

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

Master of Science in Electrical Engineering GPA: 3.7

Houston, Texas

May 2024

Narsee Monjee Institute of Management Studies (NMIMS)

Bachelor of Technology in Electronics and Telecommunications GPA: 3.75

Mumbai, India

May 2021

TECHNICAL SKILLS

- **Programming Language:** Python, R, Scala, MySQL, SQL, MATLAB, HTML, PySpark, CSS.
- **Databases:** NoSQL (MongoDB), SQL Server, Oracle, 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, Artificial Intelligence (AI).
- **Machine Learning:** Linear Regression, Logistic Regression, Neural Network, CNN, NLP, Random Forest, Image Classification, RNN.
- **Microsoft Office 365:** Word, PowerPoint, Excel, Outlook.
- **Cloud Platforms:** AWS, Azure.

WORK EXPERIENCE

Student Worker Data Analyst

University of Houston • Part-time

Sep 2023 - Present

Houston, TX

- Optimized **SQL queries** and scripted in **Python** 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** and **Python scripting**, 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.
- Enhanced team efficiency by 20% through revamped documentation processes, employing **data tracking techniques** and automated reporting for troubleshooting missing transcript errors.

Data Analyst

Dog Hugs Cat • Internship

Jun 2023 – Aug 2023

Houston, TX

- Employed advanced **data visualization techniques**, including **Python libraries** and **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 Key Performance Indicators (KPIs) during team meetings, utilizing **SQL queries** and **statistical analysis** to create data visualizations on **PowerBI dashboard**, enhancing team productivity by 15%.

Data Scientist

Trumen Technologies Pvt. Ltd. • Full-time

May 2021 – Apr 2022

Indore, India

- Collaborated with software engineers and data analysts to refine machine-learning algorithms for **image classification**, 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

- Demonstrated exceptional abilities in conducting in-depth analyses of various **Machine Learning (ML) Algorithms**, resulting in the identification of the most accurate model with an outstanding 99% success rate for exoplanet detection.
- Compiled datasets from **PHLEC**, **Kepler Data set**, and **TESS** to construct a robust database, facilitating advanced analytics, comprehensive analysis, and sophisticated modeling.
- Utilized advanced **feature engineering techniques** to enhance the performance of the **Keplerform ML Algorithm**, achieving an unprecedented 93% accuracy in exoplanet detection.

Hospital Management System

- Developed and applied industry-standard business rules to create a comprehensive and accurate **Data Model** and **Entity-Relationship Diagram (ERD)**, demonstrating adeptness in client collaboration and business analytics.
- Implemented **XAMPP server integration**, seamlessly connecting the front-end with the database and reducing data retrieval time by an impressive 30%, thereby enhancing overall system performance.
- Designed and implemented data models and optimized data processing workflows, utilizing complex **SQL queries** to extract and analyze data from multiple databases, resulting in a 20% enhancement in data accuracy and efficiency.