



Ensamblado de Modelos Boosting

<https://catboost.ai/en/docs/concepts/tutorial-1>



Boosting

En el caso del **bagging**, teníamos un conjunto de modelos independientes, cuyos outputs agregados servían para el output final.

Los modelos van en paralelo y tienen el mismo objetivo

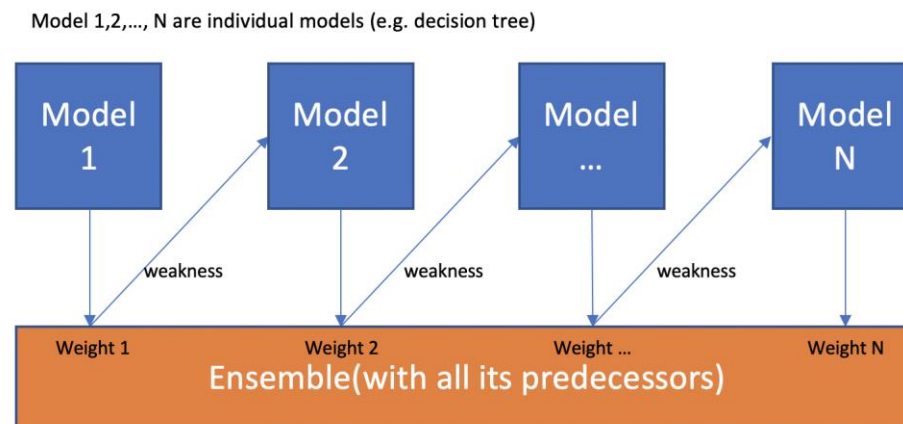
Boosting

En el caso del **bagging**, teníamos un conjunto de modelos independientes, cuyos outputs agregados servían para el output final.

Los modelos van en paralelo y tienen el mismo objetivo

En **boosting** los modelos se entrenan secuencialmente y por tanto existe una dependencia entre ellos.

Además, los modelos tienen como target el error cometido por el modelo anterior.



Boosting

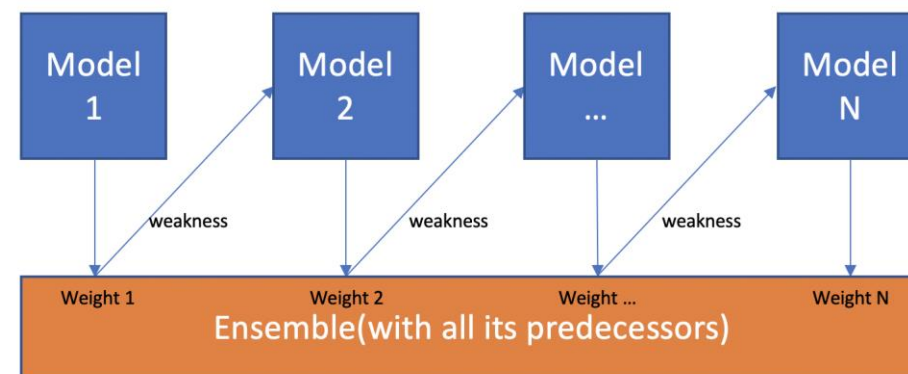
En el caso del **bagging**, teníamos un conjunto de modelos independientes, cuyos outputs agregados servían para el output final.

Los modelos van en paralelo y tienen el mismo objetivo

En **boosting** los modelos se entrenan secuencialmente y por tanto existe una dependencia entre ellos.

Además, los modelos tienen como target el error cometido por el modelo anterior.

Model 1,2,..., N are individual models (e.g. decision tree)



Básicamente en esta técnica los modelos van intentando mejorar su predecesor, recibiendo los errores del mismo, e intentando mejorar su resultado

Boosting

En el caso del **bagging**, teníamos un conjunto de modelos independientes, cuyos outputs agregados servían para el output final.

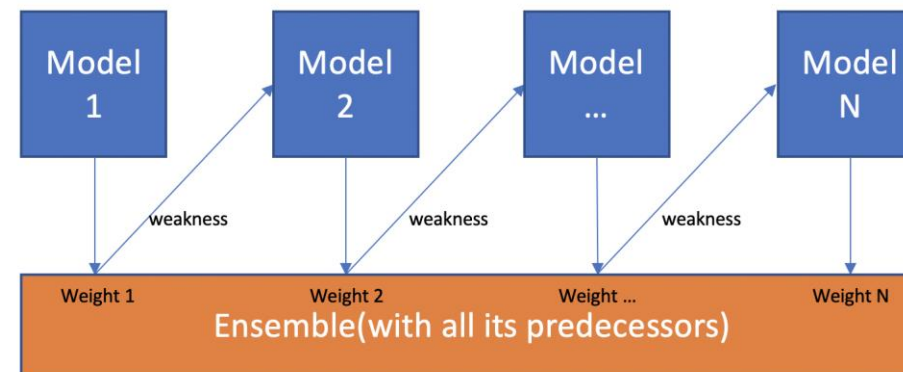
Los modelos van en paralelo y tienen el mismo objetivo

En **boosting** los modelos se entrenan secuencialmente y por tanto existe una dependencia entre ellos.

Además, los modelos tienen como target el error cometido por el modelo anterior.

También emplea, generalmente, árboles de decisión como submodelos

Model 1,2,..., N are individual models (e.g. decision tree)



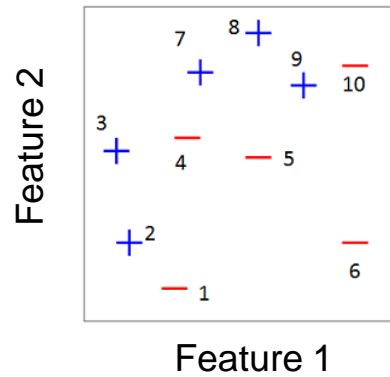
Básicamente en esta técnica los modelos van intentando mejorar su predecesor, recibiendo los errores del mismo, e intentando mejorar su resultado

Boosting: Ada Boost

Para explicar el concepto veamos además como funciona el boosting original el AdaBoost:

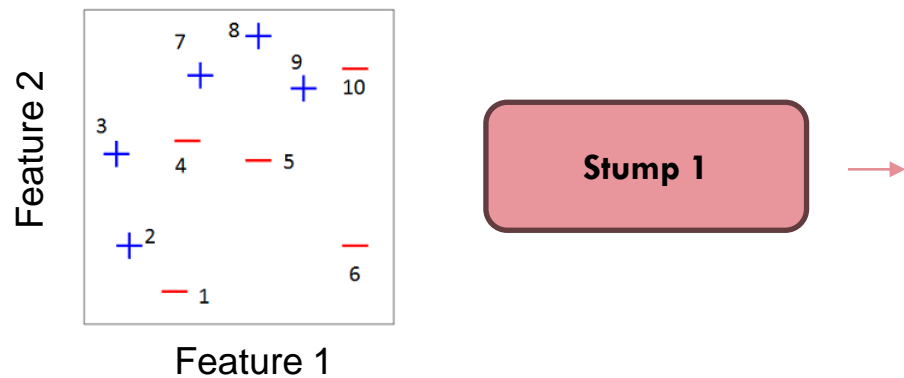
Boosting: Ada Boost

Para explicar el concepto veamos además como funciona el boosting original el AdaBoost:



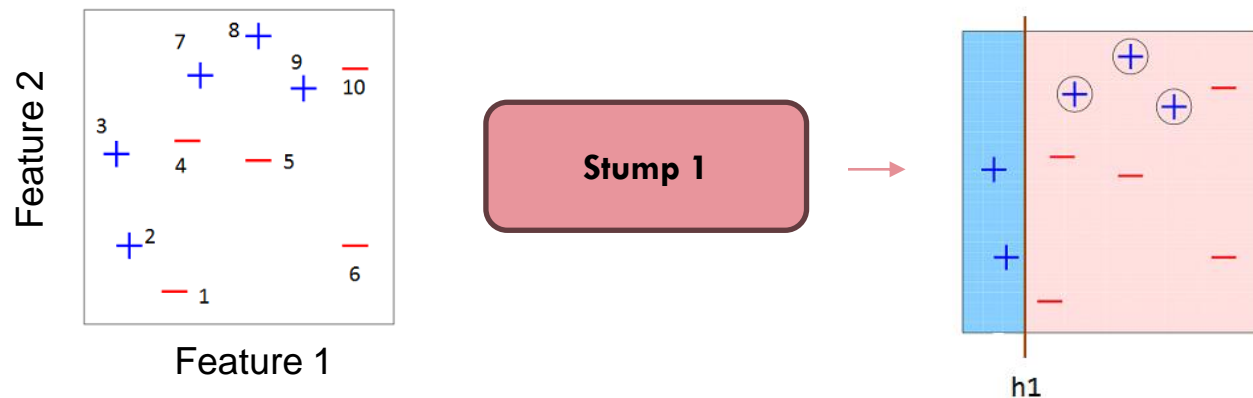
Boosting: Ada Boost

Para explicar el concepto veamos además como funciona el boosting original el AdaBoost:



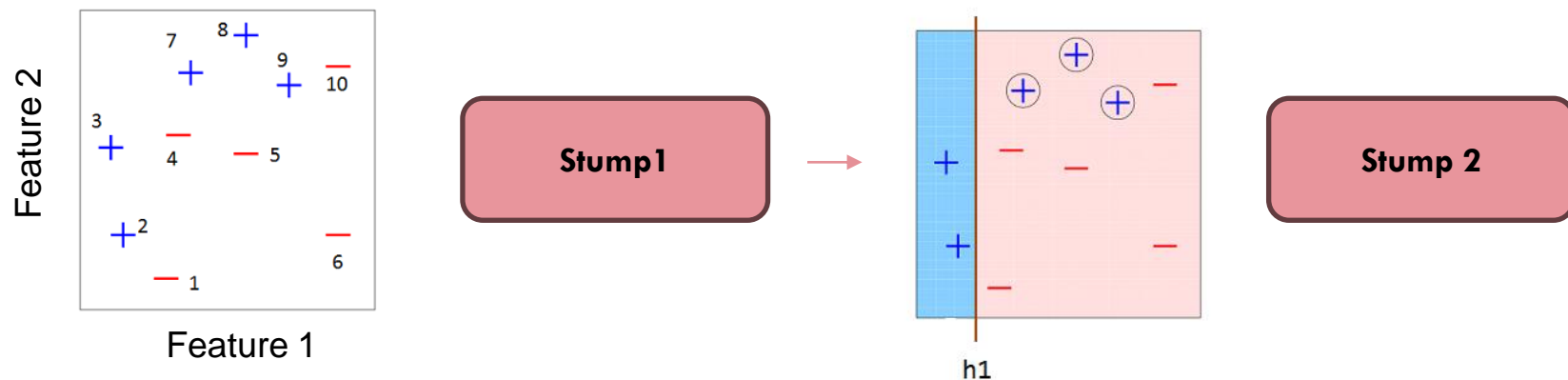
Boosting: Ada Boost

Para explicar el concepto veamos además como funciona el boosting original el AdaBoost:



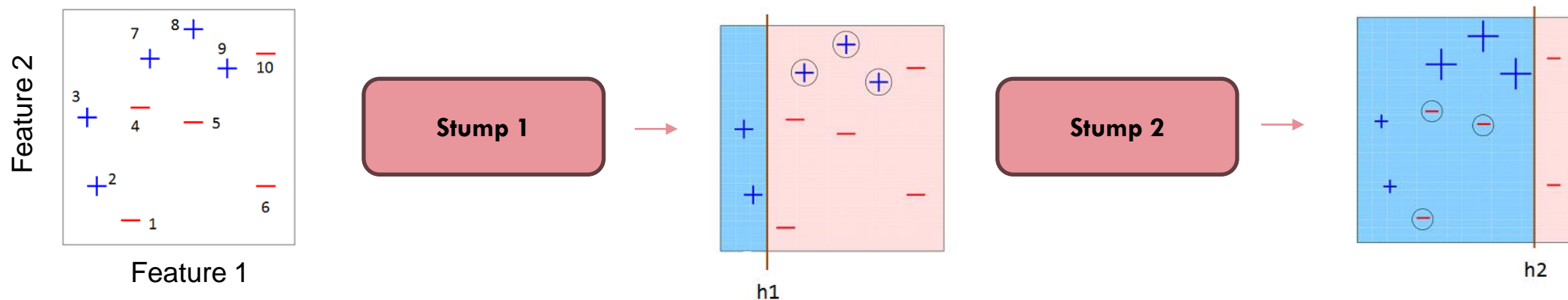
Boosting: Ada Boost

Para explicar el concepto veamos además como funciona el boosting original el AdaBoost:



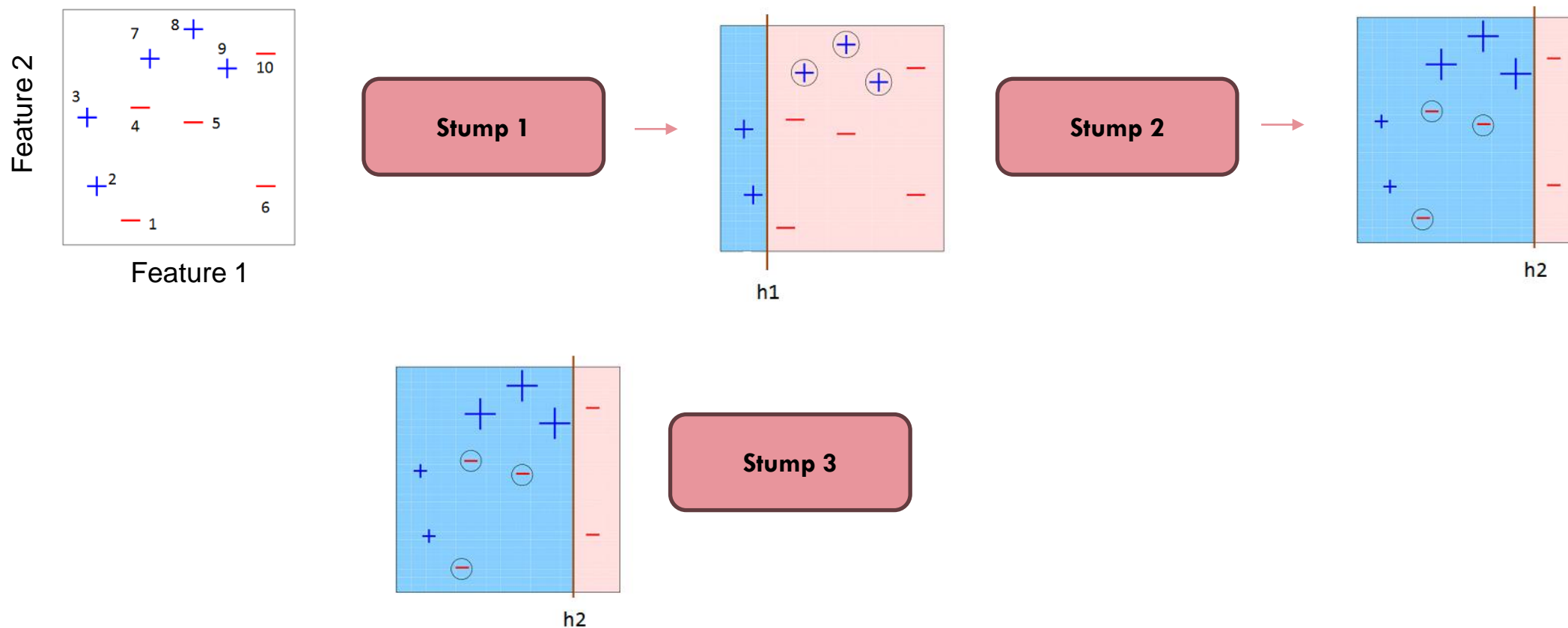
Boosting: Ada Boost

Para explicar el concepto veamos además como funciona el boosting original el AdaBoost:



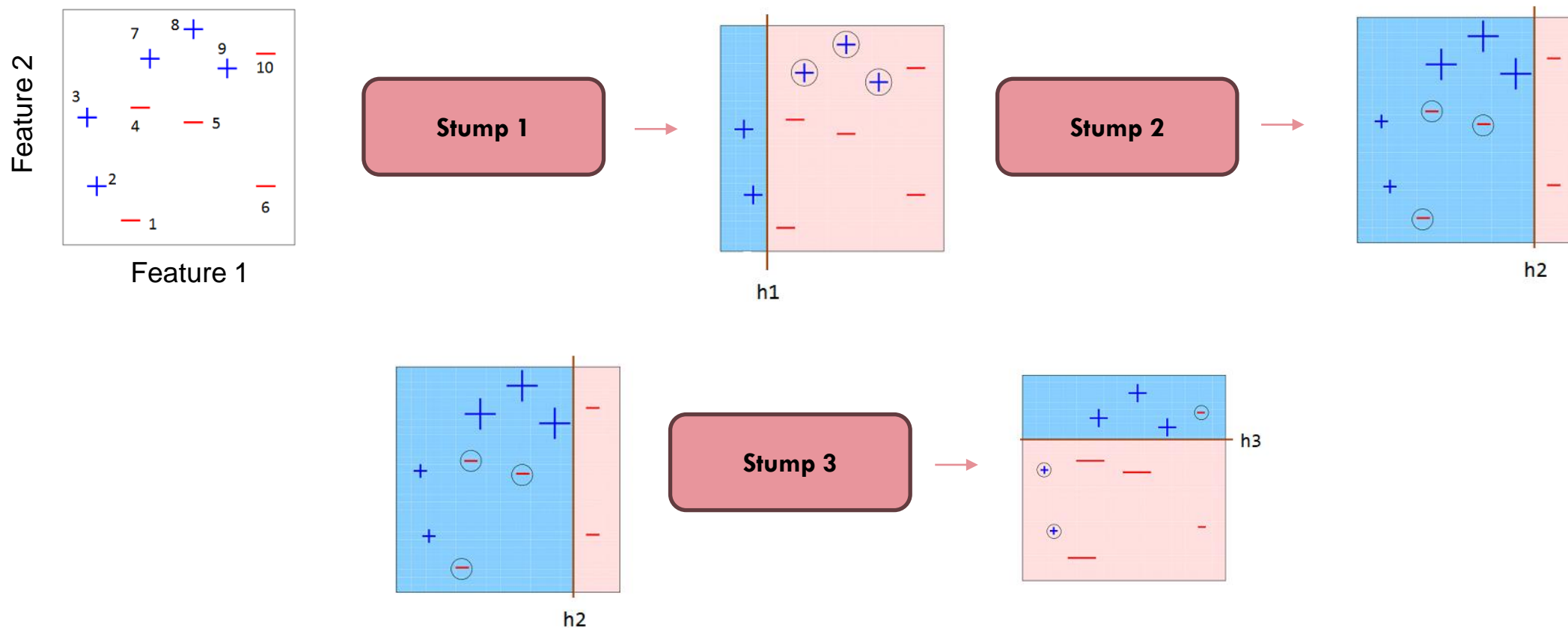
Boosting: Ada Boost

Para explicar el concepto veamos además como funciona el boosting original el AdaBoost:



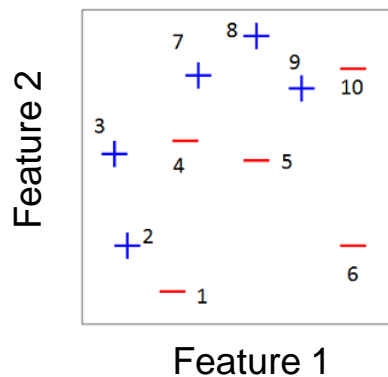
Boosting: Ada Boost

Para explicar el concepto veamos además como funciona el boosting original el AdaBoost:



Boosting: Ada Boost

Para explicar el concepto veamos además como funciona el boosting original el AdaBoost:



$$H_{final} = \text{sign} \left(\alpha_1 \begin{array}{|c|} \hline \text{blue} \\ \hline \text{red} \end{array} + \alpha_2 \begin{array}{|c|} \hline \text{blue} \\ \hline \text{red} \end{array} + \alpha_3 \begin{array}{|c|} \hline \text{blue} \\ \hline \text{red} \end{array} \right) = \begin{array}{|c|c|c|} \hline \text{blue} & \text{blue} & \text{red} \\ \hline \text{blue} & \text{red} & \text{red} \\ \hline \end{array}$$

Boosting: Gradient Boosted (Trees)

Boosting: Gradient Boosted (Trees)

Son también modelos que trabajan secuencialmente.

Boosting: Gradient Boosted (Trees)

Son también modelos que trabajan secuencialmente.

Generalmente son árboles

Boosting: Gradient Boosted (Trees)

Son también modelos que trabajan secuencialmente.

Generalmente son árboles

Al igual que AdaBoost cada modelo se centra en los errores del anterior modelo, pero a diferencia de Ada:

Boosting: Gradient Boosted (Trees)

Son también modelos que trabajan secuencialmente.

Generalmente son árboles

Al igual que AdaBoost cada modelo se centra en los errores del anterior modelo, pero a diferencia de Ada:

- En problemas de clasificación: Cada submodelo es un regresor que intenta adivinar la probabilidad que le faltó a cada clase mal clasificada del anterior modelo para haberse clasificado bien

Boosting: Gradient Boosted (Trees)

Son también modelos que trabajan secuencialmente.

Generalmente son árboles

Al igual que AdaBoost cada modelo se centra en los errores del anterior modelo, pero a diferencia de Ada:

- En problemas de clasificación: Cada submodelo es un regresor que intenta adivinar la probabilidad que le faltó a cada clase mal clasificada del anterior modelo para haberse clasificado bien
- En problemas de regression: Cada submodelo intenta estimar el error de las estimaciones hechas por el submodelo anterior

Boosting: Gradient Boosted (Trees)

Son también modelos que trabajan secuencialmente.

Generalmente son árboles

Al igual que AdaBoost cada modelo se centra en los errores del anterior modelo, pero a diferencia de Ada:

- En problemas de clasificación: Cada submodelo es un regresor que intenta adivinar la probabilidad que le faltó a cada clase mal clasificada del anterior modelo para haberse clasificado bien
- En problemas de regression: Cada submodelo intenta estimar el error de las estimaciones hechas por el submodelo anterior

La respuesta final es la combinación lineal (sumas modificadas por un coeficiente) o regresión de las respuestas de cada submodelo.

Boosting: Gradient Boosted (Trees)

Son también modelos que trabajan secuencialmente.

Generalmente son árboles

Al igual que AdaBoost cada modelo se centra en los errores del anterior modelo, pero a diferencia de Ada:

- En problemas de clasificación: Cada submodelo es un regresor que intenta adivinar la probabilidad que le faltó a cada clase mal clasificada del anterior modelo para haberse clasificado bien
- En problemas de regression: Cada submodelo intenta estimar el error de las estimaciones hechas por el submodelo anterior

La respuesta final es la combinación lineal (sumas modificadas por un coeficiente) o regresión de las respuestas de cada submodelo.

Los pesos de la combinación anterior también se entrenan (son parámetros del modelo) mediante un algoritmo de gradiente descendente.

Sí, demasiado para un ratito, así que no le des más vueltas. Porque además con el tiempo han aparecido variantes (que veremos en la parte práctica) y que usaremos como cajas negras (como casi todos los modelos). Este conocimiento debe servirte para saber que son muy muy precisos pero generalizan mejor que los árboles y ganan a los random forest en casi todo.

Boosting: Gradient Boosted (Trees)

Hay unas cuantas variantes:

- GBT (puros)
- Histogram-based Gradient Boosting
- XGBoost (Extreme Gradient Boosting)
- LightGBM (Light Gradient Boosting Machines)
- CatBoost (Categorical Boosting)

Boosting: Hiperparámetros “genéricos”

- `n_estimators`
- `max_Depth`
- `min_samples_leaf`
- `max_samples_split`
- `max_features`
- `class_Weight`
- `Learning_rate`
- `Alpha`

