CHECKPOINT - 1

Project Information:

Github link:
https://github.com/akshay-002/Accident-Risk-Prediction-using-Sparse-and-Heterogeneous-Data.git

Project name: Accident Risk Prediction using Sparse & Heterogenous Data Insights and Model Development.

Members :

Name	UID
Usha Sree Vuchidi	121195692
Sivani Mallangi	121219053
Akshay Suresh	120981975
Renu Sri Periketi	121272409

What datasets are you choosing? Cite the source(s).

Kaggle dataset ,U.S Accidents 2016 - 2023

Sources:

1. Moosavi, Sobhan, Mohammad Hossein Samavatian, Srinivasan Parthasarathy, and Rajiv Ramnath. “A Countrywide Traffic Accident Dataset.”, 2019.
2. Moosavi, Sobhan, Mohammad Hossein Samavatian, Srinivasan Parthasarathy, Radu Teodorescu, and Rajiv Ramnath. "Accident Risk Prediction based on Heterogeneous Sparse Data: New Dataset and Insights." In proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, ACM, 2019.

Why are we choosing this dataset?

The countrywide car accident dataset likely contains a variety of features that provide insights into each accident's circumstances. The main reasons we have chosen this dataset are as follows,

- Large Sample Size: Statistical significance at around 7.7 million records offers a very strong basis on which any form of statistical inference can be reliable and sound in the conclusions derived from it.

- Richness of Data: The dataset likely includes various attributes that allow for comprehensive analyses and multifaceted insights

- Temporal Insights: Covering multiple years (2016 to 2023), the dataset enables the study of trends over time, allowing for the evaluation of the effectiveness of safety initiatives or changes in driving behavior.

- Machine Learning Applications: The dataset is well-suited for developing predictive models to forecast accident likelihood, severity, or frequency, contributing to proactive safety measures.

- Real-World Impact : Findings from analyses can lead to educational campaigns, improved traffic management, and community awareness initiatives, directly benefiting the public.Car accidents have significant societal implications, affecting public safety and health. Analyzing this dataset can lead to actionable insights that can help reduce accidents and save lives.