# 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

## 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 & Analytical Tools**: Tableau, Power Bi, Apache Spark, Excel, NumPy, Pandas, Jupyter Notebook.
- **Statistical Techniques & Algorithms:** Data Analysis, Data Mining, Statistical Analysis, ETL, ELT, Data Visualization.
- **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**                                                                          **Sep 2023 - Present**
University of Houston • Part-time                                                                              Houston, TX
- Optimized **SQL queries** for the university admission website, leading to a **50% increase** in upload speed of test scores to the university database, resulting in a more efficient process for student applications.
- Developed and implemented **ETL processes** for data integration, resulting in a **40% reduction** in data processing time and a **20% increase** in data accuracy.
- Captured Missing Student Data from **Amazon Web Services** using **SQL Queries** to Ensure Precise and Comprehensive Student Records for the University Portal.
- Revamped **documentation processes** to streamline troubleshooting of missing transcript errors, resulting in a **20% increase** in team efficiency through **data tracking** techniques and **automated reporting**.

**Data Analyst**                                                                                       **Jun 2023 – Aug 2023**
Dog Hugs Cat • Internship                                                                                      Houston, TX
- Implemented advanced **data visualization** techniques to present **A/B testing** results to cross-functional teams, leading to a **15% increase** in collaboration and alignment on website design decisions.
- Developed advanced predictive **data models** by collaborating with **cross-functional** teams, allowing the company to anticipate market trends with **95% accuracy** and gain a competitive edge in strategic decision-making.
- Developed complex data models using **advanced SQL queries** to extract and analyze data from the **data warehouse**, resulting in a **30% reduction** in processing time and a **15% increase** in data visualization capabilities.
- Identified and analyzed **Key Performance Indicators (KPIs)** during **team meetings**, utilizing statistical analysis to create **data visualizations** on the **PowerBI** dashboard, resulting in a **15% increase** in team productivity.

**Data Scientist**                                                                                     **May 2021 – Apr 2022**
Trumen Technologies Pvt. Ltd. • Full-time                                                                      Indore, India
- Collaborated with software engineers and data analysts to refine the machine-learning algorithm for **image classification**, resulting in a 10% improvement in accuracy to achieve an impressive **95% success** rate in identifying sensor types.
- Analyzed company databases through comprehensive **business data mining**, identifying pivotal trends and patterns to inform product development strategies, resulting in a **15% boost** in customer satisfaction.
- Executed **data cleaning** and **data profiling** processes on a diverse range of 10,000 unstructured datasets using SQL and Excel, resulting in a 95% improvement in data accuracy and reliability for advanced analytics projects.
- Collaborated with cross-functional teams to develop and execute 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.