Prajwal Tidke

About me - Experienced cloud data engineer with 4 years of expertise in the Manufacturing, Warranty, Payroll and Finance Domain. Dedicated to crafting efficient cloud data solutions through rigorous analysis and passionate about data science, blending strong technical and interpersonal skills. My goal is to enhance my skills and create value for employers and society.

I'm currently pursuing Masters Degree in Big Data Analytics @San Diego State University

Selected Dataset - The home of the U.S. Government's open data

This website hosts various datasets collected and maintained by the U.S. Government over the years, all available in the public domain. I am particularly interested in working with financial datasets. Instead of focusing on a single dataset, I plan to utilize multiple related datasets from the website that are associated with financial fraud. For example, the "**21st Century Corporate Financial Fraud, United States, 2005-2010**" dataset can serve as the foundation for my fraud detection pipeline, while other databases, such as the "**Consumer Complaint Database**" and "**Debt to the Penny**" can act as supplementary datasets for predictive analysis in fraud detection and prevention. These supporting datasets may not be directly linked to the core fraud datasets, but they could provide crucial insights into the factors that contribute to financial fraud. Although this is an initial concept for applying these datasets, we can enhance the fraud detection pipeline by incorporating additional datasets during the detailed Knowledge Discovery in Databases (KDD) phase.

Selected Data Science Book - Forecasting: Principles and Practice

This online textbook provides a thorough introduction to forecasting methods, emphasizing practical applications over theoretical discussions. It highlights the use of R for forecasting, offering practical examples with real data through the **fpp2** and **forecast** R packages. Additionally, the book focuses on graphical methods, primarily using **ggplot2** to explore data and present forecasting outcomes. These topics will help me gain a strong understanding of forecasting methodologies, which I can apply to various data pipelines. Currently, I can use these methods to develop a pipeline for financial fraud detection utilizing the datasets available on data.gov

Languages and Tools

Connect with me

Most Used Languages