Teoria de regularizacion.

Una de las caracteristicas mas importantes de nuestro modelo es que este pueda funcionar con datos nuevos por lo tanto que generalicen o que no se sobreajusten a los datos de entrenamiento. Regularizar es el conjunto de técnicas que nos ayuda a eso.

Overfiting o sobreajuste:

Ocurre cuando un modelo de aprendizaje automatico se entrena demasiado bien con los datos de entrenamiento (memorizándolos o algo por el estilo en vez de “aprender de ellos” )

Underfitting:

Lo contrario, sucede cuando un modelo es demasiado simple para capturar la estructura de los datos de entrenamiento, resultando en un rendimiento pobre de entrenamiento y de generalización, tendríamos un alto error porque no se ha aprendido nada de los datos.

Bias :

Error por simplificación excesiva, asumiendo relaciones mas simples de lo que realmente son y es lo que conduce al underfitting. En general se suele producir el error en cada una de las muestras y básicamente es un error de simplicidad. Existe un train error bastante alto.

Variance:

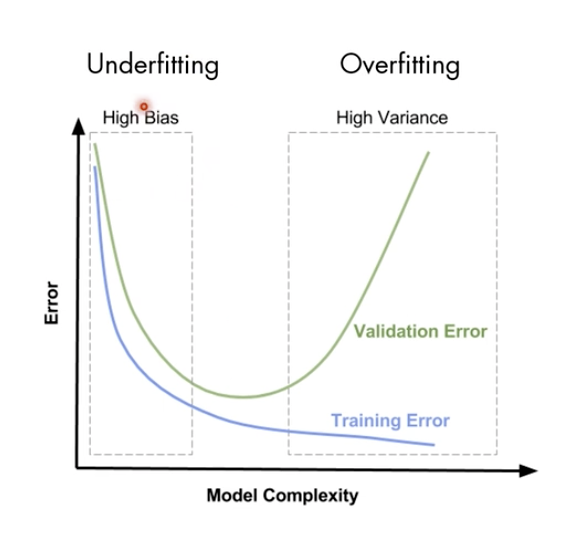
Diferencia entre el training error y el test error. ES LO QUE QUEREMOS REDUCIR. A fin de cuentas nos dice que no estamos generalizando bien. Un error que puede darse por complejidad excesiva, prestando demasiada antecion a los problemas de entrenamiento, y luego no generaliza nuestro modelo, cayendo en overfitting.

Trade-off Bias/Variance:

Compromiso entre Bias y varianza:

Es el equilibrio que tiene que haber entre los dos. En general lo que ocurre es que si quiero mejos varianza voy a tener que sacrificar el Bias.

Generalmente buscamos un modelo con bajo Bias y baja varianza que no sea excesivamente complejo, del grafico es el punto del medio.. donde no están los cuadrados.



En definitiva, el proceso de regularización es aquel que altera ligeramente la formulación matemática de un modelo, con la intención de prevenir el overfitting.

Una manera de regularizar puede ser **eliminando grados de una regresión polinómica o aplanando los pesos.**

**En concreto vamos a simplicar el modelo para que haya menos overfiting y generalicen mejor. A cambio sufriremos un pequeño aumento en los errores. De nuevo, Bias vs Varianza.**

**La regularización se suele aplicar en regresiones lineales y logísticas, también existen para otros modelos, pero serán distintas. Lo que se busca es perjudicar la función de perdida para que esta tenga mas error del que debería tener de manera que no se sobreajuste.**

**Las penalizaciones o perjudicaciones mas populares son L1 y L2:**

**L2** – Consigue que los parámetros estimados por el modelo (w) no tengan (en valor absoluto) un valor demasiado grande, de manera que aplana los pesos y evita los extremos. No vamos a encontrar ninguna feature con un valor demasiado alto o que sobresalga, sino va a estar mas a medida de las otras.

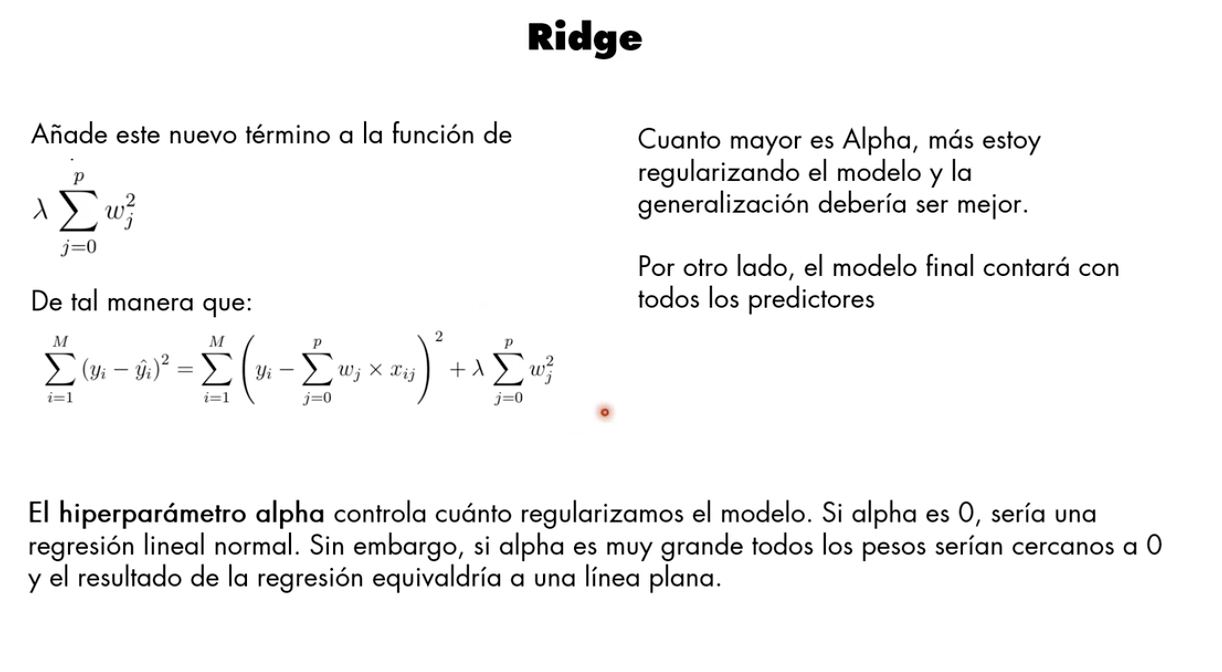
**L1 –** Tiende a eliminar los pesos con menor importancia, es como si estuviese realizando un feature selection.

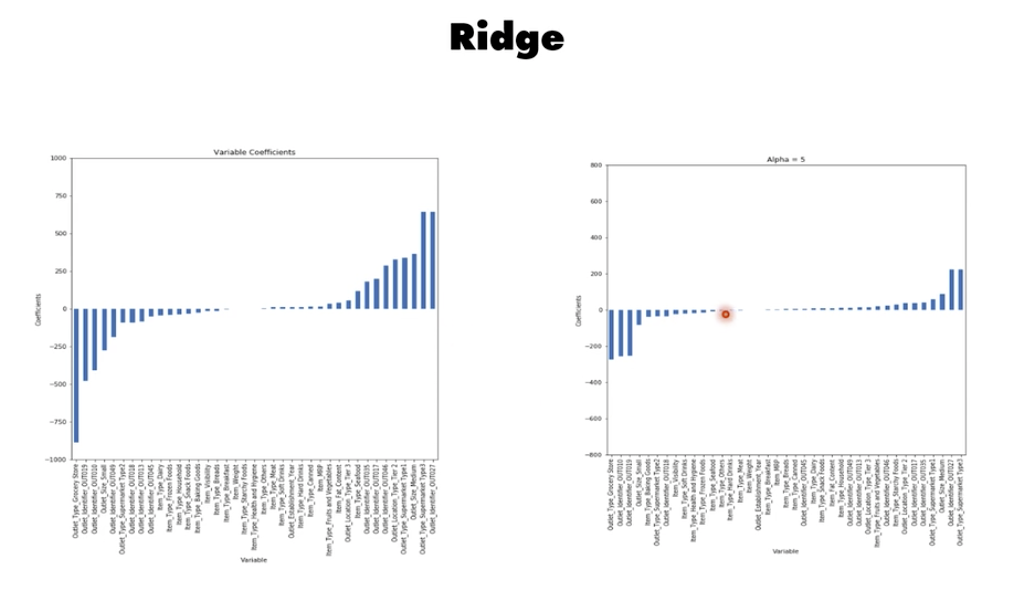
Regresiones que utilizan regularización:

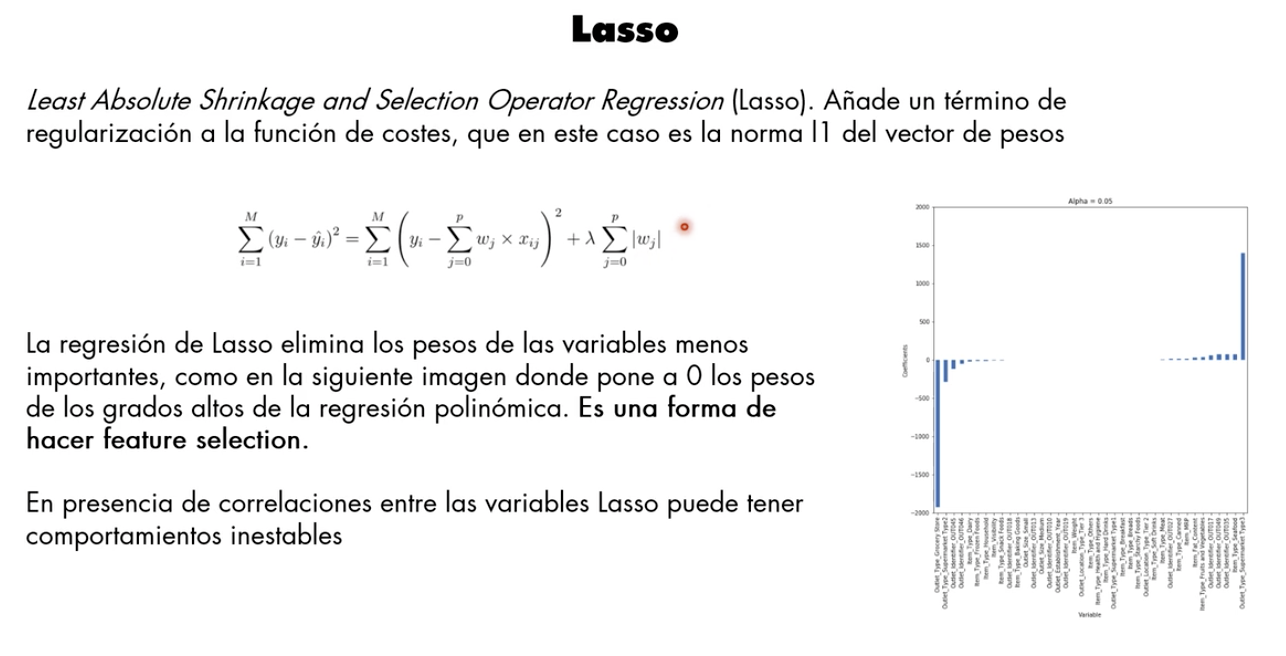
**Ridge:** Utiliza para penalizar la L2

**Lasso:** Utiliza para penalizar la L1

**Elastic Net:** Combinación lineal de L1 y L2.

****

****

****

