

Métodos de Clasificación

MsC. Edmond Géraud

Table of contents

Clasificación en ML	1
Métricas en clasificación	2
¿Qué veremos?	3
Regresión Logística (teoría básica)	3
Regresión logística binaria	4
Veamos con un ejemplo	5
La prueba de Wald	7
Los Odds, Log Odds, y Odds ratio	11
Más cosas	13
Práctica	13

Clasificación en ML

Los métodos de clasificación son técnicas de aprendizaje automático que se utilizan para predecir la clase a la que pertenece un objeto o un conjunto de datos. Estos métodos son muy útiles para una amplia variedad de aplicaciones, como el reconocimiento de imágenes, el procesamiento del lenguaje natural, la detección de fraudes y la clasificación de correos electrónicos no deseados, así como por ejemplo detección, y diagnosis de enfermedades

Algunos de los métodos de clasificación más comunes en aprendizaje automático son:

1. Árboles de decisión: es un método de aprendizaje supervisado que se utiliza para predecir la clase de un objeto mediante la construcción de un árbol de decisiones.
2. Regresión logística: es un método estadístico que se utiliza para predecir la probabilidad de que un objeto pertenezca a una clase determinada.

3. Máquinas de vectores de soporte (SVM): es un método de aprendizaje supervisado que se utiliza para clasificar objetos mediante la construcción de un hiperplano que separe las clases.
4. Redes neuronales artificiales: es un método de aprendizaje profundo que utiliza redes de neuronas para clasificar objetos.
5. Naïve Bayes: es un método estadístico que se basa en el teorema de Bayes para predecir la clase de un objeto.
6. PLS-DA (Partial Least Squares Discriminant Analysis) es otro método de clasificación que se utiliza comúnmente en la minería de datos y el análisis multivariado. PLS-DA es una técnica que combina el análisis discriminante mediante regresión por mínimos cuadrados parciales.

PLS-DA se utiliza típicamente cuando se tiene un conjunto de datos con múltiples variables independientes (características o variables explicativas) y se desea clasificar los datos en diferentes categorías (clases). PLS-DA busca identificar una combinación lineal de las variables independientes que mejor discrimina entre las diferentes categorías, de modo que los datos puedan ser clasificados en función de esta combinación lineal.

Cada uno de estos métodos tiene sus propias ventajas y desventajas, y su elección dependerá de la naturaleza del problema de clasificación y del conjunto de datos utilizado. Además, existen otros métodos de clasificación que se están utilizando en la actualidad y que son objeto de investigación constante en la comunidad de aprendizaje automático.

En la práctica, cada método posee ventajas en cuanto al análisis pero con el fin de observar cuál es mejor, de igual manera recurrimos a ciertas métricas.

Métricas en clasificación

Las métricas de clasificación son medidas que se utilizan para evaluar el rendimiento de un modelo de clasificación. Estas métricas se calculan comparando las predicciones del modelo con las etiquetas reales de las muestras de datos.

Algunas de las métricas de clasificación más comunes son:

1. Exactitud (Accuracy): es la proporción de muestras correctamente clasificadas por el modelo.
2. Precisión (Precision): es la proporción de muestras clasificadas como positivas por el modelo que son realmente positivas.
3. Sensibilidad (Recall o True Positive Rate): es la proporción de muestras positivas que son correctamente identificadas como positivas por el modelo.

4. Especificidad (Specificity o True Negative Rate): es la proporción de muestras negativas que son correctamente identificadas como negativas por el modelo.
5. F1-score: es una medida que combina la precisión y la sensibilidad en una sola métrica.

Además de estas métricas, también existen otras medidas como el AUC-ROC (área bajo la curva ROC) que se utilizan para evaluar el rendimiento de un modelo de clasificación. La elección de la métrica adecuada dependerá del problema de clasificación y del conjunto de datos utilizado

¿Qué veremos?

En esta sesión intentaremos entender tanto la regresión logística como el PLS-DA. Y si nos da tiempo la regresión logística con regularización.

Regresión Logística (teoría básica)

La regresión logística modela una relación entre variables predictoras y una variable de respuesta categórica.

Por ejemplo, después de cierto procesamiento en unas imágenes histológicas, podríamos predecir si se tiene cierta condición, como por ejemplo, leucemia o sin leucemia. Otro ejemplo, si tenemos ciertas variables clínicas las cuales están asociadas a cierta condición de salud, de forma categórica también podríamos realizar dicho análisis.

Ahora bien, cabe resaltar que este análisis es **supervisado**, quiere decir, que debemos de tener *marcada* la variable respuesta. Lo contrario que sucede con por ejemplo el análisis de componentes principales, el cual *tan solo nos sirve* para visualizar en una dimensión reducida aglomeraciones de datos.

No obstante no solamente se puede utilizar una técnica, se pueden mezclar varias para producir el resultado deseado. Me explico, supongamos que estamos estudiando el metabolismo, y estamos interesados, en cómo afecta el alcohol al sexo. No obstante no tenemos marcado el sexo, ni el consumo del alcohol. En una primera instancia, se podría realizar un PCA, y con técnicas no supervisadas, como lo sería k-means, podríamos encontrar aglomeraciones entre los pacientes. Ahora bien, este es un tipo de información y una vez obtenidos dichos conglomerados, se podría aplicar por ejemplo la regresión logística, sobre los datos originales, que han sido *marcados* con técnicas no supervisadas de ML, para observar las “diferencias” entre los metabolitos respecto al sexo y al consumo del alcohol.

Es decir, La regresión logística nos ayuda a estimar una probabilidad de caer en un determinado nivel de la respuesta categórica dado un conjunto de predictores. Podemos elegir

entre tres tipos de regresión logística, dependiendo de la naturaleza de la variable de respuesta categórica:

1. Regresión logística binaria:

Se utiliza cuando la respuesta es binaria (es decir, tiene dos resultados posibles). El ejemplo de cáncer / no cáncer, o responder en un examen, sí o no, en una encuesta, o predecir si se tiene la presión arterial alta o baja.

2. Regresión logística nominal:

Se utiliza cuando hay tres o más categorías sin un orden natural de los niveles. Ejemplos de respuestas de una empresa de una empresa (p. ej., marketing, ventas, RRHH), el tipo de motor de búsqueda utilizado (por ejemplo, Google, Yahoo!, MSN) y el color (negro, rojo, azul, naranja).

3. Regresión logística ordinal:

Se utiliza cuando hay tres o más categorías con una ordenación natural de los niveles, pero la clasificación de los niveles no significa necesariamente que los intervalos entre ellos sean iguales. Ejemplos de respuestas ordinales podrían ser cómo califican los estudiantes la eficacia de un curso universitario en una escala del 1 al 5, niveles de sabores de las alitas picantes, y estado de salud (por ejemplo, bueno, estable, grave, crítico).

Los problemas particulares de la regresión logística incluyen términos de error no normales, varianza de error no constante y restricciones en la función de respuesta (es decir, la respuesta está limitada entre 0 y 1).

Investigaremos formas de tratar estos temas solamente en la **regresión logística binaria**, aunque el concepto se puede ampliar a los demás tipos.

Regresión logística binaria

El modelo de regresión múltiple logístico es el siguiente:

$$\begin{aligned} P(Y = 1|X_1, X_2, \dots, X_p) &= \pi = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p)}} \\ &= \frac{e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p)}}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p)}} \\ &= \frac{e^{(X\beta)}}{1 + e^{(X\beta)}} = \frac{1}{1 + e^{-X\beta}} \end{aligned}$$

Tal que π denota la probabilidad y no el número. Es decir, es la probabilidad de éxito dados los predictores. Es decir, nos dice la probabilidad de un evento, no que este ocurra o no. Si se fijan, el modelo es el mismo que el lineal, tan solo que lo tenemos arreglado en una fracción.

De hecho, de las ecuaciones anteriores, podemos calcular la función de verosimilitud, es una función que describe la probabilidad de observar un conjunto de datos específico dado un modelo estadístico y un conjunto de parámetros. En otras palabras, la función de verosimilitud mide la capacidad de un modelo para explicar los datos observados en términos de la distribución de probabilidad de las variables involucradas.

En el contexto de la regresión logística, la función de verosimilitud se utiliza para estimar los coeficientes de regresión que mejor se ajustan a los datos observados. La función de verosimilitud de la regresión logística multivariante es el producto de las probabilidades condicionales de las muestras de datos para cada conjunto de valores de las variables independientes.

Matemáticamente, la función de verosimilitud se define como:

$$L(\beta_0, \beta_1, \dots, \beta_p) = \prod_{i=1}^n P(Y_i | X_{i1}, X_{i2}, \dots, X_{ip}; \beta_0, \beta_1, \dots, \beta_p)$$

Donde:

- $L(\beta_0, \beta_1, \dots, \beta_p)$ es la función de verosimilitud para los parámetros $\beta_0, \beta_1, \dots, \beta_p$
- $P(Y_i | X_{i1}, X_{i2}, \dots, X_{ip}; \beta_0, \beta_1, \dots, \beta_p)$ es la probabilidad condicional de que la i -ésima muestra tenga una etiqueta Y_i dada su combinación de variables independientes $X_{i1}, X_{i2}, \dots, X_{ip}$ y los parámetros $\beta_0, \beta_1, \dots, \beta_p$
- n es el número total de muestras de datos

El objetivo de la regresión logística es encontrar los valores óptimos de los coeficientes $\beta_0, \beta_1, \dots, \beta_p$ que maximizan la función de verosimilitud. Esto se logra mediante técnicas de optimización como el gradiente descendente o la regresión de máxima verosimilitud.

Es decir, si con el algoritmo de gradiente descendente que se utilizaba para las regresiones múltiples con penalización se buscaba el mínimo. Con esta se encuentra el máximo.

Veamos con un ejemplo

Para ilustrarlo, consideremos los datos publicados sobre $n = 27$ pacientes con leucemia. Los datos (leucemia_remisión.txt) tienen una variable de respuesta de si se produjo la remisión de la leucemia, que viene dada por un 1.

Las variables predictoras son la celularidad de la célula de la sección del coágulo de médula, el porcentaje diferencial de frotis de blastocitos, porcentaje de infiltrado absoluto de células

leucémicas en la médula ósea, índice de etiquetado porcentual de las células leucémicas de la médula ósea li, número absoluto de blastos en la sangre periférica blastos, y la mayor temperatura antes del inicio del tratamiento temp.

Cargamos los datos y realizamos la regresión logística

```
remission <-
  read.table(
    "./data/leukemia_remision.txt",
    sep = "",
    skip = 1,
    header = T,
    row.names = 1
  )
lrmodel.full <-
  glm(
    remiss ~ cell + smear + infil + li + blast + temp,
    family = binomial(link = "logit"),
    data = remission
  )
summary(lrmodel.full)
```

Call:

```
glm(formula = remiss ~ cell + smear + infil + li + blast + temp,
    family = binomial(link = "logit"), data = remission)
```

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-1.95165	-0.66491	-0.04372	0.74304	1.67069

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	58.0385	71.2364	0.815	0.4152
cell	24.6615	47.8377	0.516	0.6062
smear	19.2936	57.9500	0.333	0.7392
infil	-19.6013	61.6815	-0.318	0.7507
li	3.8960	2.3371	1.667	0.0955 .
blast	0.1511	2.2786	0.066	0.9471
temp	-87.4339	67.5735	-1.294	0.1957

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 34.372 on 26 degrees of freedom
Residual deviance: 21.751 on 20 degrees of freedom
AIC: 35.751

Number of Fisher Scoring iterations: 8

La prueba de Wald

La prueba de Wald, mira significancia de los coeficientes individuales en la regresión logística, sería un símil a los t-test que se utilizan en regresión lineal. Para los estimados de la función de máxima verosimilitud, el ratio.

$$Z = \frac{\hat{\beta}_i}{s.e(\hat{\beta}_i)}$$

Donde la hipótesis nula, es que los coeficientes son cero. Es decir los estimados no aportan nada al modelo $H_0 : \beta_i = 0$. Donde se utiliza la distribución estándar (es decir la gaussiana, con media cero), para determinar el p-valor del test. Además, los intervalos de confianza se puede construir de la siguiente manera.

$$\hat{\beta}_i \pm z_{1-\alpha/2} s.e(\hat{\beta}_i)$$

Es decir, en el modelo

```
confint(lrmodel.full)
```

Waiting for profiling to be done...

	2.5 %	97.5 %
(Intercept)	-70.9683777	222.202990
cell	-27.7332544	138.404531
smear	-60.4544868	152.174139
infil	-159.7565104	67.536927
li	0.1944541	9.526820
blast	-4.5238625	4.715064
temp	-244.7720744	24.913187

De igual manera, podemos realizar una selección de variables con el criterio AIC

```
lrmodel.step <- step(lrmodel.full)
```

Start: AIC=35.75

remiss ~ cell + smear + infil + li + blast + temp

	Df	Deviance	AIC
- blast	1	21.755	33.755
- infil	1	21.857	33.857
- smear	1	21.869	33.869
- cell	1	22.071	34.071
<none>		21.751	35.751
- temp	1	23.877	35.877
- li	1	26.095	38.095

Step: AIC=33.76

remiss ~ cell + smear + infil + li + temp

	Df	Deviance	AIC
- infil	1	21.858	31.858
- smear	1	21.869	31.869
- cell	1	22.073	32.073
<none>		21.755	33.755
- temp	1	24.198	34.199
- li	1	30.216	40.216

Step: AIC=31.86

remiss ~ cell + smear + li + temp

	Df	Deviance	AIC
- smear	1	21.953	29.953
<none>		21.858	31.858
- temp	1	24.292	32.292
- cell	1	24.477	32.477
- li	1	30.434	38.434

Step: AIC=29.95

remiss ~ cell + li + temp

	Df	Deviance	AIC
<none>		21.953	29.953
- temp	1	24.341	30.341


```
- cell 1 24.648 30.648
- li 1 30.829 36.829
```

```
summary(lrmodel.step)
```

Call:

```
glm(formula = remiss ~ cell + li + temp, family = binomial(link = "logit"),
     data = remission)
```

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-2.02043	-0.66313	-0.08323	0.81282	1.65887

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	67.634	56.888	1.189	0.2345
cell	9.652	7.751	1.245	0.2130
li	3.867	1.778	2.175	0.0297 *
temp	-82.074	61.712	-1.330	0.1835

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 34.372 on 26 degrees of freedom
 Residual deviance: 21.953 on 23 degrees of freedom
 AIC: 29.953

Number of Fisher Scoring iterations: 7

Encontramos que solamente la variable li, es la importante como predictora.

```
lrmodel.reduced <- glm(remiss ~ li, family = binomial(link = "logit"), data = remission)
summary(lrmodel.reduced)
```

Call:

```
glm(formula = remiss ~ li, family = binomial(link = "logit"),
     data = remission)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.9448	-0.6465	-0.4947	0.6571	1.6971

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-3.777	1.379	-2.740	0.00615 **
li	2.897	1.187	2.441	0.01464 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 34.372 on 26 degrees of freedom
 Residual deviance: 26.073 on 25 degrees of freedom
 AIC: 30.073

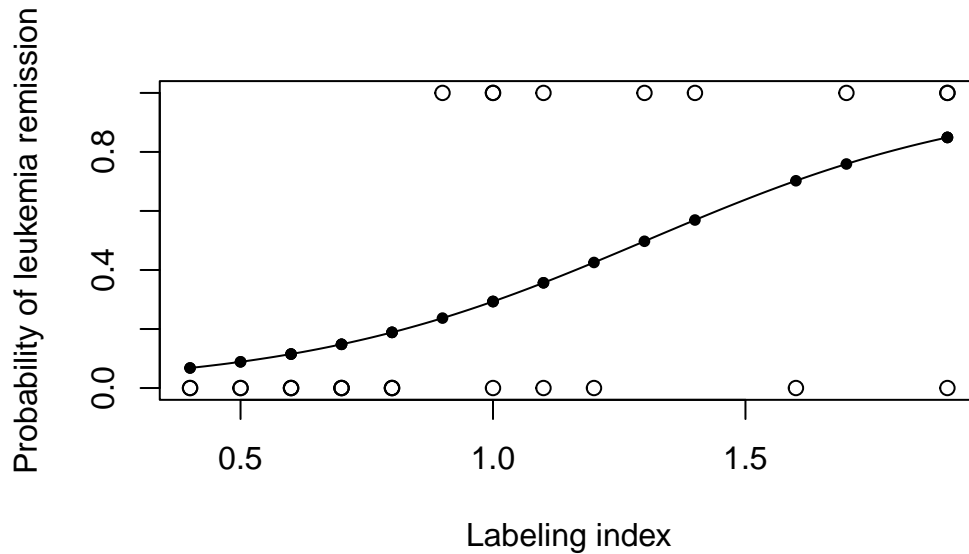
Number of Fisher Scoring iterations: 4

Donde la ecuación de la regresión es la siguiente:

$$P(1) = \exp(Y') / (1 + \exp(Y')Y' = 3.78 + 2.90 \times li$$

Como solamente tenemos un único predictor en este modelo, podemos visualizar la forma sigmoidal de la regresión.

```
attach(remission)
plot(li, remiss, xlab = "Labeling index", ylab = "Probability of leukemia remission")
curve(predict(lrmodel.reduced, data.frame(li = x), type = "resp"),
       lty = 1.5,
       add = TRUE)
points(li, fitted(lrmodel.reduced), pch = 20) # opcional
```



Los Odds, Log Odds, y Odds ratio

El término “odds” en estadística y en el contexto de la regresión logística se refiere a la relación entre la probabilidad de que ocurra un evento y la probabilidad de que no ocurra. En otras palabras, los odds representan la proporción de eventos exitosos frente a los eventos no exitosos.

En la regresión logística, los odds se utilizan para modelar la probabilidad de que una variable dependiente binaria (por ejemplo, una variable que toma valores 0 o 1) esté presente o ausente

Existen diversas maneras de escribir la regresión logística, por ejemplo la primera es.

$$\frac{\pi}{1 - \pi} = \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_{p-1} x_{p-1})$$

Que es una ecuación, la cual describe la relación entre que ocurra un evento y no ocurra. Donde π es la probabilidad de que ocurra el evento. Es decir, si estamos apostando a una carrera de caballos, por uno en especial gane la carrera, con una probabilidad del 80 porciento, el **odds**, sería $0.80/(1 - 0.80) = 4$ es decir 4 : 1

La segunda opción es la siguiente:

$$\log\left(\frac{\pi}{1 - \pi}\right) = \beta_0 + \beta_1 x_1 + \dots + \beta_{p-1} x_{p-1}$$

Donde básicamente esta representación nos dice que el logaritmo natural de los datos, es una función lineal de los datos, y se le conoce como **log odds**. A esto se le conoce como la **transformación logit** de la probabilidad de éxito de π

Por otro lado, el **odds ratio** que lo denominaremos como θ entre los odds de dos predictores, por ejemplo $X_{(1)}$ y $X_{(2)}$ es dado por :

$$\theta = \frac{((\pi/(1-\pi))|x = X_{(1)})}{((\pi/(1-\pi))|x = X_{(2)})}$$

Por otra parte, para una regresión logística los odds de éxito son:

$$\frac{\pi}{1-\pi} = \exp(X\beta)$$

Es decir, si arreglamos esta ecuación con la anterior, podemos establecer la relación entre el predictor y la respuesta. El odds ratio tiene que ser cualquier número no negativo.

Por ejemplo, cuando solo hay un predictor , X las odds de éxito son

$$\frac{\pi}{1-\pi} = \exp(\beta_0 + \beta_1 X)$$

De manera que si incrementamos X por una unidad se convierte en:

$$\theta = \frac{\exp(\beta_0 + \beta_1(X+1))}{\exp(\beta_0 + \beta_1 X)} = \exp(\beta_1)$$

Por ejemplo, el odds ratio para `li` se calcula como $\exp(2.89726)$

```
exp(summary(lrmodel.reduced)$coefficients["li",1])
```

```
[1] 18.12449
```

También podemos calcular el intervalo de confianza al 95 por ciento como:

```
exp(summary(lrmodel.reduced)$coefficients["li",1]) +qnorm(c(0.025,0.5,0.975)) * summary(lr
```

```
[1] 15.79836 18.12449 20.45061
```

La interpretación del odds ratio es que para cada un incremento de una unidad en `li` el estimado de odds para la recurrencia de la leucemia, está multiplicada por 18.1245. Sin embargo, `li` aparece caer entre 0 y 2, por lo que tendría más sentido decir que para cada incremento en 0.1 unidades en `li`, el estimado de los odds remission esta multiplicada por $\exp(2.89726 \times 0.1) = 1.336$. Entonces,

Más cosas ...

En realidad, cuando estamos tratando con una gran cantidad de datos, no sería necesario comprobar los supuestos, que básicamente serían los mismos que los de la regresión multivariada. No obstante, es una buena práctica hacerlo.

Práctica

```
library(nnet)
library(caret)
```

Loading required package: ggplot2

Loading required package: lattice

```
set.seed(123156)
datos <- read.csv("../data/BreastTissue.csv")
datos$Class <- as.factor(datos$Class)
datos[, -1] <- scale(datos[, -1])
idx <- sample(1:nrow(datos), size = nrow(datos)*0.7, replace = F)
train <- datos[idx,]
test <- datos[-idx,]

mdl <- multinom(Class ~ ., data = train)
```

```
# weights: 66 (50 variable)
initial value 132.590201
iter 10 value 50.461488
iter 20 value 38.672193
iter 30 value 36.370299
iter 40 value 34.299155
```

```

iter 50 value 33.336058
iter 60 value 32.635862
iter 70 value 30.688927
iter 80 value 28.896317
iter 90 value 28.612742
iter 100 value 28.341508
final value 28.341508
stopped after 100 iterations

```

```

res <- predict(mdl,test)

confusionMatrix(test$Class,res)

```

Confusion Matrix and Statistics

	Reference					
Prediction	adi	car	con	fad	gla	mas
adi	6	0	0	0	0	0
car	0	7	0	0	0	0
con	1	0	3	0	0	0
fad	0	1	0	4	1	0
gla	0	0	0	0	4	0
mas	0	1	0	0	2	2

Overall Statistics

```

Accuracy : 0.8125
95% CI : (0.6356, 0.9279)
No Information Rate : 0.2812
P-Value [Acc > NIR] : 6.484e-10

```

```
Kappa : 0.7728
```

```
McNemar's Test P-Value : NA
```

Statistics by Class:

	Class: adi	Class: car	Class: con	Class: fad	Class: gla
Sensitivity	0.8571	0.7778	1.00000	1.0000	0.5714
Specificity	1.0000	1.0000	0.96552	0.9286	1.0000

Pos Pred Value	1.0000	1.0000	0.75000	0.6667	1.0000
Neg Pred Value	0.9615	0.9200	1.00000	1.0000	0.8929
Prevalence	0.2188	0.2812	0.09375	0.1250	0.2188
Detection Rate	0.1875	0.2188	0.09375	0.1250	0.1250
Detection Prevalence	0.1875	0.2188	0.12500	0.1875	0.1250
Balanced Accuracy	0.9286	0.8889	0.98276	0.9643	0.7857

Class: mas

Sensitivity	1.0000
Specificity	0.9000
Pos Pred Value	0.4000
Neg Pred Value	1.0000
Prevalence	0.0625
Detection Rate	0.0625
Detection Prevalence	0.1562
Balanced Accuracy	0.9500

```

set.seed(23)
ctrl <-
  trainControl(
    method = "repeatedcv",
    number = 10,
    classProbs = TRUE,
    summaryFunction = multiClassSummary
  )
mdl <- train(
  x = train[,-1],
  # spectral data
  y = train$Class,
  # factor vector
  method = "pls",
  # pls-da algorithm
  tuneLength = 10,
  # number of components
  trControl = ctrl,
  # ctrl contained cross-validation option
  preProc = c("center", "scale"),
  # the data are centered and scaled
  metric = "Accuracy"
) # metric is ROC for 2 classes

```

Warning in nominalTrainWorkflow(x = x, y = y, wts = weights, info = trainInfo,

: There were missing values in resampled performance measures.

```
plsda
```

```
function (x, ...)
UseMethod("plsda")
<bytecode: 0x55f8908bbd48>
<environment: namespace:caret>
```

```
res <- predict(mdl,test)
confusionMatrix(test$Class,res)
```

Confusion Matrix and Statistics

	Reference					
Prediction	adi	car	con	fad	gla	mas
adi	6	0	0	0	0	0
car	0	7	0	0	0	0
con	1	0	3	0	0	0
fad	0	1	0	3	1	1
gla	0	0	0	1	3	0
mas	0	2	0	2	1	0

Overall Statistics

Accuracy : 0.6875
95% CI : (0.4999, 0.8388)
No Information Rate : 0.3125
P-Value [Acc > NIR] : 1.45e-05

Kappa : 0.6186

McNemar's Test P-Value : NA

Statistics by Class:

	Class: adi	Class: car	Class: con	Class: fad	Class: gla
Sensitivity	0.8571	0.7000	1.00000	0.50000	0.60000
Specificity	1.0000	1.0000	0.96552	0.88462	0.96296

Pos Pred Value	1.0000	1.0000	0.75000	0.50000	0.75000
Neg Pred Value	0.9615	0.8800	1.00000	0.88462	0.92857
Prevalence	0.2188	0.3125	0.09375	0.18750	0.15625
Detection Rate	0.1875	0.2188	0.09375	0.09375	0.09375
Detection Prevalence	0.1875	0.2188	0.12500	0.18750	0.12500
Balanced Accuracy	0.9286	0.8500	0.98276	0.69231	0.78148

Class: mas

Sensitivity	0.00000
Specificity	0.83871
Pos Pred Value	0.00000
Neg Pred Value	0.96296
Prevalence	0.03125
Detection Rate	0.00000
Detection Prevalence	0.15625
Balanced Accuracy	0.41935