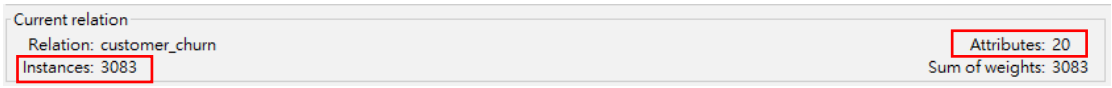
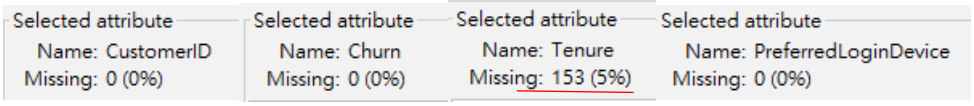
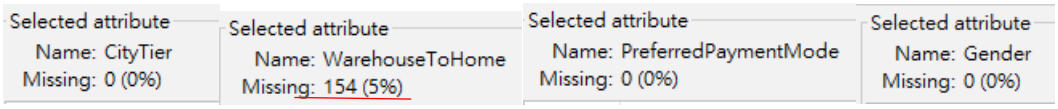
ECT_HW3_109403021_Weka 題

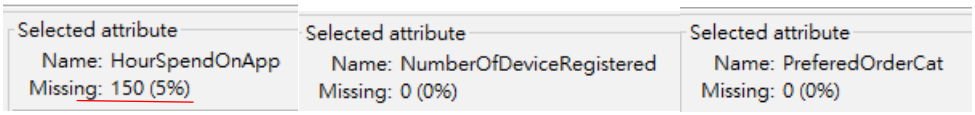
1. 資料筆數和屬性數量:

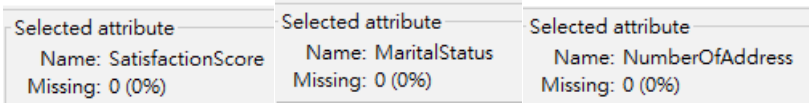


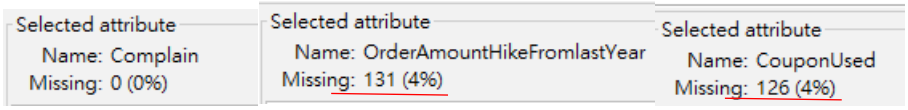
各欄位空值個數:

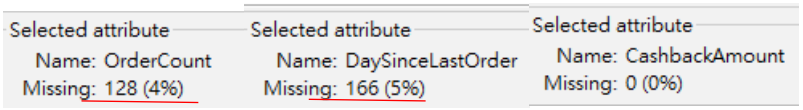




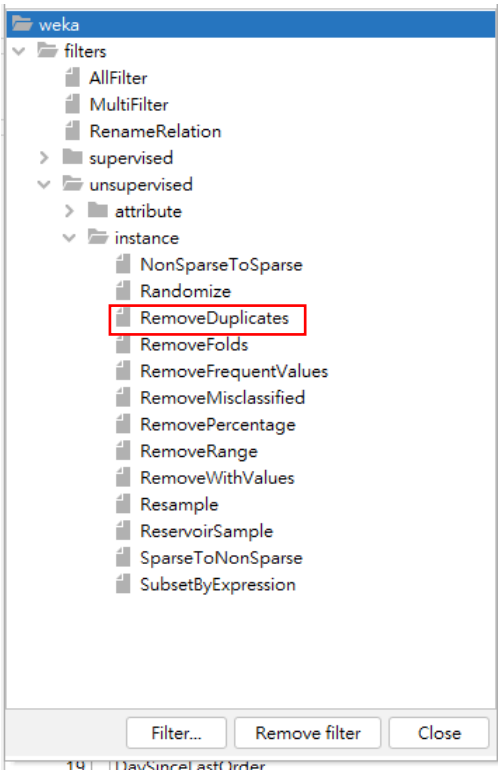




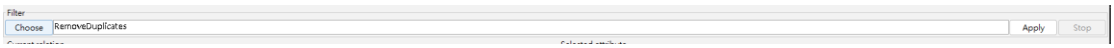




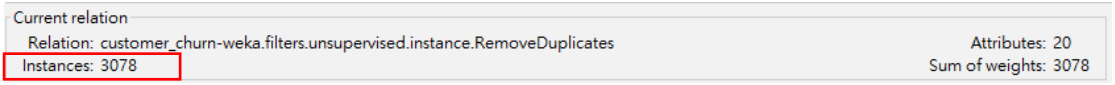
2. 選取到 **RemoveDuplicates**



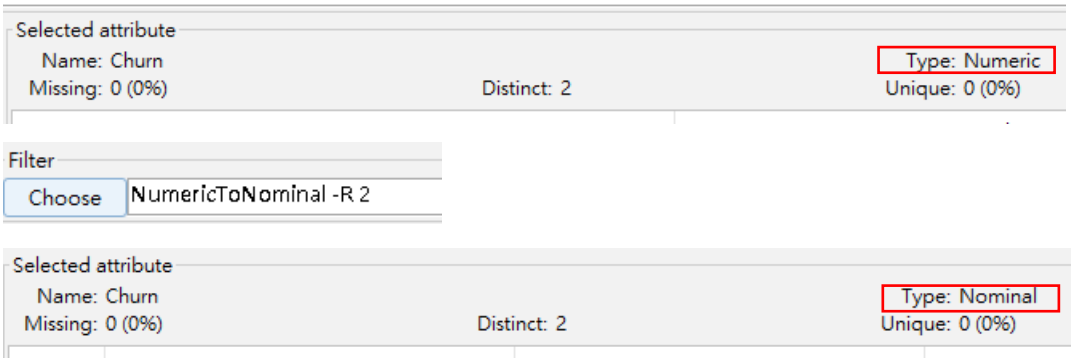
按下 **Apply**



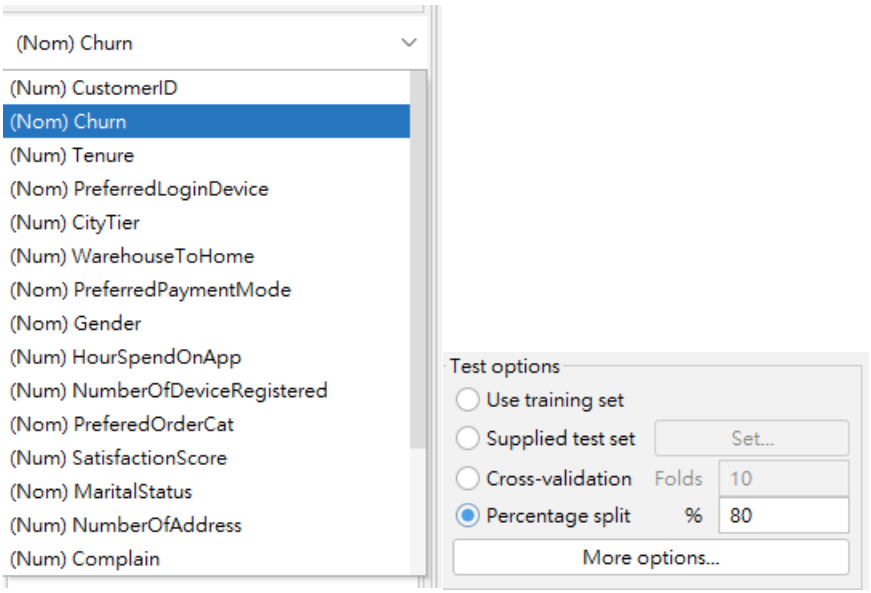
得到剩餘的資料筆數



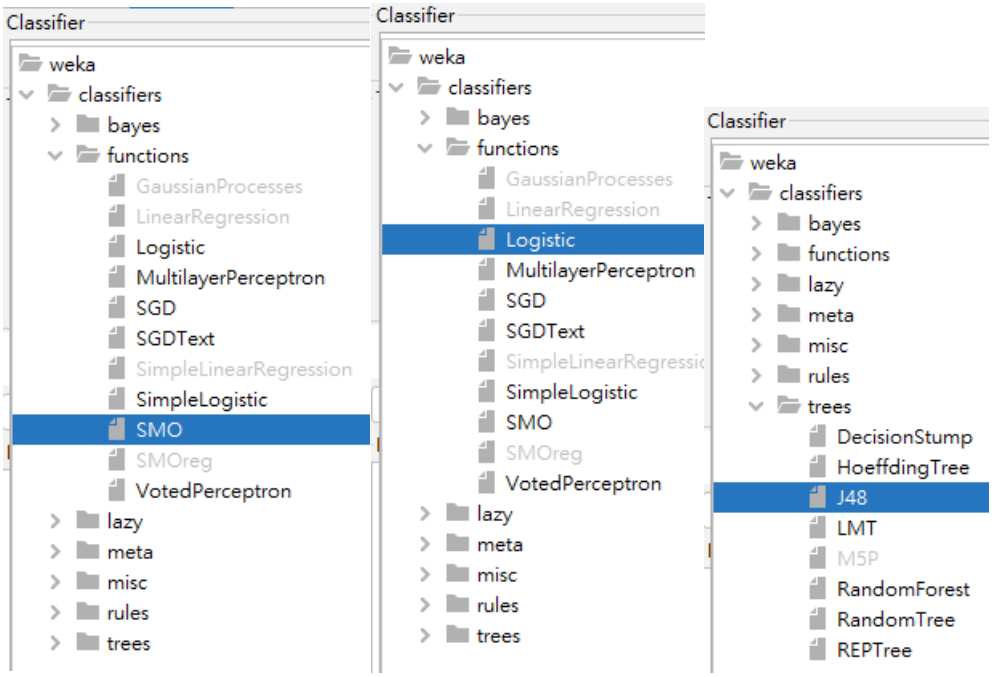
3. 將要預測的 **Churn** 轉為 **nominal**



4. 三個模型訓練、測試前都先把預測的欄位調整為 **Churn**，且三個模型我都將資料切割為 **8:2(訓練:測試)**



分別選擇 **SMO**、**Logistic** 和 **J48**，來分別訓練、測試 **SVM**、**Logistic Regression**、**Decision Tree** 模型



得到以下結果:

SVM

```
=== Summary ===

Correctly Classified Instances         497               80.6818 %
Incorrectly Classified Instances       119               19.3182 %
Kappa statistic                          0.5276
Mean absolute error                      0.1932
Root mean squared error                  0.4395
Relative absolute error                 45.7395 %
Root relative squared error             96.4154 %
Total Number of Instances              616
```

Logistic

```
=== Summary ===

Correctly Classified Instances         507               82.3052 %
Incorrectly Classified Instances       109               17.6948 %
Kappa statistic                          0.5687
Mean absolute error                      0.2472
Root mean squared error                  0.3546
Relative absolute error                 58.5385 %
Root relative squared error             77.7876 %
Total Number of Instances              616
```

Decision Tree

```
=== Summary ===

Correctly Classified Instances         556               90.2597 %
Incorrectly Classified Instances        60                9.7403 %
Kappa statistic                          0.7768
Mean absolute error                      0.1261
Root mean squared error                  0.2869
Relative absolute error                 29.8471 %
Root relative squared error             62.9301 %
Total Number of Instances              616
```

可以看到在只有調整資料集成訓練比測試 **8:2**，其他參數都使用默認沒有調整的情況下，**3** 個模型的準確率如上，**Decision Tree** 的表現最好