Los Valores SHAP en Machine Learning son una forma de explicar cómo las características (inputs) de un modelo influyen en su predicción de salida de manera individualizada y precisa. Son especialmente útiles para interpretar modelos complejos, proporcionando una visión clara de la importancia de cada característica en la decisión final del modelo.

**¿Qué son los Valores SHAP?**

Los Valores SHAP (SHapley Additive exPlanations) se basan en la teoría de juegos, específicamente en los valores de Shapley, que distribuyen "justamente" la contribución de cada jugador (en este caso, cada característica de entrada) al éxito total del equipo (la predicción del modelo). En términos sencillos, te dicen cuánto contribuye cada característica a la predicción final, teniendo en cuenta la interacción con otras características.

**¿Cómo se Calculan?**

Para calcular los Valores SHAP de una característica:

1. Se evalúa el modelo múltiples veces con conjuntos de características ligeramente diferentes, incluyendo y excluyendo la característica en cuestión.
2. Se observa cómo cambia la predicción del modelo con y sin esa característica.
3. Estos cambios se promedian de manera que se cumplan tres propiedades: precisión local (la suma de los valores SHAP de todas las características es igual a la predicción del modelo), ausencia (características no presentes reciben un valor de cero) y consistencia (si una característica se vuelve más importante, su valor SHAP no disminuye).

**Diferencias con la Importancia de Características en Scikit-Learn**

La importancia de las características que viene por default en los métodos de scikit-learn (como árboles de decisión y bosques aleatorios) generalmente se basa en criterios como la ganancia de información o el conteo de divisiones. Estos métodos:

* Pueden dar una visión global de qué características son importantes en general para el modelo.
* No consideran la interacción entre características.
* No son consistentes; es decir, una característica puede volverse más importante para el modelo, pero su medida de importancia podría no reflejar este cambio.

En contraste:

* Los Valores SHAP proporcionan una medida individualizada y precisa de la importancia de las características, considerando las interacciones entre ellas.
* Son consistentes, garantizando que si la importancia real de una característica aumenta, su valor SHAP también lo hace.
* Ofrecen tanto explicaciones globales (importancia general de las características en todo el modelo) como locales (importancia de las características para una única predicción).

En resumen, los Valores SHAP ofrecen una aproximación más detallada y teóricamente sólida para entender cómo cada característica influye en las predicciones de un modelo de Machine Learning, superando algunas de las limitaciones de los métodos de importancia de características por defecto en paquetes como scikit-learn.

**DIFERENCIAS**

La diferencia fundamental entre el cálculo de los **Valores SHAP** para una variable y el cálculo de la **Importancia de las Características** (por ejemplo, mediante el índice Gini) radica en cómo se mide la contribución de cada característica a la predicción del modelo y qué aspectos de la predicción se están tratando de explicar.

**Valores SHAP**

* **Fundamento Teórico:** Se basan en la teoría de juegos, específicamente en los valores de Shapley, para distribuir la "contribución" de cada característica a la predicción del modelo de manera justa entre todas las características.
* **Cálculo:** Implica evaluar el impacto de agregar o quitar una característica del conjunto de características, observando cómo cambia la predicción del modelo. Requiere calcular el valor esperado de la predicción con y sin la característica, sobre todas las posibles combinaciones de otras características.
* **Propósito:** Proporcionan una medida precisa de la importancia de cada característica a nivel individual (para una sola predicción) y global (para todas las predicciones), teniendo en cuenta las interacciones entre características.
* **Ventajas:** Ofrece explicaciones detalladas y precisas, es consistente (si una característica se vuelve más influyente, su valor SHAP aumenta), y aplica tanto a nivel de predicción individual como en el modelo completo.

**Importancia de las Características (Gini)**

* **Fundamento Teórico:** En el contexto de los árboles de decisión y bosques aleatorios, la importancia se mide a menudo por el índice Gini, que indica cuánto contribuye una característica a la homogeneidad de los nodos y las divisiones en el árbol.
* **Cálculo:** Se calcula basándose en cuánto reduce una característica la impureza de Gini cuando se utiliza para dividir los datos en un árbol. Se considera la suma total de la disminución de la impureza traída por esa característica, a menudo normalizada por el número de divisiones.
* **Propósito:** Proporciona una medida global de la importancia de una característica en un modelo, indicando cuánto mejora la característica las divisiones del modelo en promedio.
* **Ventajas:** Fácil de calcular y entender en el contexto de los modelos basados en árboles, pero no capta la complejidad de las interacciones entre características ni proporciona importancia a nivel de predicción individual.

**Diferencias Clave**

* **Perspectiva:** Los valores SHAP ofrecen una vista más detallada y personalizada de la contribución de las características, mientras que el índice Gini proporciona una vista más general y menos detallada.
* **Interacciones:** Los valores SHAP pueden capturar y explicar las interacciones entre características, mientras que la importancia de las características basada en Gini no lo hace de manera directa.
* **Consistencia:** Los valores SHAP son consistentes en su asignación de importancia, mientras que la importancia basada en Gini puede no reflejar siempre adecuadamente el cambio en la importancia de las características.

En resumen, los valores SHAP proporcionan un enfoque más granular y teóricamente fundado para entender la importancia de las características, adecuado para análisis detallados y explicaciones precisas, mientras que la importancia de las características basada en Gini ofrece una visión más general de la utilidad de las características en modelos basados en árboles.

**From text**

SHAP stands for SHapely Additive exPlanations and combines additive feature attribution methods with Shapley framework from game theory. On the one hand, additive feature attribution methods explain the outcome of a model as a sum of the individual contribution of each feature and have three properties: local accuracy, missingness and consistency (Lundberg et al., 2017). On the other hand, in cooperative game theory Shapley framework is used to determine the fair contribution of each player to the outcome by examining all possible combinations of players (Michalak et al., 2013). For binary classification, SHAP values can be interpreted as the direct contribution of features to the outcome in either way.