ANKUR SHAH

work.ankurshah@gmail.com | Houston, TX | 346-542-2320

LinkedIn/ankurshah02/ tableau/ankurshah02/ GitHub/214SF website/ankurshah02.github.io/

WORK EXPERIENCE

Data Analyst **Sep 2023 - Present**
University of Houston Houston, TX

- Optimized SQL queries for the university admission website, resulting in a 50% increase in the 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.
- Implemented advanced SQL queries to extract missing student data from Amazon Web Services, ensuring comprehensive student records for the University Portal.
- Developed a comprehensive documentation process for missing student records, enabling analysis of system performance and identification of areas for improvement, resulting in a 15% increase in data accuracy.

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

- Deployed advanced data visualization techniques, including SQL queries, to present A/B testing results, fostering collaboration and alignment on website design decisions.
- Teamed up 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.
- Identified key trends and insights from the MapReduce analysis of Google Keywords, enabling the development of targeted SEO strategies that led to a 50% improvement in click-through rates on search engine results pages.
- Employed SQL queries and statistical analysis to identify key performance indicators (KPIs) in weekly team meetings, leading to interactive Power BI data visualizations and a 15% increase in actionable insights empowering informed decision-making.

Data Scientist **May 2021 – Apr 2022**
Trumen Technologies Pvt. Ltd. Indore, IN

- Utilized advanced machine learning techniques to develop an object detection algorithm that accurately identified various types of sensors, aiding workers in selecting the correct leveling sensor for specific locations within containers.
- 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 data 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.

ACADEMIC PROJECTS

Exoplanet Detection Using Machine Learning

- Constructed a new Database for Machine Learning (ML) algorithms to detect exoplanets by combining datasets from PHLEC, Kepler Data set, and TESS, leveraging a multidisciplinary approach to enhance accuracy and robustness.
- Employed various machine learning algorithms to determine the most effective method for exoplanet detection, achieving 99% efficiency, enhancing capabilities and scientific understanding in exoplanetary research.
- Created Keplerformer, a transformer model optimized for time series data, achieving a validation accuracy of 93.3%. This model was tailored to address the specific challenges of time series analysis, enhancing accuracy and performance.

Hospital Management System

- Applied advanced algorithms and techniques to optimize business rules for data modeling, resulting in the creation of a highly efficient relational database system with a 98% decrease in query response time.
- Formulated complex SQL queries to efficiently create a hospital database, allowing for the upload and retrieval of data from the system, resulting in a 40% increase in data accuracy and availability.
- Enhanced data retrieval processes by 40% through the development of a seamless integration between the database and frontend using XAMPP, optimizing efficiency and reducing loading times.

EDUCATION

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

TECHNICAL SKILLS & CERTIFICATION

- **Programming Languages**: Python, R, SQL, HTML, PySpark, CSS, Scala, Git
- **Databases:** NoSQL, Data Warehousing, MongoDB, Big Data
- **Data Processing & Visualization Tools**: Tableau, Power BI, Apache Spark, Excel, Pandas, Jupyter Notebook
- **Statistical Techniques & Algorithms:** Data Analysis, Data Mining, Statistical Analysis, ETL, ELT, Data Visualization, Hadoop
- **Python Packages**: TensorFlow, NumPy, SciPy, Matplotlib, Keras, XGBoost, OpenCV, PyTorch, Scikit-Learn
- **Machine Learning:** Neural Networks, Computer Vision, Natural Language Processing, Random Forest, RNN, Artificial Intelligence (AI)
- **Cloud Platforms**: AWS, Azure
- **Data Visualization Tableau and Paraview** by Hewlett Packard Enterprise Data Science **Expires May 2030**
- **Scientific Programming with Python/R** by Hewlett Packard Enterprise Data Science **Expires Mar 2030**