

Hyperparameters tuning methods

Metodologías

Grid search

El Grid Search es una técnica común para ajustar hiperparámetros en modelos de aprendizaje automático. Consiste en definir un conjunto de valores para cada hiperparámetro que se desea ajustar y luego entrenar el modelo con cada combinación posible de valores. El modelo se evalúa mediante alguna métrica de rendimiento, como la precisión o el error cuadrático medio, y se selecciona la combinación de hiperparámetros que mejor se desempeña en los datos de validación.

Existen dos tipos principales de Grid Search:

1. **Exhaustive Search Grid Search:** Este método prueba todas las combinaciones posibles de hiperparámetros dentro del rango especificado. Por lo tanto, se considera exhaustivo y garantiza encontrar la mejor combinación de hiperparámetros dentro del rango dado. Sin embargo, este enfoque puede ser muy costoso computacionalmente y puede llevar mucho tiempo si hay muchos hiperparámetros o si cada hiperparámetro tiene un rango amplio.
2. **Random Grid Search:** Este método es similar al Grid Search Exhaustivo, pero en lugar de probar todas las combinaciones posibles de hiperparámetros, se selecciona un subconjunto aleatorio de combinaciones. Esto puede ser mucho más eficiente en términos de tiempo de cómputo, especialmente cuando hay muchos hiperparámetros. A menudo se encuentra que un subconjunto relativamente pequeño de combinaciones de hiperparámetros es suficiente para encontrar una buena configuración.

Optimización bayesiana (usualmente más usado al ser menos costoso computacionalmente y con mejores resultados en ciertas ocasiones)

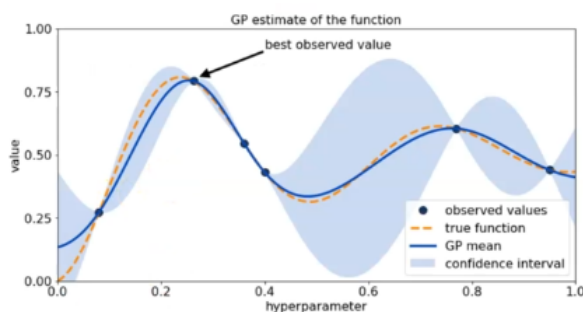
El método de optimización bayesiana es una técnica avanzada para ajustar hiperparámetros en modelos de aprendizaje automático. A diferencia del Grid Search y Random Search, que exploran sistemáticamente el espacio de hiperparámetros, la optimización bayesiana utiliza una estrategia de búsqueda más inteligente y basada en modelos para encontrar la combinación óptima de hiperparámetros.

La optimización bayesiana comienza con una función de pérdida (o función objetivo) que mide el rendimiento del modelo en los datos de validación en función de los hiperparámetros. A medida que se exploran diferentes combinaciones de hiperparámetros, el objetivo es encontrar la combinación que minimiza esta función de pérdida.

En lugar de probar todas las combinaciones posibles de hiperparámetros, la optimización bayesiana utiliza un modelo probabilístico para modelar la función objetivo y determinar qué combinación de hiperparámetros probar a continuación. El modelo probabilístico se actualiza a medida que se prueban nuevas combinaciones de hiperparámetros, lo que permite que la optimización bayesiana sea altamente eficiente y se adapte mejor a la función objetivo.

El proceso de optimización bayesiana es iterativo. En cada iteración, el modelo probabilístico se ajusta a los datos observados y se utiliza para seleccionar una nueva combinación de hiperparámetros para evaluar. Después de entrenar y evaluar el modelo con la nueva combinación de hiperparámetros, se observa el rendimiento en los datos de validación y se actualiza el modelo probabilístico con esta información.

Algoritmo de optimización



1. Definir una función objetivo $f(x)$ que se desea minimizar. En el contexto de ajuste de hiperparámetros, esta función objetivo es la métrica de evaluación, como la precisión o el error cuadrático medio, del modelo entrenado con un conjunto específico de hiperparámetros.
2. Seleccionar una distribución de probabilidad a priori $P(x)$ que represente la incertidumbre inicial sobre los valores de los hiperparámetros. Esta distribución puede ser uniforme o basada en información previa.
3. Actualizar la distribución de probabilidad después de cada evaluación del modelo con una distribución de probabilidad a posteriori $P(x|D)$, donde D representa los resultados de la evaluación. La distribución a posteriori se actualiza mediante el teorema de Bayes: $P(x|D) \propto P(D|x)P(x)$
4. Seleccionar la siguiente combinación de hiperparámetros x para evaluar utilizando una estrategia de selección de adquisición, como la incertidumbre esperada (Expected Improvement) o la entropía gaussiana (Gaussian entropy). La estrategia de selección de adquisición equilibra la exploración (probar valores de hiperparámetros con mayor incertidumbre) y la explotación (probar valores de hiperparámetros que se esperan que mejoren la función objetivo).
5. Evaluar el modelo con la combinación de hiperparámetros seleccionada en el paso 4 y registrar los resultados.
6. Repetir los pasos 3-5 hasta que se alcance un criterio de parada, como un número máximo de evaluaciones del modelo o una convergencia satisfactoria.

**Expected improvement*

El Expected Improvement (EI) es una medida utilizada en el método de optimización bayesiana para seleccionar la siguiente combinación de hiperparámetros a evaluar. El se define como la cantidad esperada de mejora en la función objetivo con respecto a la mejor observación hasta el momento. Matemáticamente, se define como:

$$EI(x) = E[\max(f(x^*) - f(x), 0)]$$

donde $f(x^*)$ es el valor de la mejor muestra hasta el momento, x^* es la locación de la mejor muestra.