ANKUR SHAH

Data Scientist | Data Analyst

• 346-542-2320 • work.ankurshah@gmail.com • Houston TX • https://www.linkedin.com/in/ankurshah02/ • https://github.com/214SF •
• https://public.tableau.com/app/profile/ankurshah02 • https://214sf.github.io/ankurshah02.github.io/ •

EDUCATION

University of Houston	**Houston, Texas**
Master of Science in Electrical Engineering **GPA: 3.7**	May 2024
Narsee Monjee Institute of Management Studies (NMIMS)	**Mumbai, India**
Bachelor of Technology in Electronics and Telecommunications **GPA: 3.75**	May 2021

WORK EXPERIENCE

Student Worker Data Analyst	**Sep 2023 - Present**
University of Houston • Part-time	Houston, TX

- Optimized **SQL queries** for the university admission website, resulting in a 50% increase in upload speed of test scores to the university database, enhancing efficiency for student applications.
- Designed and executed **ETL processes**, leveraging advanced **SQL techniques**, leading to a 40% reduction in data processing time and a 20% improvement in data accuracy.
- Utilized advanced **SQL queries** to extract Missing Student Data from **Amazon Web Services**, ensuring comprehensive student records for the University Portal.
- Optimized team workflow by 20% through redesigned **documentation processes**, employing data tracking techniques and automated reporting for troubleshooting missing transcript errors.

Data Analyst	**Jun 2023 – Aug 2023**
Dog Hugs Cat • Internship	Houston, TX

- Employed advanced **data visualization techniques**, including **SQL queries**, to present A/B testing results, fostering collaboration and alignment on website design decisions.
- Collaborated with cross-functional teams to develop advanced predictive data models, utilizing **SQL queries** for data extraction and analysis, resulting in 95% accurate market trend anticipation.
- Reduced processing time by 30% and enhanced **data visualization** capabilities by 15% through the development of complex **data models** using advanced **SQL queries**.
- Identified and analyzed KPIs in team meetings, leveraging **SQL queries** and statistical analysis to create **PowerBI** data visualizations, resulting in actionable insights for informed decision-making.

Data Scientist	**May 2021 – Apr 2022**
Trumen Technologies Pvt. Ltd. • Full-time	Indore, India

- Collaborated with software engineers and data analysts to refine **machine-learning algorithms** for **image processing**, achieving a 10% improvement in accuracy and a 95% success rate in identifying sensor types.
- Conducted comprehensive business data mining using **SQL**, resulting in a 15% boost in customer satisfaction by identifying pivotal trends and patterns to inform product development strategies.
- Executed **data cleaning** and **profiling processes** on 10,000 unstructured datasets using SQL and Excel, resulting in a 95% improvement in data accuracy and reliability for advanced analytics projects.
- Developed and executed a data-driven **Tableau dashboard** highlighting sensor performance, employing **statistical analysis techniques** to enhance data insights and facilitate decision-making.

PROJECTS

Exoplanet Detection using Machine Learning
- Achieved a 99% success rate in exoplanet detection by conducting in-depth analyses of various **Machine Learning (ML) algorithms**.
- Constructed a robust database by compiling datasets from **PHLEC, Kepler Data set, and TESS**, facilitating advanced analytics and modeling.
- Enhanced **Keplerform ML Algorithm's** performance to achieve 93% accuracy in exoplanet detection through advanced feature engineering.

Hospital Management System
- Developed comprehensive **Data Model and Entity-Relationship Diagram (ERD)** using industry-standard business rules, showcasing adeptness in client collaboration and business analytics.
- Integrated **XAMPP server**, reducing data retrieval time by 30% and enhancing system performance.
- Optimized data processing workflows with complex **SQL queries**, resulting in a 20% enhancement in data accuracy and efficiency.

TECHNICAL SKILLS

- **Programming Language**: Python, R, Scala, SQL, MATLAB, HTML, PySpark, CSS
- **Databases:** NoSQL, Data Warehouse, SQL Server Analysis Services (SSAS), SQL Server Reporting Services (SSRS)
- **Data Processing & Visualization Tool**: Tableau, Power Bi, Apache Spark, Excel, NumPy, Pandas, Jupyter Notebook
- **Statistical Techniques & Algorithms:** Data Analysis, Data Mining, Statistical Analysis, ETL, ELT, Data Visualization, Hadoop
- **Python Package**: TensorFlow, NumPy, SciPy, Matplotlib, Scikit-Learn, Keras, XGBoost, OpenCV, PyTorch
- **Machine Learning:** Linear\Logistic Regression, Computer Vision, Natural Language Processing, Random Forest, RNN, Artificial Intelligence (AI)
- **Microsoft Office 365:** Word, PowerPoint, Excel, Outlook
- **Cloud Platforms**: AWS, Azure, GCP

CERTIFICATION

• Data Visualization Operating Tableau and Paraview by Hewlett Packard Enterprise Data Science	**May 2023**
• Scientific Programming with Python/R by Hewlett Packard Enterprise Data Science	**Mar 2023**