NEKKANTI SATYA PRASANT SHOURI

nsatyaprasantshouri@gmail.com|+1 (224) 877- 4453|LinkedIn|https://shouri.netlify.app/

SUMMARY

Data Analyst with over 2 years of professional experience. Experienced with SQL Server, PowerBI, and Tableau for data analysis and visualization. Skilled in using advanced analytical and problem-solving abilities to derive actionable insights and support data-driven decision-making. Skilled in deploying scalable, cost-effective cloud solutions using AWS and Azure. Expertise in data storage, processing, analysis, and reporting and enhancing processes through innovative solutions.

EDUCATION

Master of Science in Management Information Systems ***Aug'2022 - May2024***
Northern Illinois University
Bachelor of Technology in Computer Science ***June2016 - June2020***
GITAM(Deemed to be) University

TECHNICAL SKILLS

Methodologies:	SDLC, Agile, Waterfall
Programming Languages:	Python, HTML, MySQL, CSS, R programming
Packages:	NumPy, Pandas, Matplotlib, SciPy, ggplot2, Scikit learn, PySpark, Keras, TensorFlow
Database:	MySQL, MongoDB, PostgreSQL, Azure Cosmos DB
Visualization:	Tableau, PowerBI, Microsoft Excel, Seaborn, SAS Viya
Cloud computing:	AWS Services: EC2, S3, IAM, RDS, Lambda, Redshift, QuickSight, Azure: Blob storage, Data Lake, Azure key vault
Cloud Technologies:	AWS, Microsoft Azure
Version Control Tools:	Git, GitHub
Machine Learning Techniques:	Linear & Logistic Regression, Random Forest, k-nearest Neighbour (KNN)
Data Analysis:	Data Analysis, SAS, Data Quality Checks, Statistical Techniques, Statistical Regression

PROFESSIONAL EXPERIENCE

Data Analyst Intern ***Aug'2024 - Current***
Advansoft International Inc.**,**

- Utilized **SAS Enterprise Miner** to perform **Chi-square analysis**, ensuring the validity of the data and detecting the significant patterns in demographic studies.
- Leveraged **SAS** capabilities for large-scale Chi-squared analyses, uncovering critical associations between clinical trial outcomes and treatment groups in healthcare research.
- Visualize the data into various charts to understand the trend, patterns, and correlations using **Power BI**
- Mapped data between source systems, **data warehouses**, and data marts, facilitating seamless integration for customer usage analysis, network optimization, and performance reporting.
- Applied various **Excel** functions such as **pivot tables**, and **VLOOKUP** to aggregate data

Survey Analyst/ Phone Interviewer ***Mar'2023 - Dec'2023***
Northern Illinois University

- Worked on behalf of the **IDPH** to perform a **BRFSS** study by gathering comprehensive data on respondents' health conditions and other relevant information through structured phone interviews.
- Leveraged **MS Excel** to create well-structured spreadsheets, ensuring accurate and organized data recording.
- Utilized **SQL** to store, manage, and retrieve data efficiently, ensuring data integrity and accessibility for analysis.
- Created Pivot Tables in **MS Excel** to summarize the frequency of various health conditions by zip code, providing a clear and concise overview of the data.
- Developed visual reports using **MS Excel** to illustrate the findings, highlighting key health trends and patterns across different communities.
- Compiled and submitted detailed reports to the **IDPH**, enabling them to implement effective and essential health programs tailored to the needs of specific communities based on zip code data.

Business Data Analyst ***Feb'2021 - July2022***
Sudharma Infra Pvt Ltd., India

- Gathered data from various sources including project management software, ERP systems, field reports, IoT devices, and sensors, ensuring data integrity by cleaning and preprocessing data using **Python** and **SQL** to improve accuracy by 80%.
- Maintained **Azure SQL** Database to manage large volumes of data like material usage, labor productivity, and equipment utilization
- Utilized statistical models and machine learning techniques in **Python**, including the Critical Path Method (CPM), to analyze dependencies and optimize project schedules, which helped to effectively predict future trends, enhance project timelines, and mitigate potential issues, resulting in an 80% increase in forecasting accuracy.

- Analyzed cost data using **MS Excel** to identify areas of potential savings and cost overruns, aiding in budget management and reducing costs by 15%.
- Created visualizations such as charts and graphs in **Power BI** to streamline data analysis enhancing data visualization and stakeholder decision-making, resulting in a 20% reduction in reporting time.
- Conducted ad-hoc analyses using **Python** and **Microsoft Excel** to support decision-making processes for project managers, executives, and other stakeholders.
- Assessed data to identify potential risks and issues in projects, providing insights to mitigate these risks, resulting in a 20% reduction in project delays.
- Collaborated with project managers, engineers, architects, and other stakeholders to understand their data needs and deliver actionable insights.

Data Analyst Intern ***Aug'2020 - Jan'2021***
DISYS Pvt Ltd., India

- Leveraged **SQL Server** and **PowerBI** to streamline data analysis for finance, sales, and marketing, enhancing data visualization and stakeholder decision-making, resulting in a 20% reduction in reporting time.
- Increased sales conversion rates by 15% through data analysis and targeted marketing campaigns using **SQL Server** and **PowerBI.**
- Developed an 85% accurate crop yield prediction model using data analysis of environmental factors and advanced **algorithms**, results led to a 10% reduction in crop waste and a projected annual saving for the company.
- Created dynamic dashboards and reports in **PowerBI** and **Tableau** for tracking model performance and interactive data exploration.
- Developed custom functions and algorithms in **Python** to extract meaningful insights from complex data structures, contributing to data-driven decision-making.
- Worked closely with cross-functional teams, including developers, business analysts, and project managers, to ensure that all technical and business requirements were met.

ACADEMIC PROJECTS

Vehicle Reliability Dashboard

- Created a dashboard using **Tableau** to evaluate the dependability of various automakers using information from social media, consumer reports, and other data sources integrated with **SQL databases**.
- Incorporated interactive elements for easier navigation, such as tooltips and filters designed with Tableau's UI capabilities.
- Included user review sentiment analysis performed using **Python** (NLTK and TextBlob) to determine perceived dependability.
- Proven expertise in user interface design with **Tableau**, data integration using **SQL** and **Python,** and advanced data visualization.

AWS Data Pipeline Implementation

- Implemented **AWS Lambda** function to trigger automated ingestion whenever new data files were uploaded to **Amazon S3**.
- Integrated **AWS Kinesis** Data Streams to handle and process the streaming data in real-time, ensuring timely updates and minimal latency in data processing.
- Set up and maintained data schemas and metadata using **AWS Glue** Data Catalog, ensuring consistency and organization across datasets stored in Amazon Redshift.
- Connected **Amazon QuickSight** to **Amazon Redshift** to create interactive dashboards and visualizations.
- Developed reports and dashboards to provide insights into key metrics and trends, enabling stakeholders to make informed decisions based on the analyzed data.

Tourism Data Analysis and Predictive Modeling

- Created models using **scikit-learn** to predict how tourists rate destinations and classify them by importance.
- Applied analytics techniques like linear regression with **NumPy** and decision tree classification using **pandas** to derive actionable insights in the tourism sector.
- Streamlined resource allocation and prioritized investments based on destination significance, leveraging **SQL** databases for data management and **Matplotlib** for visualizing results, maximizing efficiency and visitor satisfaction.
- Utilized **Matplotlib** and **Seaborn** to create bar plots, scatter plots, and histograms, enhancing data analysis by providing exploration, storytelling, and decision-making.

Crop yield Prediction

- Leveraged big data processing techniques using **Apache Hadoop** to improve accuracy and scalability in predicting crop yields based on soil information.
- Applied machine learning algorithms such as **Random Forest**, **Apriori**, and **K-Nearest Neighbors (KNN)** for soil classification and fertility prediction, ensuring precise recommendations for crop cultivation.
- Utilized **Hadoop's Distributed File System (HDFS)** for efficient data storage and MapReduce framework for parallel and distributed data processing.
- Processed datasets sourced from **Kaggle**, containing parameters like land type, rainfall, weather conditions, and fertilizer content.

- Worked with **Java** on **Ubuntu** using VirtualBox as the development environment.

CERTIFICATIONS

- Microsoft Azure Fundamentals (AZ-900) Certified
- Microsoft Azure Administrator (AZ-104) Certified
- PCEP - Python programmer Certified
- Introduction to Generative AI (Google Certified)
- Microsoft Power BI Data Analyst: Course PL-300T00-A