

Diferencias entre Boosting y Bagging:

Un clasificador más fuerte se puede formar por el conjunto de votos ponderados es decir de clasificadores y muestreo ponderado, es decir los ejemplos. Los clasificadores ponderados se utilizan en los clasificadores con el mismo peso en el voto.

Por un lado tenemos Boosting, este no combina los clasificadores con el mismo peso en el voto, cada nuevo modelo está influenciado por el comportamiento de los anteriores, puede dañar su rendimiento con los datasets ruidosos, este será entrenado hasta que consiga clasificar bien el máximo de ejemplos posibles y se lleva acabo con Adaboost escogiendo un valor de confianza α_t , además boosting concederá menor peso a las muestras que están clasificadas correctamente el más alto a los mal clasificados. Boosting realizará un muestreo ponderado mientras que Bagging no lo realizará. Bagging será capaz de ayudar a mejorar en redes neuronales y árboles de decisión, en Bagging los elementos del conjunto de entrenamiento clasificados por $h(t)$ se pueden usar en $h(t+1)$ y la hipótesis final no se selecciona el clasificador que mejor haya evaluado el conjunto de entrenamiento.

Las diferencias más notables son que el boosting asignará los pesos a cada registro de entrenamiento y Bagging elige aleatoriamente los registros para formar los subconjuntos, ambos mejoran los clasificadores inestables, boosting pondera y da más pesos a los ejemplos que más cuestan de clasificar y bagging entrena un clasificador débil con el subconjunto cogido T veces.