

# Untitled

## sasmarketing

SAS Macro Programs used for Marketing purpose

This repository contains a list of SAS macro-programs that I have developed for marketing purpose, such as : - formating and preparing data for modeling - applying cross-validation modeling - analysing models goodness-of-fit - projecting modeling results from new data

## Installation

## Example

Chargement de la librairie ;

```
filename x '\\srvfic4\DDOC\DMCD\MARK\REFERENTIEL ORGANISATIONNEL\REFERENTIEL_SAS\SAS-BOITE-OUTILS\pgm\sasmarketing\source2\source2.sas';  
%include x/source2;
```

Chargement du jeu de données Titanic ;

```
proc import datafile = "&DIR_DSRC\train.csv"  
    out = titanic  
    replace  
    dbms = CSV;  
run;  
  
data titanic;  
    set titanic;  
  
    if Survived = 1 then  
        Survived2="Survécu";  
    else Survived2 = "Mort";  
run;
```

Construction du modèle de validation croisée avec sélection de variables ;

```
%let var_x= Age Pclass Fare SibSp Parch PassengerId;

* Run AIC-based stepwise variable selection;
%SYMDEL COVARFIN;
%AICoptSW(
    *Jeu de données en entrée;
    indat=titanic,
    *la variable cible binaire (1/0) à prédire;
    y=Survived,

    *Variables prédicteurs/descripteurs ;
    x=&var_x,

    * nombre qui détermine la reproductibilité de l'échantillonnage de validation croisée ;
    seed=1,

    * spécifie le nombre de sous-échantillons distincts pour réaliser la validation croisée ;
    fold=5,

    * nombre de fois que le processus de validation croisée est répété ;
    repeats=1
);
```

Retrieve the best set of variables candidates according to the given frequency threshold;

```
%SYMDEL COVARFIN;
data DTRV.varfreq_fin;
    length cat $10000;
    do until (last._name_);
        set DTRV.varfreq_wide(where =(freq > 0.9));
        by _name_ notsorted;
        cat=catx(' ',cat,varlist);
    end;
run;

proc sql noprint;
    select cat
    into :COVARFIN separated by ' '
    from DTRV.VARFREQ_FIN;
quit;
%put &COVARFIN;
```

Calibration des modèles de ciblage;

```
%cvAUC (y=Survived, covars=&COVARFIN, fold=5, repeats=1);
```

```

* Représentation graphique des performance des modèles;
ods graphics / width=700px height=480px;
proc sgplot data=dsas.res_dec;
  vbox tx_cible / category=rang group=target clusterwidth=0.5;
  xaxis display=(noline nolabel noticks);
  yaxis display=(noline noticks) grid;
run;

ods graphics / width=700px height=480px;
proc sgplot data=dsas.res_auc;
  vbox auc / category=target clusterwidth=0.5;
  xaxis display=(noline nolabel noticks);
  yaxis display=(noline noticks) grid;
run;

```

Scoring sur un nouveau jeu de données;

```

data datapred; set titanic; run;
%mod_pred(y=Survived, covars=&COVARFIN, newdata=datapred, repeats=1, fold=5, by=PassengerId);

```