

Clasificación de Múltiples Clases con Clasificadores Binarios

Pontificia Universidad Javeriana
Francisco Carlos Calderon Ph.D
2020

Objetivos

Usar un método de clasificación binaria para clasificar entre múltiples clases.

Problema

Clasificadores como la regresión logística solo son capaces de lidiar con dos clases.

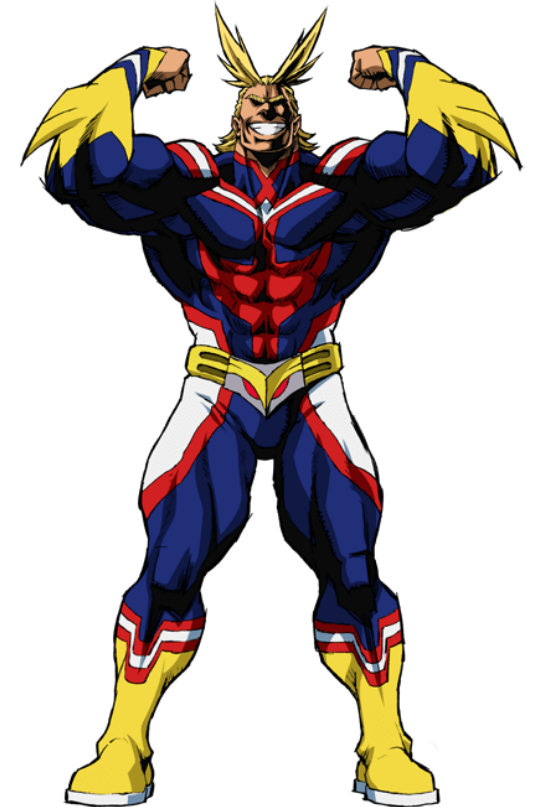
Entre estos clasificadores binarios:

- Algunos entregan $y \in \{0,1\}$
- Otros entregan $h_{\theta}(x) = P(y = 1|x; \theta)$ y a partir de esta se calcula y .

One vs all, One-vs-Rest, one-against-all

Se entrena un clasificador por cada clase.

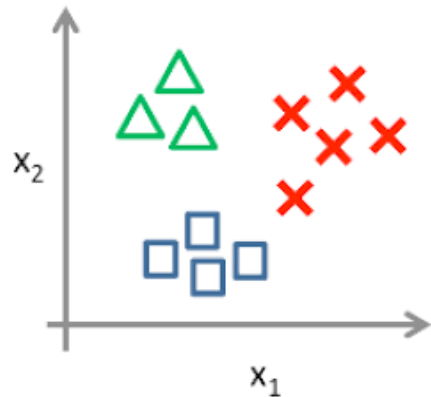
- Cada clase contra el resto de clases de manera binaria.
- Cada clasificador debe generar un valor de confianza. $h_{\theta,k}(x) = P(y = 1|x; \theta, k) | k = \{1, m\}$



https://bokunoheroacademia.fandom.com/es/wiki/One_For_All

One vs all, One-vs-Rest, one-against-all

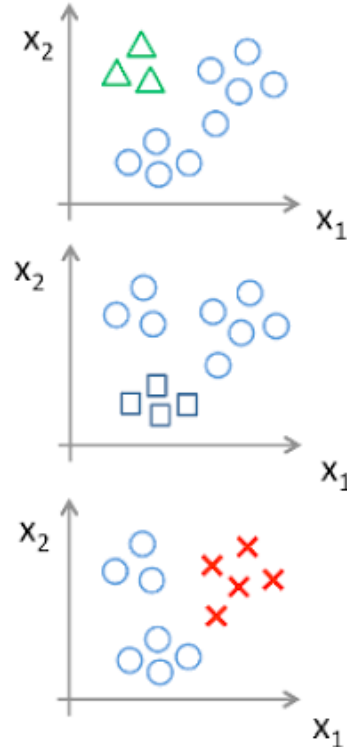
One-vs-all (one-vs-rest):



Class 1: **Green**

Class 2: **Blue**

Class 3: **Red**



Si se tienen K clases diferentes se entrenan K clasificadores.

Cuando ya se tengan los K clasificadores entrenados y llegue una nueva muestra x :

1. Se crea un vector con los resultados de confianza de cada clasificador: $\mathbf{h}_k(x) = \{h_1(x), h_2(x), \dots, h_k(x)\} / k=\{1, 2, \dots, K\}$.
2. Encontramos cual es el valor más alto de confianza entre todos los $\mathbf{h}_k(x)$, y el índice corresponderá a la etiqueta de predicción:

$$y = \underset{k \in \{1 \dots K\}}{\operatorname{argmax}} \mathbf{h}_k(x)$$

https://www.cc.gatech.edu/classes/AY2016/cs4476_fall/results/proj4/html/jnanda3/index.htm

One vs all, One-vs-Rest, one-against-all

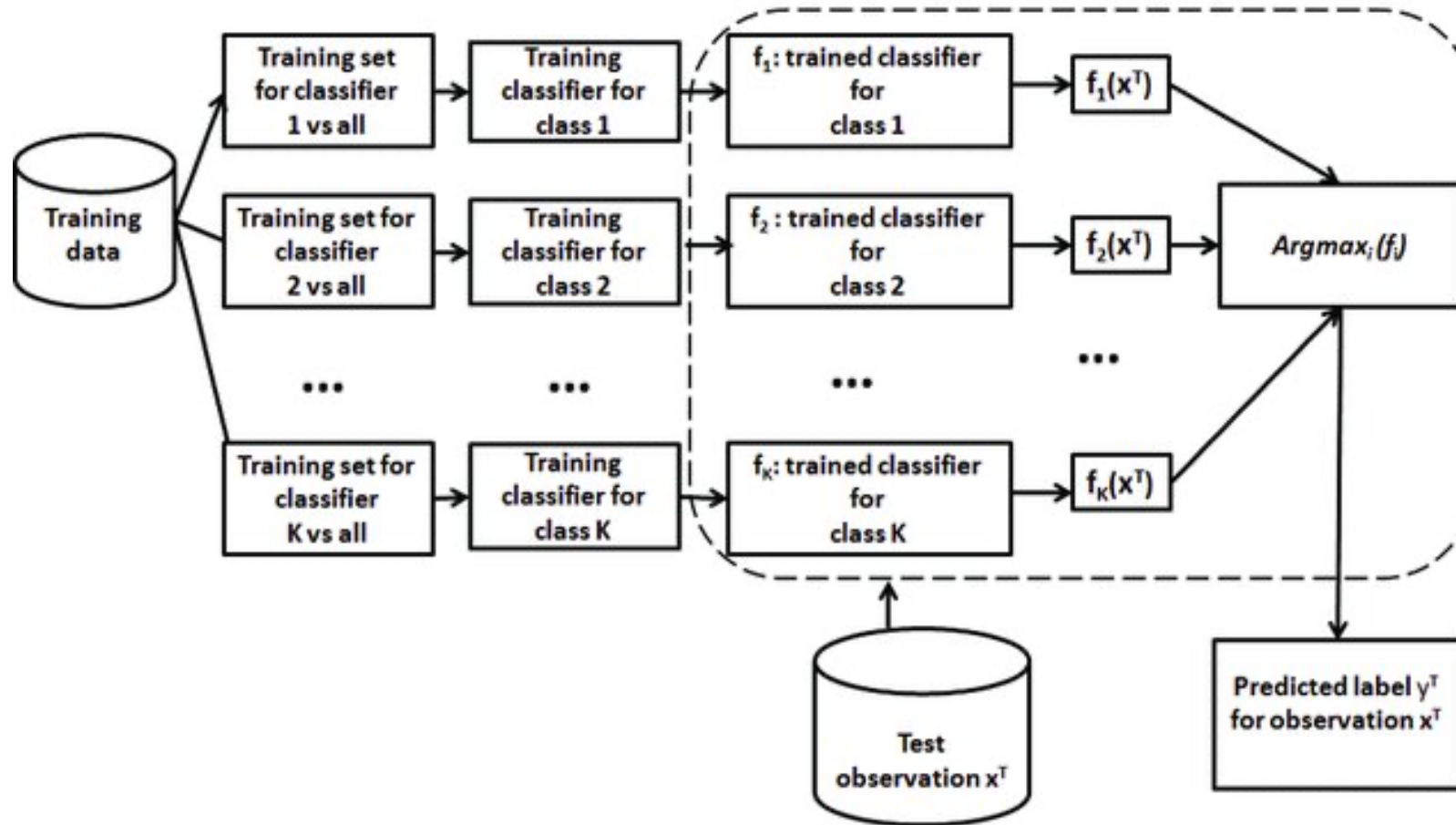


Imagen de : Sow-activity classification from acceleration patterns: A machine learning approach

Escalante et al. <https://www.sciencedirect.com/science/article/abs/pii/S0168169913000082?via%3Dihub>

Desventajas

la escala de los valores de confianza puede diferir entre los clasificadores binarios.

La distribución de clases está casi siempre desequilibrada en el conjunto de entrenamiento.



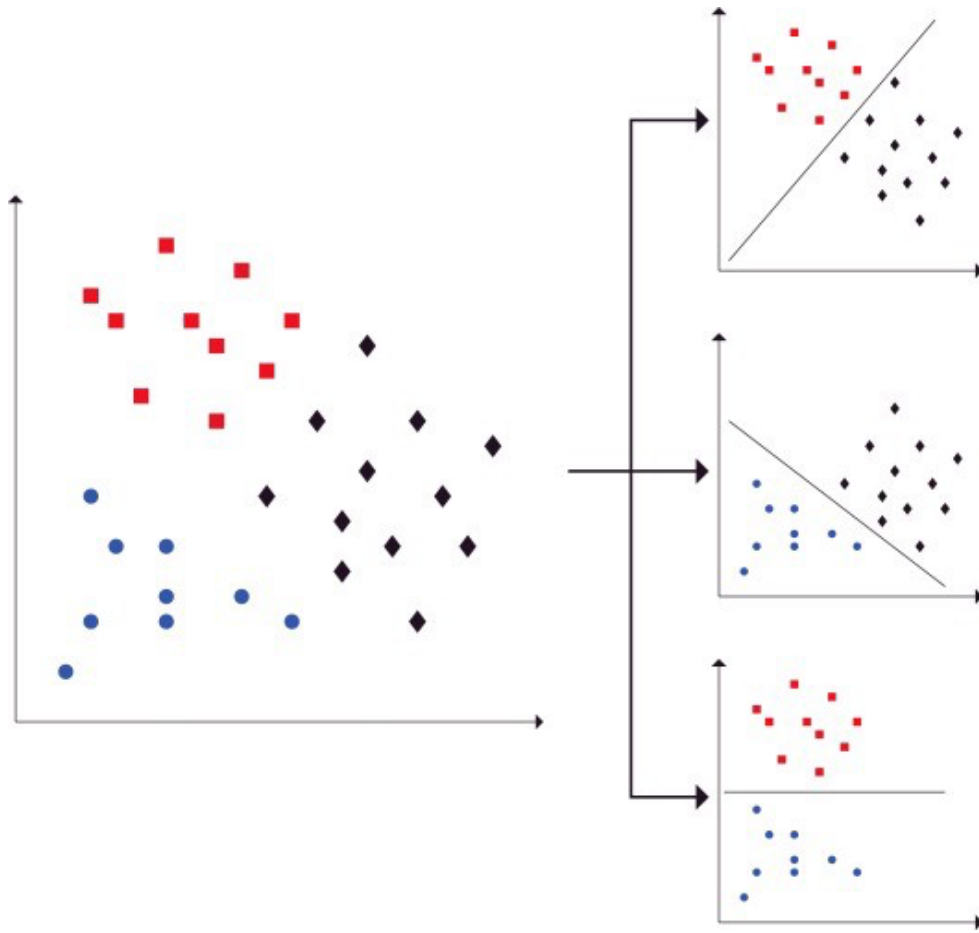
https://bokunoheroacademia.fandom.com/es/wiki/One_For_All

One vs One también llamado All vs All o All-Pairs



- Se entrenan $(K(K-1))/2$ clasificadores. Uno por cada combinación posible de clasificadores binarios.
- Cada clasificador recibe las muestras de un par de clases del conjunto de entrenamiento.
- En la predicción aplica la muestra a cada clasificador y la clase con mayor número de predicciones gana.
- Otras variantes toman la suma de la predicción para cada clase

One vs One



Construir $K(K-1)/2$ clasificadores.

Sea $h_{i,j}$ la salida del clasificador:

$$h(x) = \underset{i}{\operatorname{argmax}} \sum_j h_{i,j}(x)$$

Tenga en cuenta que:

$$h_{i,j} = -h_{j,i}$$

Desventajas

- la escala de los valores de confianza puede diferir entre los clasificadores binarios.
- El numero de clasificadores aumenta con respecto a OvA pero son clasificadores más simples para un K “grande”.
- Dependiendo de la implementación pueden llegarse a dar soluciones ambiguas.

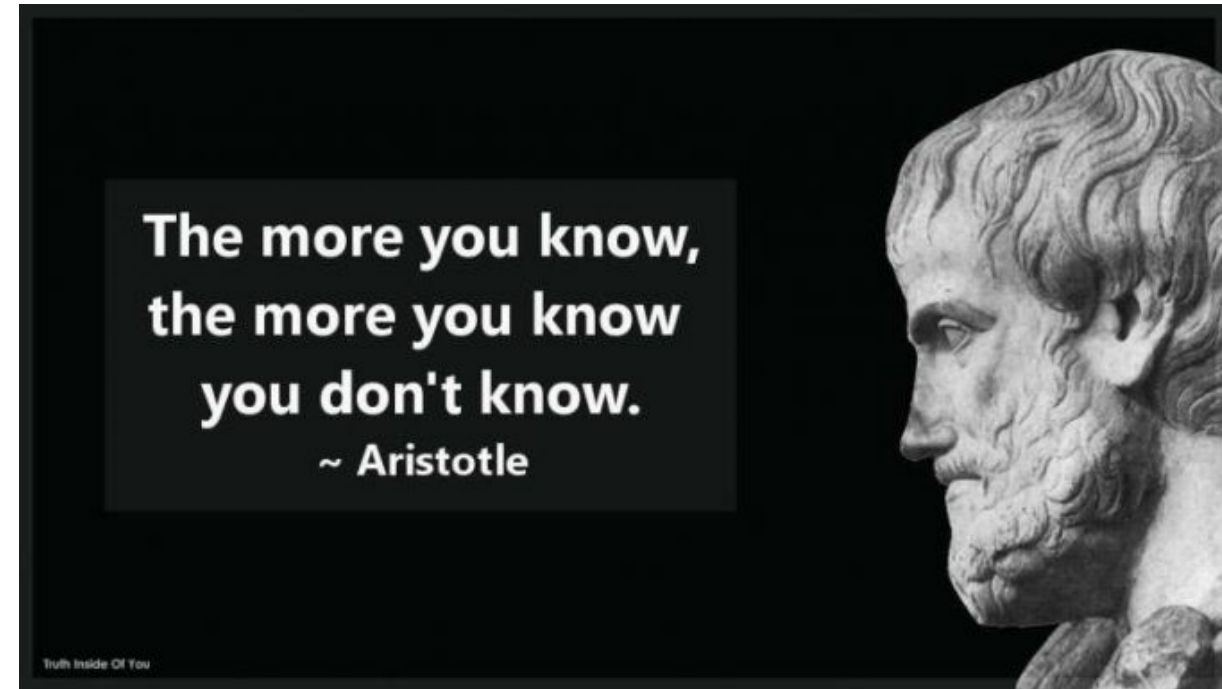


Ejercicio en clase

- Correr una par de ejemplos en sklearn usando los métodos tratados.

Lecturas extras:

1. <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=655F2BC08074BA0C900810868EE0C3DC?doi=10.1.1.175.107&rep=rep1&type=pdf> (Survey on Multiclass Classification Methods)
2. <https://www.sciencedirect.com/science/article/abs/pii/S0950705116301459> (Empowering one-vs-one decomposition with ensemble learning for multi-class imbalanced data)
3. <https://scikit-learn.org/stable/modules/multiclass.html> (1.12. Multiclass and multilabel algorithms)
4. Mirar en la lectura anterior la diferencia entre multiclase y multietiqueta.



[Esta foto](#) de Autor desconocido está bajo licencia [CC BY](#)