Santhanakrishnan Ramani

SID: 105720585

**Assignment 0**

***Task 1: Bayes Classification***

|  |  |
| --- | --- |
| P(Death) | 0.676965 |
| P(Class=1st,Age=Child,Gender=Male) | 0.002272 |
| P(Class=1st,Age=Child,Gender=Female) | 0.000454 |
| P(Class=1st,Age=Adult,Gender=Male) | 0.079509 |
| P(Class=1st,Age=Adult,Gender=Female) | 0.065425 |
|  |  |
| P(Class=2nd,Age=Child,Gender=Male) | 0.004998 |
| P(Class=2nd,Age=Child,Gender=Female) | 0.005906 |
| P(Class=2nd,Age=Adult,Gender=Male) | 0.076329 |
| P(Class=2nd,Age=Adult,Gender=Female) | 0.042254 |
|  |  |
| P(Class=3rd,Age=Child,Gender=Male) | 0.021808 |
| P(Class=3rd,Age=Child,Gender=Female) | 0.014085 |
| P(Class=3rd,Age=Adult,Gender=Male) | 0.209905 |
| P(Class=3rd,Age=Adult,Gender=Female) | 0.074966 |
|  |  |
| P(Class=Crew,Age=Child,Gender=Male) | 0 |
| P(Class=Crew,Age=Child,Gender=Female) | 0 |
| P(Class=Crew,Age=Adult,Gender=Male) | 0.39164 |
| P(Class=Crew,Age=Adult,Gender=Female) | 0.01045 |

|  |  |
| --- | --- |
| P(Class=1st,Age=Child,Gender=Male|Death) | 0 |
| P(Class=1st,Age=Child,Gender=Female|Death) | 0 |
| P(Class=1st,Age=Adult,Gender=Male|Death) | 0.079195 |
| P(Class=1st,Age=Adult,Gender=Female|Death) | 0.002685 |
|  |  |
| P(Class=2nd,Age=Child,Gender=Male|Death) | 0 |
| P(Class=2nd,Age=Child,Gender=Female|Death) | 0 |
| P(Class=2nd,Age=Adult,Gender=Male|Death) | 0.103356 |
| P(Class=2nd,Age=Adult,Gender=Female|Death) | 0.008725 |
|  |  |
| P(Class=3rd,Age=Child,Gender=Male|Death) | 0.02349 |
| P(Class=3rd,Age=Child,Gender=Female|Death) | 0.011409 |
| P(Class=3rd,Age=Adult,Gender=Male|Death) | 0.259732 |
| P(Class=3rd,Age=Adult,Gender=Female|Death) | 0.059732 |
|  |  |
| P(Class=Crew,Age=Child,Gender=Male|Death) | 0 |
| P(Class=Crew,Age=Child,Gender=Female|Death) | 0 |
| P(Class=Crew,Age=Adult,Gender=Male|Death) | 0.449664 |
| P(Class=Crew,Age=Adult,Gender=Female|Death) | 0.002013 |

|  |
| --- |
| ***Probability Table*** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Male** | | **Female** | |
|  | **Child** | **Adult** | **Child** | **Adult** |
| **First** | 0 | 0.674286 | 0 | 0.027778 |
| **Second** | 0 | 0.916667 | 0 | 0.139785 |
| **Third** | 0.729167 | 0.837662 | 0.548387 | 0.539394 |
| **Crew** | ND | 0.777262 | ND | 0.130435 |

***Classification Table***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Male** | | **Female** | |
|  | **Child** | **Adult** | **Child** | **Adult** |
| **First** | Survival | Death | Survival | Survival |
| **Second** | Survival | Death | Survival | Survival |
| **Third** | Death | Death | Death | Death |
| **Crew** | NC | Death | NC | Survival |

***Task 2: Naïve Bayes Classification***

|  |  |
| --- | --- |
| P(Death=Y) | 0.676965 |
|  |  |
| P(Class=1st|Death) | 0.0818792 |
| P(Class=2nd|Death) | 0.1120805 |
| P(Class=3rd|Death) | 0.3543624 |
| P(Class=crew|Death) | 0.4516779 |
|  |  |
| P(Age=Child|Death) | 0.0348993 |
| P(Age=Adult|Death) | 0.9651007 |
|  |  |
| P(Gender=Male|Death) | 0.9154362 |
| P(Gender=Female|Death) | 0.0845638 |

|  |  |
| --- | --- |
| P(Death=N) | 0.323034984 |
|  |  |
| P(Class=1st|Survived) | 0.285513361 |
| P(Class=2nd|Survived) | 0.165963432 |
| P(Class=3rd|Survived) | 0.250351617 |
| P(Class=crew|Survived) | 0.298171589 |
|  |  |
| P(Age=Child|Survived) | 0.080168776 |
| P(Age=Adult|Survived) | 0.919831224 |
|  |  |
| P(Gender=Male|Survived) | 0.516174402 |
| P(Gender=Female|Survived) | 0.483825598 |

***Probability Table***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Male** | | **Female** | |
|  | **Child** | **Adult** | **Child** | **Adult** |
| **First** | 0.316935 | 0.527924 | 0.043727 | 0.09927 |
| **Second** | 0.522135 | 0.724782 | 0.097214 | 0.206056 |
| **Third** | 0.696059 | 0.846617 | 0.184136 | 0.352318 |
| **Crew** | 0.710219 | 0.855222 | 0.194548 | 0.367951 |

***Classification Table***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Male** | | **Female** | |
|  | **Child** | **Adult** | **Child** | **Adult** |
| **First** | Survival | Death | Survival | Survival |
| **Second** | Death | Death | Survival | Survival |
| **Third** | Death | Death | Survival | Survival |
| **Crew** | Death | Death | Survival | Survival |

***Task 3:***

*Advantages of Task 1:*

* The classification provided will be mostly accurate as it considers all the dependence between the attributes into consideration.

*Disadvantages of Task 1:*

* Difficult to compute the values when many attributes have a dependence on each other, as the no of iterations to compute the required values increases.
* There can be some cases where the prediction can’t be made.

*Advantages of Task 2:*

* Computationally it is easier.
* It is scalable.

*Disadvantages of Task 2:*

* As it considers all the attributes are independent of each other, the prediction provided by it might not be true for all cases.

*Circumstances when an empirical table provide better predictions:*

* The no of attributes involved are limited to 4 or 5.
* Attributes involved are dependent on each other.
* Attributes involved mostly have a Boolean value.

*Circumstances when the naïve Bayes table provide better predictions:*

* The no of attributes involved are more.
* Dependency of attributes doesn’t matter much.
* When the attributes are multi valued.