



I N S E A









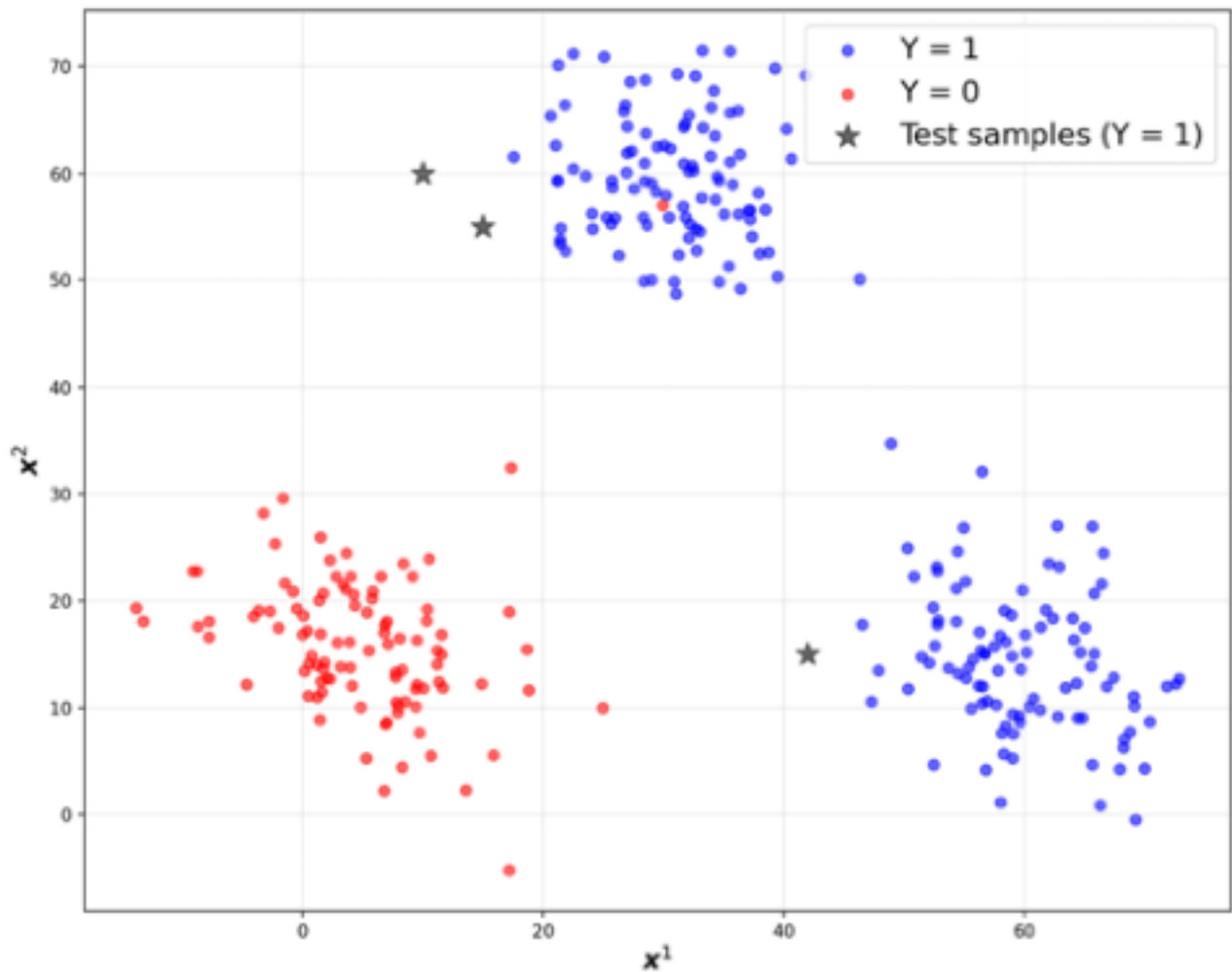
Quelle meilleure séparation linéaire sur ces données?

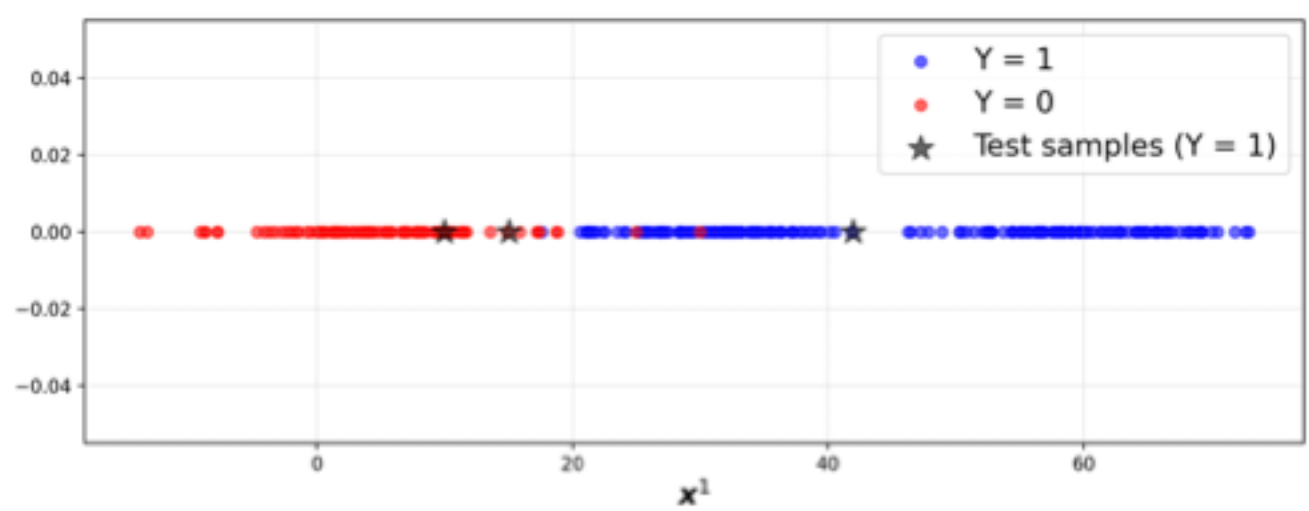
Calculus' is a harder train de test.

$d=2$ and n are independent d test $=0$

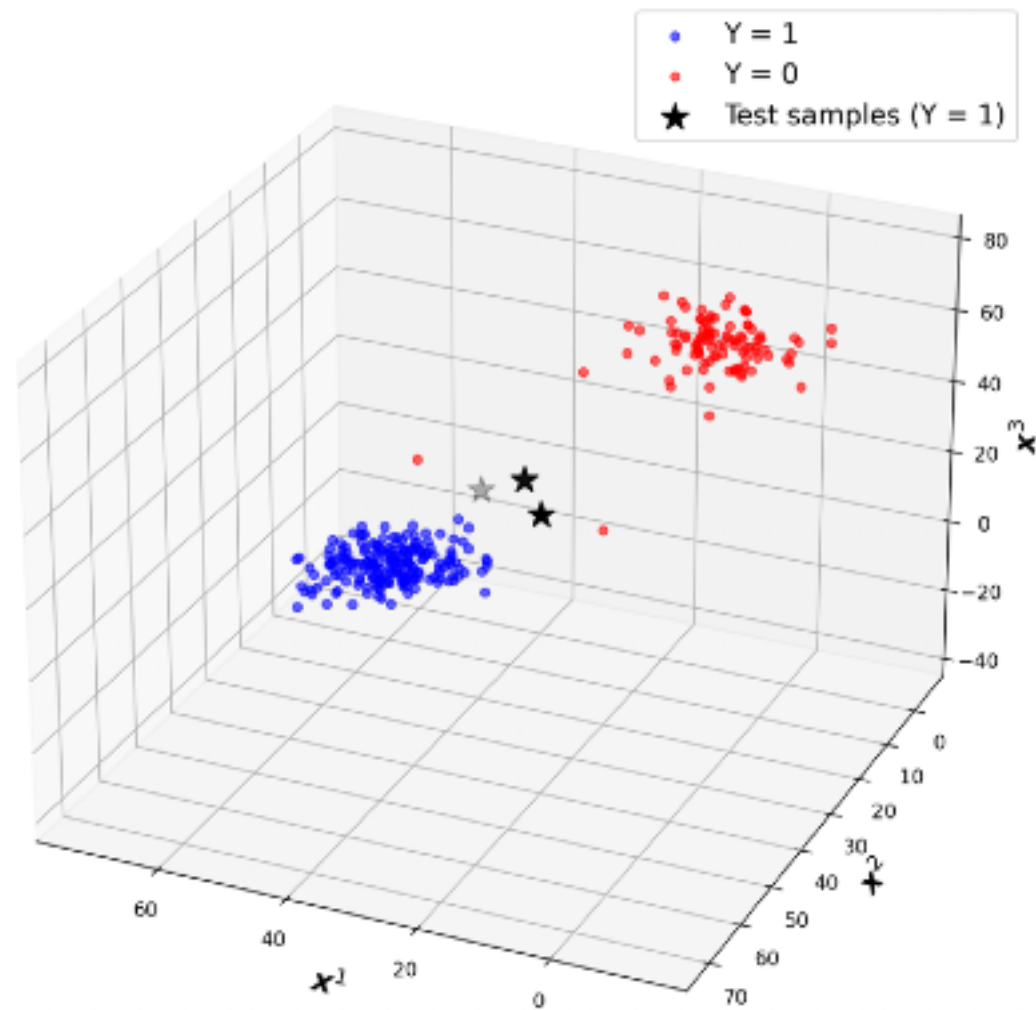
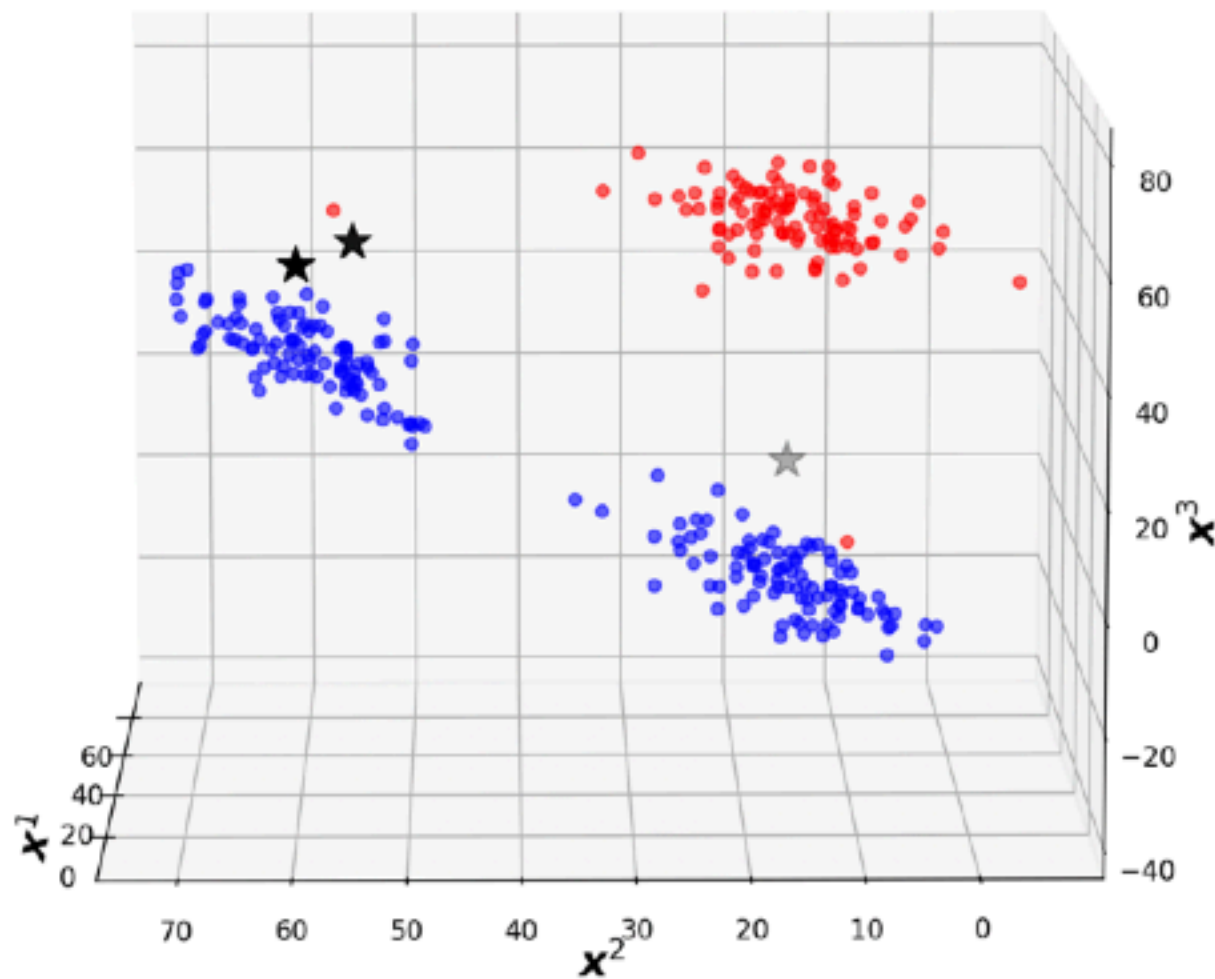
La méthode d'entraînement linéaire est biaisée par la méthode d'entraînement: elle est biaisée par la méthode d'entraînement.

$d = 2$

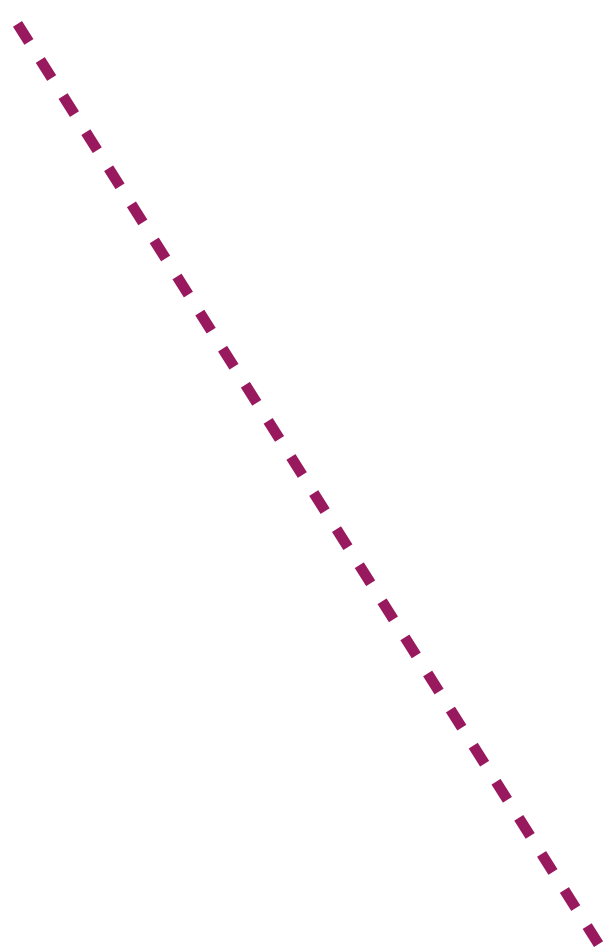


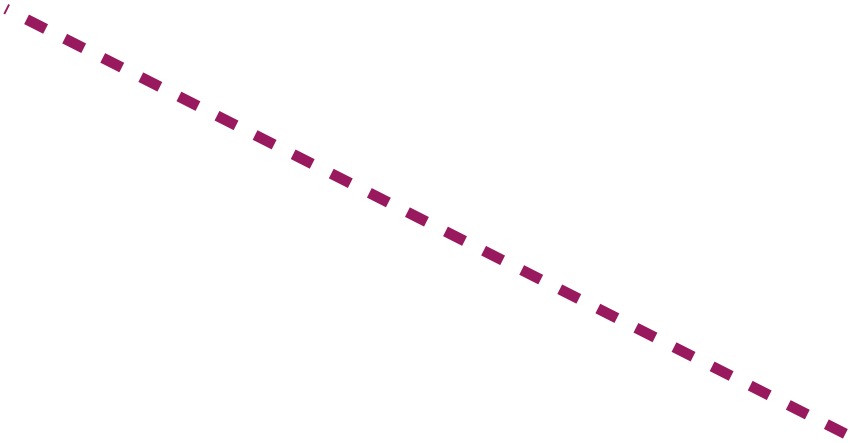


$d = 3$









$d = 3$ don't need a neural network $= 0$

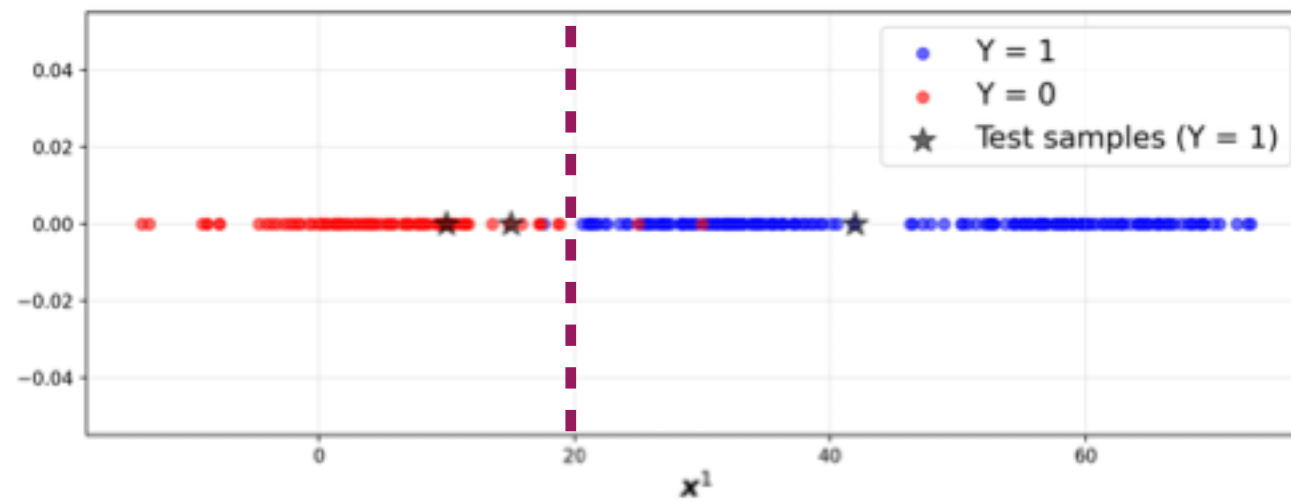
Test error makes a U shape

Machine learning classic: zero-to-hero

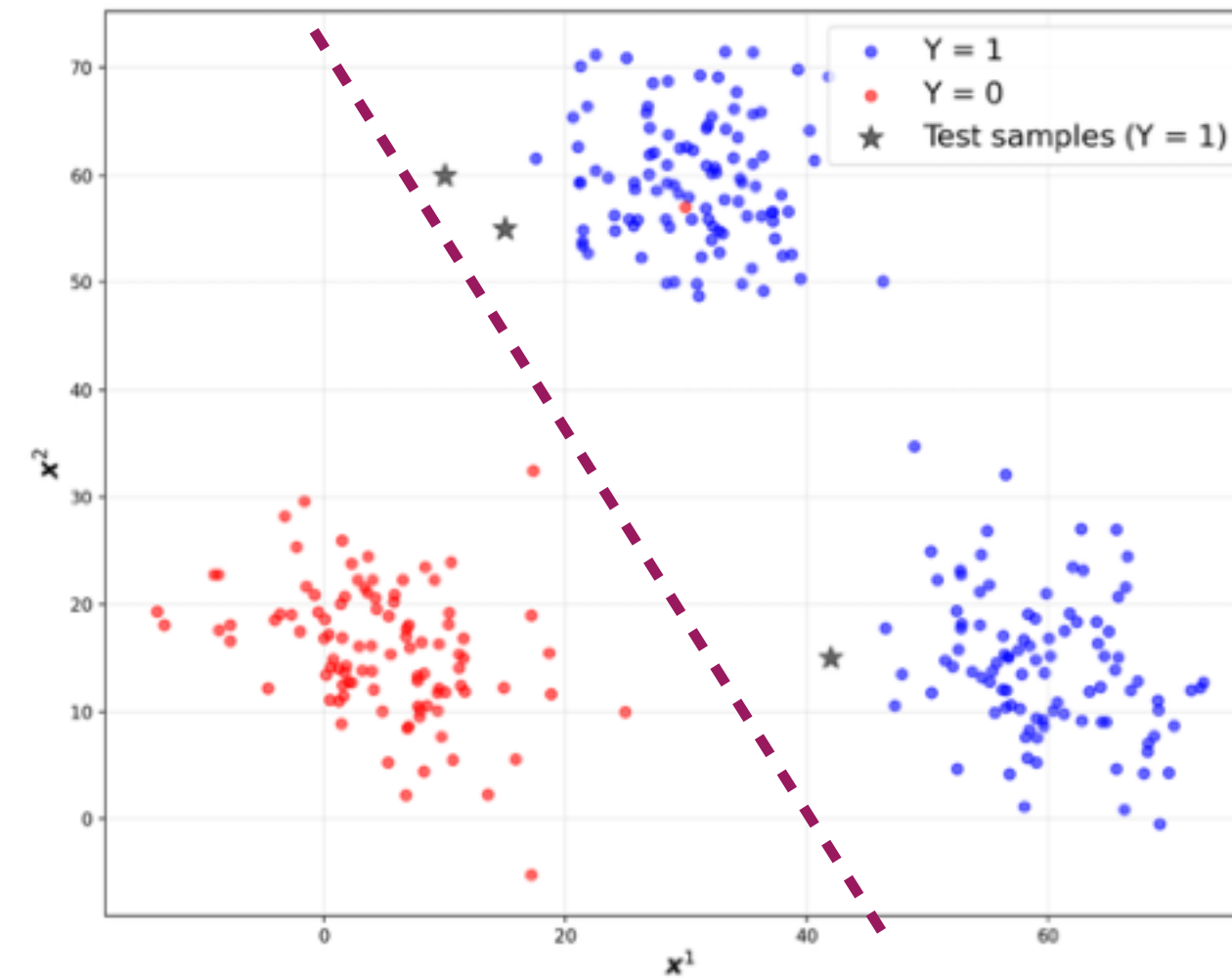
Undergraduate dimension put a user's verification

Quelle est la meilleure séparation linéaire sur ces données ?

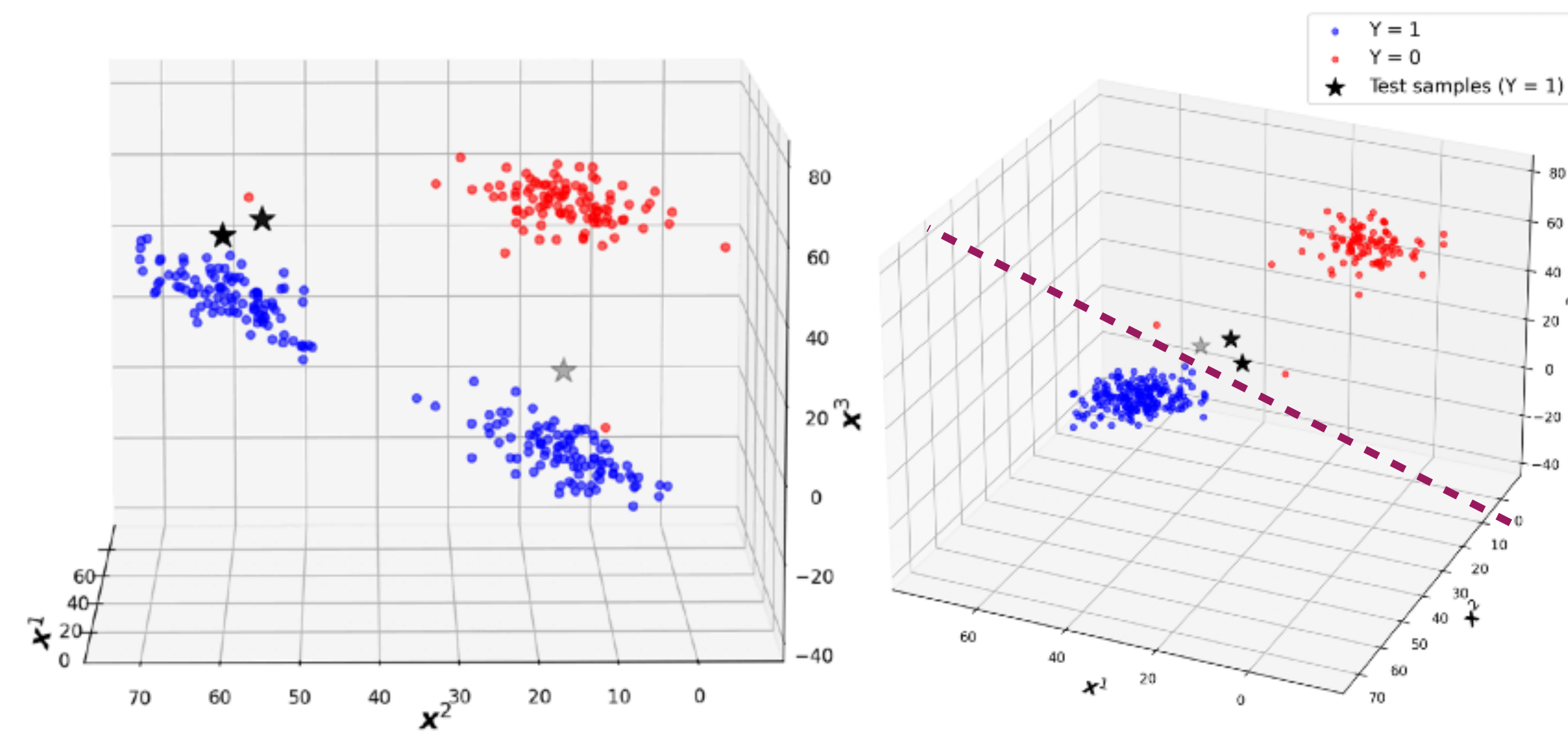
$d = 1$



$d = 2$



$d = 3$



Calculer l'erreur de train et de test.

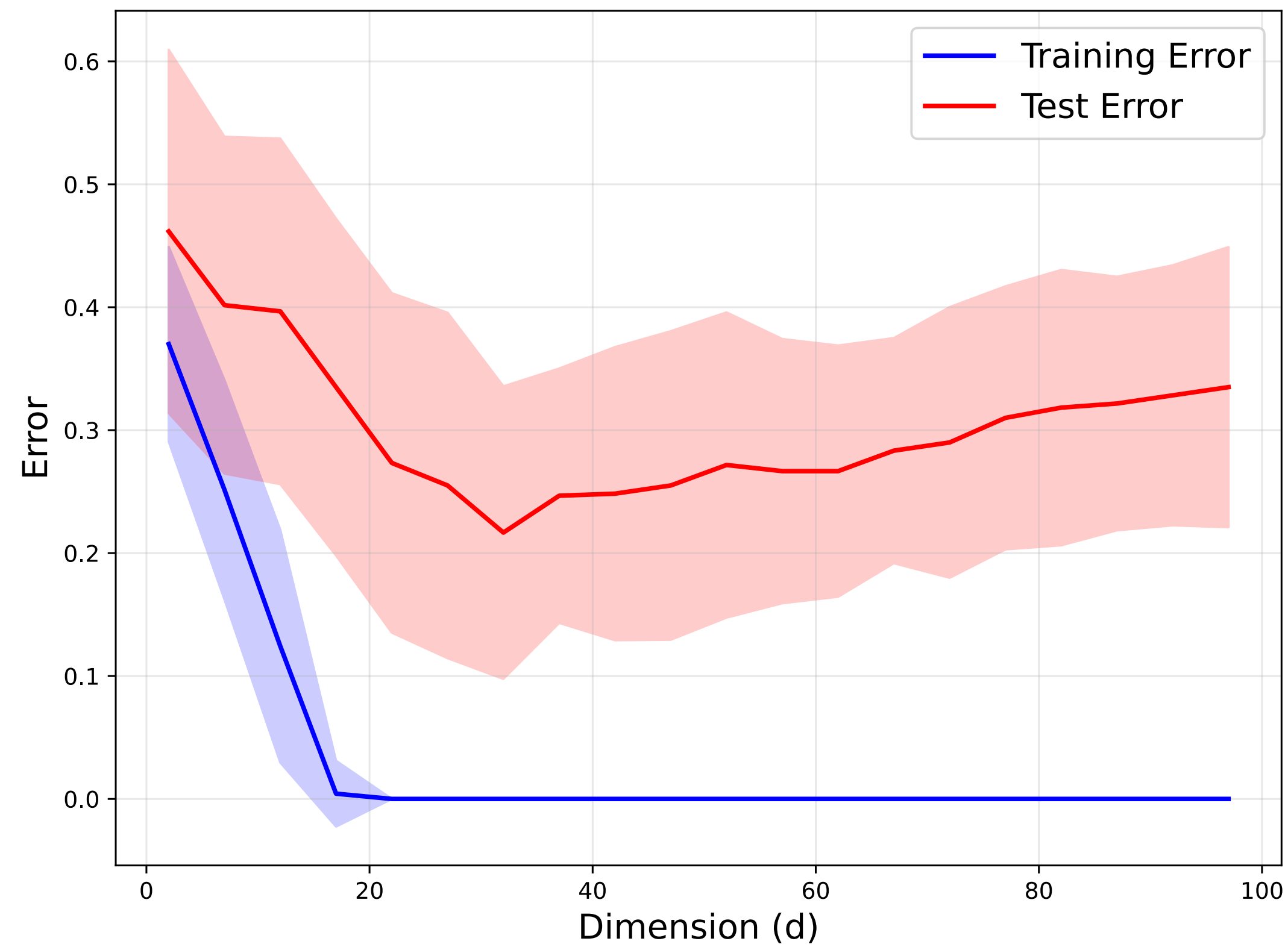
$d = 3$ donne la meilleure erreur de train = 0

$d = 2$ donne la meilleure erreur de test = 0

“La meilleure” séparation linéaire sur le train n’est pas la meilleure sur le test: elle est biaisée par les outliers

Une grande dimension peut causer l'**overfitting**





Variance nulle = prédiction constante
= underfitting