



Hiperparámetros Optimización

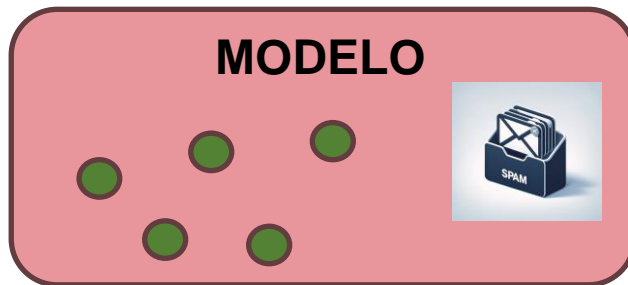


Modelos: Parâmetros e hiperparâmetros



Modelos: Parámetros e hiperparámetros

PARAMETROS

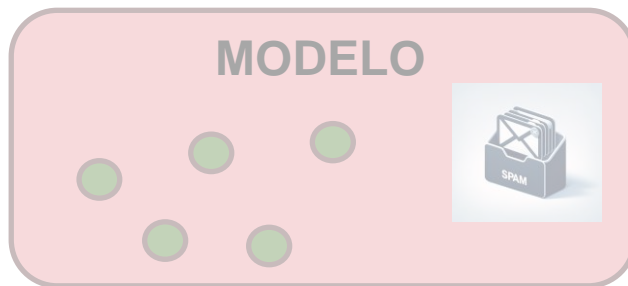


Un **parámetro** en un modelo de Machine Learning es una variable interna **cuyo valor es aprendido** a partir de los datos durante el entrenamiento del modelo. Estos parámetros son la esencia del modelo, ya que definen la relación entre los datos de entrada y la predicción.



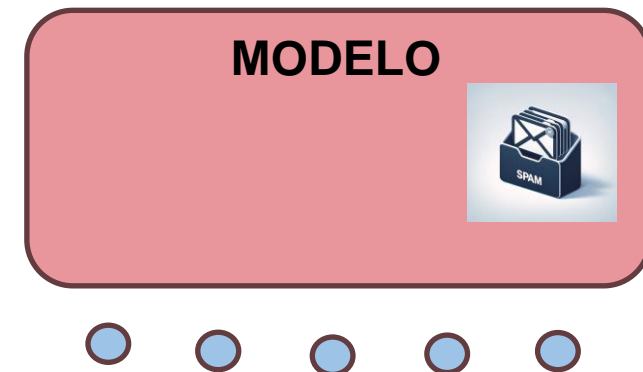
Modelos: Parámetros e hiperparámetros

PARAMETROS



Un **parámetro** en un modelo de Machine Learning es una variable interna cuyo valor es aprendido a partir de los datos durante el entrenamiento del modelo. Estos parámetros son la esencia del modelo, ya que definen la relación entre los datos de entrada y la predicción.

HIPERPARAMETROS



Un **hiperparámetro**, en cambio, es una configuración externa del modelo que no se aprende a partir de los datos, sino que **se establece antes del proceso de entrenamiento.**

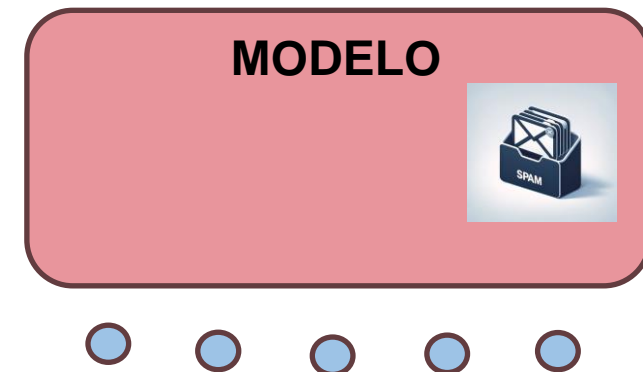


Modelos: Hiperparámetros

“Genéricos”:

- Asociados a regularización por “castigo” de la función de pérdida (Alpha, l1_ratio, etc)
- Asociados a utilizar algoritmos de optimización de la función de pérdida (learning_rate, tolerance, etc)

HIPERPARAMETROS



Modelos: Hiperparámetros

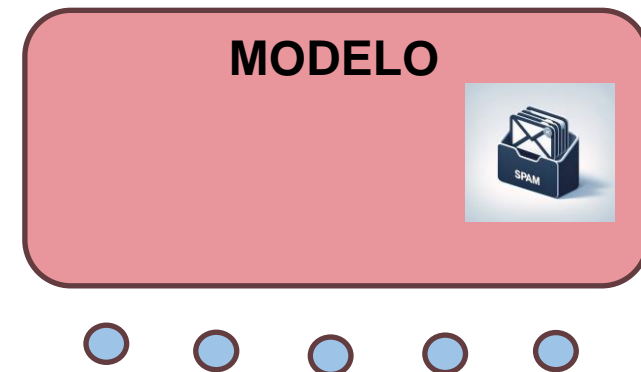
“Genéricos”:

- Asociados a regularización por “castigo” de la función de pérdida (Alpha, l1_ratio, etc)
- Asociados a utilizar algoritmos de optimización de la función de pérdida (learning_rate, tolerance, etc)

“Específicos”:

- Árboles: Profundidad, número mínimo de muestras por hoja, etc

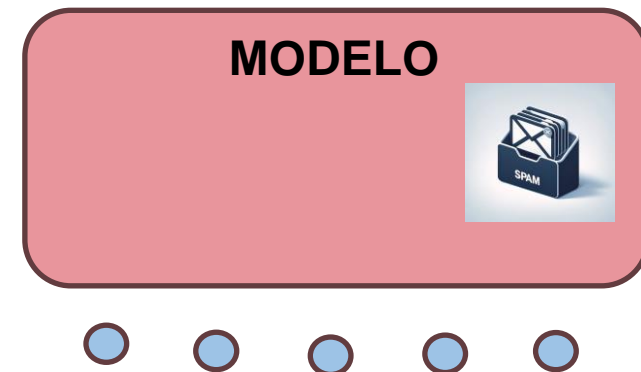
HIPERPARAMETROS



Hiperparámetros: ¿Cómo escogerlos?

- ¿Fuerza bruta? Todas las posibles combinaciones...

HIPERPARAMETROS



Hiperparámetros: ¿Cómo escogerlos?

- ¿Fuerza bruta? Todas las posibles combinaciones...

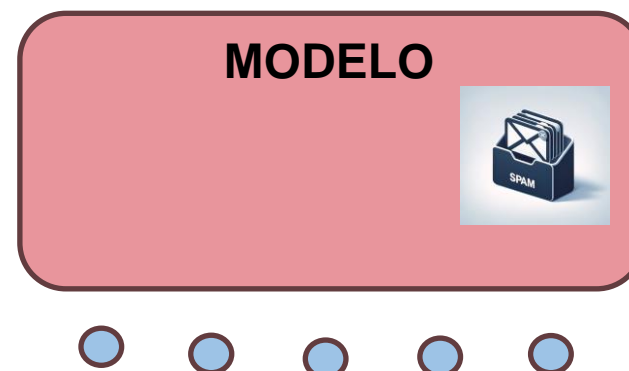
Arboles:

Max_Depth, probar valores de 1 a 10

Min_leaf_samples probar de 1 a 100

Tendría que probar 1000 combinaciones distintas... No es mucho... Salvo que el modelo tarde por ejemplo 1 hora en entrenarse (no es raro)... 1000 horas para encontrar el modelo y si tengo posibles clases con 1000 muestras mínimo,...10000 horas hmm,

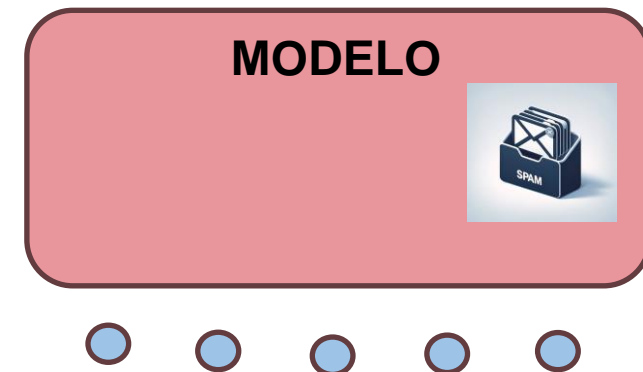
HIPERPARAMETROS



Hiperparámetros: ¿Cómo escogerlos?

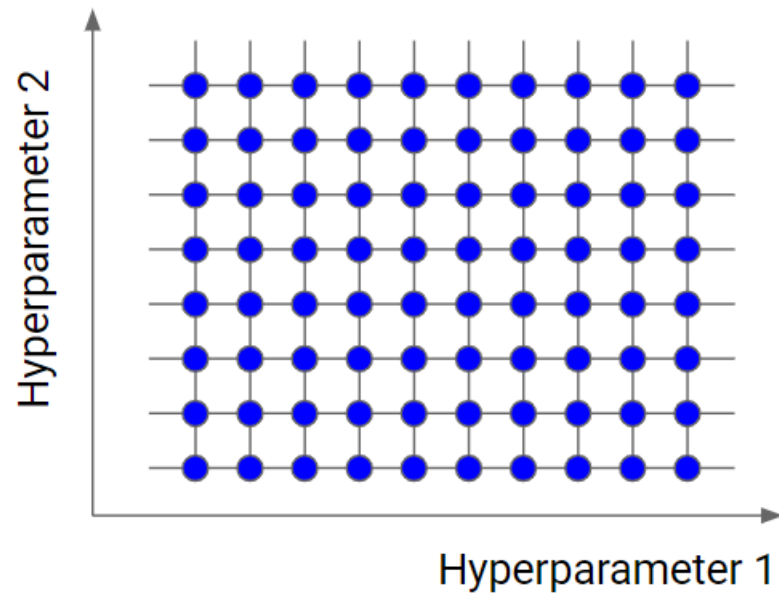
- No hacemos fuerza bruta:
 1. En general, seleccionamos un conjunto de hiperparámetros
Luego un rango de valores de los mismos
 2. A partir de ahí hacemos las combinaciones. Vamos a ver tres formas
 - Grid Search
 - Random Search
 - Bayesian Search

HIPERPARAMETROS



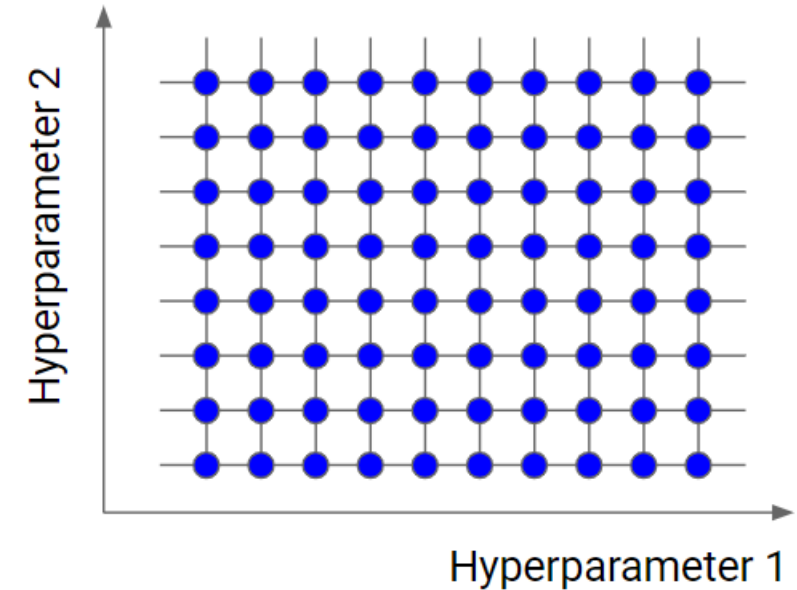
Hiperparámetros: Grid Search

- Supongamos que hemos escogidos dos hiperparámetros para optimizar



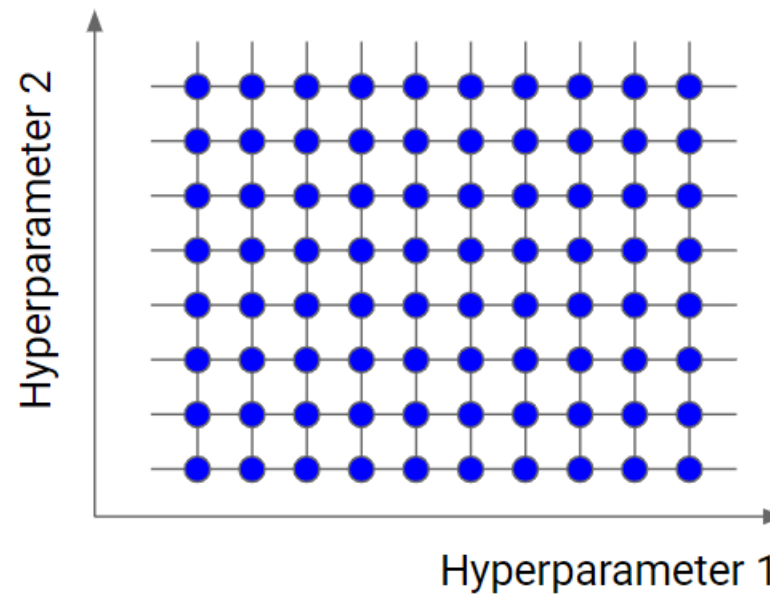
Hiperparámetros: Grid Search

- Supongamos que hemos escogidos dos hiperparámetros para optimizar
- Y un conjunto de rangos de valores para cada uno (representados por las líneas horizontales en el caso de hiperparámetro 1 y las verticales en el caso de hiperparámetro 2)



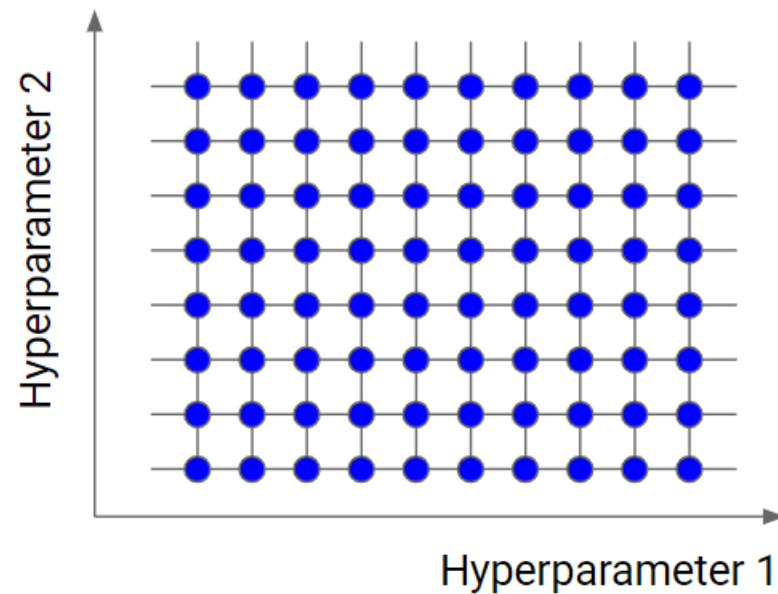
Hierparámetros: Grid Search

- Supongamos que hemos escogidos dos hiperparámetros para optimizar
- Y un conjunto de rangos de valores para cada uno (representados por las líneas horizontales en el caso de hiperparámetro 1 y las verticales en el caso de hiperparámetro 2). A esto se le denomina grid (aunque para más hiperparámetros es difícil visualizarlo)
- Grid Search prueba todas las combinaciones y escoge la que mejor resultado da



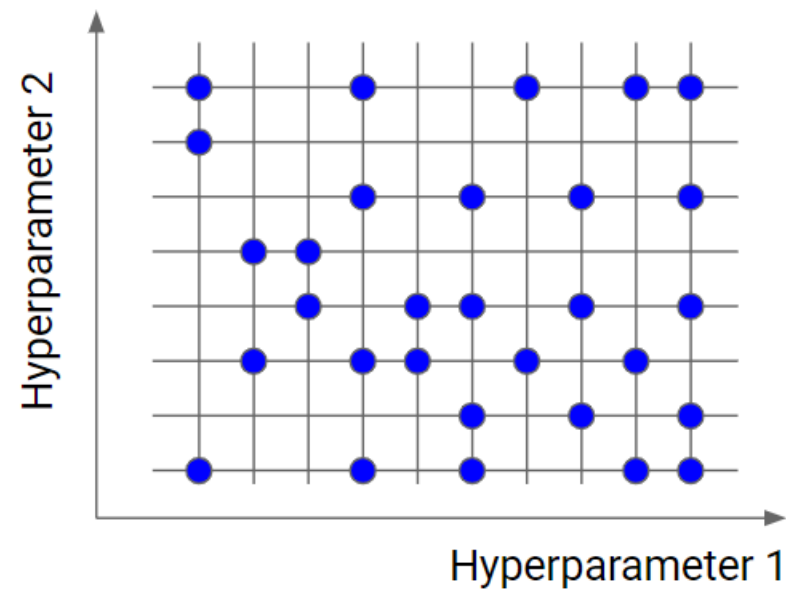
Hiperparámetros: Grid Search

- Supongamos que hemos escogidos dos hiperparámetros para optimizar
- Y un conjunto de rangos de valores para cada uno (representados por las líneas horizontales en el caso de hiperparámetro 1 y las verticales en el caso de hiperparámetro 2). A esto se le denomina grid (aunque para más hiperparámetros es difícil visualizarlo)
- Grid Search prueba todas las combinaciones y escoge la que mejor resultado da
- Lo usamos con pocos hiperparámetros y rangos muy acotados (es una fuerza bruta-especial)



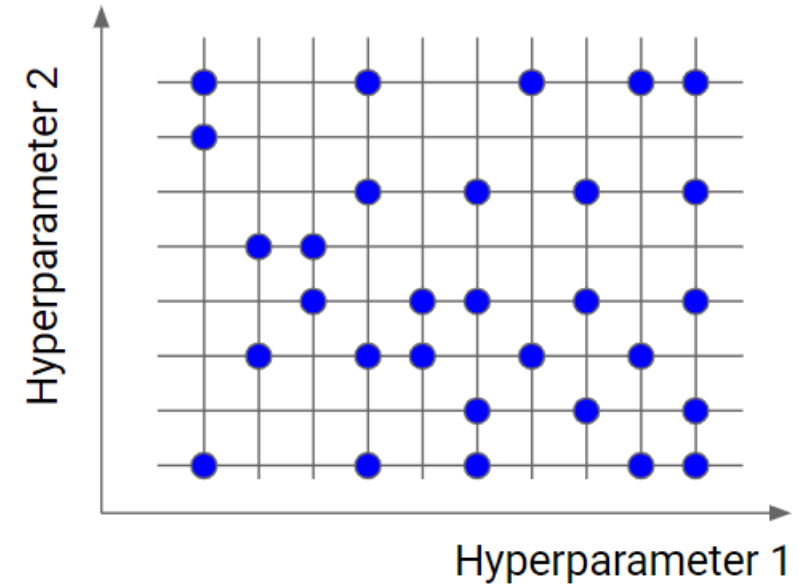
Hiperparámetros: Random Search

- Escogemos un grid (un conjunto de hiperparámetros y unos rangos)



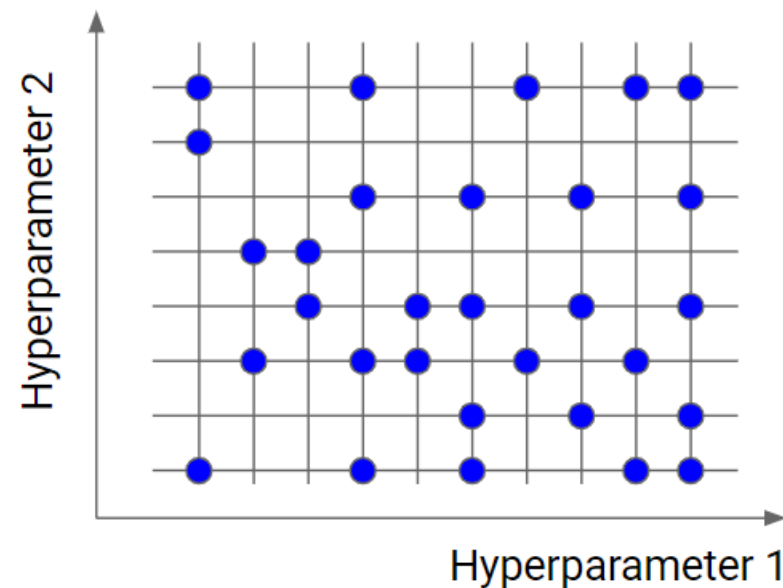
Hiperparámetros: Random Search

- Escogemos un grid (un conjunto de hiperparámetros y unos rangos)
- La diferencia es que ahora se seleccionan (es decir el método de sklearn, selecciona) combinaciones de forma aleatorias y no todas



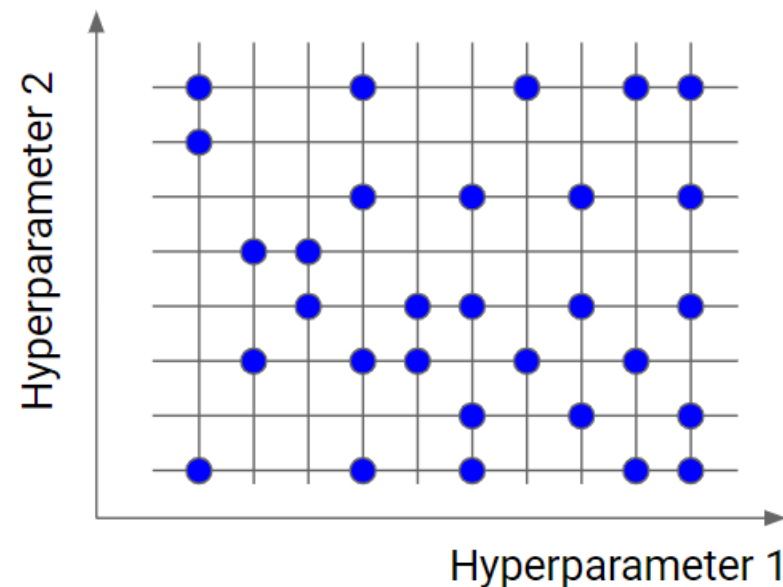
Hiperparámetros: Random Search

- Escogemos un grid (un conjunto de hiperparámetros y unos rangos)
- La diferencia es que ahora se seleccionan (es decir el método de sklearn, selecciona) combinaciones de forma aleatorias y no todas
- Para esas combinaciones se escoge la que genere el modelo con la mejor evaluación en la métrica de evaluación (otra vez medida empleando cross-validation)



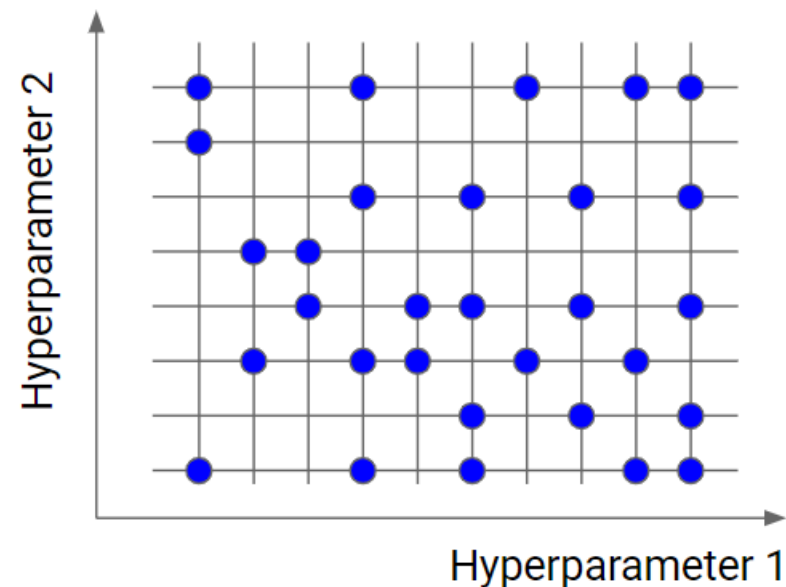
Hiperparámetros: Random Search

- Escogemos un grid (un conjunto de hiperparámetros y unos rangos)
- La diferencia es que ahora se seleccionan (es decir el método de sklearn, selecciona) combinaciones de forma aleatorias y no todas
- Para esas combinaciones se escoge la que genere el modelo con la mejor evaluación en la métrica de evaluación (otra vez medida empleando cross-validation)
- Es eficiente para muchos hiperparámetros y/o muchos posibles valores de estos



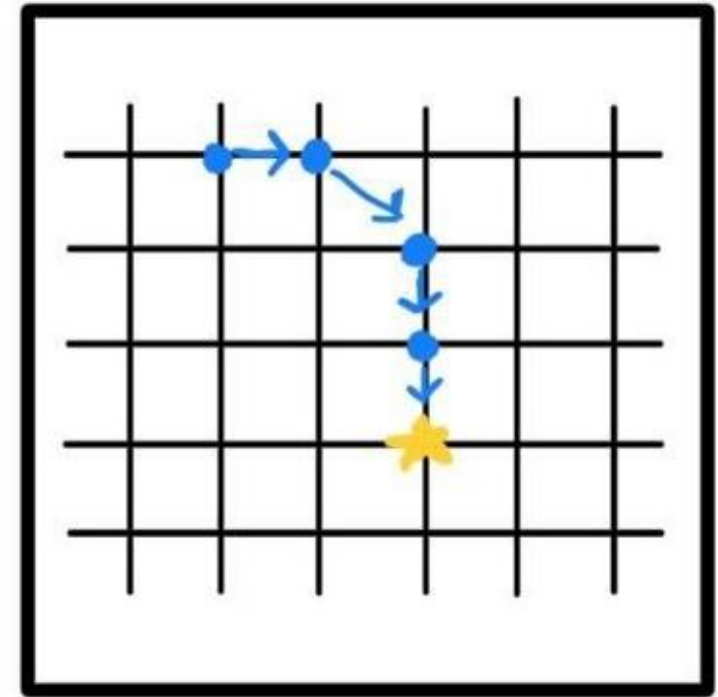
Hiperparámetros: Random Search

- Escogemos un grid (un conjunto de hiperparámetros y unos rangos)
- La diferencia es que ahora se seleccionan (es decir el método de sklearn, selecciona) combinaciones de forma aleatorias y no todas
- Para esas combinaciones se escoge la que genere el modelo con la mejor evaluación en la métrica de evaluación (otra vez medida empleando cross-validation)
- Es eficiente para muchos hiperparámetros y/o muchos posibles valores de estos
- Es ineficiente porque puede no llegar a probar con la combinación verdaderamente óptima



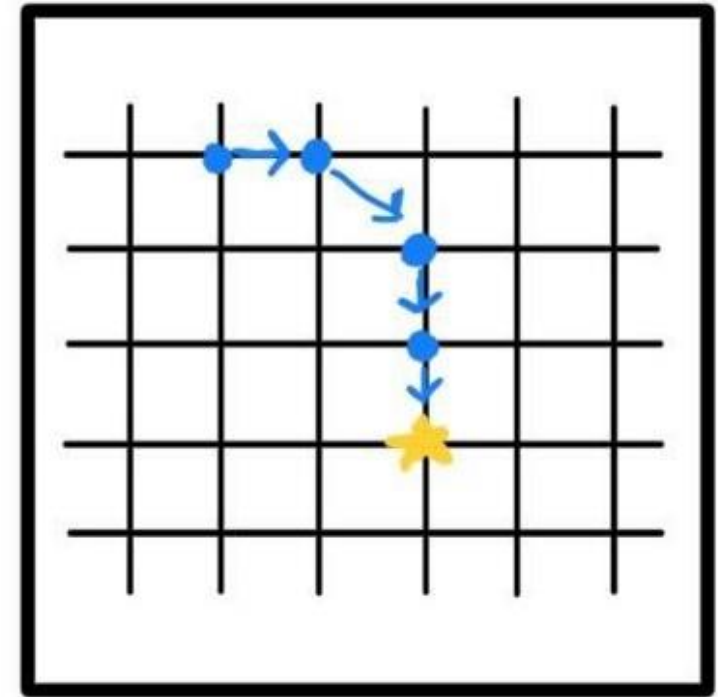
Hiperparámetros: Bayesian Search

- Escogemos un grid (un conjunto de hiperparámetros y unos rangos)



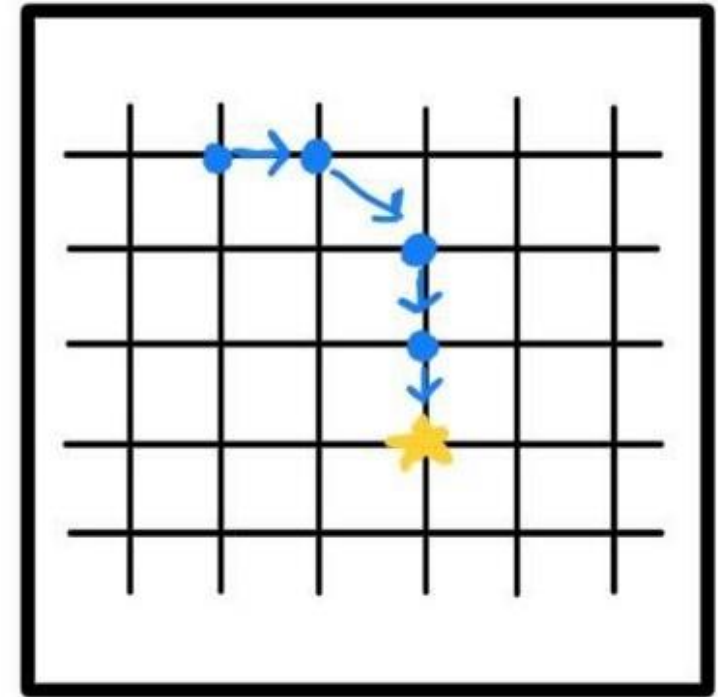
Hiperparámetros: Bayesian Search

- Escogemos un grid (un conjunto de hiperparámetros y unos rangos)
- La selección ahora no se hace ni completa (grid search), ni aleatoria sino que se utilizan métodos probabilísticos para (partiendo de una combinación) escoger la siguiente.



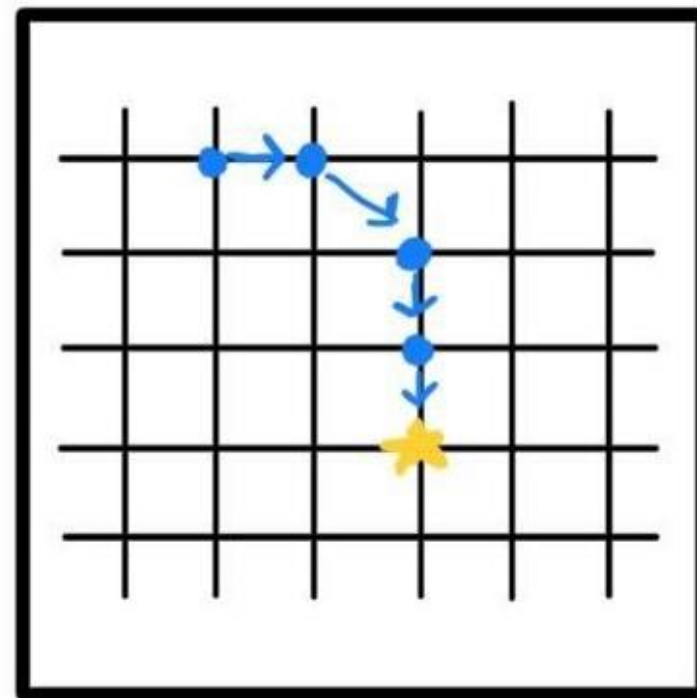
Hiperparámetros: Bayesian Search

- Escogemos un grid (un conjunto de hiperparámetros y unos rangos)
- La selección ahora no se hace ni completa (grid search), ni aleatoria sino que se utilizan métodos probabilísticos para (partiendo de una combinación) escoger la siguiente.
- Intenta aunar la ventaja de no probar todas las combinaciones mitigando la desventaja de que pueda no probar la mejor de las combinaciones

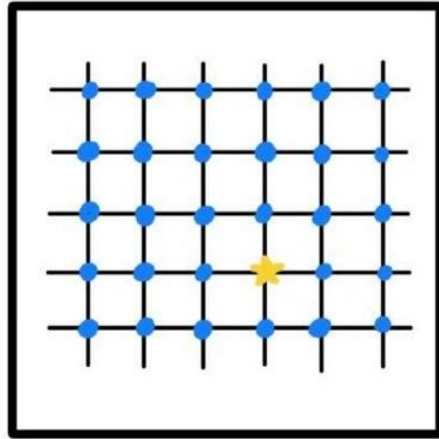


Hiperparámetros: Bayesian Search

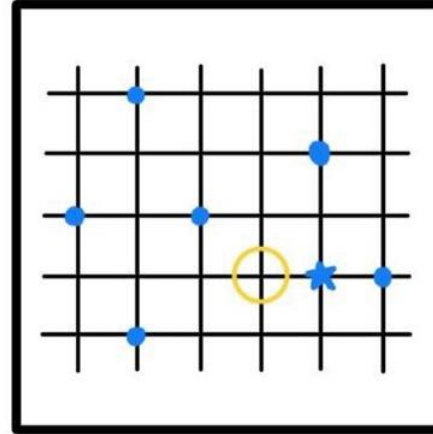
- Escogemos un grid (un conjunto de hiperparámetros y unos rangos)
- La selección ahora no se hace ni completa (grid search), ni aleatoria sino que se utilizan métodos probabilísticos para (partiendo de una combinación) escoger la siguiente.
- Intenta aunar la ventaja de no probar todas las combinaciones mitigando la desventaja de que pueda no probar la mejor de las combinaciones
- Cuando quiero probar varios hiperparámetros con muchos valores o muchos hiperparámetros con pocos valores o en sustitución de la random



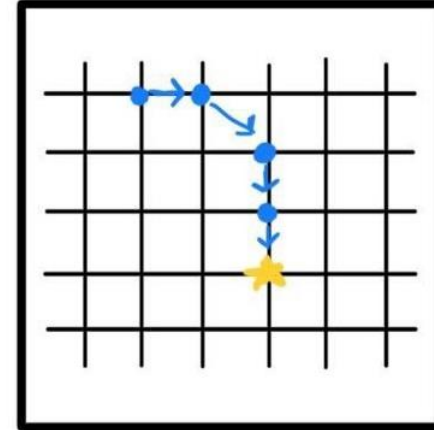
Hiperparámetros: Comparación



Grid Search



Random Search



Bayesian Optimization

- Evaluation points
- ★ Optimal parameters
- ★ Local optimal parameters



Hiperparámetros: Regresión Lineal

- `'fit_intercept'`
- `'normalize'`
- `'n_jobs'`



Hiperparámetros: Regresión Lineal

- ``fit_intercept``
- ``normalize``
- ``n_jobs``
-
- Ridge -> alpha
- Lasso -> alpha
- ElasticNet -> alpha y l1_ratio



Hiperparámetros: Regresión Logística

- Penalty
- C
- L1_ratio



Hiperparámetros: Regresión Logística

- Penalty
- C
- L1_ratio
- max_iter
- tol



Hiperparámetros: Regresión Logística

- Penalty
- C
- L1_ratio
- max_iter
- tol
- class_weight



Hiperparámetros: Árboles decisión



Hiperparámetros: Árboles decisión

- Max_Depth
- Min_samples_leaf
- Min_samples_split
- Max_Features
- Class_Weight



