Package ‘modelimpact’

October 13, 2022

Type Package
Title Functions to Assess the Business Impact of Churn Prediction Models
Version 1.0.0
Author Peer Christensen
Maintainer Peer Christensen <hr.pchristensen@gmail.com>
Description Calculate the financial impact of using a churn model in terms of cost, revenue, profit and return on investment.
License MIT + file LICENSE
URL https://github.com/PeerChristensen/modelimpact
BugReports https://github.com/PeerChristensen/modelimpact/issues
Encoding UTF-8
LazyData true
RoxygenNote 7.1.1
Imports dplyr, magrittr, utils
Depends R (>= 2.10)
NeedsCompilation no
Repository CRAN
Date/Publication 2021-05-06 08:50:02 UTC

R topics documented:
cost_revenue ................................................. 2
predictions .................................................. 3
profit .......................................................... 3
profit_thresholds .......................................... 4
roi ............................................................. 5

Index 7
cost_revenue

**Calculate cost and revenue**

### Description

Calculates cost and revenue after sorting observations.

### Usage

```r
cost_revenue(
  x,
  fixed_cost = 0,
  var_cost = 0,
  tp_val = 0,
  prob_col = NA,
  truth_col = NA
)
```

### Arguments

- **x**: A data frame containing predicted probabilities of a target event and the actual outcome/class.
- **fixed_cost**: Fixed cost (e.g. of a campaign)
- **var_cost**: Variable cost (e.g. discount offered)
- **tp_val**: The average value of a True Positive
- **prob_col**: The unquoted name of the column with probabilities of the event of interest.
- **truth_col**: The unquoted name of the column with the actual outcome/class. Possible values are `Yes` and `No`.

### Value

A data frame with the following columns:

- **row**: row numbers
- **pct**: percentiles
- **cost_sum**: cumulated costs
- **cum_rev**: cumulated revenue

### Examples

```r
cost_revenue(predictions,
  fixed_cost = 1000,
  var_cost = 100,
  tp_val = 2000,
  prob_col = Yes,
  truth_col = Churn)
```
predictions

Predictions from a customer churn model.

Description

A dataset containing 2145 observations with four columns specifying predicted probabilities and predicted and actual class.

Usage

predictions

Format

A data frame with 2145 rows and 4 variables:

predict Predicted class
No Predicted probability of class 'No'
Yes Predicted probability of class 'Yes'
Churn Actual class ...

profit

Calculate profit

Description

Calculates profit after sorting observations.

Usage

profit(
x,
fixed_cost = 0,
var_cost = 0,
 tp_val = 0,
prob_col = NA,
 truth_col = NA
)
Arguments

x A data frame containing predicted probabilities of a target event and the actual outcome/class.

fixed_cost Fixed cost (e.g. of a campaign)

var_cost Variable cost (e.g. discount offered)

tp_val The average value of a True Positive

prob_col The unquoted name of the column with probabilities of the event of interest.

truth_col The unquoted name of the column with the actual outcome/class. Possible values are 'Yes' and 'No'.

Value

A data frame with the following columns:

row = row numbers

pct = percentiles

profit = profit for number of rows selected

Examples

profit(predictions,
       fixed_cost = 1000,
       var_cost = 100,
       tp_val = 2000,
       prob_col = Yes,
       truth_col = Churn)

profit_thresholds

Find optimal threshold for churn prediction (class)

Description

Finds the optimal threshold (from a business perspective) for classifying churners.

Usage

profit_thresholds(
  x,
  var_cost = 0,
  prob_accept = 1,
  tp_val = 0,
  fp_val = 0,
  tn_val = 0,
  fn_val = 0,
  prob_col = NA,
  truth_col = NA
)
Arguments

x  A data frame containing predicted probabilities of a target event and the actual outcome/class.
var_cost  Variable cost (e.g. of a campaign offer)
prob_accept  Probability of offer being accepted. Defaults to 1.
tp_val  The average value of a True Positive. ‘var_cost’ is automatically subtracted.
fp_val  The average cost of a False Positive. ‘var_cost’ is automatically subtracted.
tn_val  The average value of a True Negative.
fn_val  The average cost of a False Negative.
prob_col  The unquoted name of the column with probabilities of the event of interest.
truth_col  The unquoted name of the column with the actual outcome/class. Possible values are ‘Yes’ and ‘No’.

@return A data frame with the following columns:
threshold = prediction thresholds
payoff = calculated profit for each threshold

Examples

profit_thresholds(predictions,
  var_cost = 100,
  prob_accept = .8,
  tp_val = 2000,
  fp_val = 0,
  tn_val = 0,
  fn_val = -2000,
  prob_col = Yes,
  truth_col = Churn)

roi  Calculate Return on investment (ROI)

Description

Calculates ROI after sorting observations with ROI defined as (Current Value - Start Value) / Start Value

Usage

roi(x, fixed_cost = 0, var_cost = 0, tp_val = 0, prob_col = NA, truth_col = NA)
Arguments

- **x**: A data frame containing predicted probabilities of a target event and the actual outcome/class.
- **fixed_cost**: Fixed cost (e.g. of a campaign)
- **var_cost**: Variable cost (e.g. discount offered)
- **tp_val**: The average value of a True Positive
- **prob_col**: The unquoted name of the column with probabilities of the event of interest.
- **truth_col**: The unquoted name of the column with the actual outcome/class. Possible values are 'Yes' and 'No'.

Value

A data frame with the following columns:

- **row**: row numbers
- **pct**: percentiles
- **cum_rev**: cumulated revenue
- **cost_sum**: cumulated costs
- **roi**: return on investment

Examples

```r
roi(predictions,  
    fixed_cost = 1000,  
    var_cost = 100,  
    tp_val = 2000,  
    prob_col = Yes,  
    truth_col = Churn)
```
Index

* datasets
  predictions, 3

cost_revenue, 2

predictions, 3
profit, 3
profit_thresholds, 4

roi, 5