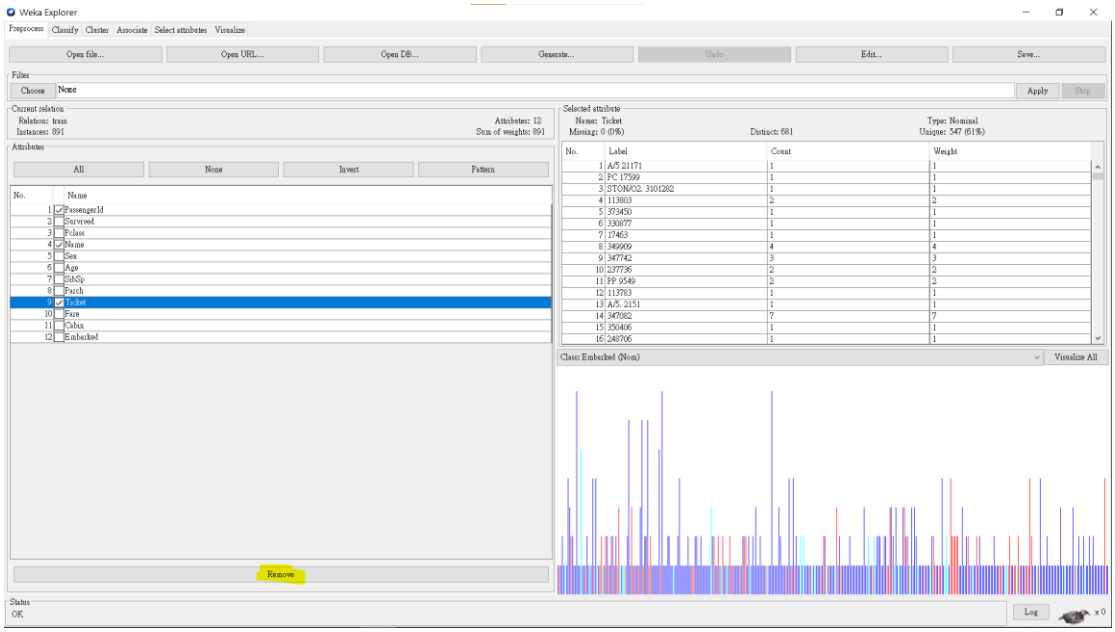
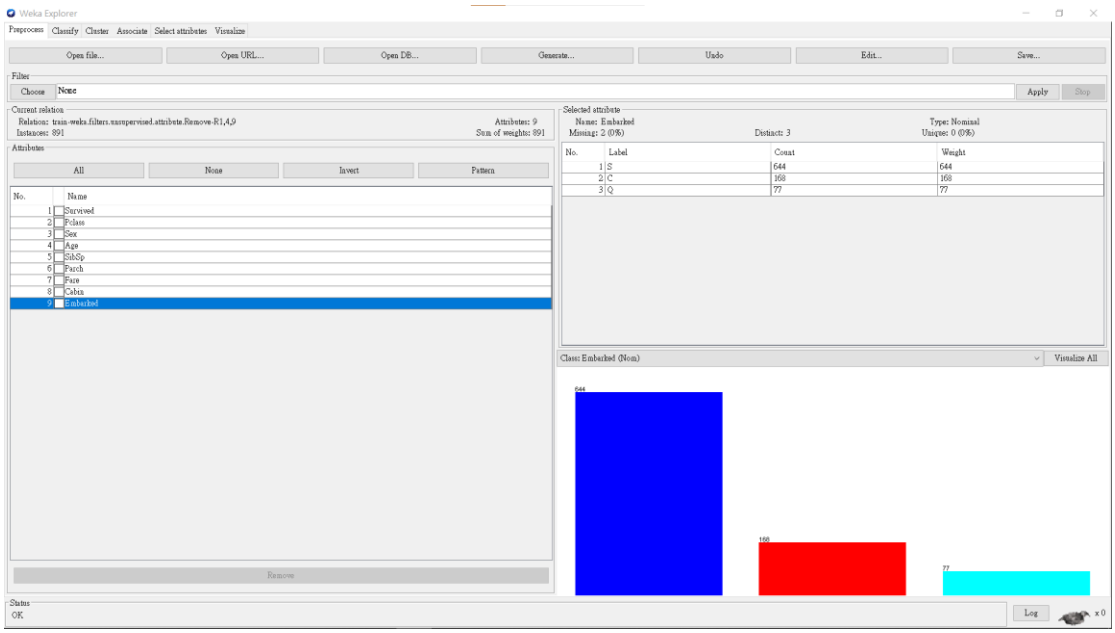
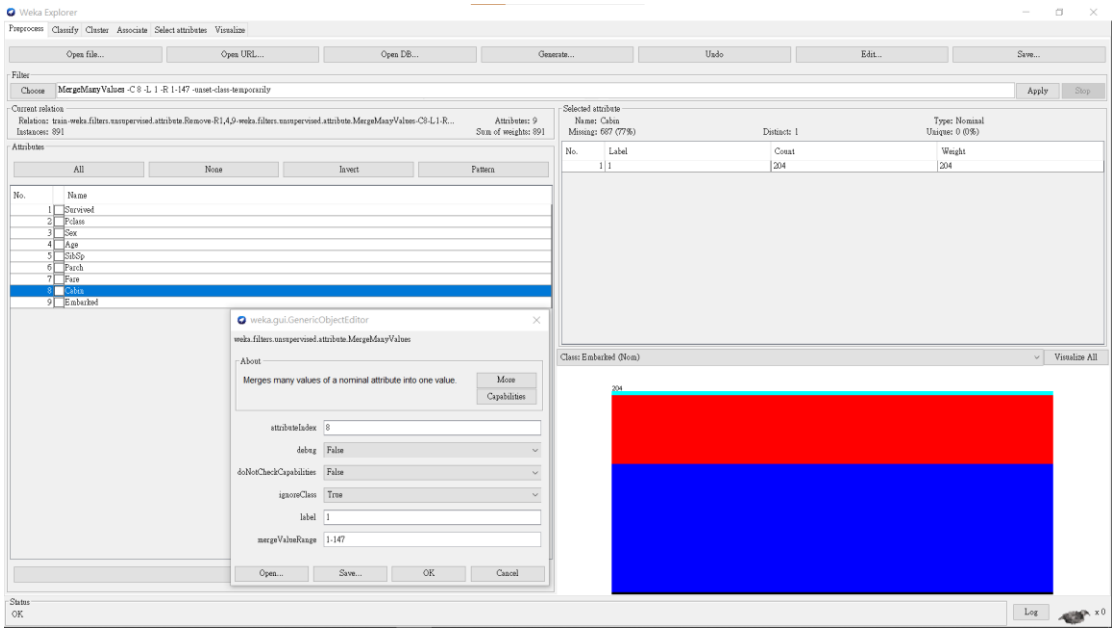
1.1 載入 train.arff 後，選擇 PassengerId、Name、Ticket 點 remove

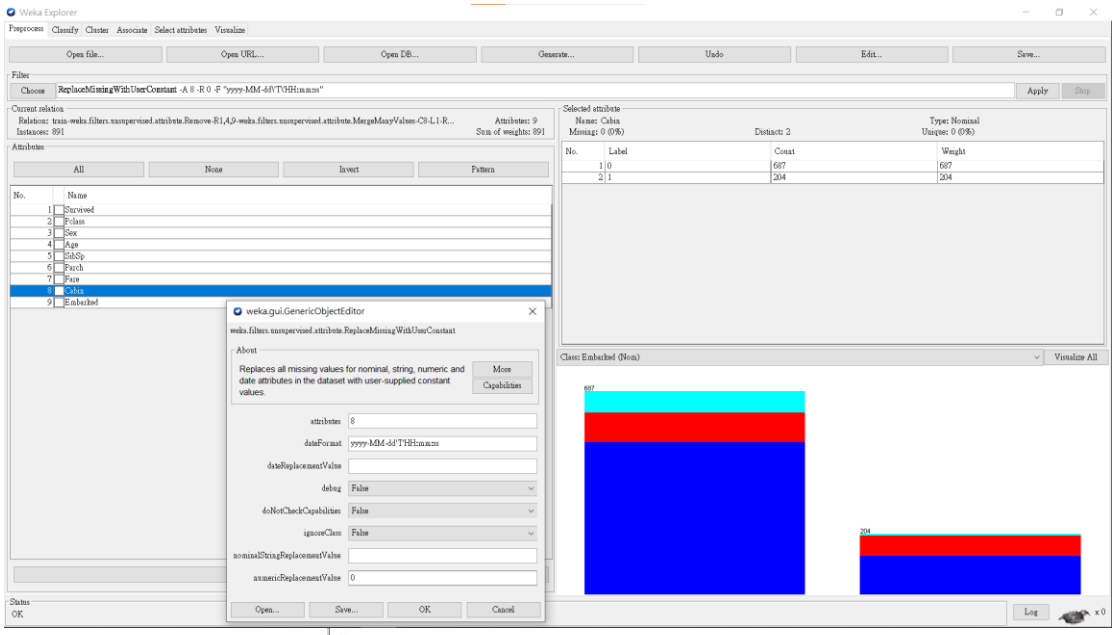




1.2 使用 MergeManyValues 將 Cabin 內有值的 No.1~147 用 1 取代，如下圖



再用 ReplaceMissingWithUserConstant 將空值取代為 0，如下圖

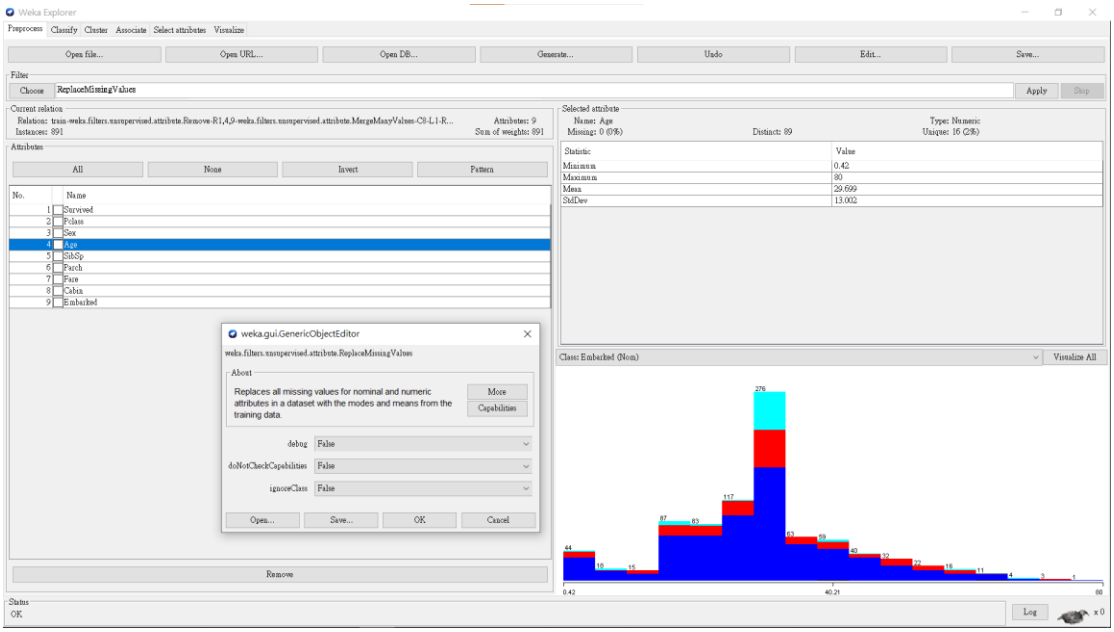


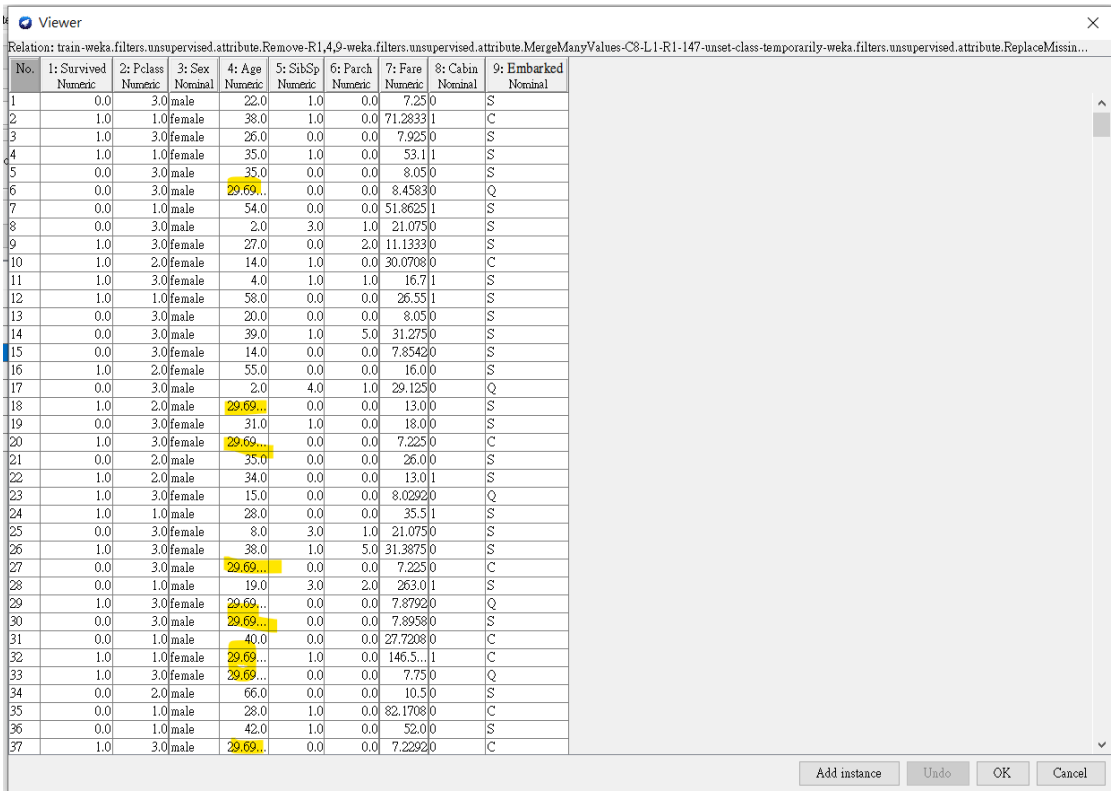
Viewer

Relation: train-weka.filters.unsupervised.attribute.Remove-R1,4,9-weka.filters.unsupervised.attribute.MergeMan

No.	1: Survived Numeric	2: Pclass Numeric	3: Sex Nominal	4: Age Numeric	5: SibSp Numeric	6: Parch Numeric	7: Fare Numeric	8: Cabin Nominal	9: Embarked Nominal
1	0.0	3.0	male	22.0	1.0	0.0	7.25	0	S
2	1.0	3.0	female	26.0	0.0	0.0	7.925	0	S
3	0.0	3.0	male	35.0	0.0	0.0	8.05	0	S
4	0.0	3.0	male		0.0	0.0	8.4583	0	Q
5	0.0	3.0	male	2.0	3.0	1.0	21.075	0	S
6	1.0	3.0	female	27.0	0.0	2.0	11.1333	0	S
7	1.0	2.0	female	14.0	1.0	0.0	30.0708	0	C
8	0.0	3.0	male	20.0	0.0	0.0	8.05	0	S
9	0.0	3.0	male	39.0	1.0	5.0	31.275	0	S
10	0.0	3.0	female	14.0	0.0	0.0	7.8542	0	S
11	1.0	2.0	female	55.0	0.0	0.0	16.0	0	S
12	0.0	3.0	male	2.0	4.0	1.0	29.125	0	Q
13	1.0	2.0	male		0.0	0.0	13.0	0	S
14	0.0	3.0	female	31.0	1.0	0.0	18.0	0	S
15	1.0	3.0	female		0.0	0.0	7.225	0	C
16	0.0	2.0	male	35.0	0.0	0.0	26.0	0	S
17	1.0	3.0	female	15.0	0.0	0.0	8.0292	0	Q
18	0.0	3.0	female	8.0	3.0	1.0	21.075	0	S
19	1.0	3.0	female	38.0	1.0	5.0	31.3875	0	S
20	0.0	3.0	male		0.0	0.0	7.225	0	C
21	1.0	3.0	female		0.0	0.0	7.8792	0	Q
22	0.0	3.0	male		0.0	0.0	7.8958	0	S
23	0.0	1.0	male	40.0	0.0	0.0	27.7208	0	C
24	1.0	3.0	female		0.0	0.0	7.75	0	Q
25	0.0	2.0	male	66.0	0.0	0.0	10.5	0	S
26	0.0	1.0	male	28.0	1.0	0.0	82.1708	0	C
27	0.0	1.0	male	42.0	1.0	0.0	52.0	0	S
28	1.0	3.0	male		0.0	0.0	7.2292	0	C
29	0.0	3.0	male	21.0	0.0	0.0	8.05	0	S
30	0.0	3.0	female	18.0	2.0	0.0	18.0	0	S

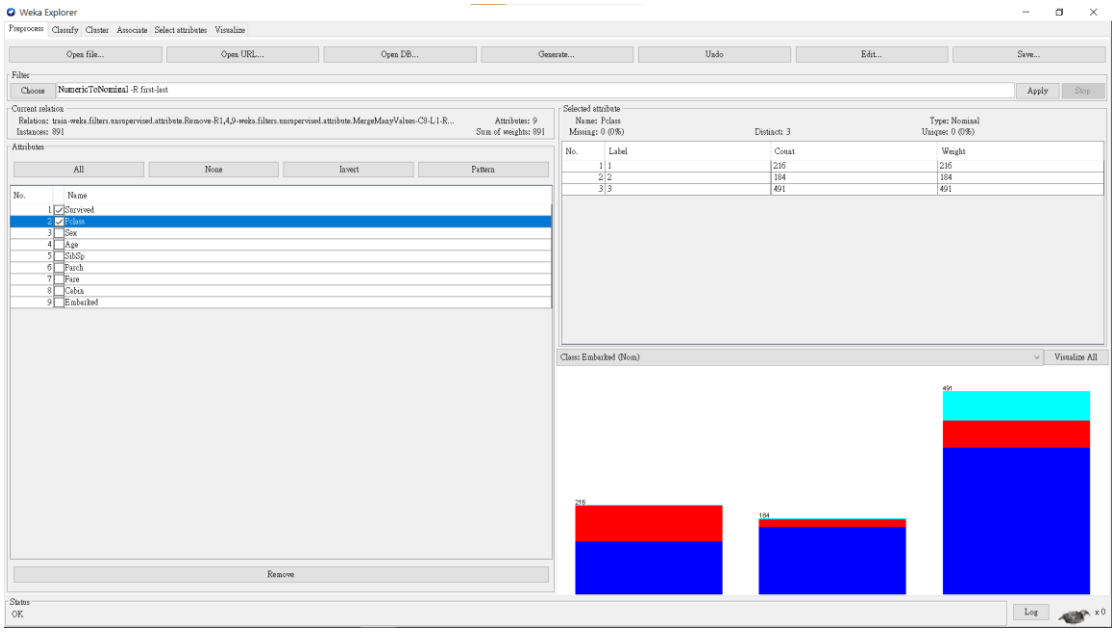
1.3 直接使用 ReplaceMissingValues 這個函數，將平均值 29.699 帶入全部空值





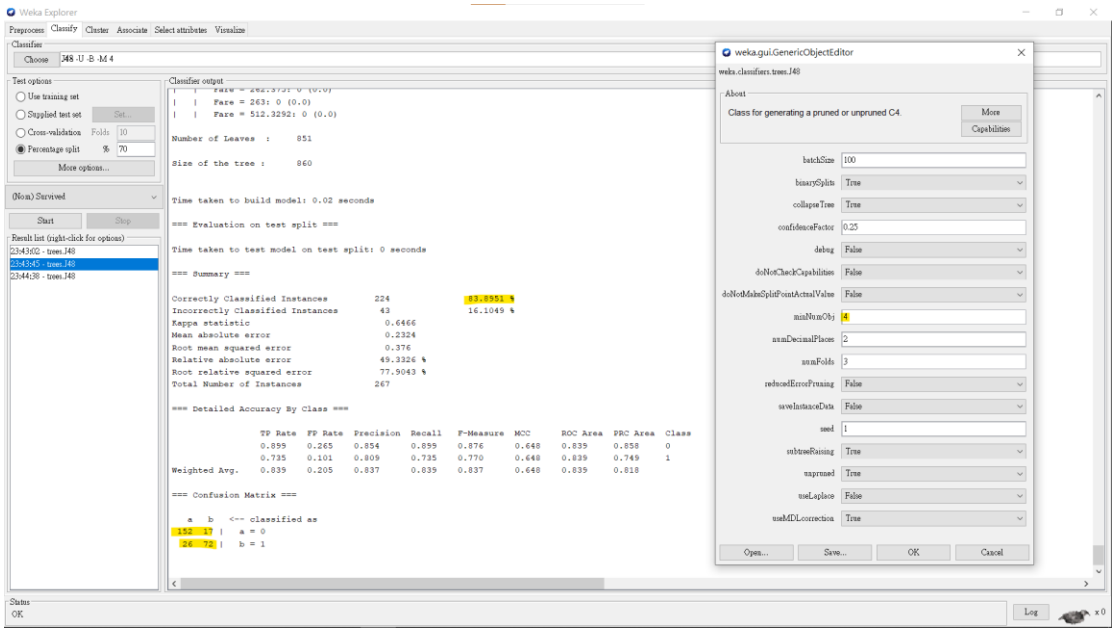
No.	1: Survived Numeric	2: Pclass Numeric	3: Sex Nominal	4: Age Numeric	5: SibSp Numeric	6: Parch Numeric	7: Fare Numeric	8: Cabin Nominal	9: Embarked Nominal
1	0.0	3.0	male	22.0	1.0	0.0	7.25	0	S
2	1.0	1.0	female	38.0	1.0	0.0	71.2833	1	C
3	1.0	3.0	female	26.0	0.0	0.0	7.925	0	S
4	1.0	1.0	female	35.0	1.0	0.0	53.1	1	S
5	0.0	3.0	male	35.0	0.0	0.0	8.05	0	S
6	0.0	3.0	male	29.69...	0.0	0.0	8.4583	0	Q
7	0.0	1.0	male	54.0	0.0	0.0	51.8625	1	S
8	0.0	3.0	male	2.0	3.0	1.0	21.075	0	S
9	1.0	3.0	female	27.0	0.0	2.0	11.1333	0	S
10	1.0	2.0	female	14.0	1.0	0.0	30.0708	0	C
11	1.0	3.0	female	4.0	1.0	1.0	16.7	1	S
12	1.0	1.0	female	58.0	0.0	0.0	26.55	1	S
13	0.0	3.0	male	20.0	0.0	0.0	8.05	0	S
14	0.0	3.0	male	39.0	1.0	5.0	31.275	0	S
15	0.0	3.0	female	14.0	0.0	0.0	7.8542	0	S
16	1.0	2.0	female	55.0	0.0	0.0	16.0	0	S
17	0.0	3.0	male	2.0	4.0	1.0	29.125	0	Q
18	1.0	2.0	male	29.69...	0.0	0.0	13.0	0	S
19	0.0	3.0	female	31.0	1.0	0.0	18.0	0	S
20	1.0	3.0	female	29.69...	0.0	0.0	7.225	0	C
21	0.0	2.0	male	35.0	0.0	0.0	26.0	0	S
22	1.0	2.0	male	34.0	0.0	0.0	13.0	1	S
23	1.0	3.0	female	15.0	0.0	0.0	8.0292	0	Q
24	1.0	1.0	male	28.0	0.0	0.0	35.5	1	S
25	0.0	3.0	female	8.0	3.0	1.0	21.075	0	S
26	1.0	3.0	female	38.0	1.0	5.0	31.3875	0	S
27	0.0	3.0	male	29.69...	0.0	0.0	7.225	0	C
28	0.0	1.0	male	19.0	3.0	2.0	263.0	1	S
29	1.0	3.0	female	29.69...	0.0	0.0	7.8792	0	Q
30	0.0	3.0	male	29.69...	0.0	0.0	7.8958	0	S
31	0.0	1.0	male	40.0	0.0	0.0	27.7208	0	C
32	1.0	1.0	female	29.69...	1.0	0.0	146.5...	1	C
33	1.0	3.0	female	29.69...	0.0	0.0	7.75	0	Q
34	0.0	2.0	male	66.0	0.0	0.0	10.5	0	S
35	0.0	1.0	male	28.0	1.0	0.0	82.1708	0	C
36	0.0	1.0	male	42.0	1.0	0.0	52.0	0	S
37	1.0	3.0	male	29.69...	0.0	0.0	7.2292	0	C

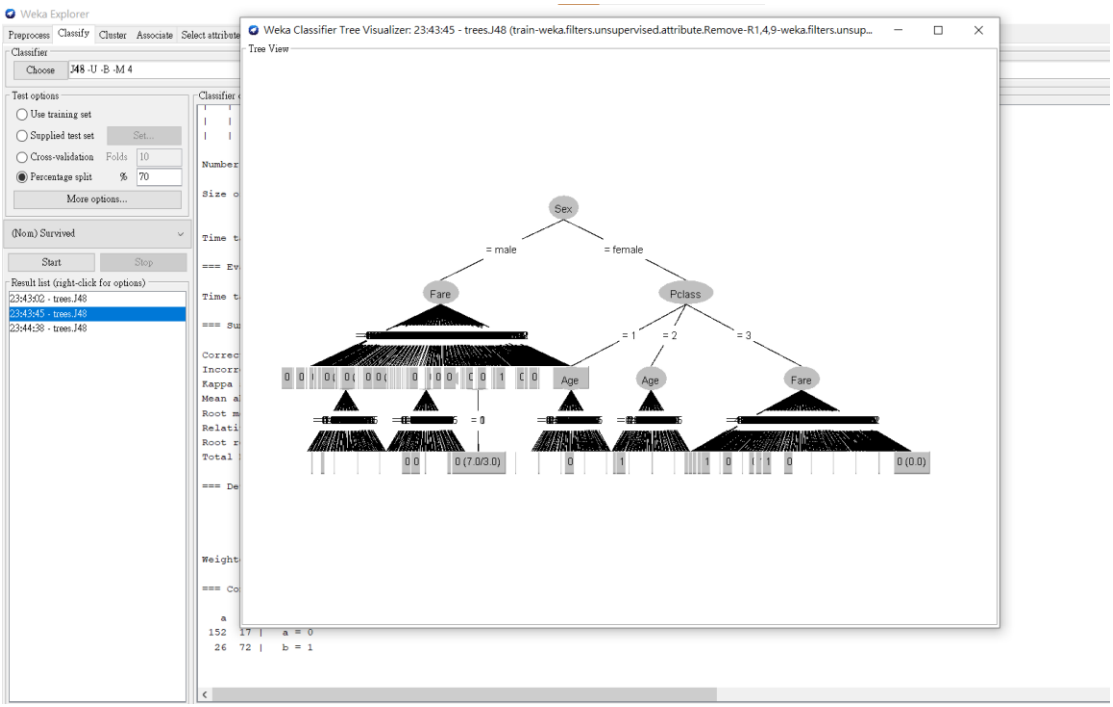
1.4 直接對 Survived 與 Pclass 使用 NominalToNumeric，如下圖



因 survival 以 0 和 1 代表生存與否，僅為一個數字代號，數字本身沒有意義，於是將型態改為 nominal 較恰當；Pclass 則代表傳票等級，與 survival 意義相同。

1.5




Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attribute
Classifier
Choose
J48 -U -B -M 4
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
70
More options...
(Nom) Survived
Start
Stop
Result list (right-click for options)
23:43:02 - trees.J48
23:43:45 - trees.J48
23:44:28 - trees.J48
Weka Classifier Tree Visualizer: 23:43:45 - trees.J48 (train-weka.filters.unsupervised.attribute.Remove-R1,4,9-weka.filters.unsup...
Tree View
Sex
= male
= female
Fare
Pclass
= 1
= 2
= 3
Age
Age
Fare
0 (7.0/3.0)
0 (0.0)
a
152 17 | a = 0
26 72 | b = 1