

Karan Ajay Pisay

Austin, TX (Willing to relocate) | (443)-593-8409 | karan.pisay7@gmail.com | [GitHub](#) | [LinkedIn](#) | [Portfolio](#)

WORK EXPERIENCE:

Enterprise Applications Data Analyst

September 2023 - Present

Tri-County Electric Cooperative, Inc., Texas, USA

- Automated data extraction with **Alteryx** and crafted forecasting visualizations in **Tableau**, reduced energy procurement costs by 60%.
- Engineered high-accuracy Linear Regression models for power consumption forecasting, cutting resource wastage by 15%.
- Enhanced financial efficiency by 30% through ad-hoc analysis using **Python** and process optimization resulting in enterprise cost savings.
- Boosted resource reliability by **geospatially** plotting high consumption areas using GIS tools like Catalyst and Leaflet to identify patterns and potential problem zones and implemented fault-tolerant measures.
- Integrated data from various sources, including smart meters, SCADA systems, CIS, and FIS, into a unified **Azure data warehouse**.
- Collaborated with a cross-functional team for data collection and data cleaning using **SQL** and **Apache Spark**, ensuring high-quality data.

Marketing Data Analyst

June 2023 - September 2023

LeveragAI, California, USA

- Established a monthly reporting mechanism utilizing **PowerBI** to produce Key Performance Indicator Dashboards (KPI), Management Reports, Comprehensive Financial Reports, and Year-end Projection Reports, to support strategic decision-making.
- Constructed **predictive models** to enhance targeting precision and analyze customer behavior, increasing campaign engagement by 25%.
- Leveraged data visualization and reporting tools like **AWS QuickSight** and pivots in **Excel** to deliver real-time campaign performance dashboards and engagement analytics, facilitating strategic decision-making and driving business intelligence.
- Partnered with cross-functional teams, leveraging communication skills and critical thinking to identify key metrics, optimize marketing strategies, and enhance customer acquisition by 15%.

Data Engineer

July 2018 - June 2021

Triangular Automation Pvt Ltd, India

- Designed and implemented **ETL pipelines** by leveraging tools like **Apache Airflow** and **Apache Spark** to ingest, process, and store large volumes of customer credit data, ensured data governance, enhanced data integration and data quality by 40%.
- Created and refined **XGBoost** models for customer segmentation and predictive analytics, improving credit risk assessments and decision-making processes.
- Analyzed credit data to produce dashboards for ARR, MRR, and lead conversion, utilizing **data mining** techniques to provide insights.
- Conducted **feature engineering** to enhance credit scoring models and risk assessment processes, supporting strategic financial decisions and improving credit evaluation accuracy for leadership and stakeholders.
- Communicated analysis results to various management levels, summarizing key insights and recommendations.

EDUCATION:

Master of Professional Studies in Data Science

August 2021 - May 2023

University of Maryland, Baltimore County, Baltimore, MD

Bachelor of Engineering in Computer Science

August 2016 - June 2019

Pune University, India

TECHNICAL SKILLS:

Programming Languages: Python, R, Java, Scala, C++, JavaScript

Database Technologies: SQL, PostgreSQL, MySQL, NoSQL (MongoDB), AWS S3, Neptune, Big Query

Statistical Analysis: Hypothesis testing, Descriptive/Predictive Analytics, ANOVA, A/B Testing, Chi-Square

Analytical Tools: Alteryx, Microsoft Excel, Tableau, VBA, Power BI, Qlik, Splunk, Tableau Prep

Machine Learning (AI/ML) Algorithms: Regression, Classification, Clustering, Decision Trees, Random Forest, XGboost, KNN, LLM

Data Processing: Data Cleaning, Data Wrangling, Data Modelling, Feature engineering, Data normalization and scaling, ETL/ELT

Libraries: Pandas, Seaborn, NLTK, Beautiful Soup, Scrapy, OpenCV, Scikit-learn, PyTorch, TensorFlow, SQLAlchemy, Flask

DevOps and Tools: JIRA, CI/CD, Slack, Asana, HubSpot, Git, GitHub, Databricks, Workday, Snowflake, Ivanti, Dbeaver, SSMS, Docker

Big Data Technologies and Cloud Platforms: AWS, Apache Spark, Microsoft Azure, Google Cloud Platform (GCP), Hadoop

PROJECTS:

Real-time Sentiment Analysis on Cryptocurrency Market | AWS (SageMaker, EMR, S3, Lambda), Python, PowerBI, LangChain

- Engineered a web application to preprocess and ingest data from Twitter and Google News, achieving a 91% accuracy rate with NLTK's VADER and fine-tuned BERT Transformer. Leveraged AWS SageMaker, EMR, and Lambda for efficient data handling and processing.
- Deployed advanced Natural Language Processing models and dynamic dashboards using AWS S3 and PowerBI, significantly enhancing the sentiment analysis and insights into cryptocurrency market trends.

Real-time Data Pipeline for Retail Analytics | Apache Kafka, Apache Spark, PySpark, Python, MySQL, Tableau

- Designed and implemented a real-time data pipeline using Apache Kafka and Spark, ensuring seamless data ingestion, low-latency processing, and data storage in MySQL, resulting in a 70% improvement in data processing speeds.
- Created interactive Tableau dashboards for real-time retail analytics, enhancing decision-making, improving inventory management accuracy by 25%, and increasing overall sales by 10% through actionable insights into customer behavior and sales performance.

Automotive Customer Insights and Sales Forecasting | Snowflake, Principal Component Analysis, CRISP-DM, EDA, Jupyter Notebook

- Conducted extensive time series analysis and feature engineering using Snowflake and Prophet, delivering actionable insights into customer behavior, vehicle preferences, and market trends, leading to a significant increase in customer retention and sales performance.

CERTIFICATIONS:

- PMI: Disciplined Agile SCRUM Master (DASM)
 - AWS Associate Solutions Architect
 - AWS Cloud Practitioner
 - Alteryx Machine Learning Fundamentals
 - Microsoft Essentials in Data Analysis
 - Certified Associate Project Management (CAPM)
-