

# Combinación de clasificadores

Abdelmalik Moujahid

Grupo de Inteligencia Computacional  
Universidad del País Vasco  
Curso 2014-2015

# Índice

- 1 Clasificadores híbridos
- 2 Combinado del mismo clasificador base
- 3 Combinado distintos clasificadores base

# Índice

- 1 Clasificadores híbridos
- 2 Combinado del mismo clasificador base
- 3 Combinado distintos clasificadores base

## Clasificadores híbridos

### Motivación

- Combinación de opiniones de expertos (modelos) antes de tomar una decisión final
- Cada paradigma asociado a una región de decisión de un determinado tipo
- Al combinar paradigmas tratar de obtener la región de decisión idónea para el problema

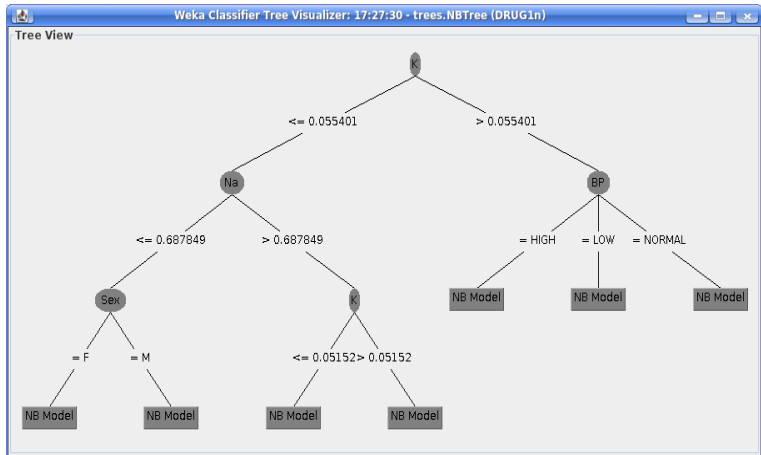
## Clasificadores híbridos

El clasificador se induce teniendo en cuenta dos o más paradigmas clasificatorios:

### Algunos ejemplos

- **Lazy Bayesian Rules** (Zheng y Webb, 2000): para cada ejemplo el clasificar obtiene un naive Bayes con aquellos casos mas cercanos
- **Naive Bayes Tree** (Kohavi, 1996): induce árboles cuyas hojas son clasificadores naive Bayes, los cuales se usarán en los ejemplos que alcancen dichas hojas
- **Logistic Model Trees** (Landwerhr y col., 2003): árboles en cuyas hojas se lleva a cabo la clasificación, de aquellas ejemplos que los alcancen, por medio de una regresión logística

## Clasificadores híbridos: Naive Bayes Tree



# Índice

- 1 Clasificadores híbridos
- 2 Combinado del mismo clasificador base**
- 3 Combinado distintos clasificadores base

## Bagging (Breiman, 1994)

### Bagging (Bootstrap AGGREGating)

- Se combinan clasificadores obtenidos de **réplicas bootstrap** del conjunto de entrenamiento
- Para cada ejemplo se selecciona la **clase más votada**
- Bagging buenos resultados con **clasificadores inestables** (pequeña variación en los datos produce gran variación en la clasificación)
- Árboles (inestables);  $k$ -NN (estables)



## Bagging (Breiman, 1994)

### 1 Generación del modelo

- Sea  $N$  el número de ejemplos en el conjunto de entrenamiento
- Para cada una de las  $t$  iteraciones:
  - Muestrar  $N$  ejemplos con remplazamiento del conjunto de entrenamiento
  - Inducir el modelo usando la muestra
  - Almacenar el modelo resultante

### 2 Clasificación

- Para cada uno de los  $t$  modelos:
  - Predecir la clase del ejemplo usando el modelo
- Devolver la clase que ha sido predicha más veces

## Random forest (Breiman, 2001)

- Variante de bagging con **árboles de clasificación como clasificador de base**
- Clasificador consistente en una colección de  $L$  clasificadores con estructura de árbol, donde cada árbol se ha construido a partir de un **vector aleatorio**
- El vector aleatorio puede consistir en:
  - **Aleatoriedad en los datos**
  - **Aleatoriedad en las variables**

## Boosting (Freund y Shapire, 1997)

- El conjunto de clasificador se desarrolla de manera **incremental**, añadiendo uno en cada iteración
- El clasificador en el paso  $t$  se entrena sobre un conjunto de entrenamiento muestreado de manera selectiva
- La distribución de muestreo comienza siendo uniforme
- Los objetos mal clasificados en la iteración anterior ven aumentada su probabilidad de ser seleccionados

# Boosting: AdaBoost

## Fase de Entrenamiento

- 1 Inicialización de parámetros: peso de cada uno de los  $N$  ejemplos:  $w_j^1 = \frac{1}{N}$ ; número de iteraciones  $L$ ; conjunto de clasificadores  $\mathcal{D} = \emptyset$
- 2 Para  $k = 1 \dots L$ 
  - Extraer una muestra  $S_k$  del conjunto de datos usando la distribución  $w_j^k$
  - Inducir el clasificador  $D_k$  usando  $S_k$  como conjunto de entrenamiento
  - Calcular el error en la iteración  $k$ :  $\varepsilon_k = \sum_{j=1}^N w_j^k \delta_j^k$
  - If  $\varepsilon_k = 0$  o  $\varepsilon_k \geq 0,5$ , ignorar  $D_k$ , reinicializar  $w_j^k$  a  $\frac{1}{N}$  y continuar
  - Else, computar:  $\beta_k = \frac{\varepsilon_k}{1-\varepsilon_k}$
  - Actualizar los pesos:  $w_j^{k+1} = w_j^k \beta_k^{(1-\delta_j^k)} / \sum_{j=1}^N w_j^k \beta_k^{(1-\delta_j^k)}$
- 3 Devolver  $\mathcal{D}$  y  $\beta_1, \dots, \beta_L$

## Fase de clasificación

- Calcular el soporte que recibe cada clase  $c_r$  como  $\sum_{D_k(\mathbf{x})=c_r} \ln(\frac{1}{\beta_k})$
- Clasificar el ejemplo  $\mathbf{x}$  como la clase que recibe el mayor soporte

## Boosting versus bagging

- Boosting combina múltiples modelos que se complementan unos a otros
- Al igual que bagging, usa voto por la mayoría para combinar las salidas de los modelos individuales
- Al igual que bagging, combina modelos del mismo paradigma
- En contraposición con bagging, donde cada modelo individual se induce de manera separada, en boosting cada nuevo modelo está influenciado por el comportamiento de los anteriores
- En boosting los nuevos modelos se convierten en expertos para ejemplos mal clasificados por los anteriores modelos
- En boosting la contribución final de cada modelo se basa en su bondad

# Índice

- 1 Clasificadores híbridos
- 2 Combinado del mismo clasificador base
- 3 Combinado distintos clasificadores base**

## Combinado distintos clasificadores base

### Métodos básicos

- Fusión
- Stacking

1- Ludmila I. Kuncheva (2004). *Combining Pattern Classifiers: Methods and Algorithms*. John Wiley and Sons, Inc.

2- J. Kittler, M. Hatef, Robert P.W. Duin, J. Matas (1998). *On combining classifiers*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 20(3):226-239.

3- David H. Wolpert (1992). *Stacked generalization*. *Neural Networks*. 5:241-259.

## Stacking (Wolpert, 1992)

- Diversas capas (o niveles) de clasificadores
- Cada capa utiliza los resultados obtenidos por la capa anterior
- La última capa está formada por un único clasificador que es el que toma la decisión final



## Conclusiones

- Combinando clasificadores tratar de encontrar superficies de decisión adecuadas al problema
  - *Hibridaciones*
  - Combinación de decisiones obtenidas con el mismo clasificador de base
  - Combinación de decisiones obtenidas con distintos clasificadores de base
- Campo activo de investigación empírica

## Combinando clasificadores en Weka

Lazy Bayesian Rules	Lazy	LBR
Naive Bayes Tree	Trees	NBTree
Logistic Model Trees	Trees	LMT
Bagging	Meta	Bagging
Random Forest	Trees	RandomForest + RandomTree
Fusion	Meta	Vote
Stacked	Meta	Stacking + StackingC