

Hyperparameters of Algorithms for the whole dataset are listed as follows.

1. Decision tree

Set as default

weka.gui.GenericObjectEditor

weka.classifiers.trees.J48

About

Class for generating a pruned or unpruned C4.

More

Capabilities

batchSize 100

binarySplits False

collapseTree True

confidenceFactor 0.25

debug False

doNotCheckCapabilities False

doNotMakeSplitPointActualValue False

minNumObj 2

numDecimalPlaces 2

numFolds 3

reducedErrorPruning False

saveInstanceData False

seed 1

subtreeRaising True

unpruned False

useLaplace False

useMDLcorrection True

Open... Save... OK Cancel

2. Bagging with decision tree

Classifier Choose J48 (orange arrow)

weka.gui.GenericObjectEditor

weka.classifiers.meta.Bagging

About

Class for bagging a classifier to reduce variance.

More

Capabilities

bagSizePercent 100

batchSize 100

calcOutOfBag False

classifier Choose J48 -C 0.25 -M 2

debug False

doNotCheckCapabilities False

numDecimalPlaces 2

numExecutionSlots 1

numIterations 10

outputOutOfBagComplexityStatistics False

printClassifiers False

representCopiesUsingWeights False

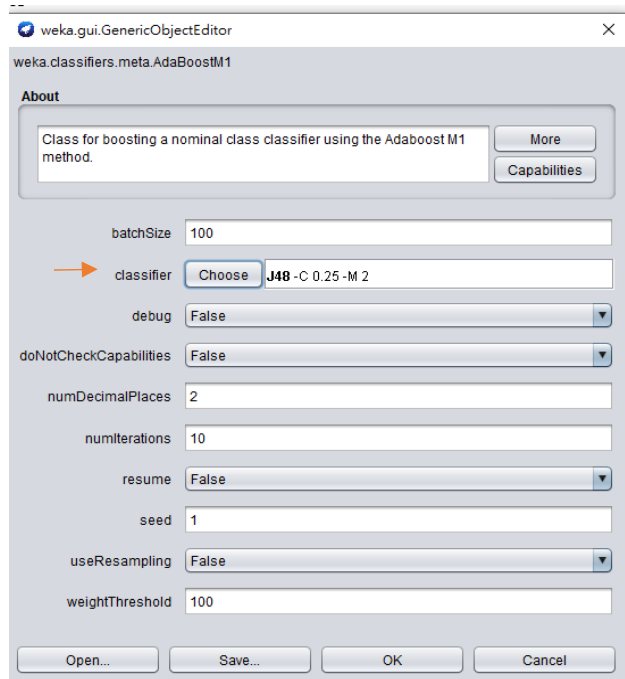
seed 1

storeOutOfBagPredictions False

Open... Save... OK Cancel

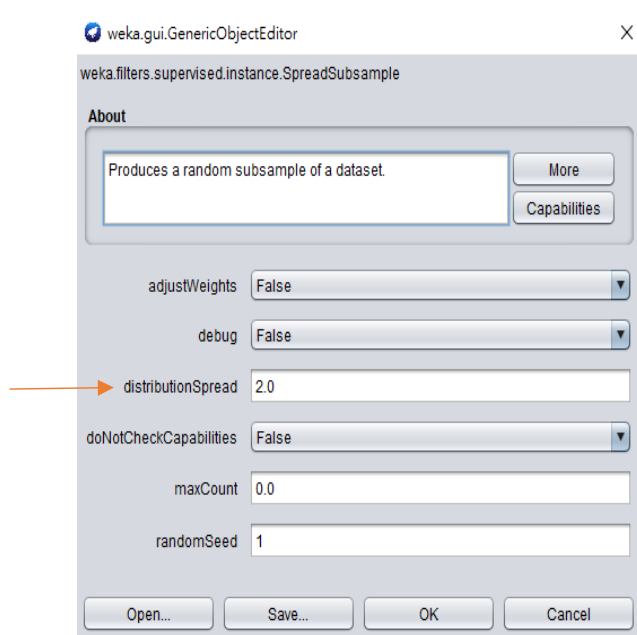
3. AdaBoost with decision tree

Classifier Choose J48 (orange arrow)



4. SpreadSubsample with Decision tree

The value of distributionSpread was adjusted as 2.0



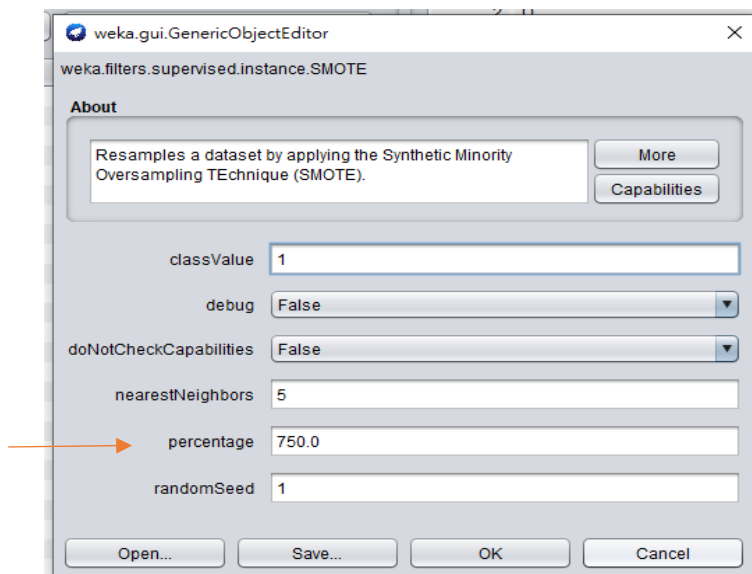
5. SMOTE with Decision tree

Class A has the highest frequency and its frequency is not adjusted. The frequencies of other classes were adjusted to nearly the same with that of Class A, by adjusting the percentage (orange arrow). The following figures are listed by the order of classValue. The classValue of Class A is 4,

For Class C

classValue: 1

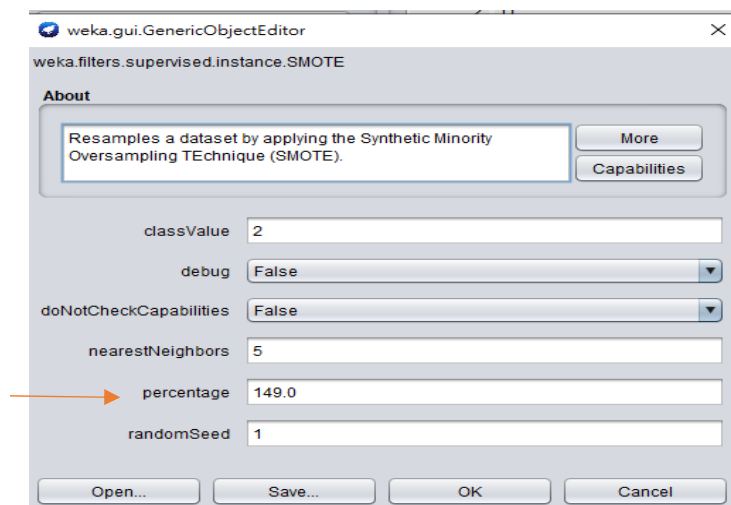
percentage: 750



For Class D

classValue: 2

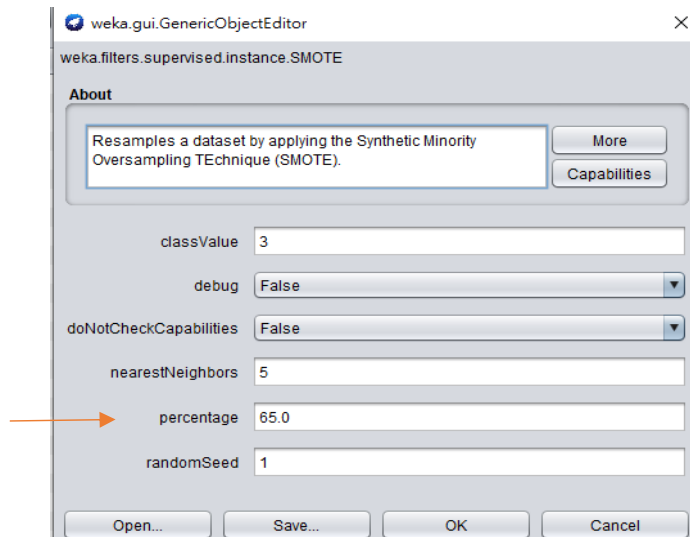
percentage: 149



For Class B

classValue: 3

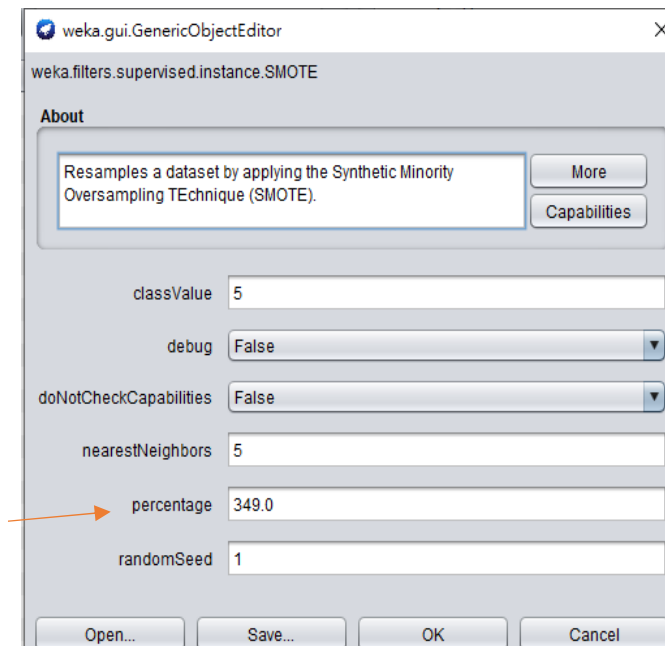
percentage: 65



Class E

classValue: 5

percentage: 349



6. Multilayer perceptron

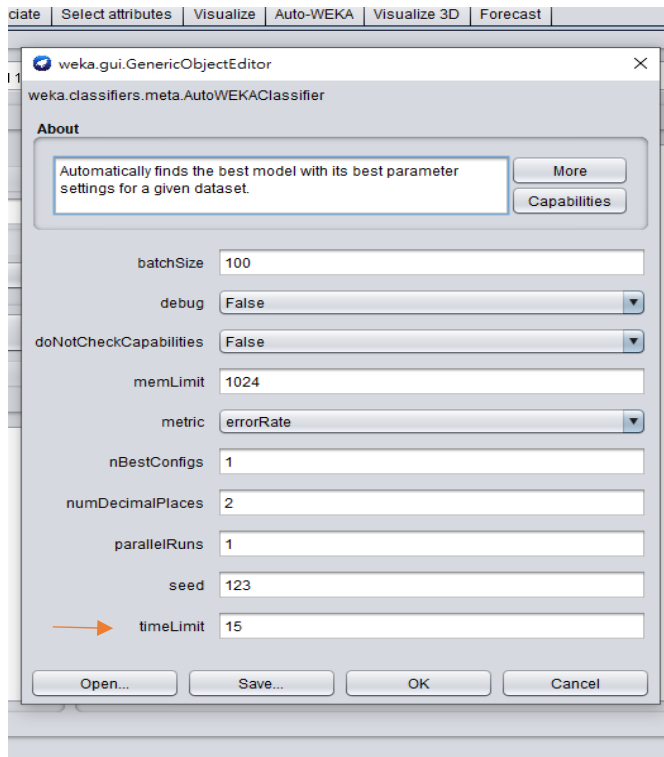
Set as default

The screenshot shows the 'weka.gui.GenericObjectEditor' window for the 'weka.classifiers.functions.MultilayerPerceptron' classifier. The 'About' tab is active, displaying a description: 'A classifier that uses backpropagation to learn a multi-layer perceptron to classify instances.' Below this, there are two buttons: 'More' and 'Capabilities'. The main area contains various configuration options, each with a label and a value field (either a text box or a dropdown menu). At the bottom, there are four buttons: 'Open...', 'Save...', 'OK', and 'Cancel'.

Property	Value
GUI	False
autoBuild	True
batchSize	100
debug	False
decay	False
doNotCheckCapabilities	False
hiddenLayers	a
learningRate	0.3
momentum	0.2
nominalToBinaryFilter	True
normalizeAttributes	True
normalizeNumericClass	True
numDecimalPlaces	2
reset	True
resume	False
seed	0
trainingTime	500
validationSetSize	0
validationThreshold	20

7. AutoWEKA

Set as default. For better performance, try giving Auto-WEKA more time.
(orange arrow)



Hyperparameters of Algorithm for General Surgery dataset

Only the hyperparameters of SMOTE with decision tree are listed here. The hyperparameters of other algorithms are set as default.

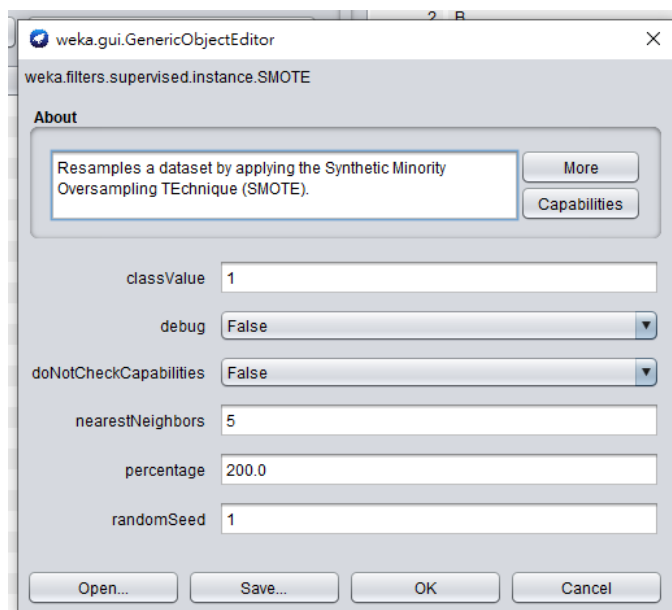
SMOTE with Decision tree:

Class A has the highest frequency and its frequency is not adjusted. The frequencies of other classes were adjusted to nearly the same with that of Class A, by adjusting the percentage (orange arrow). The following figures are listed by the order of classValue. The classValue of Class A is 3,

For class D:

class Value: 1

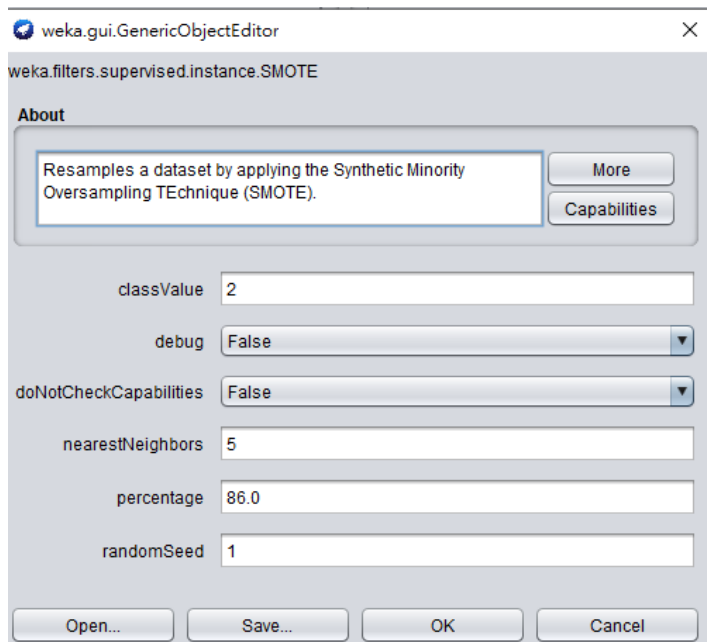
Percentage: 200



For Class B

classValue: 2

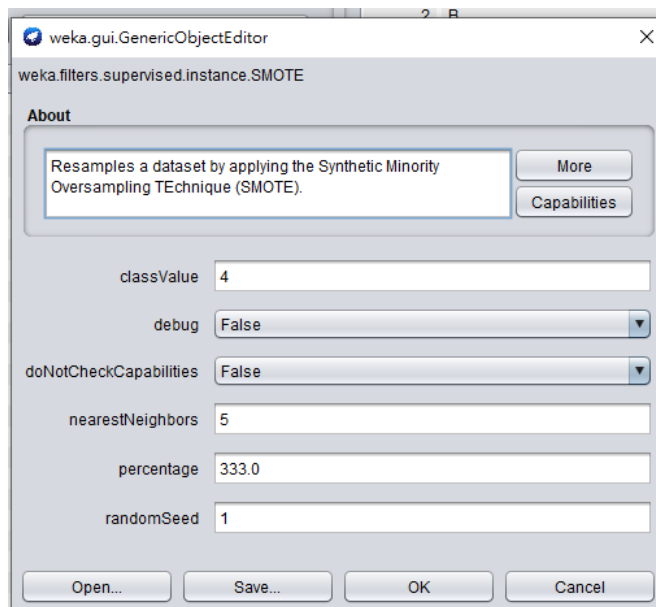
percentage: 86



For Class E

classValue: 4

percentage: 333

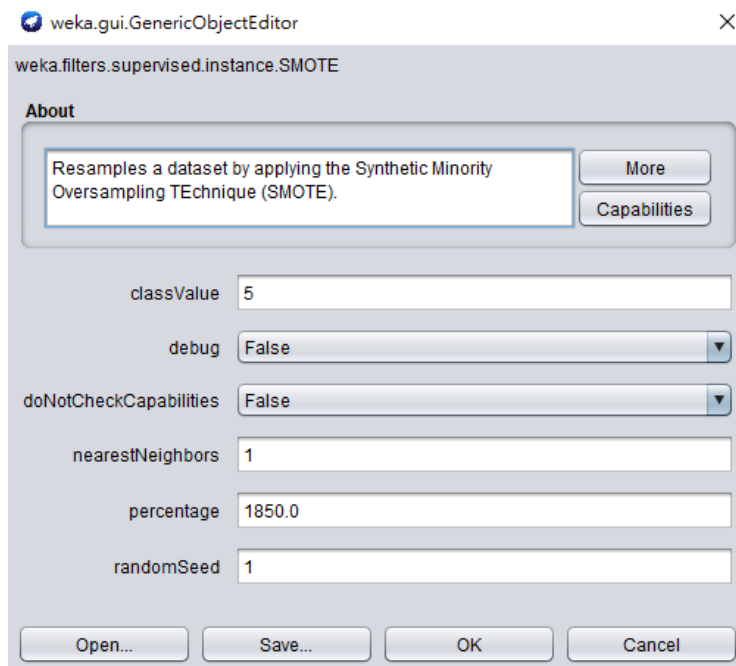


For Class C

classValue: 5

nearestNeighbors: 1

percentage: 1850



The hyperparameters of SimpleLogistic are as default setting.

Hyperparameters of Algorithm for General Urology dataset

Only the hyperparameters of SMOTE with decision tree are listed here. The hyperparameters of other algorithms are set as default.

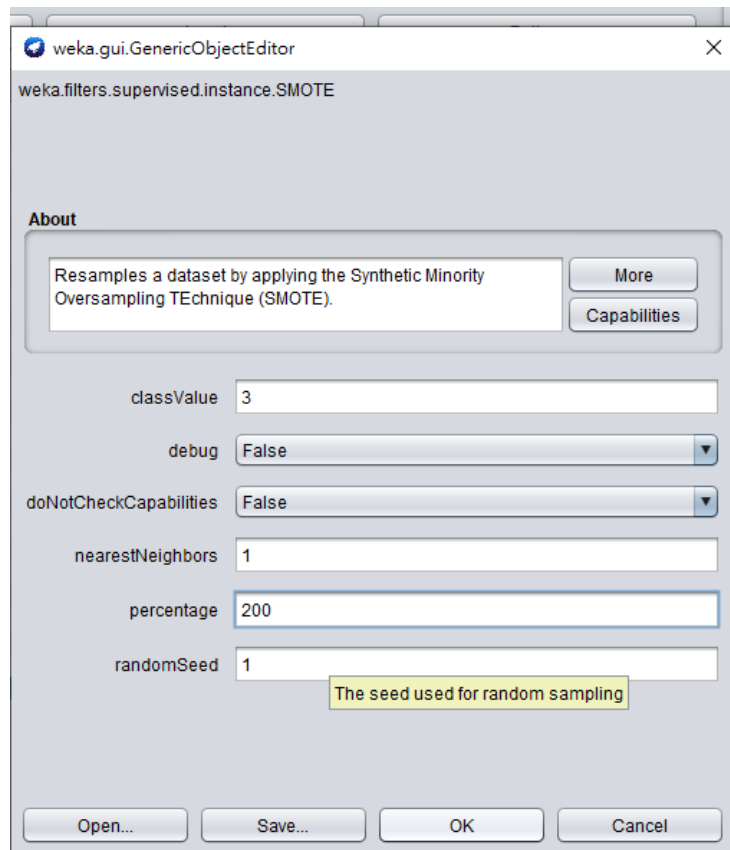
SMOTE with Decision tree:

Class A has the highest frequency and its frequency is not adjusted. The frequencies of other classes were adjusted to nearly the same with that of Class A, by adjusting the percentage (orange arrow). The following figures are listed by the order of classValue. The classValue of Class A is 1,

Class D

classValue: 3

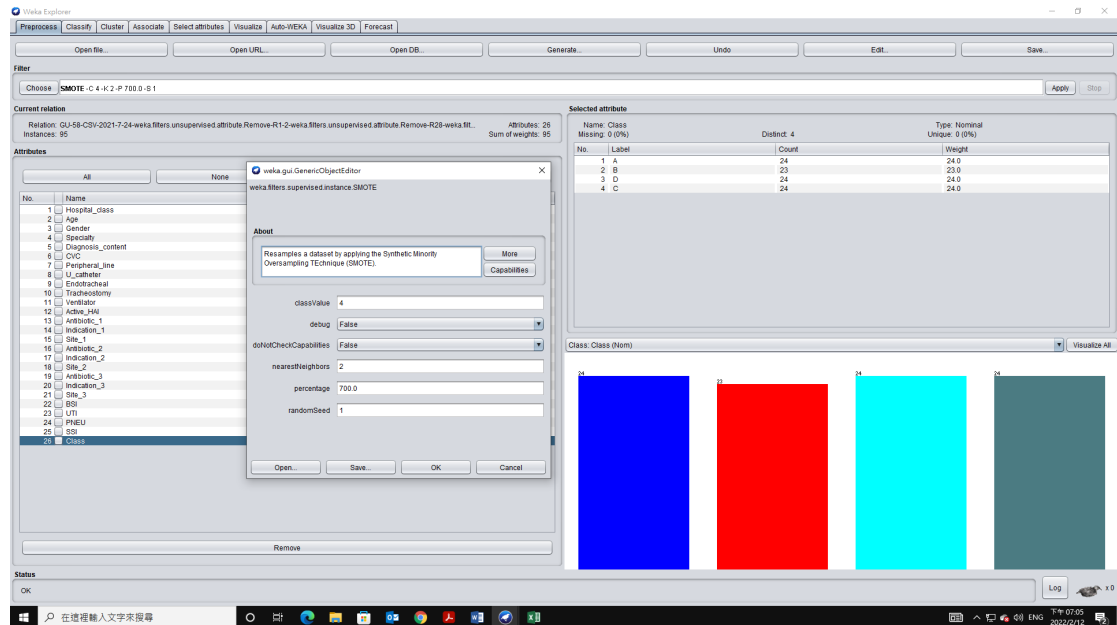
percentage: 200



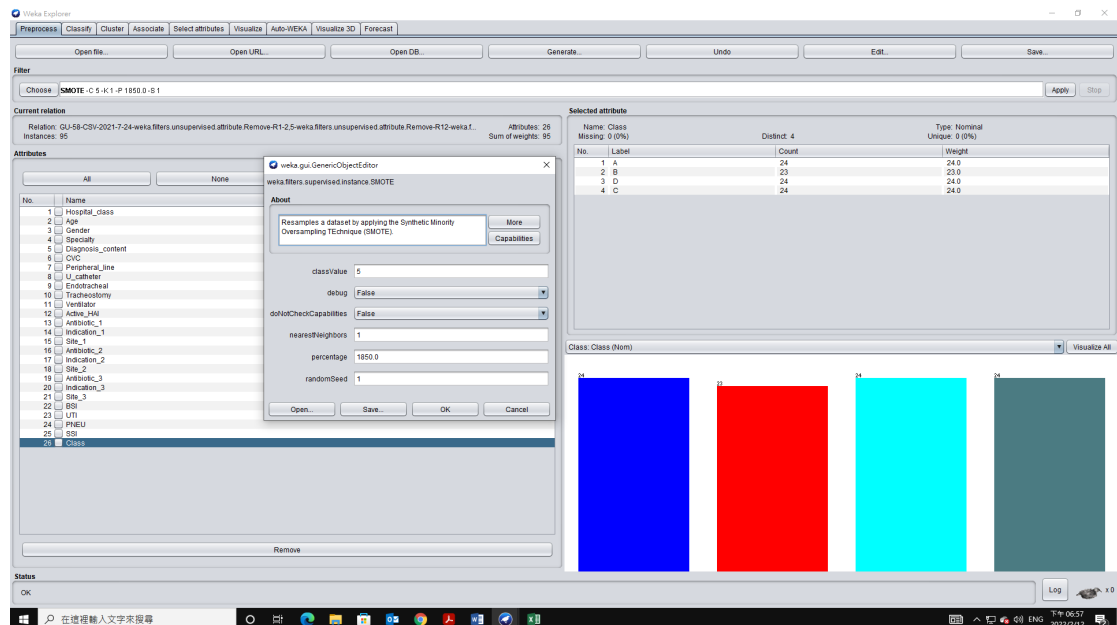
Class C:

calssValue: 4

percentage: 200

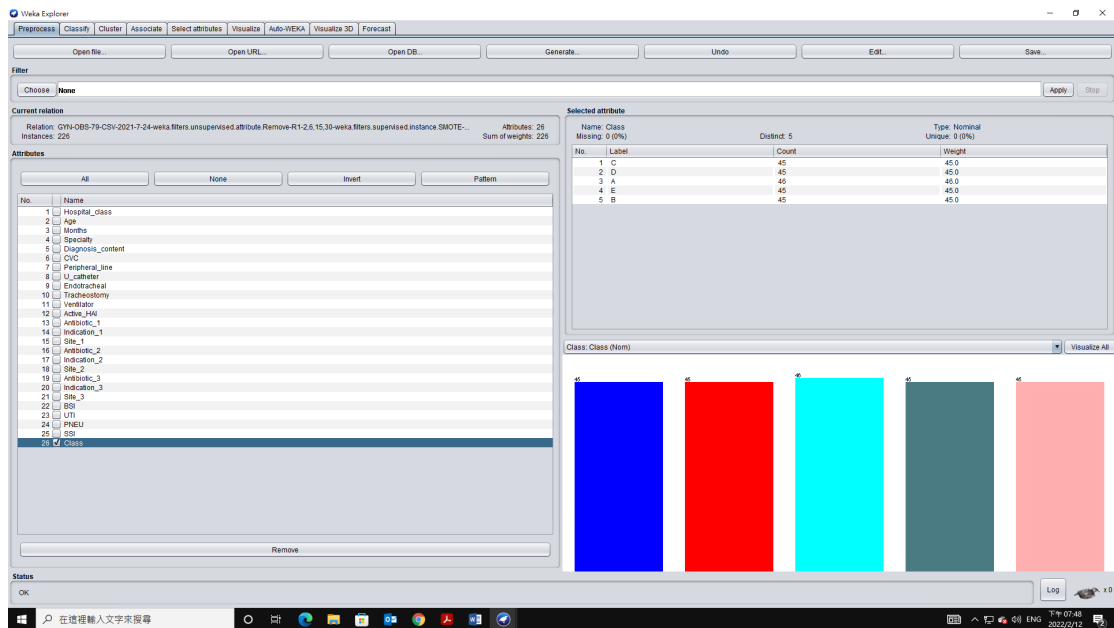


The final frequencies of the classes were adjusted as the following figure.



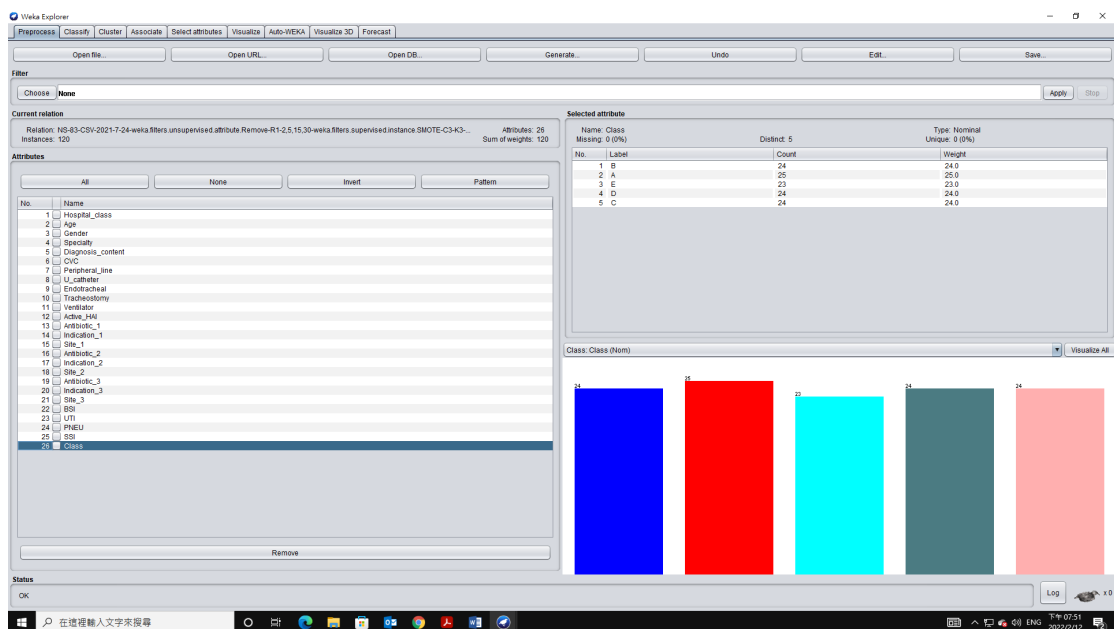
SMOTE for Obstetrics and Gynecology:

The final frequencies of the classes were adjusted as the following figure.



SMOTE for Neurosurgery:

The final frequencies of the classes were adjusted as the following figure.



SMOTE for Orthopedics:

The final frequencies of the classes were adjusted as the following figure.

