Lecture 5 Homework 3a

Targeted Marketing Campaign

In this problem we will use historical data from past customer responses to build a classification model. The model will then be applied to a new set of prospects to whom we may want extend an offer for a PEP. Rather than doing a mass marketing campaign to all new prospects, we would like to target those that are likely to respond positively to our offer (according to our classification model).

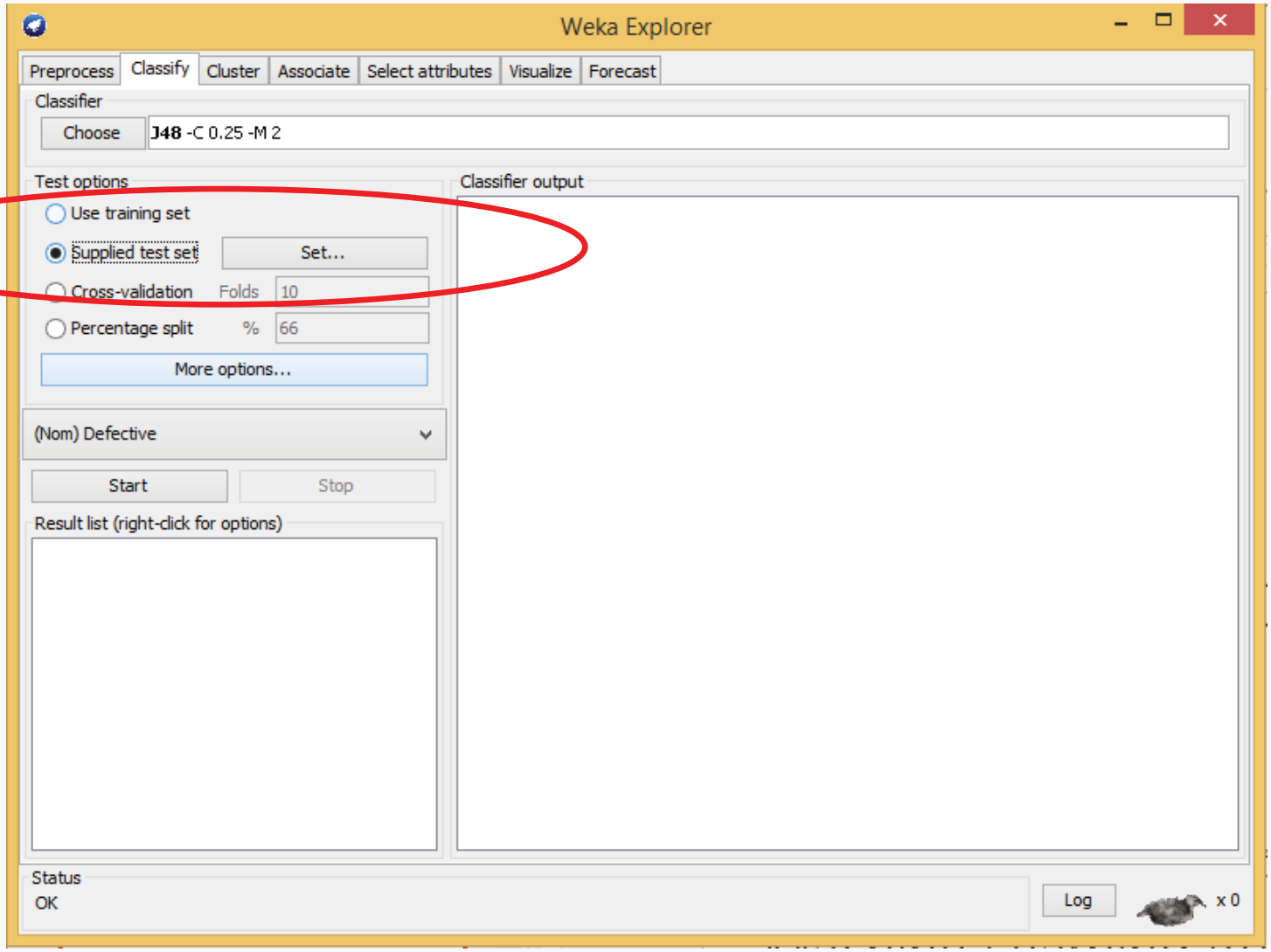
There are two data sets available (the data sets are comma delimited, and the first row contains the field names):

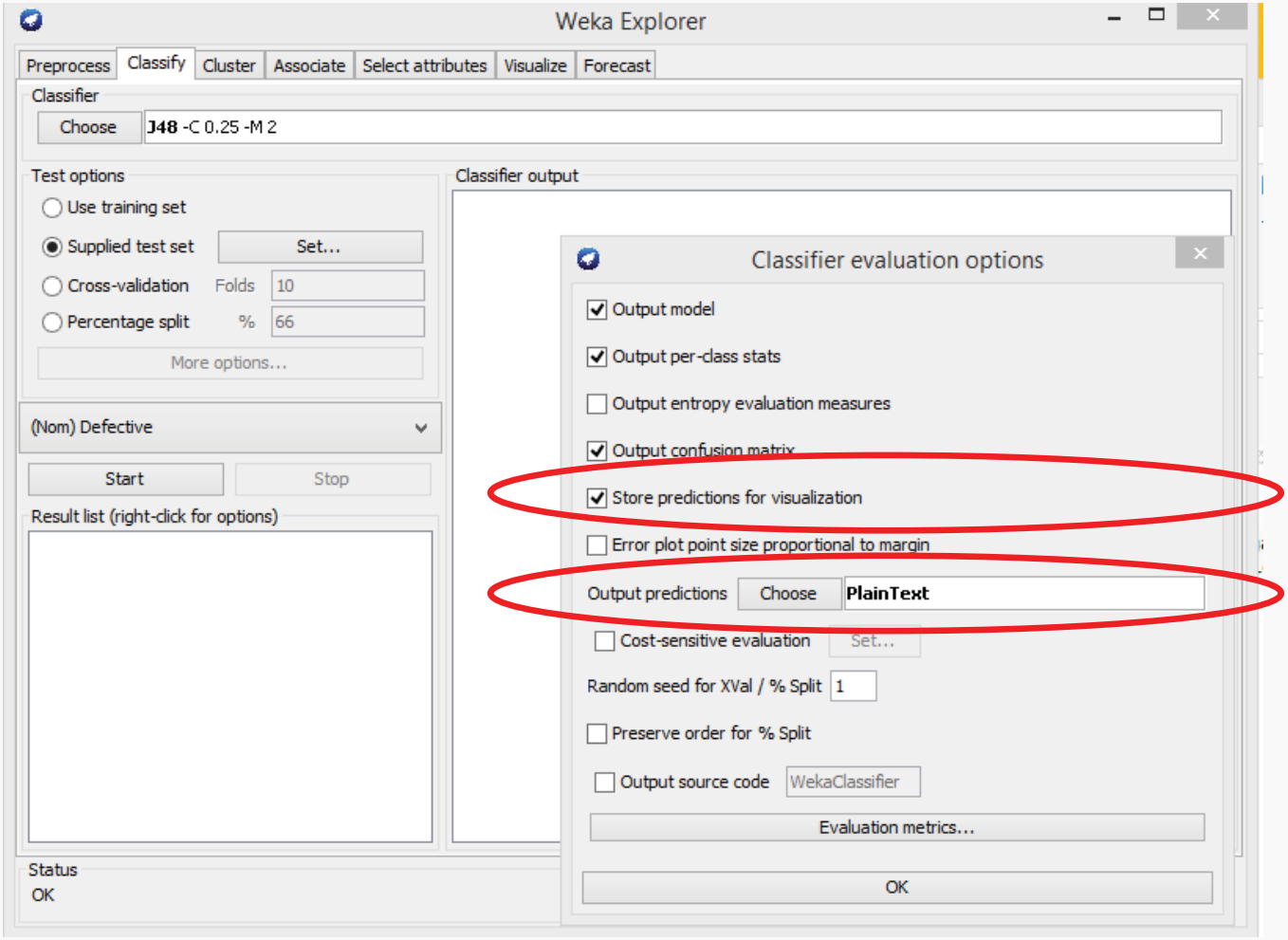
- bank-data.csv - Labelled training data set for Building a Model
- bank-new.csv - A set of new customers from which to find the "hot prospects" for the next mailing, using the profiles built from the training set.

- Decision Tree, 10-fold cross validation;
- Make predictions with the model against test data set;
- False Positive costs $10, False Negative results in a $1000 loss;
- Lift charts over predictive data;

From model construction to prediction to impact…

1. Build our predictive model in WEKA Explorer;
2. Use our model to score (predict) which new customers to target in our upcoming advertising campaign;
   - ARFF file manipulation (hacking), all too common pita…
   - Excel manipulation to join model output with our customers list
3. Compute the lift chart to assess business impact of our predictive model on the advertising campaign
   - How are Lift charts built, of all the charts and/or performance measures from a model this one is 'on you' to construct;
   - Where is the business 'bang for the buck'?


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Forecast
Classifier
Choose
J48 -C 0.25 -M 2
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
Classifier output
(Nom) Defective
Start
Stop
Result list (right-click for options)
Status
OK
Log
x 0


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Forecast
Classifier
Choose
J48 -C 0.25 -M 2
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
Classifier output
(Nom) Defective
Start
Stop
Result list (right-click for options)
Status
OK
Classifier evaluation options
Output model
Output per-class stats
Output entropy evaluation measures
Output confusion matrix
Store predictions for visualization
Error plot point size proportional to margin
Output predictions
Choose
PlainText
Cost-sensitive evaluation
Set...
Random seed for XVal / % Split
1
Preserve order for % Split
Output source code
WekaClassifier
Evaluation metrics...
OK