A. Introduction:

This project decides the severity of the accidents in order to provide the right required service to the accident victims.

This project might interest the departments responsible for providing emergency medical and transportation facilities in case of accidents.

As this project determines the severity of accidents, authorities will able to provide correct service according to surrounding condition and severity such as if there is traffic jam due to accident and victim need immediate medical care then instead of ambulance, authorities can provide helicopters. In case of less sever accidents, ambulance can be used for emergency. In that case it will distribute the service correctly as per requirement without waste of resources.

B. Data Description:

Data Source: Data is taken from Kaggle, open source data.

Data consist of 37 attributes and 194673 observations.

It consists of categorical and numerical variables. Label of the data is SEVERITYCODE which is categorical variable with code for each severity level.

Additionally, several other variables are included to determine speed of the vehicle, type of accidents, fatality and injuries in accidents, weather conditions, location where accident happened.

All detailed description can be found in metadata provided on Kaggle with data set.

C. Methodology:

This problem is a classification problem where model will classify the accident scenario in several severity levels.

To obtain the best model for solution, serval algorithms are studied as follow:

Logistic Regression

Decision Tree

Artificial Neural Network

1. Logistic Regression:

Logistic regression performs very well on training and testing data as well.

Score on validation data= 1.00

Therefore, logistic regression is a potential model for classifying the accident severity.

2. Decision Tree:

Decision Trees also perform perfect on testing data as well

Score on validation data: 1.00

Decision trees can also be used as good performing model for classification.

3. Artificial Neural Network:

Artificial neural networks perform well with 70.35 training accuracy and 70.34 validation accuracy.

D. Results:

Algorithm/ Model	Validation Accuracy
1. Logistic Regression	1.00
2. Decision Trees	1.00
3. Artificial Neural Network	70.34

E. Discussion:

In this problem solving, surprisingly logistic regression and decision tree works better than neural network. Therefore, both good performing techniques can be used as final potential models.

F. Conclusion:

For accident severity classification problem logistic regression or decision tree models can be used with perfect accuracy of 1.