

```

> #Titanic Dataset with 1310 obs. of 14 variables:
> rm(list=ls()) #deleting all variables from workspace
> ls()
character(0)
> data=read.csv("titanic_data.csv",header=TRUE,stringsAsFactors = T) #Loading data
> nrow(data)
[1] 1310
> str(data)
'data.frame': 1310 obs. of 14 variables:
 $ pclass : int 1 1 1 1 1 1 1 1 1 1 ...
 $ survived : int 1 1 0 0 0 1 1 0 1 0 ...
 $ name : Factor w/ 1308 levels "", "Abbing, Mr. Anthony",...: 23 25 26 27 28 32
47 48 52 56 ...
 $ sex : Factor w/ 3 levels "", "female", "male": 2 3 2 3 2 3 2 3 2 3 ...
 $ age : num 29 0.917 2 30 25 ...
 $ sibsp : int 0 1 1 1 1 0 1 0 2 0 ...
 $ parch : int 0 2 2 2 2 0 0 0 0 0 ...
 $ ticket : Factor w/ 930 levels "", "110152", "110413",...: 189 51 51 51 51 126 94
17 78 827 ...
 $ fare : num 211 152 152 152 152 ...
 $ cabin : Factor w/ 187 levels "", "A10", "A11",...: 45 81 81 81 81 151 147 17 63
1 ...
 $ embarked : Factor w/ 4 levels "", "C", "Q", "S": 4 4 4 4 4 4 4 4 2 ...
 $ boat : Factor w/ 28 levels "", "1", "10", "11",...: 13 4 1 1 1 14 3 1 28 1 ...
 $ body : int NA NA NA 135 NA NA NA NA NA 22 ...
 $ home.dest: Factor w/ 370 levels "", "?Havana, Cuba",...: 310 232 232 232 232 238
163 25 23 230 ...
> sum(is.na(data)) #total 1459 NA values
[1] 1459
> #I will use CARET package for preprocessing of data:
> library(caret)
Loading required package: lattice
Loading required package: ggplot2
> preprocvalues=preProcess(data,method=c("medianImpute","center","scale")) #taking
median for all NA with respective variables & adjusting scale
> library(RANN)
> data_pro=predict(preprocvalues,data)
> sum(is.na(data_pro)) #total 0 NA values
[1] 0
> dv=dummyVars("~.",data_pro,fullRank = T) # creating dummy variable to handle
factors
> data_tran=data.frame(predict(dv,data_pro))
> str(data_tran)
'data.frame': 1310 obs. of 2830 variables:
 $ pclass
: num -1.55 -1.55 -1.55 -1.55 -1.55 ...

```

```

$ survived
: num  1.272 1.272 -0.786 -0.786 -0.786 ...
$ name.Abbing..Mr..Anthony
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Abbott..Master..Eugene.Joseph
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Abbott..Mr..Rossmore.Edward
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Abbott..Mrs..Stanton..Rosa.Hunt.
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Abelseth..Miss..Karen.Marie
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Abelseth..Mr..Olaus.Jorgensen
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Abelson..Mr..Samuel
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Abelson..Mrs..Samuel..Hannah.Wizosky.
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Abrahamsson..Mr..Abraham.August.Johannes
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Abrahim..Mrs..Joseph..Sophie.Halaut.Easu.
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Adahl..Mr..Mauritz.Nils.Martin
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Adams..Mr..John
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Ahlin..Mrs..Johan..Johanna.Persdotter.Larsson.
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Aks..Master..Philip.Frank
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Aks..Mrs..Sam..Leah.Rosen.
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Albimona..Mr..Nassef.Cassem
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Aldworth..Mr..Charles.Augustus
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Alexander..Mr..William
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Alhomaki..Mr..Ilmari.Rudolf
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Ali..Mr..Ahmed
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Ali..Mr..William
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Allen..Miss..Elisabeth.Walton
: num  1 0 0 0 0 0 0 0 0 0 ...
$ name.Allen..Mr..William.Henry
: num  0 0 0 0 0 0 0 0 0 0 ...

```

```

$ name.Allison..Master..Hudson.Trevor
: num  0 1 0 0 0 0 0 0 0 0 ...
$ name.Allison..Miss..Helen.Loraine
: num  0 0 1 0 0 0 0 0 0 0 ...
$ name.Allison..Mr..Hudson.Joshua.Creighton
: num  0 0 0 1 0 0 0 0 0 0 ...
$ name.Allison..Mrs..Hudson.J.C..Bessie.Waldo.Daniels.
: num  0 0 0 0 1 0 0 0 0 0 ...
$ name.Allum..Mr..Owen.George
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Andersen.Jensen..Miss..Carla.Christine.Nielsine
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Andersen..Mr..Albert.Karvin
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Anderson..Mr..Harry
: num  0 0 0 0 0 1 0 0 0 0 ...
$ name.Andersson..Master..Sigvard.Harald.Elias
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Andersson..Miss..Ebba.Iris.Alfrida
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Andersson..Miss..Ellis.Anna.Maria
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Andersson..Miss..Erna.Alexandra
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Andersson..Miss..Ida.Augusta.Margareta
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Andersson..Miss..Ingeborg.Constanzia
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Andersson..Miss..Sigrid.Elisabeth
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Andersson..Mr..Anders.Johan
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Andersson..Mr..August.Edvard...Wennerstrom..
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Andersson..Mr..Johan.Samuel
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Andersson..Mrs..Anders.Johan..Alfrida.Konstantia.Brogren.
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Andreasson..Mr..Paul.Edvin
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Andrew..Mr..Edgardo.Samuel
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Andrew..Mr..Frank.Thomas
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Andrews..Miss..Kornelia.Theodosia
: num  0 0 0 0 0 0 1 0 0 0 ...
$ name.Andrews..Mr..Thomas.Jr
: num  0 0 0 0 0 0 0 1 0 0 ...

```

```

$ name.Angheloff..Mr..Minko
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Angle..Mr..William.A
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Angle..Mrs..William.A..Florence..Mary..Agnes.Hughes.
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Appleton..Mrs..Edward.Dale..Charlotte.Lamson.
: num  0 0 0 0 0 0 0 0 1 0 ...
$ name.Arnold.Franchi..Mr..Josef
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Arnold.Franchi..Mrs..Josef..Josefine.Franchi.
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Aronsson..Mr..Ernst.Axel.Algot
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Artagaveytia..Mr..Ramon
: num  0 0 0 0 0 0 0 0 0 1 ...
$ name.Ashby..Mr..John
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Asim..Mr..Adola
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Asplund..Master..Carl.Edgar
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Asplund..Master..Clarence.Gustaf.Hugo
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Asplund..Master..Edvin.Rojj.Felix
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Asplund..Master..Filip.Oscar
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Asplund..Miss..Lillian.Gertrud
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Asplund..Mr..Carl.Oscar.Vilhelm.Gustafsson
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Asplund..Mr..Johan.Charles
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Asplund..Mrs..Carl.Oscar..Selma.Augusta.Emilia.Johansson.
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Assaf.Khalil..Mrs..Mariana...Miriam..
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Assaf..Mr..Gerios
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Assam..Mr..Ali
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Astor..Col..John.Jacob
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Astor..Mrs..John.Jacob..Madeleine.Talmadge.Force.
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Attalah..Miss..Malake
: num  0 0 0 0 0 0 0 0 0 0 ...

```

\$ name.Attalah..Mr..Sleiman
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Aubart..Mme..Leontine.Pauline
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Augustsson..Mr..Albert
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Ayoub..Miss..Banoura
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Baccos..Mr..Raffull
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Backstrom..Mr..Karl.Alfred
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Backstrom..Mrs..Karl.Alfred..Maria.Mathilda.Gustafsson.
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Baclini..Miss..Eugenie
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Baclini..Miss..Helene.Barbara
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Baclini..Miss..Marie.Catherine
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Baclini..Mrs..Solomon..Latifa.Qurban.
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Badman..Miss..Emily.Louisa
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Badt..Mr..Mohamed
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Bailey..Mr..Percy.Andrew
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Baimbrigge..Mr..Charles.Robert
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Balkic..Mr..Cerin
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Ball..Mrs...Ada.E.Hall.
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Banfield..Mr..Frederick.James
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Barah..Mr..Hanna.Assi
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Barbara..Miss..Saiide
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Barbara..Mrs...Catherine.David.
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Barber..Miss..Ellen..Nellie.
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Barkworth..Mr..Algernon.Henry.Wilson
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Barry..Miss..Julia
 : num 0 0 0 0 0 0 0 0 0 0 ...

```

$ name.Barton..Mr..David.John
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Bateman..Rev..Robert.James
: num  0 0 0 0 0 0 0 0 0 0 ...
[list output truncated]
> data_tran$survived=as.factor(data_tran$survived) # converting response variable
in factor
> set.seed(5)
> index <- createDataPartition(data_tran$survived, p=0.75, list=FALSE) #data parti
tion
> train <- data_tran[ index,]           #Traning data=75%
> test<- data_tran[-index,]           #Test data=25%
> str(train)
'data.frame':   983 obs. of  2830 variables:
 $ pclass
: num  -1.55 -1.55 -1.55 -1.55 -1.55 ...
 $ survived
: Factor w/ 2 levels "-0.785859287383634",...: 2 2 1 1 1 2 1 2 2 2 ...
 $ name.Abbing..Mr..Anthony
: num  0 0 0 0 0 0 0 0 0 0 ...
 $ name.Abbott..Master..Eugene.Joseph
: num  0 0 0 0 0 0 0 0 0 0 ...
 $ name.Abbott..Mr..Rossmore.Edward
: num  0 0 0 0 0 0 0 0 0 0 ...
 $ name.Abbott..Mrs..Stanton..Rosa.Hunt.
: num  0 0 0 0 0 0 0 0 0 0 ...
 $ name.Abelseth..Miss..Karen.Marie
: num  0 0 0 0 0 0 0 0 0 0 ...
 $ name.Abelseth..Mr..Olaus.Jorgensen
: num  0 0 0 0 0 0 0 0 0 0 ...
 $ name.Abelson..Mr..Samuel
: num  0 0 0 0 0 0 0 0 0 0 ...
 $ name.Abelson..Mrs..Samuel..Hannah.Wizosky.
: num  0 0 0 0 0 0 0 0 0 0 ...
 $ name.Abrahamsson..Mr..Abraham.August.Johannes
: num  0 0 0 0 0 0 0 0 0 0 ...
 $ name.Abrahim..Mrs..Joseph..Sophie.Halaut.Easu.
: num  0 0 0 0 0 0 0 0 0 0 ...
 $ name.Adahl..Mr..Mauritz.Nils.Martin
: num  0 0 0 0 0 0 0 0 0 0 ...
 $ name.Adams..Mr..John
: num  0 0 0 0 0 0 0 0 0 0 ...
 $ name.Ahlin..Mrs..Johan..Johanna.Persdotter.Larsson.
: num  0 0 0 0 0 0 0 0 0 0 ...
 $ name.Aks..Master..Philip.Frank
: num  0 0 0 0 0 0 0 0 0 0 ...
 $ name.Aks..Mrs..Sam..Leah.Rosen.
: num  0 0 0 0 0 0 0 0 0 0 ...

```

```

$ name.Albimona..Mr..Nassef.Cassem
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Aldworth..Mr..Charles.Augustus
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Alexander..Mr..William
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Alhomaki..Mr..Ilmari.Rudolf
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Ali..Mr..Ahmed
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Ali..Mr..William
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Allen..Miss..Elisabeth.Walton
: num  1 0 0 0 0 0 0 0 0 0 ...
$ name.Allen..Mr..William.Henry
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Allison..Master..Hudson.Trevor
: num  0 1 0 0 0 0 0 0 0 0 ...
$ name.Allison..Miss..Helen.Lorraine
: num  0 0 1 0 0 0 0 0 0 0 ...
$ name.Allison..Mr..Hudson.Joshua.Creighton
: num  0 0 0 1 0 0 0 0 0 0 ...
$ name.Allison..Mrs..Hudson.J.C..Bessie.Waldo.Daniels.
: num  0 0 0 0 1 0 0 0 0 0 ...
$ name.Allum..Mr..Owen.George
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Andersen.Jensen..Miss..Carla.Christine.Nielsine
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Andersen..Mr..Albert.Karvin
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Anderson..Mr..Harry
: num  0 0 0 0 0 1 0 0 0 0 ...
$ name.Andersson..Master..Sigvard.Harald.Elias
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Andersson..Miss..Ebba.Iris.Alfrida
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Andersson..Miss..Ellis.Anna.Maria
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Andersson..Miss..Erna.Alexandra
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Andersson..Miss..Ida.Augusta.Margareta
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Andersson..Miss..Ingeborg.Constanzia
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Andersson..Miss..Sigrid.Elisabeth
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Andersson..Mr..Anders.Johan
: num  0 0 0 0 0 0 0 0 0 0 ...

```

\$ name.Andersson..Mr..August.Edvard...Wennerstrom..
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Andersson..Mr..Johan.Samuel
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Andersson..Mrs..Anders.Johan..Alfrida.Konstantia.Brogren..
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Andreasson..Mr..Paul.Edvin
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Andrew..Mr..Edgardo.Samuel
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Andrew..Mr..Frank.Thomas
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Andrews..Miss..Kornelia.Theodosia
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Andrews..Mr..Thomas.Jr
 : num 0 0 0 0 0 0 1 0 0 0 ...
 \$ name.Angheloff..Mr..Minko
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Angle..Mr..William.A
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Angle..Mrs..William.A..Florence..Mary..Agnes.Hughes..
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Appleton..Mrs..Edward.Dale..Charlotte.Lamson..
 : num 0 0 0 0 0 0 0 1 0 0 ...
 \$ name.Arnold.Franchi..Mr..Josef
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Arnold.Franchi..Mrs..Josef..Josefine.Franchi..
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Aronsson..Mr..Ernst.Axel.Algot
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Artagaveytia..Mr..Ramon
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Ashby..Mr..John
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Asim..Mr..Adola
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Asplund..Master..Carl.Edgar
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Asplund..Master..Clarence.Gustaf.Hugo
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Asplund..Master..Edvin.Rojj.Felix
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Asplund..Master..Filip.Oscar
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Asplund..Miss..Lillian.Gertrud
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Asplund..Mr..Carl.Oscar.Vilhelm.Gustafsson
 : num 0 0 0 0 0 0 0 0 0 0 ...

\$ name.Asplund..Mr..Johan.Charles
 : num 0 0 0 0 0 0 0 0 0 ...
 \$ name.Asplund..Mrs..Carl.Oscar..Selma.Augusta.Emilia.Johansson.
 : num 0 0 0 0 0 0 0 0 0 ...
 \$ name.Assaf.Khalil..Mrs..Mariana...Miriam..
 : num 0 0 0 0 0 0 0 0 0 ...
 \$ name.Assaf..Mr..Gerios
 : num 0 0 0 0 0 0 0 0 0 ...
 \$ name.Assam..Mr..Ali
 : num 0 0 0 0 0 0 0 0 0 ...
 \$ name.Astor..Col..John.Jacob
 : num 0 0 0 0 0 0 0 0 0 ...
 \$ name.Astor..Mrs..John.Jacob..Madeleine.Talmadge.Force..
 : num 0 0 0 0 0 0 0 0 1 0 ...
 \$ name.Attalah..Miss..Malake
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Attalah..Mr..Sleiman
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Aubart..Mme..Leontine.Pauline
 : num 0 0 0 0 0 0 0 0 0 1 ...
 \$ name.Augustsson..Mr..Albert
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Ayoub..Miss..Banoura
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Baccos..Mr..Raffull
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Backstrom..Mr..Karl.Alfred
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Backstrom..Mrs..Karl.Alfred..Maria.Mathilda.Gustafsson..
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Baclini..Miss..Eugenie
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Baclini..Miss..Helene.Barbara
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Baclini..Miss..Marie.Catherine
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Baclini..Mrs..Solomon..Latifa.Qurban..
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Badman..Miss..Emily.Louisa
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Badt..Mr..Mohamed
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Bailey..Mr..Percy.Andrew
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Baimbrigge..Mr..Charles.Robert
 : num 0 0 0 0 0 0 0 0 0 0 ...
 \$ name.Balkic..Mr..Cerin
 : num 0 0 0 0 0 0 0 0 0 0 ...

```

$ name.Ball..Mrs...Ada.E.Hall.
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Banfield..Mr..Frederick.James
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Barah..Mr..Hanna.Assi
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Barbara..Miss..Saiide
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Barbara..Mrs...Catherine.David.
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Barber..Miss..Ellen..Nellie.
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Barkworth..Mr..Algernon.Henry.Wilson
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Barry..Miss..Julia
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Barton..Mr..David.John
: num  0 0 0 0 0 0 0 0 0 0 ...
$ name.Bateman..Rev..Robert.James
: num  0 0 0 0 0 0 0 0 0 0 ...
[list output truncated]
> #####Decision Tree#####
#####
> set.seed(3)
> library(rpart)
> m=rpart(survived~.,data=train,method="class",control=rpart.control(minsplit=20,
+                               minbucket=7,maxdepth=10,usesurrogate = 2,xval=10))#
pre-proned method
> pre=predict(m,test)
> printcp(m)

```

Classification tree:

```

rpart(formula = survived ~ ., data = train, method = "class",
      control = rpart.control(minsplit = 20, minbucket = 7, maxdepth = 10,
        usesurrogate = 2, xval = 10))

```

Variables actually used in tree construction:

```

[1] age      boat.13    boat.15    boat.16    boat.3     boat.5     boat.7
boat.A     boat.C     pclass
[11] sex.female sibsp

```

Root node error: 375/983 = 0.38149

n= 983

	CP	nsplit	rel error	xerror	xstd
1	0.458667	0	1.00000	1.00000	0.040612
2	0.045333	1	0.54133	0.54400	0.033906

```

3 0.032000      2    0.49600 0.49867 0.032815
4 0.024000      3    0.46400 0.44267 0.031323
5 0.021333      5    0.41600 0.34933 0.028415
6 0.020000      7    0.37333 0.31733 0.027272
7 0.014667     12    0.22933 0.29333 0.026357
8 0.013333     14    0.20000 0.26667 0.025274
9 0.010000     15    0.18667 0.26133 0.025048
> bestcp=m$cptable[which.min(m$cptable[, "xerror"]), "CP"]
> bestcp                                     #Evaluating best cp
[1] 0.01
> pruned=prune(m, cp=bestcp)
> t=table(train$survived, predict(pruned, type="class"))
> prop.table(table(train$survived, predict(pruned, type="class")))

                -0.785859287383634  1.27152032698672
-0.785859287383634                0.59816887        0.02034588
1.27152032698672                0.05086470        0.33062055
> rownames(t)=paste("Actual", rownames(t), sep=":")
> colnames(t)=paste("predicted", colnames(t), sep=":")
> t

                predicted:-0.785859287383634 predicted:1.2715203269867
2
Actual:-0.785859287383634                588                2
0
Actual:1.27152032698672                50                32
5
> prop.table(t)

                predicted:-0.785859287383634 predicted:1.2715203269867
2
Actual:-0.785859287383634                0.59816887        0.0203458
8
Actual:1.27152032698672                0.05086470        0.3306205
5
> accuracy=sum(diag(t))/sum(t)
> accuracy                                     ###Accuracy on traning data=0.9287894
[1] 0.9287894
> t=predict(m, test, type="class")
> s=prop.table(table(t, test$survived))
> s

t                -0.785859287383634  1.27152032698672
-0.785859287383634                0.59633028        0.08868502
1.27152032698672                0.02140673        0.29357798
> accuracy=sum(diag(s))/sum(s)
> accuracy                                     ### Accuracy on test data=0.8899083
[1] 0.8899083

```

```
> ##### ROC #####
> for_auc=predict(pruned,test,type="prob")
> library(pROC)
Type 'citation("pROC")' for a citation.
```

Attaching package: 'pROC'

The following objects are masked from 'package:stats':

cov, smooth, var

```
> a=auc(test$survived,for_auc[,2])
> a #Area under the curve: 0.8977
Area under the curve: 0.8977
> #Ex:90-100,Good:80-90,fair:70-80,poor:60-70,Fail:50-60
> plot(roc(test$survived,for_auc[,2]))
> gini_coeff=2*a-1
> gini_coeff # Gini Coeff=0.7954851
[1] 0.79548
```