

## 14:prediction of categorical data using SMO

The screenshot shows the Weka Explorer interface with the 'Classify' tab selected. The classifier chosen is 'SMO - C 1.0 - L 0.001 - P 1.0E-12 - N 0 - V -1 - W 1 - K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4"'. The test options are set to 'Cross-validation' with 'Folds' set to 10. The classifier output shows a list of attributes and their weights, indicating a linear model. The status is 'OK'.

Classifier output:

```
property_magnitude
age
other_payment_plans
housing
existing_credits
job
num_dependents
own_telephone
foreign_worker
class

Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

SMO

Kernel used:
Linear Kernel: K(x,y) = <x,y>

Classifier for classes: good, bad

BinarySMO

Machine linear: showing attribute weights, not support vectors.

0.6805 * (normalized) checking_status=<0
0.3347 * (normalized) checking_status=0<=K<200
-0.4616 * (normalized) checking_status=>=200
-0.5537 * (normalized) checking_status=no checking
1.6987 * (normalized) duration
0.5398 * (normalized) credit_history=no credits/all paid
0.6015 * (normalized) credit_history=all paid
-0.109 * (normalized) credit_history=existing paid
-0.3182 * (normalized) credit_history=delayed previously
-0.7141 * (normalized) credit_history=critical/other existing credit
0.5673 * (normalized) purpose=new car
-0.5613 * (normalized) purpose=used car
```

The screenshot shows the Weka Explorer interface with the 'Classify' tab selected. The classifier chosen is 'SMO - C 1.0 - L 0.001 - P 1.0E-12 - N 0 - V -1 - W 1 - K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4"'. The test options are set to 'Cross-validation' with 'Folds' set to 10. The classifier output shows a list of attributes and their weights, indicating a linear model. The status is 'OK'.

Classifier output:

```
+ 0.4441 * (normalized) purpose=education
+ -0.3951 * (normalized) purpose=retraining
+ -0.0823 * (normalized) purpose=business
+ -0.2919 * (normalized) purpose=other
+ 1.1473 * (normalized) credit_amount
+ 0.4056 * (normalized) savings_status=<100
+ 0.115 * (normalized) savings_status=100<=K<500
+ 0.1378 * (normalized) savings_status=500<=K<1000
+ -0.3775 * (normalized) savings_status=>=1000
+ -0.2809 * (normalized) savings_status=no known savings
+ 0.2887 * (normalized) employment=unemployed
+ 0.1663 * (normalized) employment=<1
+ 0.0021 * (normalized) employment=1<=K<4
+ -0.3348 * (normalized) employment=4<=K<7
+ -0.1223 * (normalized) employment=>=7
+ 0.4503 * (normalized) installment_commitment
+ 0.3335 * (normalized) personal_status=female div/wid
+ 0.1177 * (normalized) personal_status=female mar/wid
+ -0.3697 * (normalized) personal_status=male single
+ -0.0815 * (normalized) personal_status=male mar/wid
+ 0.0514 * (normalized) other_parties=none
+ 0.5697 * (normalized) other_parties=no applicant
+ -0.6211 * (normalized) other_parties=guarantor
+ -0.0001 * (normalized) residence_since
+ -0.2247 * (normalized) property_magnitude=real estate
+ -0.0544 * (normalized) property_magnitude=life insurance
+ -0.0795 * (normalized) property_magnitude=car
+ 0.3586 * (normalized) property_magnitude=no known property
+ -0.4191 * (normalized) age
+ 0.0697 * (normalized) other_payment_plans=bank
+ 0.159 * (normalized) other_payment_plans=stores
+ -0.2287 * (normalized) other_payment_plans=none
+ 0.3271 * (normalized) housing=rent
+ -0.0702 * (normalized) housing=own
+ -0.257 * (normalized) housing=for free
+ 0.4503 * (normalized) existing_credits
+ -0.2026 * (normalized) job=unemp/unskilled non res
```

Weka Explorer

Preprocess

Classify

Cluster

Associate

Select attributes

Visualize

Classifier

Choose

SMO - C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4"

Test options

Use training set

Supplied test set

Cross-validation

Percentage split

Folds

10

%

66

More options...

(Nom) class

Start

Stop

Result list (right-click for options)

193806 - functions.SMO

Classifier output

+ 0.0190 \* (normalized) num\_dependents

+ -0.1394 \* (normalized) own\_telephone=yes

+ -0.9888 \* (normalized) foreign\_worker=no

- 1.5398

Number of kernel evaluations: 436644 (90.558% cached)

Time taken to build model: 0.33 seconds

==== Stratified cross-validation ====

==== Summary ====

Correctly Classified Instances

751

75.1 %

Incorrectly Classified Instances

249

24.9 %

Kappa statistic

0.3654

Mean absolute error

0.249

Root mean squared error

0.499

Relative absolute error

59.2607 %

Root relative squared error

108.8905 %

Total Number of Instances

1000

==== Detailed Accuracy By Class ====

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	ROC Area	PRC Area	Class
	0.471	0.530	0.793	0.471	0.630	0.371	0.471	0.781	good
	0.470	0.129	0.610	0.470	0.531	0.371	0.671	0.446	bad
Weighted Avg.	0.751	0.410	0.738	0.751	0.741	0.371	0.671	0.681	

==== Confusion Matrix ====

a b <-- classified as

610 90 | a = good

159 141 | b = bad

Status

OK

Log

x0