Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

Programme	:	**BTech – ECE and ECM**	Semester	:	**Win 2022**
Course	:	**Essentials of Data Analytics Lab**	Code	:	**CSE3506**
Faculty	:	**Gobinath N**	Slot	:	**L51 + L52**

Ex.10_Random Forest

Aim: : To solve the given dataset glass.csv using random forest.

PROGRAM:

```
rm(list=ls())  install.packages("stats")
install.packages("dplyr")
install.packages("randomForest")
library(stats)  library(dplyr)
library(randomForest)  mydata
<- read.csv("glass.csv")
View(mydata)
str(mydata)  set.seed(120)
index = sample(2,nrow(mydata),replace=TRUE,prob=c(0.75,0.25))
training <- subset(mydata, index == 1)
testing <- subset(mydata, index == 2)  RFM
= randomForest(Type~.,data=training)
species_pred = predict(RFM,testing)
testing$species_pred = species_pred
View(testing)
CFM = table(testing$species,testing$species_pred)
Classification_Accuracy = sum(diag(CFM))/sum(CFM)  Classification_Accuracy
```

OUTPUT:

Given Data:

	RI	Na	Mg	Al	Si	K	Ca	Ba	Fe	Type
1	1.52101	13.64	4.49	1.10	71.78	0.06	8.75	0.00	0.00	1
2	1.51761	13.89	3.60	1.36	72.73	0.48	7.83	0.00	0.00	1
3	1.51618	13.53	3.55	1.54	72.99	0.39	7.78	0.00	0.00	1
4	1.51766	13.21	3.69	1.29	72.61	0.57	8.22	0.00	0.00	1
5	1.51742	13.27	3.62	1.24	73.08	0.55	8.07	0.00	0.00	1
6	1.51596	12.79	3.61	1.62	72.97	0.64	8.07	0.00	0.26	1
7	1.51743	13.30	3.60	1.14	73.09	0.58	8.17	0.00	0.00	1
8	1.51756	13.15	3.61	1.05	73.24	0.57	8.24	0.00	0.00	1
9	1.51918	14.04	3.58	1.37	72.08	0.56	8.30	0.00	0.00	1
10	1.51755	13.00	3.60	1.36	72.99	0.57	8.40	0.00	0.11	1
11	1.51571	12.72	3.46	1.56	73.20	0.67	8.09	0.00	0.24	1
12	1.51763	12.80	3.66	1.27	73.01	0.60	8.56	0.00	0.00	1
13	1.51589	12.88	3.43	1.40	73.28	0.69	8.05	0.00	0.24	1
14	1.51748	12.86	3.56	1.27	73.21	0.54	8.38	0.00	0.17	1
15	1.51763	12.61	3.59	1.31	73.29	0.58	8.50	0.00	0.00	1
16	1.51761	12.81	3.54	1.23	73.24	0.58	8.39	0.00	0.00	1
17	1.51784	12.68	3.67	1.16	73.11	0.61	8.70	0.00	0.00	1
18	1.52196	14.36	3.85	0.89	71.36	0.15	9.15	0.00	0.00	1
19	1.51911	13.90	3.73	1.18	72.12	0.06	8.89	0.00	0.00	1
20	1.51735	13.02	3.54	1.69	72.73	0.54	8.44	0.00	0.07	1
21	1.51750	12.82	3.55	1.49	72.75	0.54	8.52	0.00	0.19	1

Training Data:

	RI	Na	Mg	Al	Si	K	Ca	Ba	Fe	Type	species_pred
4	1.51766	13.21	3.69	1.29	72.61	0.57	8.22	0.00	0.00	1	1.638367
7	1.51743	13.30	3.60	1.14	73.09	0.58	8.17	0.00	0.00	1	1.143433
9	1.51918	14.04	3.58	1.37	72.08	0.56	8.30	0.00	0.00	1	1.861367
10	1.51755	13.00	3.60	1.36	72.99	0.57	8.40	0.00	0.11	1	1.239767
12	1.51763	12.80	3.66	1.27	73.01	0.60	8.56	0.00	0.00	1	1.171900
15	1.51763	12.61	3.59	1.31	73.29	0.58	8.50	0.00	0.00	1	1.174933
19	1.51911	13.90	3.73	1.18	72.12	0.06	8.89	0.00	0.00	1	1.853900
24	1.51751	12.81	3.57	1.35	73.02	0.62	8.59	0.00	0.00	1	1.143667
33	1.51775	12.85	3.48	1.23	72.97	0.61	8.56	0.09	0.22	1	1.275467
34	1.51753	12.57	3.47	1.38	73.39	0.60	8.55	0.00	0.06	1	1.605033
37	1.51909	13.89	3.53	1.32	71.81	0.51	8.78	0.11	0.00	1	1.733367
41	1.51793	12.79	3.50	1.12	73.03	0.64	8.77	0.00	0.00	1	1.146267
44	1.52210	13.73	3.84	0.72	71.76	0.17	9.74	0.00	0.00	1	1.561067
46	1.51900	13.49	3.48	1.35	71.95	0.55	9.00	0.00	0.00	1	1.745400
48	1.52667	13.99	3.70	0.71	71.57	0.02	9.82	0.00	0.10	1	1.957400
49	1.52223	13.21	3.77	0.79	71.99	0.13	10.02	0.00	0.00	1	1.472000
54	1.51837	13.14	2.84	1.28	72.85	0.55	9.07	0.00	0.00	1	1.553733
56	1.51769	12.45	2.71	1.29	73.70	0.56	9.06	0.00	0.24	1	2.090433
61	1.51905	13.60	3.62	1.11	72.64	0.14	8.76	0.00	0.00	1	1.845667
63	1.52172	13.51	3.86	0.88	71.79	0.23	9.54	0.00	0.11	1	1.600000
65	1.52172	13.48	3.74	0.90	72.01	0.18	9.61	0.00	0.07	1	1.550600

Accuracy of the model:

```
> CFM = table(testing$species,testing$species_pred)
> Classification_Accuracy = sum(diag(CFM))/sum(CFM)
> Classification_Accuracy
[1] 1
```

Result:

Hence the accuracy of the data is calculated as 1 so the data is balanced.