Classification Homework
Points: 100 points

Automobile accidents.

The dataset contains information on 42,183 actual automobile accidents in 2001 in the United States that involved one of three levels of injury (MAX_SEV_IR): No Injury (MAX_SEV_IR=0), Injury (MAX_SEV_IR=1) or Fatality (MAX_SEV_IR=2). For each accident, additional information is recorded, such as day of week, weather conditions, and road type. A firm might be interested in developing a system for quickly classifying the severity of an accident based on initial reports and associated data in the system (some of which rely on GPS-assisted reporting).

Our goal here is to predict whether an accident just reported will involve an injury or will not. For this purpose, there are two classes: No Injury and Injury which includes Fatality.

Using the given 12 records in the following table, only considering two features of the records: Weather_Related and Traffic_Condition_Related:
1. Train a Naïve Bayes Classifier.
2. Predict the probability of an injury given Weather_Related = 1 and Traffic_Condition_Related = 1.

ID	Speed_Limit	Traffic_Condition_Related	Weather_Related	MAX_SEV_IR
1	40	0	1	1
2	70	0	2	0
3	35	1	2	0
4	35	1	1	0
5	25	0	1	0
6	70	0	2	1
7	70	0	2	0
8	35	0	1	1
9	30	0	2	0
10	25	0	2	0
11	55	0	2	0
12	40	2	1	0