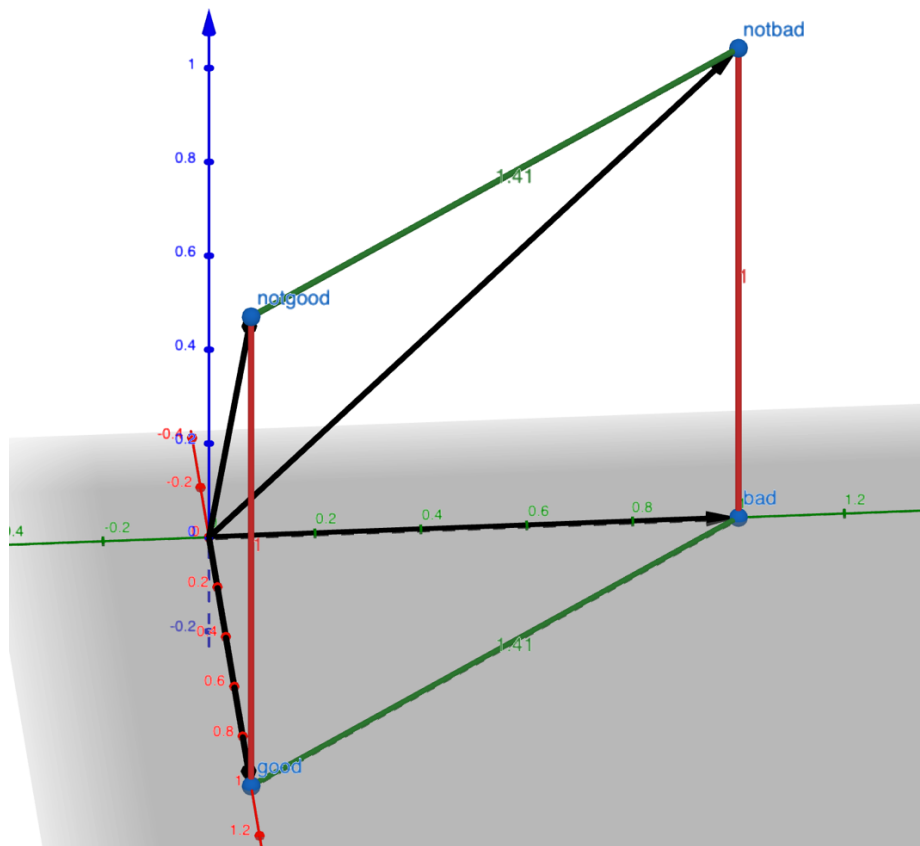


b. Dado el siguiente conjunto de datos de reseñas:

1. (-1) bad
2. (+1) good
3. (+1) not bad
4. (-1) not good

Muestra que no hay clasificador lineal que utilice características de palabras que tenga cero error sobre este conjunto de datos. Recuerda que esta es una pregunta sobre clasificadores, no sobre algoritmos de optimización; tu demostración debe ser correcta para cualquier clasificador lineal, sin importar en como se aprenden los pesos.

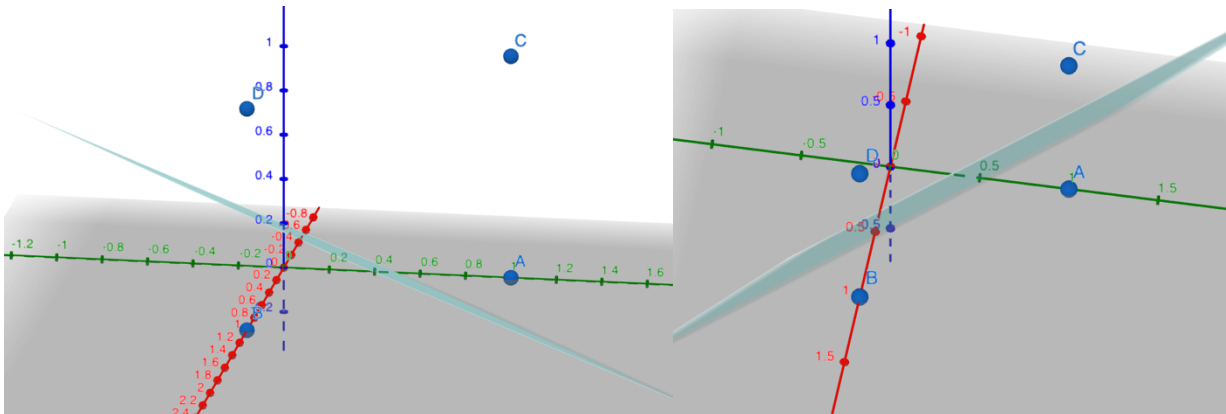
Si graficamos cualquier permutación ( $P_3 = 3! = 6$ ) de estos vectores podemos observar que existen ambigüedades en el etiquetado de las palabras como se muestra en la **img1**:



img1

donde en una de sus permutaciones quedó así: bad = [0,1,0], good = [1, 0, 0], not bad = [0, 1, 1] y not good = [1, 0, 1].

Semanticamente en el inglés 'not good' (No bueno) se parece más a 'bad' (malo) que a 'good' (bueno). Pero en la gráfica podemos ver que la distancia entre 'not good' y 'bad' ( $\sqrt{1.41^2 + 1^2} \approx 1.73$ ) es mayor que la distancia entre 'not good' y 'good' que vale 1.



Por lo cual los datos no pueden ser clasificados correctamente utilizando unicamente un hiperplano.

**Propón una sola característica adicional para tu conjunto de datos con la que pudieramos arreglar este problema.**

Para solucionar este problema lo que podemos hacer es agregar una característica binaria para saber diferenciar cuando se trata de una negación.

**Ej. Bad, good, not, not good.**

Bad = [1, 0, 0, 0]

good = [0, 1, 0, 0]

not bad = [1, 0, 1, 0]

not good = [0, 1, 1, 0]