

# Combinación de Modelos

---

# Ensamble Learning

---

**Se tiene un conjunto de modelos para la clasificación**

**Un comité**

**Se pretende obtener un resultado basado en la opinión de varios expertos**

**Los resultados pueden ser difíciles de interpretar**

# Ensamble Learning

---

**Bagging**

**Aleatorización**

**Boosting**

**Satcking**

# Bagging

---

**Elegir de forma aleatoria k conjuntos de muestras del mismo tamaño, con reemplazo**

**Construir k modelos clasificadores**

**Combinar los resultados por votación**

- para regresión por ej. el promedio

**Bueno para métodos inestables**

- muy sensibles a los datos
- Árboles de decisión vs Naive Bayes

# Aleatorización

---

**Se agrega aleatoriedad al método no en la selección de las muestras**

**Depende del método**

**Por ejemplo para Árboles de decisión**

- **Se selección al azar entre los mejores atributos para la raíz de cada subárbol, no nec. el mejor**

**En weka "Random Forests"**

# Boosting

---

**Se crean modelos de forma iterativa, de forma que cada nuevo modelo ponga mayor peso en los datos que fueron mal clasificados por los anteriores**

**Cada modelo tiene un peso que indica su efectividad**

**Combinar los resultados por votación y pesos**

**En weka "AdaBoostM1"**

# Stacking – meta aprendizaje

---

**Se combinan modelos mediante otro método de aprendizaje en lugar de votación**

**Los modelos básicos se obtienen con métodos de nivel 0**

**El modelo final se obtiene con un método de nivel 1**

**En weka “Stacking”**

# Evaluación y Uso de Modelos

---



# Validación

---

Se busca que los modelos descubiertos tengan las siguientes cualidades:

- Precisos
- Comprensibles
- Interesantes/Relevantes

# Validación

---

El conjunto de datos de donde se intenta extraer conocimiento se llama conjunto de entrenamiento

La meta es obtener conocimiento válido no solo para conjunto de entrenamiento considerado sino para datos similares

El conocimiento puede/debe ser probado con otro conjunto, conjunto de prueba

# Validación Simple

---

Validación con datos conocidos, pero distintos a los usados en el entrenamiento

Validación simple - dividir el conjunto de datos en dos: un conjunto de entrenamiento y uno pequeño de prueba, los de prueba no se deben usar en el entrenamiento

# Validación

---

El problema de la partición entrenamiento/prueba, es que los datos que se usan para entrenamiento no se usan para comprobación y viceversa

En el caso de no disponer de muchos datos o bien el modelo se entrena con pocos datos (o menos de lo que sería deseable) o bien se valida con pocos datos (o menos de lo que sería deseable)

# Cross-validation

---

¿Existe alguna manera de utilizar gran parte de los datos y aún así tener una referencia para saber lo bueno que es el modelo?

Una herramienta sencilla es el método de validación cruzada o “Cross-validation”

# Cross-validation

---

Validación cruzada, con pocos datos

El conjunto se divide en dos: A y B

División aleatoria y los conjuntos son del mismo tamaño

Primero se entrena el modelo con los datos de A y se validan con B, se calcula el error; se entrenan los datos con B, se validan con A y se calcula el error

Se usa el modelo con el menor error

# k-fold Cross-validation

---

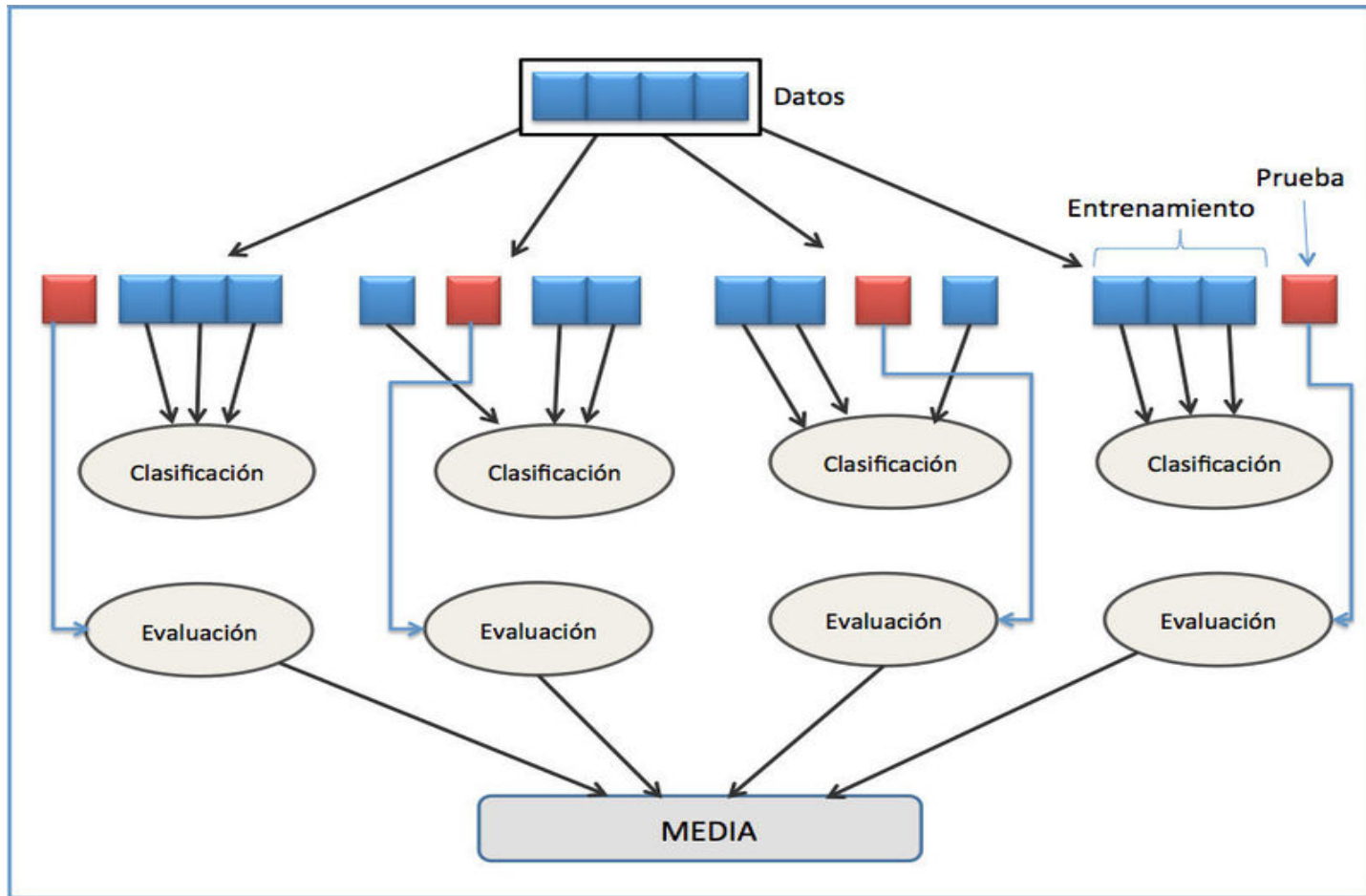
Validación cruzada con k pliegues o k-fold cross-validation

Es una variación de la cruzada, y consiste en dividir el conjunto en n grupos

Se usan k-1 grupos para pruebas y el restante para el entrenamiento, se calcula el error y se repite el proceso k veces cambiando el conjunto de entrenamiento

Se toma la media aritmética de los resultados de cada iteración para obtener un único resultado

# 4-fold cross-validation





# k-fold Cross-validation

---

¿Desventajas?

# k-fold Cross-validation

---

¿La validación cruzada permite aprovechar mejor los datos para obtener?

- Mejores modelos
- Mejores validaciones de los modelos
- Mejores modelos y mejores validaciones de los modelos al mismo tiempo

# k-fold Cross-validation

---

Si aprendiéramos un modelo con sólo un 50% de los 25 datos muestreados, ¿es de esperar que el modelo fuera mejor o peor que el de la validación cruzada?

Probar en weka

# Evaluación

---

En la clasificación se contabilizan los errores en el conjunto de prueba

Evaluación basada en costos

- Errores ponderados por el costo

Regresión

- Distancia entre el valor real y predicho (ev. cuadrático)

Clustering

- Complejidad del modelo y medidas de intra e inter cluster

# Comparación

---

Comparar el rendimiento relativo de dos modelos de clasificación alternativos

# Comparación

---

**Matriz de confusión  
(confusion matrix)**

	Predicción	
Clase real	<b>TP:</b> True positive	<b>FN:</b> False negative
	<b>FP:</b> False positive	<b>TN:</b> True negative

**Evaluación del clasificador**

**accuracy = (TP+TN)/(TP+TN+FP+FN) *Exactitud***

# Comparación

---

- ◆ **precision =  $TP / (TP + FP)$**
- ◆ **recall = sensitivity =  $TP / (TP + FN)$  *True positive recognition rate***
- ◆ **specificity =  $TN / (TN + FP)$  *True negative recognition rate***
  
- ◆  **$F = 2TP / (2TP + FP + FN)$  **F-measure****  
 **$F = 2 * \text{precision} * \text{recall} / (\text{precision} + \text{recall})$**
  
- ◆  **$F_b = (1 + b_2) \text{precision} * \text{recall} / (b_2 * \text{precision} + \text{recall})$**

# Comparación

## Accuracy

	Predicción	
Real	TP	FN
	FP	TN

## Recall

	Predicción	
Real	TP	FN
	FP	TN

## Precision

	Predicción	
Real	TP	FN
	FP	TN

## F-measure

	Predicción	
Real	TP	FN
	FP	TN



# Análisis ROC

---

Desarrolladas en los años 50 para analizar señales con ruido: caracterizar el compromiso entre aciertos y falsas alarmas

Permiten comparar visualmente distintos modelos de clasificación

El área que queda bajo la curva es una medida de la precisión (accuracy) del clasificador:

- Cuanto más cerca estemos de la diagonal (área cercana a 0.5), menos preciso será el modelo
- Un modelo "perfecto" tendrá área 1

# Análisis ROC

---

Se grafican los valores normalizados:

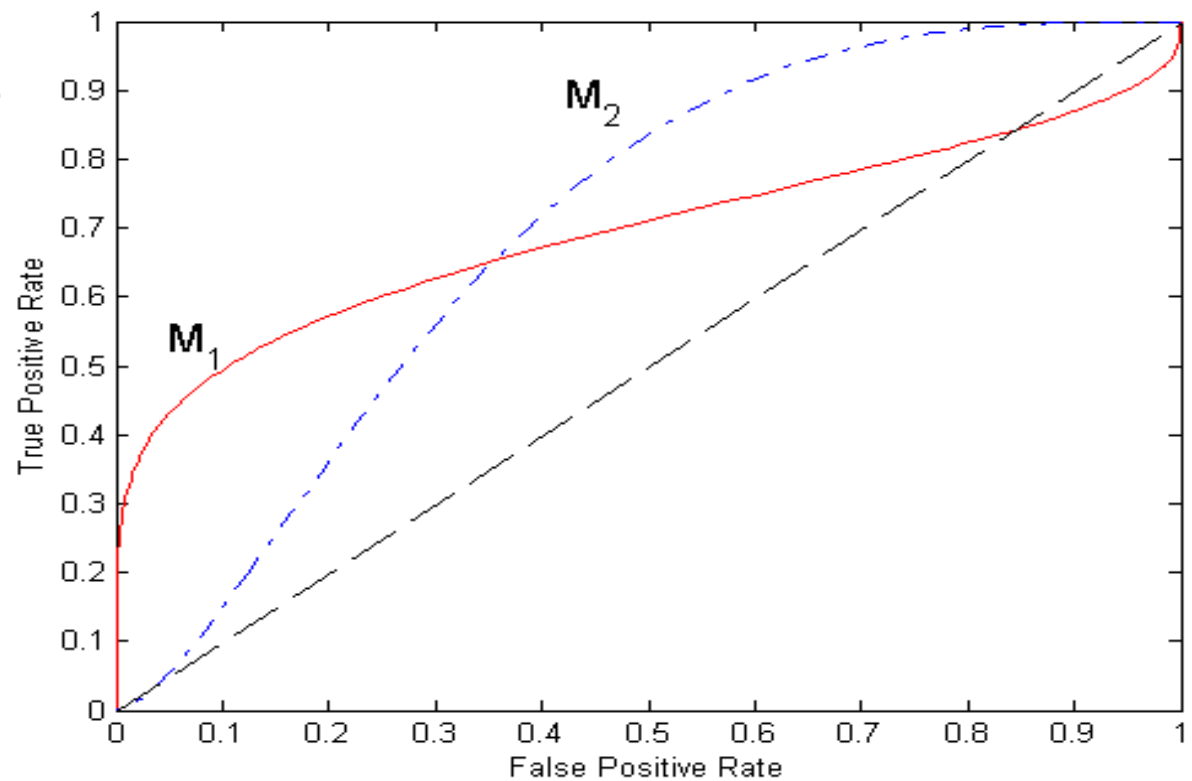
**TPR = TP/(TP+FN) True Positive Rate**

**FNR = FN/(TP+FN) False Negative Rate**

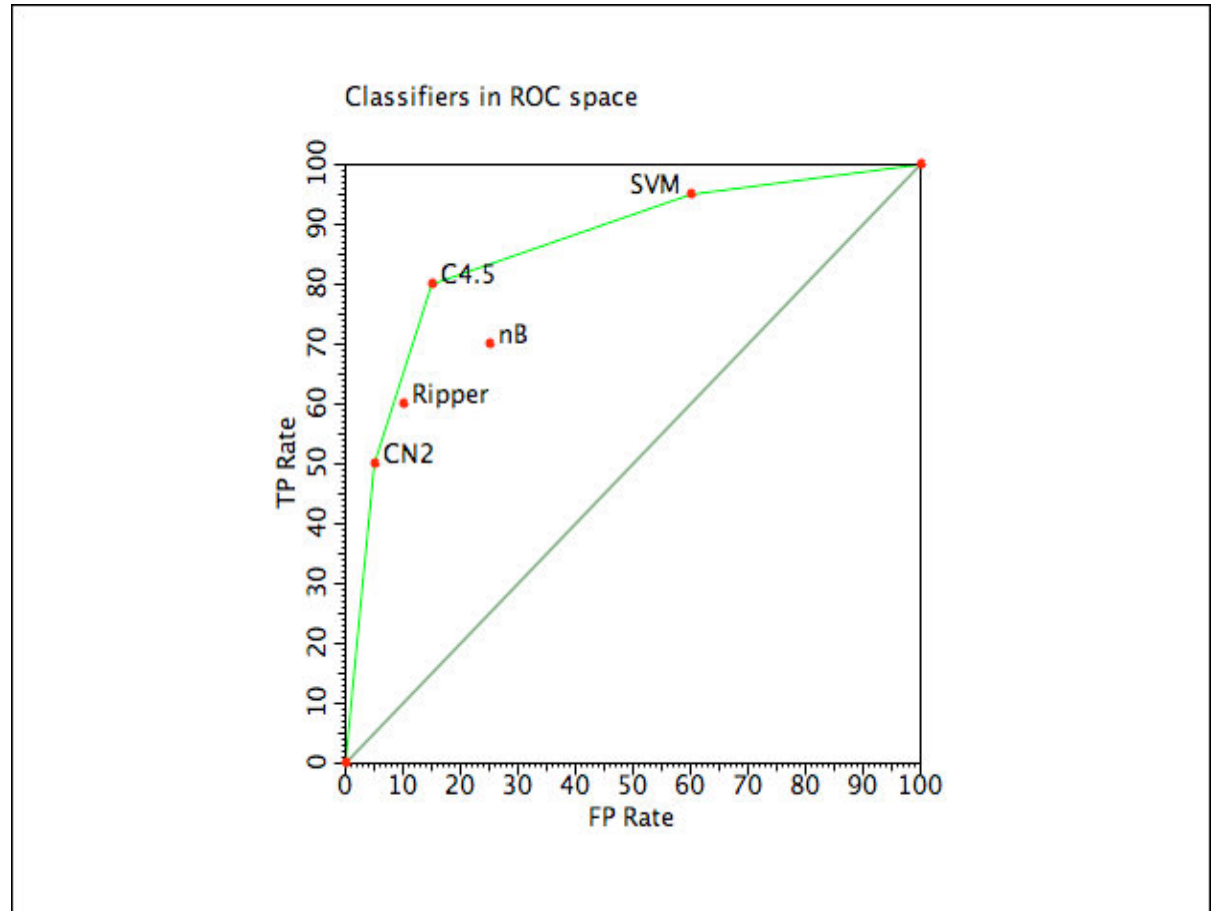
	Predicción	
Clase real	<b>TP:</b> True positive	<b>FN:</b> False negative
	<b>FP:</b> False positive	<b>TN:</b> True negative

# Análisis ROC

Ningún modelo es consistentemente mejor que el otro: M1 es mejor para FPR bajos, M2 para FPR altos.



# Análisis ROC



# Análisis ROC

