

# Analisis de resultados

Yeison Estiben Botero Rivera.

Enero 2025

Como vimos en el anterior seccion vimos la aplicacion que hicimos de dos problemas de la vida real y ahora analizaremos los resultados que arrojaron, limitaciones, posibles mejoras, y analizar el modelo en si.

Como se ve en la regresion lineal de ventas reales vs predichas para el primer problema a pesar de ser una buena aproximacion con los datos que habia sucede que en muchos puntos de las ventas predicha la distancia con las ventas reales es bastante grande lo cual genera un error relativo alto de puntos a puntos y por tanto podria generar problemas a la hora de usarlo en la empresa, esto se dio debido a la poca cantidad de datos que habia, por tanto para mejorar este modelo y la aproximacion correspondiente lo mejor seria que la empresa proporcionara una mayor cantidad de datos, esta misma situacion tambien es observable en el plano de regresion lineal multiple.

Para el segundo problema que hicimos que es el de la relacion de horas de estudio y calificacion en la primera grafica que vemos podemos notar que hay puntos que estan bien aproximados y otros que no estan tan bien aproximados respecto a la curva esta situacion se debe de nuevo a la poca cantidad de datos que tenemos pero aun asi parece que no esta tan mal la aproximacion, ahora analizemos la comparacion de la curva con la linea y la parabola, se hace evidente que la parabola es la mejor aproximacion que hay de los datos lo cual mejora enormemente lo hecho anteriormente, la lineal no parece darnos una mejor aproximacion que la curva, por tanto podemos ver que con los modelos polinomiales del metodo de minimos cuadrados al hacer una mejor aproximacion es preferible este metodo para una mayor eficacia y aunque son 3 coeficientes que hallar en vez de 2 en general no deberia incrementar tanto el coste computacional.

Un problema que tiene nuestro modelo de minimos cudrados es que solo sirven para nuestros problemas elegidos y por tanto una mejora que se podria considerar necesaria es que sea mas flexible para otros problemas aunque obviamente mientras mayor sea el numero de datos el coste computacional ira incrementando.