

Data Requirements

Data used is "**Example Dataset**" provided by the IBM Data Science course on Coursera.org ,
Because of the following reasons:

1. It's rich with parameters which might be extremely useful for future research.
2. It's real life data recorded by Seattle state GIS.
3. It has special records where only behavior of drivers/pedestrians plays the role of making the collision (drugs - alcohol - etc.).
4. physical details of collisions are recorded so it's highly valuable for machine learning projects.

Classification approach will be selected from different classifier models ;built using Scikit-learn on JupyterNotebook; based on best accuracy metrics. Data pre-processing is handled depending on "Metadata.pdf" file attached.