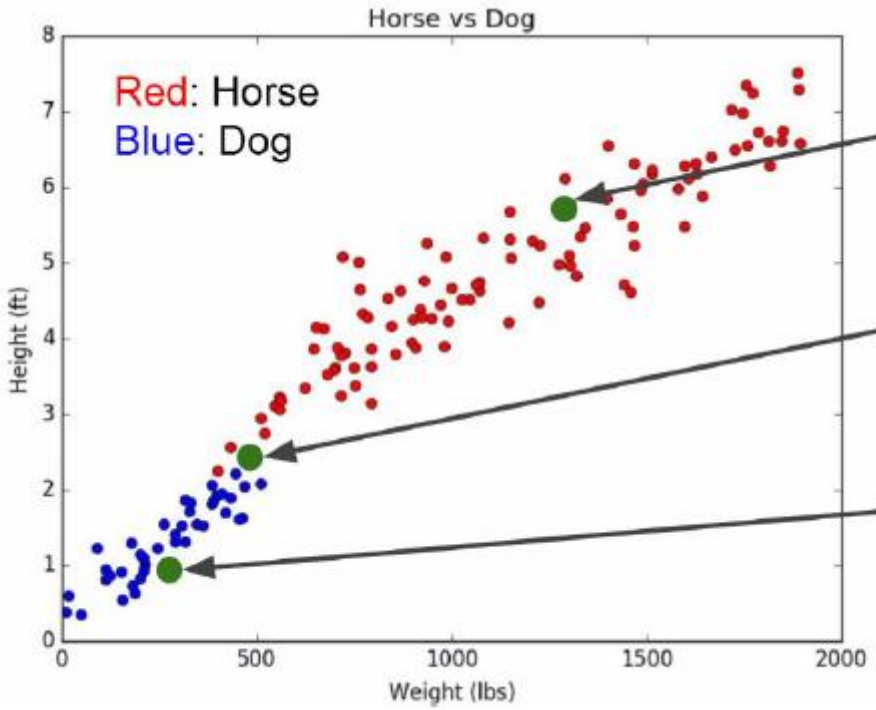
Introduction to K Nearest Neighbors

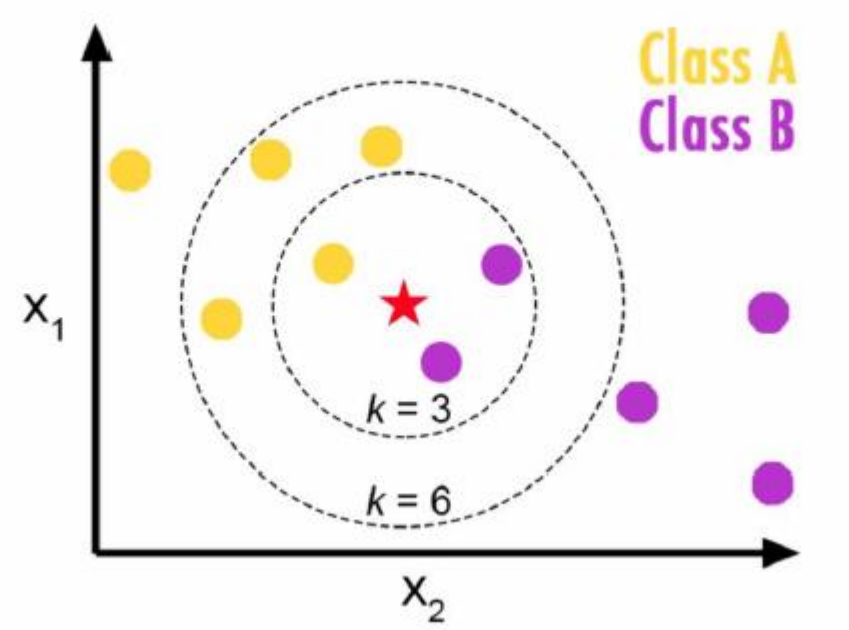


New datapoint:
Is it a horse or a dog?

New datapoint:
Is it a horse or a dog?

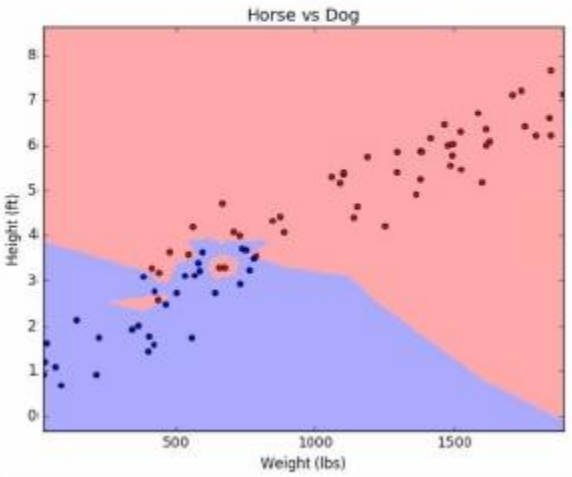
New datapoint:
Is it a horse or a dog?

Choosing a K will affect what class a new point is assigned to:

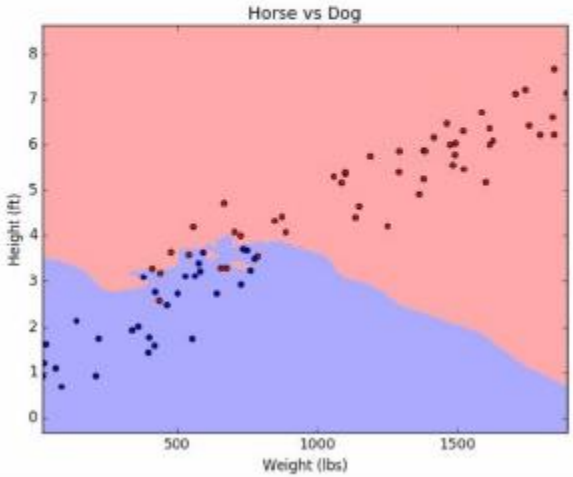


Choosing a K will affect what class a new point is assigned to:

k=1

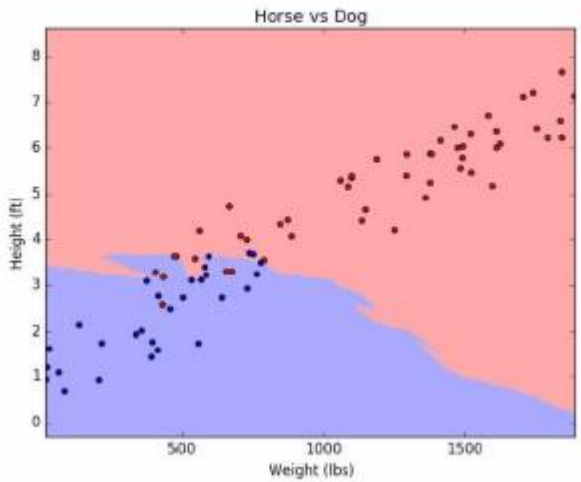


k=5

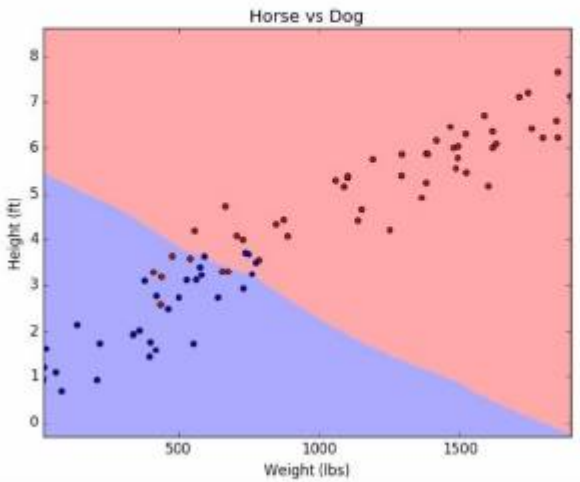


Choosing a K will affect what class a new point is assigned to:

k=10



k=50



Pros

- Very simple
- Training is trivial
- Works with any number of classes
- Easy to add more data
- Few parameters
  - K
  - Distance Metric

Cons

- High Prediction Cost (worse for large data sets)
- Not good with high dimensional data
- Categorical Features don’t work well