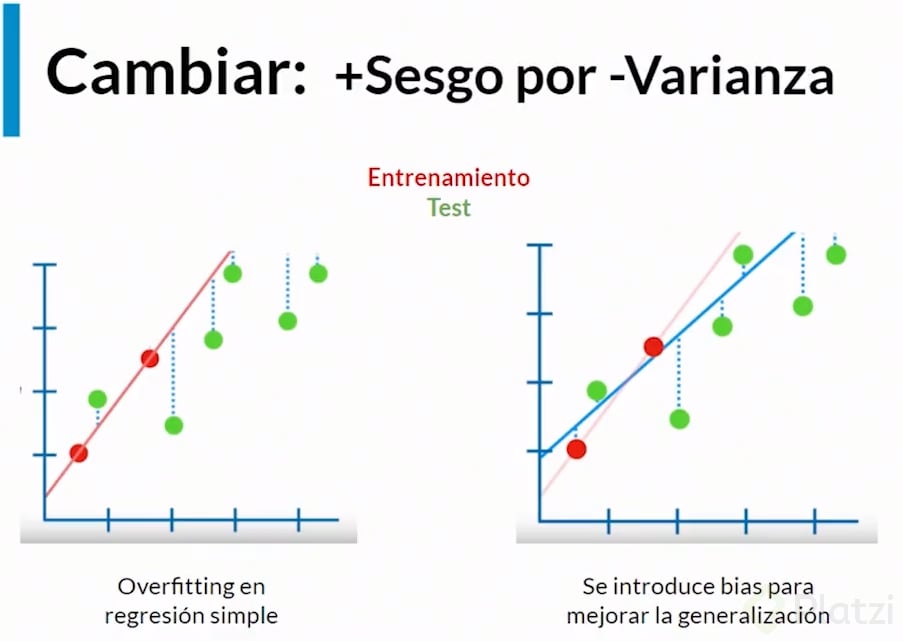
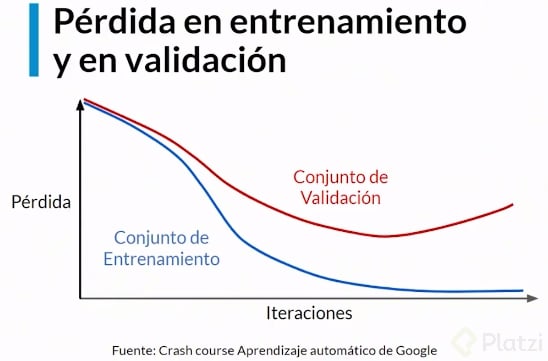
<h1>¿Qué es la regularización y cómo aplicarla?</h1>

Esta técnica consiste en disminuir la complejidad de nuestro modelo a través de una penalización aplicada a sus variables más irrelevantes.



Como podemos apreciar en la gráfica 1, hay un sub ajuste, ya que la linea roja se acopla muy bien para los datos de prueba, pero no para los datos de entrenamiento. La linea roja en los datos de prueba da una mala generalización, una mala aproximación.  
Entonces, la regularización consiste en introducir un poco de sesgo para introducir la varianza de los datos.

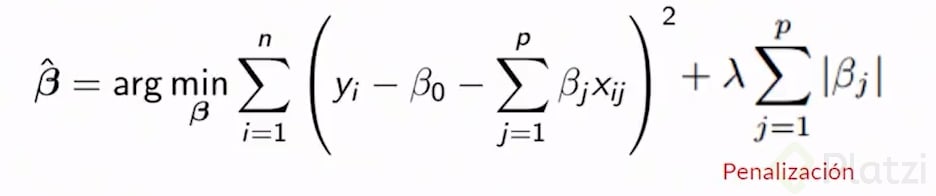
Pero para poder aplicar regularización necesitamos un termino adicional el concepto de perdida. El concepto de perdida nos dice que tan lejos están nuestras predicciones de los datos reales, esto quiere decir que entre menor sea la perdida mejor será nuestro modelo.



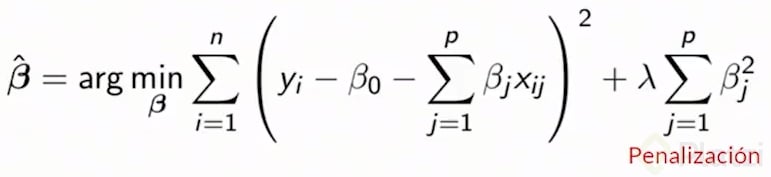
Como podemos ver en la gráfica que la perdida tiende a disminuir, porque en algún momento van a ser vistos, van a ser operados y el modelo va a tender a ajustarse a esos datos de entrenamiento, pero lo que tenemos que mirar es cómo se va a comportar en el mundo real. En el conjunto de validación o pruebas es muy normal que nuestra perdida comience a disminuir porque hay una buena generalización, pero llega un punto donde nuevos valores comienza a introducirse donde esa perdida vuelve a comenzar a subir ese es el punto donde en general se considera que comienza a haber sobreajuste. Es la perdida la medida que vamos a utilizar para poder utilizar la regularización.

**¿Cuántos tipos de regularización existen?**

* **L1 Lasso:** Reducir la complejidad a través de eliminación de features que no aportan demasiado al modelo.  
  Penaliza a los features que aporta poca información volviéndolos cero, eliminado el ruido que producen en el modelo.



* **L2 Ridge:** Reducir la complejidad disminuyendo el impacto de ciertos features a nuestro modelo.  
  Penaliza los features poco relevantes, pero no los vuelve cero. Solamente limita la información que aportan a nuestro modelo.



* **ElasticNet:** Es una combinación de las dos anteriores.

**Lasso vs Ridge.**

1.- No hay un campeón definitivo para todos los problemas.  
2.- Si hay pocos features que se relacionen directamente con la variable a predecir: Probar Lasso.  
3.- Si hay varios features relacionados con la variable a predecir: Probar Ridge.