

# Sesión 8-Tema 6: Boosting y Adaboost

## Tema 6: Boosting y Adaboost – Resumen

### ¿Qué es el aprendizaje automático o *machine learning*?

Programar una computadora para que mejore en la realización de una tarea a partir de datos de ejemplo o de la experiencia.

Existen varios tipos de aprendizaje:

#### Supervisado:

- Los datos de entrada, imágenes, por ejemplo, vienen con una etiqueta indicando a la clase que pertenecen, así de este modo se puede saber si se ha equivocado o no. Se usa para reconocer patrones, imágenes, etc.
- Regresión: Predecir (calcular) el valor de salida según una serie de parámetros.

#### No supervisados:

- Los datos de entradas vienen solos, sin etiquetas.
- Se agrupan los datos en diferentes *clusters*.
- Los grupos deben ser homogéneos entre sí, distintos con los demás.

#### Por refuerzo:

- Función  $f(n)$ , valor elevado si funciona bien (no se equivoca en clasificar), valor bajo si clasifica mal.

## Boosting y Adaboost

La siguiente notación se usa en los algoritmos de aprendizaje supervisado:

- Patrones: Vectores de datos  $\rightarrow \mathbb{R}^d$
- Clases a las que pertenecen los valores
- Conjunto de entrenamiento del algoritmo
- Función aprendida
- Clasificador: Si una variable puede elegir entre '+' o '-', el clasificador es una de ambas opciones

Los patrones o vectores de datos se representan en hiperplanos definidos por la dimensión del patrón, si la dimensión,  $d = 5$ , se generará un hiperplano de  $n-1$  dimensiones para representar el espacio donde se establecerán las variables del patrón.

En cuanto a los clasificadores, según su porcentaje de acierto, se les considera clasificadores débiles o fuertes. Un clasificador  $h_1$  con un 50% de aciertos es débil, y si tenemos varios clasificadores débiles de manera independiente, el resultado puede no ser fiable.

Si tenemos varios, podemos hacer una combinación lineal de los diferentes clasificadores débiles, de este modo obtendremos un mejor % de acierto. Para ello existe el bagging o boosting.

- **Bagging:** Para un valor  $X$ , aplicamos todos los clasificadores, contamos cada tipo de respuesta dada por cada uno, y aquella que tenga más votos, es la ganadora. Cada clasificador tiene un voto, el que más votos tenga, es la buena.
- **Boosting:** Clasificadores fuertes, % de aciertos mayores del 75%, tienen más peso en la votación y disminuye inversamente proporcional el peso de los demás clasificadores, los votos vienen ponderados por una variable alfa, la cual representa el grado de confianza en ese clasificador ( $a_1 \cdot h_1 + a_2 \cdot h_2 + \dots + a_n \cdot h_n$ ).



#### **Adaboost** (adaptive boosting):

1º Asigna a todos los ejemplos la misma probabilidad ( $D_t$ )

2º Obtiene los clasificadores débiles

3º Entrena  $h_1$  teniendo en cuenta los ejemplos

4º bloque start-end:

Bucle:

- Realiza  $A$  pruebas random de distintos clasificadores débiles.
- Genera hiperplano a partir de cada  $A$  que prueba (clasificador)
- Calcula el error de cada clasificador (prob. de fallar), con el objetivo de minimizar el error
- Escoge el mejor  $A$  (menor error tenga)

5º bloque start-end:

- $h_1 = \text{mejor}(A)$ . Obtenemos valor de confianza  $a_1$  (según error presentado por  $h_1$ ).
- A partir del  $\epsilon$  de  $h_1$ , calculamos  $a_t$

6º bloque start-end:

- Calcula la nueva probabilidad de los ejemplos dados según el grado de confianza de cada clasificador ( $a_1 \cdot h_1 + a_2 \cdot h_2 + \dots$ )
- Calculamos  $Z_t$ , constante de normalización para que la suma de las masas distribuidas sea igual a 1.

7º repetimos algoritmo hasta que no queden clasificadores débiles por usar

AdaBoost

Índice Wiki individual