



# Aprendizaje Conjunto

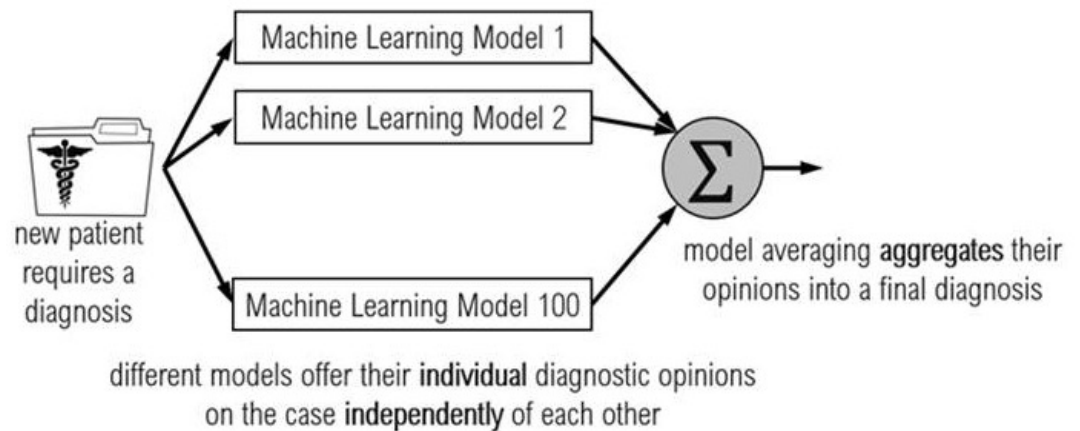
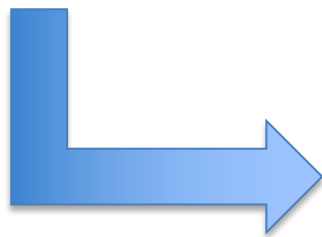
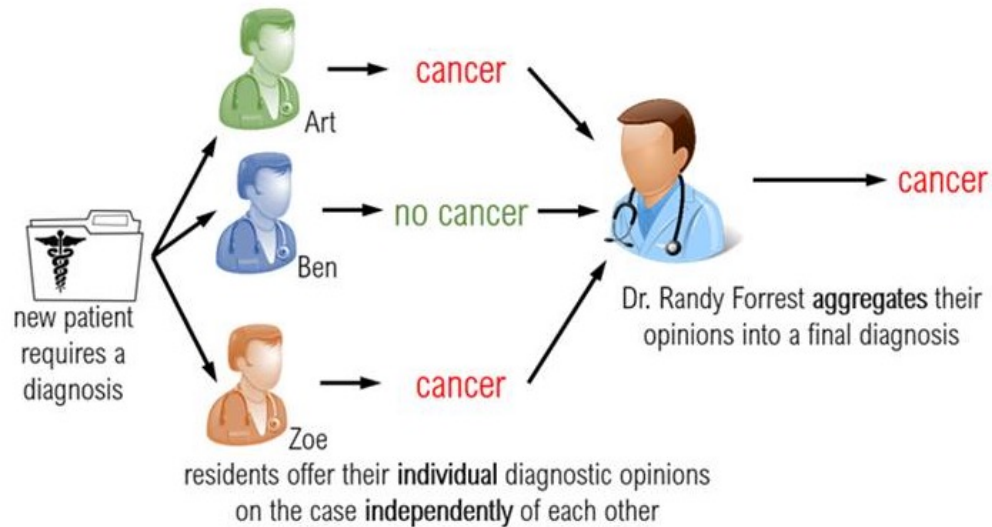
- Introducción
- Fundamentos
- Método Bagging (Bootstrap aggregating)

- **Inspiración.** El aprendizaje conjunto se basa en la *segunda naturaleza* de las personas para buscar diversas opiniones antes de tomar una decisión importante.
- **Idea básica.** Combinar varios clasificadores de patrones individuales para obtener un clasificador que los supere a cada uno de ellos.

- **Antecedentes:**

- Teorema del jurado de Condorcet en 1785 en el ensayo sobre la aplicación del análisis a la probabilidad de decisiones mayoritarias.
- El concurso sobre estimación del peso de un hacha en la visita de Sir Francis Galton (1822-1911) a la feria ganadera.
- El estudio de James M. Surowiecki en su libro de 2004, “The wisdom of the Crowds”. Donde se indican los criterios para que una multitud sea “*sabia*”:
  - Diversidad de opinión
  - Independencia
  - Decentralización
  - Agregación

- Idea gráfica



- **Conceptos:**

- Aprendiz fuerte. Es un método de aprendizaje que produce un clasificador que puede ser arbitrariamente exacto.
- Aprendiz débil. Es un método de aprendizaje que produce un clasificador el cual es ligeramente más exacto que una clasificación aleatoria.

- Es un método homogéneo simple pero efectivo para generar un ensamble de clasificadores.
- El clasificador conjunto generado consolida las salidas de varios clasificadores (homogéneos) entrenados en una sola clasificación.
- Cada clasificador en el conjunto es entrenado con una muestra con reemplazo de las instancias del conjunto de entrenamiento.

- Es un método homogéneo simple pero efectivo para generar un ensamble de clasificadores.
- El clasificador conjunto generado consolida las salidas de varios clasificadores (homogéneos) entrenados en una sola clasificación.
- Cada clasificador en el conjunto es entrenado con una muestra con reemplazo de las instancias del conjunto de entrenamiento.

## Bagging Training

**Require:**  $I$  (a base inducer),  $T$  (number of iterations),  $S$  (the original training set),  $\mu$  (the sample size).

- 1:  $t \leftarrow 1$
- 2: **repeat**
- 3:    $S_t \leftarrow$  a sample of  $\mu$  instances from  $S$  with replacement.
- 4:   Construct classifier  $M_t$  using  $I$  with  $S_t$  as the training set
- 5:    $t \leftarrow t + 1$
- 6: **until**  $t > T$

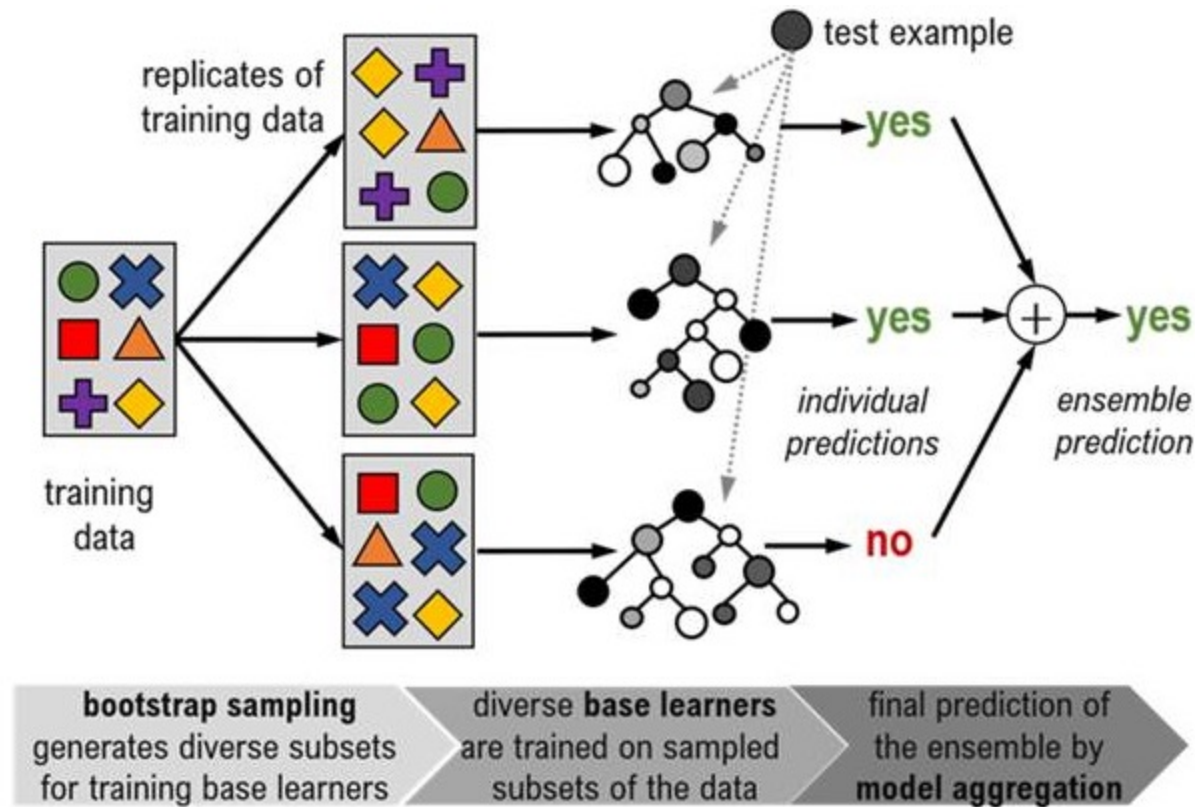


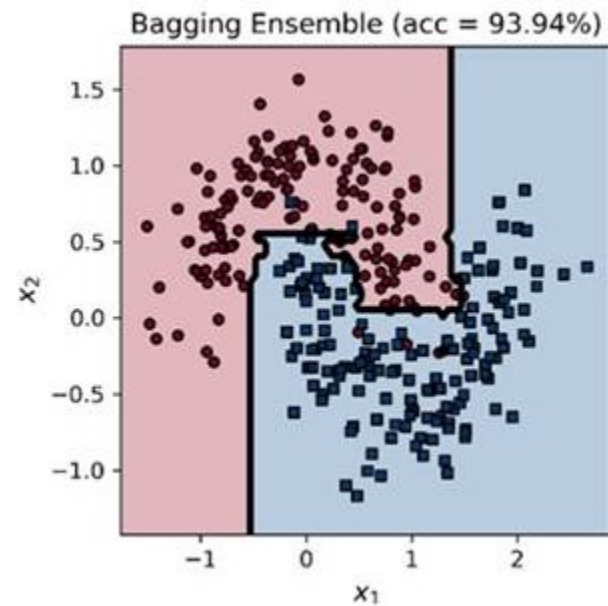
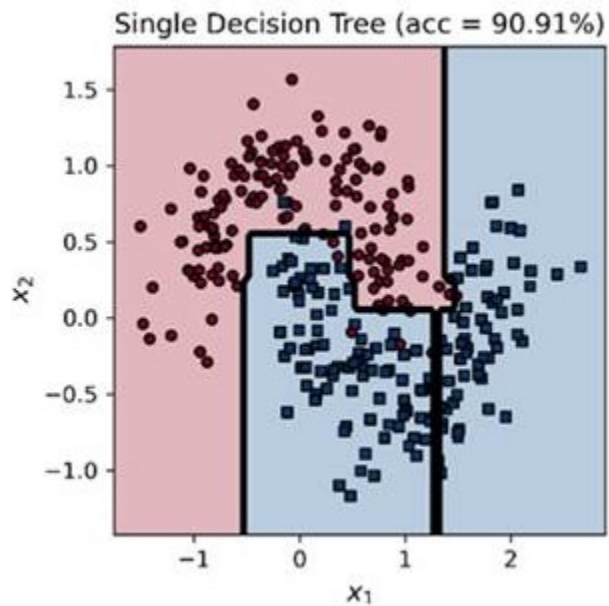
## Bagging Classification

**Require:**  $x$  (an instance to be classified)

**Ensure:**  $C$  (predicted class)

- 1:  $Counter_1, \dots, Counter_{|dom(y)|} \leftarrow 0$  {initializes class votes counters}
- 2: **for**  $i = 1$  **to**  $T$  **do**
- 3:    $vote_i \leftarrow M_i(x)$  {get predicted class from member  $i$ }
- 4:    $Counter_{vote_i} \leftarrow Counter_{vote_i} + 1$  {increase by 1 the counter of the corresponding class}
- 5: **end for**
- 6:  $C \leftarrow$  the class with the largest number votes
- 7: **Return**  $C$





# UNIVERSIDAD DE GUANAJUATO

