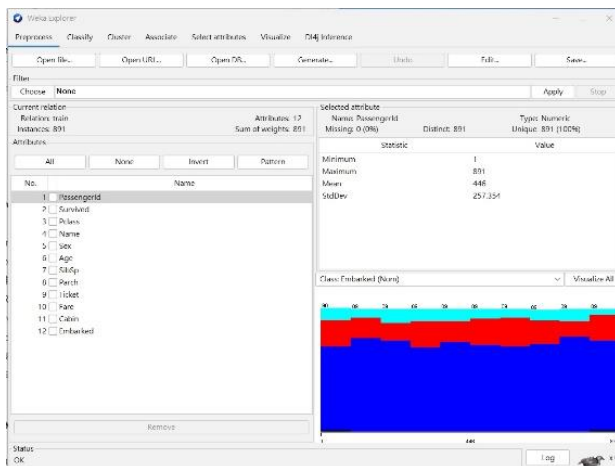


## 電子商務技術\_HW2

108403201 資管三 A 黃名揚

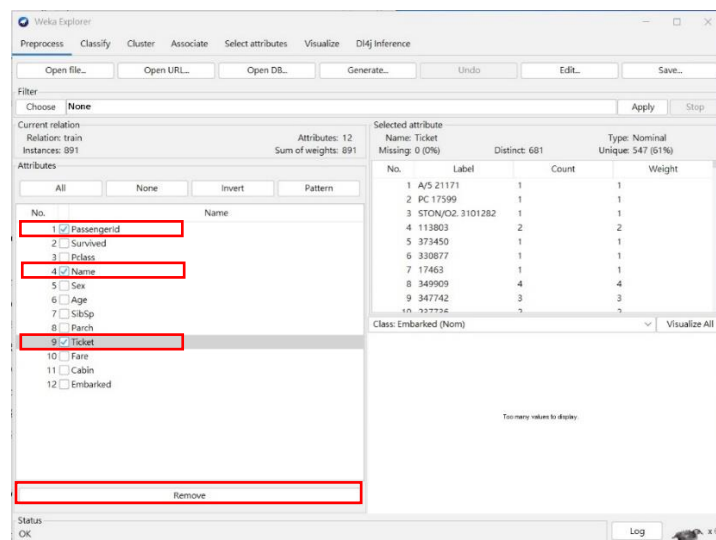
壹、請使用 weka 完成以下題目，並截圖結果附上適當說明

### 1.1 載入 train.arff，將 PassengerId、Name、Ticket 欄位刪除

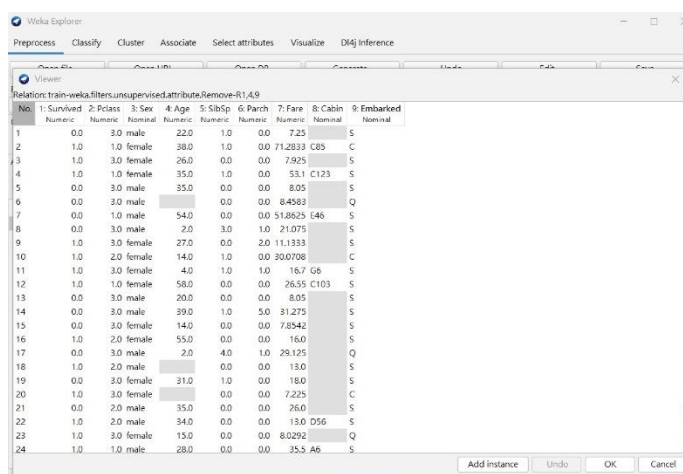
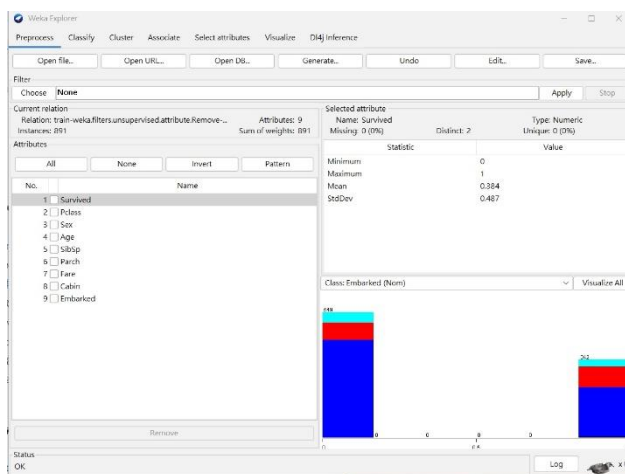


No.	1: PassengerId	2: Survived	3: Pclass	4: Name	5: Sex	6: Age	7: SibSp	8: Parch	9: Ticket	10: Fare	11: Cabin	12: Embarked
1	1.0	0.0	3.0	Braund, _ male	male	22.0	1.0	0.0	A/5 211...	7.25		S
2	2.0	1.0	1.0	Cumings, _ female	female	38.0	1.0	0.0	PC 175...	71.2833	C85	C
3	3.0	1.0	3.0	Heikkinen, _ female	female	26.0	0.0	0.0	STON/O...	7.925		S
4	4.0	1.0	1.0	Futrelle, _ female	female	35.0	1.0	0.0	113803	53.1	C123	S
5	5.0	0.0	3.0	Allen, M. _ male	male	35.0	0.0	0.0	373450	8.05		S
6	6.0	0.0	3.0	Moran, _ male	male		0.0	0.0	330877	8.4583		Q
7	7.0	0.0	1.0	McCart, _ male	male	54.0	0.0	0.0	17463	51.8625	E46	S
8	8.0	0.0	3.0	Palsson, _ male	male	2.0	3.0	1.0	349909	21.075		S
9	9.0	1.0	3.0	Johnson, _ female	female	27.0	0.0	2.0	347742	11.1333		S
10	10.0	1.0	2.0	Nasser, _ female	female	14.0	1.0	0.0	237736	30.0708		C
11	11.0	1.0	3.0	Sandstr, _ female	female	4.0	1.0	1.0	PP 9549	16.7	G6	S
12	12.0	1.0	1.0	Bonnell, _ female	female	58.0	0.0	0.0	113783	26.55	C103	S
13	13.0	0.0	3.0	Saunders, _ male	male	20.0	0.0	0.0	A/5 21...	8.05		S
14	14.0	0.0	3.0	Anderss, _ male	male	39.0	1.0	5.0	347082	31.275		S
15	15.0	0.0	3.0	Vestro, _ female	female	14.0	0.0	0.0	350406	7.8542		S
16	16.0	1.0	2.0	Hewlett, _ female	female	55.0	0.0	0.0	248706	16.0		S
17	17.0	0.0	3.0	Rice, Ma, _ male	male	2.0	4.0	1.0	382652	29.125		Q
18	18.0	1.0	2.0	William, _ male	male		0.0	0.0	244373	13.0		S
19	19.0	0.0	3.0	Vander, _ female	female	31.0	1.0	0.0	345763	18.0		S
20	20.0	1.0	3.0	Masell, _ female	female	0.0	0.0	0.0	2649	7.225		C
21	21.0	0.0	2.0	Fynney, _ male	male	35.0	0.0	0.0	239865	26.0		S
22	22.0	1.0	2.0	Beesley, _ male	male	34.0	0.0	0.0	248698	13.0	D56	S
23	23.0	1.0	3.0	McGow, _ female	female	15.0	0.0	0.0	330923	8.0292		Q
24	24.0	1.0	1.0	Sloper, _ male	male	28.0	0.0	0.0	113788	35.5	A6	S

### Step(1)載入 train.arff



### Step(2) 選取 PassengerId、Name、Ticket 欄位並且刪除



### Step(3)刪除完成後資料

## 1.2 將 Cabin 的非空值以 1 替代，空值以 0 填入（先用 MergeManyValues

取代非空值，再用 `ReplaceMissingWithUserConstant` 取代空值)

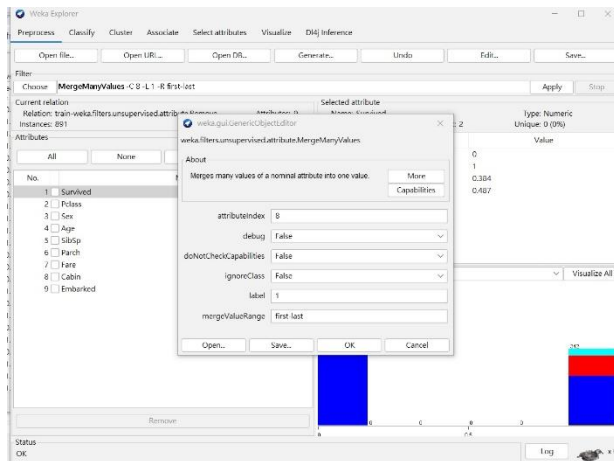
Viewer

Relation: train-weka.filters.unsupervised.attribute.Remove-R1.4.9-weka.filters.unsupervised.attribute.Copy-weka.filters.unsupervised.attribute.Copy-weka.filters.unsupervised.attribute.Copy-weka.filters.unsupervised.attribute.Copy

No.	1: Survived	2: Pclass	3: Sex	4: Age	5: SibSp	6: Parch	7: Fare	8: Cabin	9: Embarked
	Numeric	Numeric	Nominal	Numeric	Numeric	Numeric	Numeric	Nominal	Nominal
1	0.0	3.0	male	22.0	1.0	0.0	7.25	S	
2	1.0	1.0	female	38.0	1.0	0.0	71.2833	C85	C
3	1.0	3.0	female	26.0	0.0	0.0	7.925		S
4	1.0	1.0	female	35.0	1.0	0.0	53.1	C123	S
5	0.0	3.0	male	35.0	0.0	0.0	8.05		S
6	0.0	3.0	male		0.0	0.0	8.4583		Q
7	0.0	1.0	male	54.0	0.0	0.0	51.8625	E46	S
8	0.0	3.0	male	2.0	3.0	1.0	21.075		S
9	1.0	3.0	female	27.0	0.0	2.0	11.1333		S
10	1.0	2.0	female	14.0	1.0	0.0	30.0708		C
11	1.0	3.0	female	4.0	1.0	1.0	16.7	G6	S
12	1.0	1.0	female	58.0	0.0	0.0	26.55	C103	S
13	0.0	3.0	male	20.0	0.0	0.0	8.05		S
14	0.0	3.0	male	39.0	1.0	5.0	31.275		S
15	0.0	3.0	female	14.0	0.0	0.0	78.542		S
16	1.0	2.0	female	55.0	0.0	0.0	16.0		S
17	0.0	3.0	male	2.0	4.0	1.0	29.125		Q
18	1.0	2.0	male		0.0	0.0	13.0		S
19	0.0	3.0	female	31.0	1.0	0.0	18.0		S
20	1.0	3.0	female		0.0	0.0	7.225		C
21	0.0	2.0	male	35.0	0.0	0.0	26.0		S
22	1.0	2.0	male	34.0	0.0	0.0	13.0	D56	S
23	1.0	3.0	female	15.0	0.0	0.0	8.0292		Q
24	1.0	1.0	male	28.0	0.0	0.0	35.5	A6	S

Add instance Undo OK Cancel

### Step(1)處理前資料

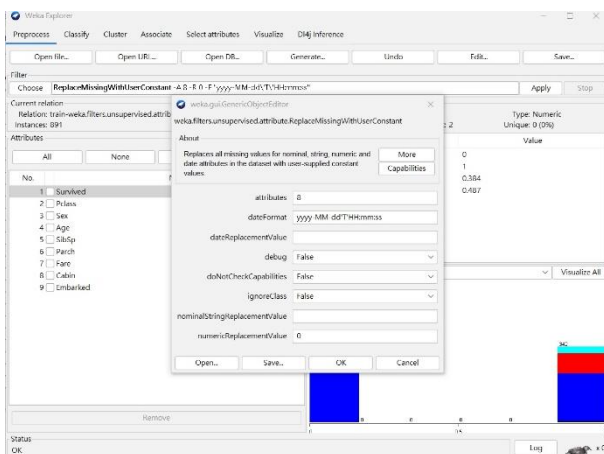


Viewer

Relation: train-weka.filters.unsupervised.attribute.Remove-R1.4.9-weka.filters.unsupervised.attribute.MergeManyValues-C8-L1-Rfirst-last

No.	1: Survived	2: Pclass	3: Sex	4: Age	5: SibSp	6: Parch	7: Fare	8: Cabin	9: Embarked
	Numeric	Numeric	Nominal	Numeric	Numeric	Numeric	Numeric	Nominal	Nominal
1	0.0	3.0	male	22.0	1.0	0.0	7.25	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	S	
4	1.0	1.0	female	35.0	1.0	0.0	53.1	S	
5	0.0	3.0	male	35.0	0.0	0.0	8.05	S	
6	0.0	3.0	male		0.0	0.0	8.4583	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	S	
9	1.0	3.0	female	27.0	0.0	2.0	11.1333	S	
10	1.0	2.0	female	14.0	1.0	0.0	30.0708	C	
11	1.0	1.0	female	4.0	1.0	1.0	16.7	S	
12	1.0	1.0	female	58.0	0.0	0.0	26.55	S	
13	0.0	3.0	male	20.0	0.0	0.0	8.05	S	
14	0.0	3.0	male	39.0	1.0	5.0	31.275	S	
15	0.0	3.0	female	14.0	0.0	0.0	78.542	S	
16	1.0	2.0	female	55.0	0.0	0.0	16.0	S	
17	0.0	3.0	male	2.0	4.0	1.0	29.125	Q	
18	1.0	2.0	male		0.0	0.0	13.0	S	
19	0.0	3.0	female	31.0	1.0	0.0	18.0	S	
20	1.0	3.0	female		0.0	0.0	7.225	C	
21	0.0	2.0	male	35.0	0.0	0.0	26.0	S	
22	1.0	2.0	male	34.0	0.0	0.0	13.0	S	
23	1.0	3.0	female	15.0	0.0	0.0	8.0292	Q	
24	1.0	1.0	male	28.0	0.0	0.0	35.5	1	S

Step(2) 用 MergeManyValues 取代非空值



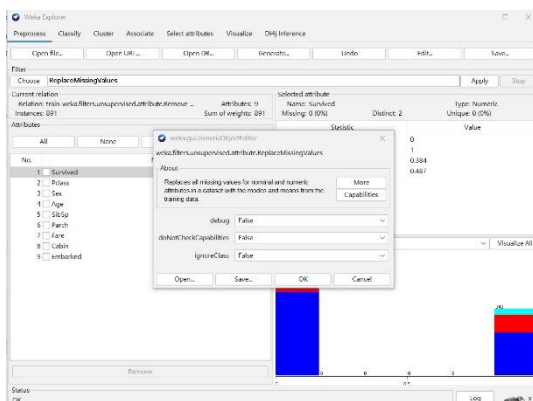
Viewer

Relation: train-weka.filters.unsupervised.attribute.Remove-R1.4.9-weka.filters.unsupervised.attribute.MergeManyValues-C8-L1-Rfirst-last-weka.filters.unsupervised.attribute.ReplaceMissingWithUserConstant

No.	1: Survived	2: Pclass	3: Sex	4: Age	5: SibSp	6: Parch	7: Fare	8: Cabin	9: Embarked
	Numeric	Numeric	Nominal	Numeric	Numeric	Numeric	Numeric	Nominal	Nominal
1	0.0	3.0	male	22.0	1.0	0.0	7.25	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	S	
4	1.0	1.0	female	35.0	1.0	0.0	53.1	S	
5	0.0	3.0	male	35.0	0.0	0.0	8.05	S	
6	0.0	3.0	male		0.0	0.0	8.4583	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	S	
9	1.0	3.0	female	27.0	0.0	2.0	11.1333	S	
10	1.0	2.0	female	14.0	1.0	0.0	30.0708	C	
11	1.0	1.0	female	4.0	1.0	1.0	16.7	S	
12	1.0	1.0	female	58.0	0.0	0.0	26.55	S	
13	0.0	3.0	male	20.0	0.0	0.0	8.05	S	
14	0.0	3.0	male	39.0	1.0	5.0	31.275	S	
15	0.0	3.0	female	14.0	0.0	0.0	78.542	S	
16	1.0	2.0	female	55.0	0.0	0.0	16.0	S	
17	0.0	3.0	male	2.0	4.0	1.0	29.125	Q	
18	1.0	2.0	male		0.0	0.0	13.0	S	
19	0.0	3.0	female	31.0	1.0	0.0	18.0	S	
20	1.0	3.0	female		0.0	0.0	7.225	C	
21	0.0	2.0	male	35.0	0.0	0.0	26.0	S	
22	1.0	2.0	male	34.0	0.0	0.0	13.0	S	
23	1.0	3.0	female	15.0	0.0	0.0	8.0292	Q	
24	1.0	1.0	male	28.0	0.0	0.0	35.5	1	S

Step(3) 再用 ReplaceMissingWithUserConstant 取代空值

### 1.3 使用 ReplaceMissingValues 將 Age 的空值以 Age 平均數填入



Viewer

Relation: train-weka.filters.unsupervised.attribute.Remove-R1.4.9-weka.filters.unsupervised.attribute.MergeManyValues-C8-L1-Rfirst-last-weka.filters.unsupervised.attribute.ReplaceMissingValues

No.	1: Survived	2: Pclass	3: Sex	4: Age	5: SibSp	6: Parch	7: Fare	8: Cabin	9: Embarked
	Numeric	Numeric	Nominal	Numeric	Numeric	Numeric	Numeric	Nominal	Nominal
1	0.0	3.0	male	22.0	1.0	0.0	7.25	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	S	
4	1.0	1.0	female	35.0	1.0	0.0	53.1	S	
5	0.0	3.0	male	35.0	0.0	0.0	8.05	S	
6	0.0	3.0	male	29.6999	0.0	0.0	8.4583	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	S	
9	1.0	3.0	female	27.0	0.0	2.0	11.1333	S	
10	1.0	2.0	female	14.0	1.0	0.0	30.0708	C	
11	1.0	1.0	female	4.0	1.0	1.0	16.7	S	
12	1.0	1.0	female	58.0	0.0	0.0	26.55	S	
13	0.0	3.0	male	20.0	0.0	0.0	8.05	S	
14	0.0	3.0	male	39.0	1.0	5.0	31.275	S	
15	0.0	3.0	female	14.0	0.0	0.0	78.542	S	
16	1.0	2.0	female	55.0	0.0	0.0	16.0	S	
17	0.0	3.0	male	2.0	4.0	1.0	29.125	Q	
18	1.0	2.0	male	29.6999	0.0	0.0	13.0	S	
19	0.0	3.0	female	31.0	1.0	0.0	18.0	S	
20	1.0	3.0	female	29.6999	0.0	0.0	7.225	C	
21	0.0	2.0	male	35.0	0.0	0.0	26.0	S	
22	1.0	2.0	male	34.0	0.0	0.0	13.0	S	
23	1.0	3.0	female	15.0	0.0	0.0	8.0292	Q	
24	1.0	1.0	male	28.0	0.0	0.0	35.5	1	S

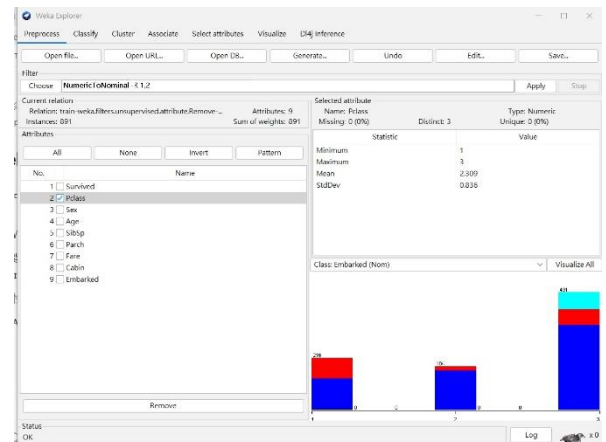
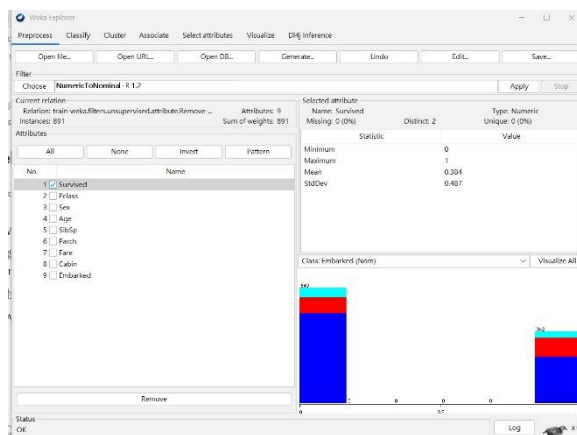
Step(1) 使用 ReplaceMissingValues 將 Age 的空值以 Age 平均數填入

## 1.4 將 Survived 與 Pclass 轉為 Nominal，並說明為何 Numeric 無法使用

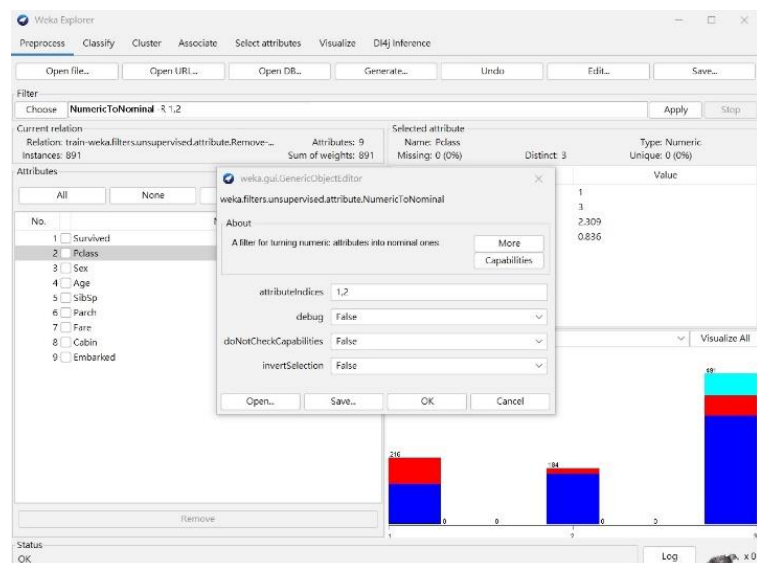
在 Decision tree

Answer: Numeric 為數字型連續型的變數，無法進行分類，如果要配合使用決

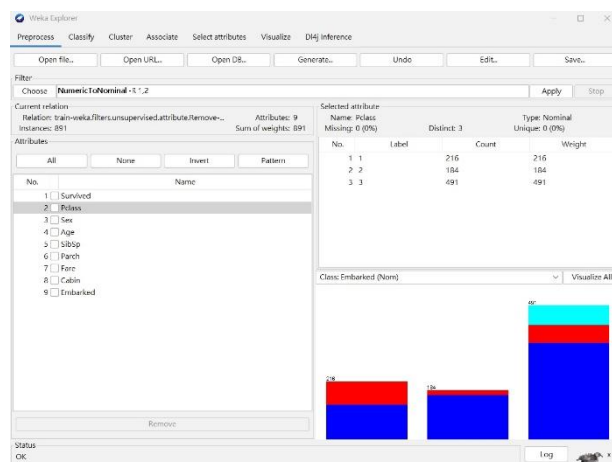
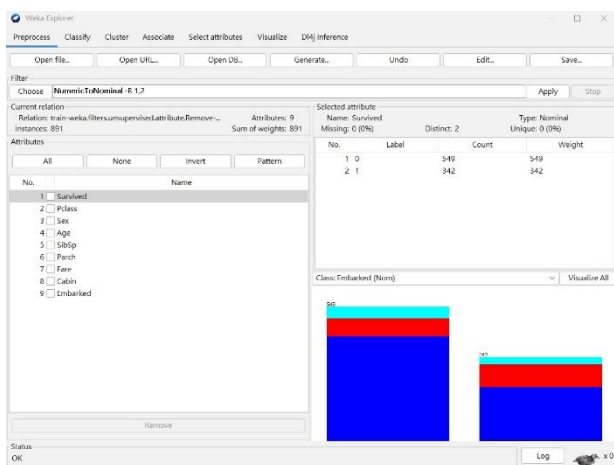
策樹，需要進行離散化。



Step(1)處理前資料



Step(2) 將 Survived 與 Pclass 轉為 Nominal



step(3)處理後資料

1.5 以 70% 切割訓練資料，使用 J48 對 Survived 進行分類，並截圖分類

準確率、混淆矩陣及視覺化的 Decision tree

