring customer protection.
- Utilized a dataset from Kaggle, consisting of historical information about fraudulent transactions, to train a machine learning model for classifying fraudulent and non-fraudulent payments.
- Implemented a Decision Tree classifier to train the fraud detection model, using features such as transaction type, amount, old balance, and new balance.
- Achieved a high accuracy score of 99.97% on the test set, demonstrating the effectiveness of the machine learning model in accurately classifying fraudulent transactions.
- Applied the trained model to predict whether a transaction is fraudulent or not, providing valuable insights to credit card companies for real-time fraud detection and prevention.

Future Sales prediction with machine learning Dec 2019 – Jan 2020

- Analysed the dataset containing information about advertising costs on TV, radio, and newspapers, along with the corresponding sales data.
- Conducted data preprocessing, including checking for null values and ensuring data quality.
- Visualized the relationships between advertising costs on different platforms (TV, radio, and newspapers) and units sold using scatter plots.
- Assessed the correlation between each feature and the target variable (sales) to identify the most influential factors.
- Used the trained model to predict future sales based on given advertising cost inputs.
- Demonstrated the value of future sales prediction in managing manufacturing and advertising costs for businesses.

Movie Recommendation system using Machine Learning Sept 2019 – Oct 2019

- Performed exploratory data analysis and Pre-processed the raw data consisting of 105339 ratings applied over 10329 movies in R.
- Normalized the data, did feature selection and found out the correlation between them using heatmap.
- Developed an item-based collaborative filtering system which finds similarity in items based on people's rating of them.