

Data:

The dataset is the collisions data in Seattle provided by SPD and recorded by Traffic records. This includes all types of collisions and is updated weekly. The dataset consists of 37 attributes which describe the location geometry, road conditions and user information.

The attributes describe the location of the collision, collision type, total number of people involved, number of pedestrians involved, number of vehicles involved, total injuries involved, number of serious injuries in the collision, number of fatalities, date and time of the incident, category of the junction, code of the collision and description of the code, whether the collision was due to inattention, whether the driver was under the influence of drugs, the weather conditions during the time of collision, the road conditions during the time of collision, the light conditions during the time of collision, whether the pedestrian was in the right of way, whether speeding was a factor, the state provided code for the collision and its description, the lane segment of the collision, the crosswalk at the collision and whether the collision involved in hitting a parked car.

The label is the severity of the collision and it describes the level of severity from 0 – 3. Where 0 is unknown, 1 is for property damage, 2 is for injury, 2b is for serious injury and 3 is for fatality. This data can be used in taking precautionary measures.

For example, if the supervised learning predicts that the accidents at the specific crosswalk or location is highly fatal, warning signs or alternative pedestrian walkways can be installed.