

Basic Concepts of Learning

Yan Yan

Deep learning? What is learning?



What is in this picture?

Deep learning? What is **learning**?



Past data

Learning system



Future data

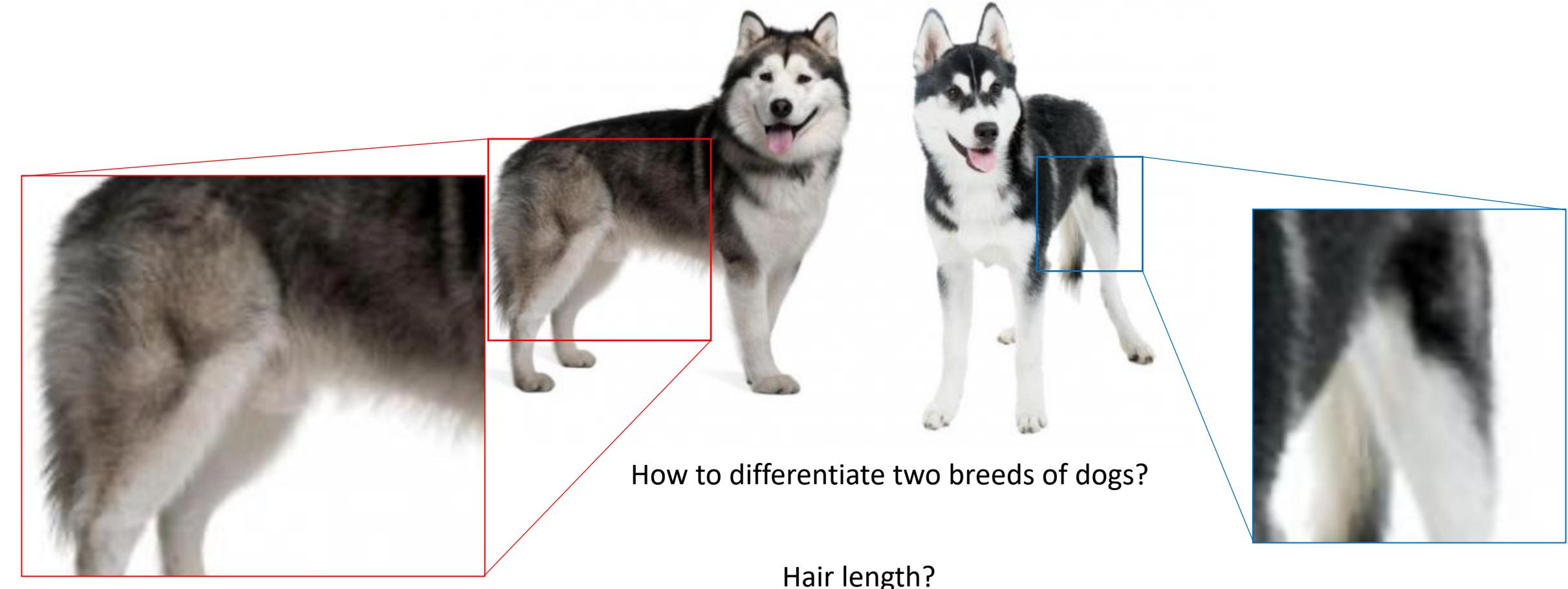
What is it?

Husky vs Malamute

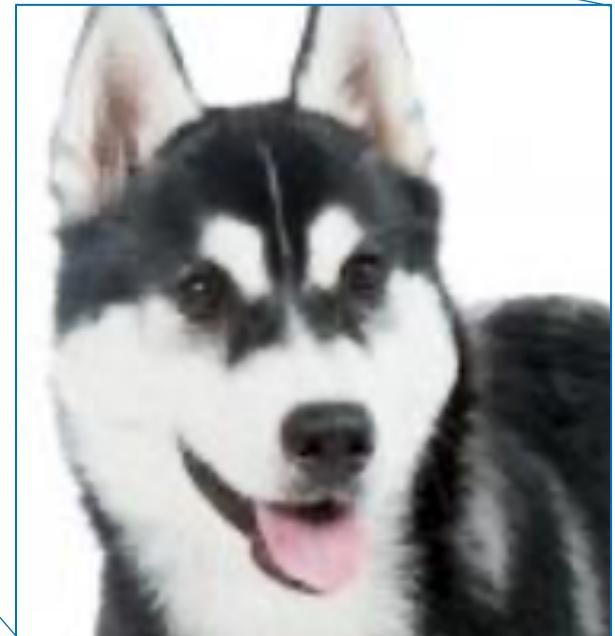
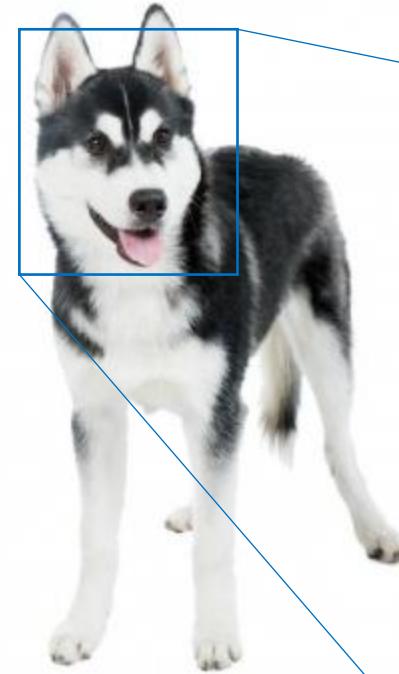
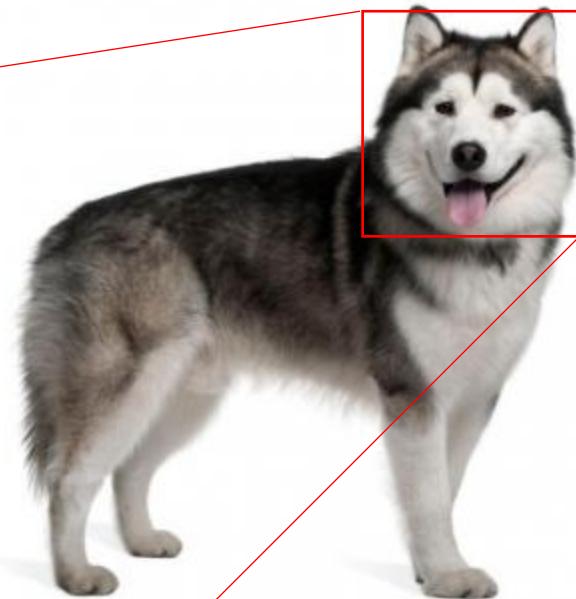


How to differentiate two breeds of dogs?

Husky vs Malamute



Husky vs Malamute



How to differentiate two breeds of dogs?

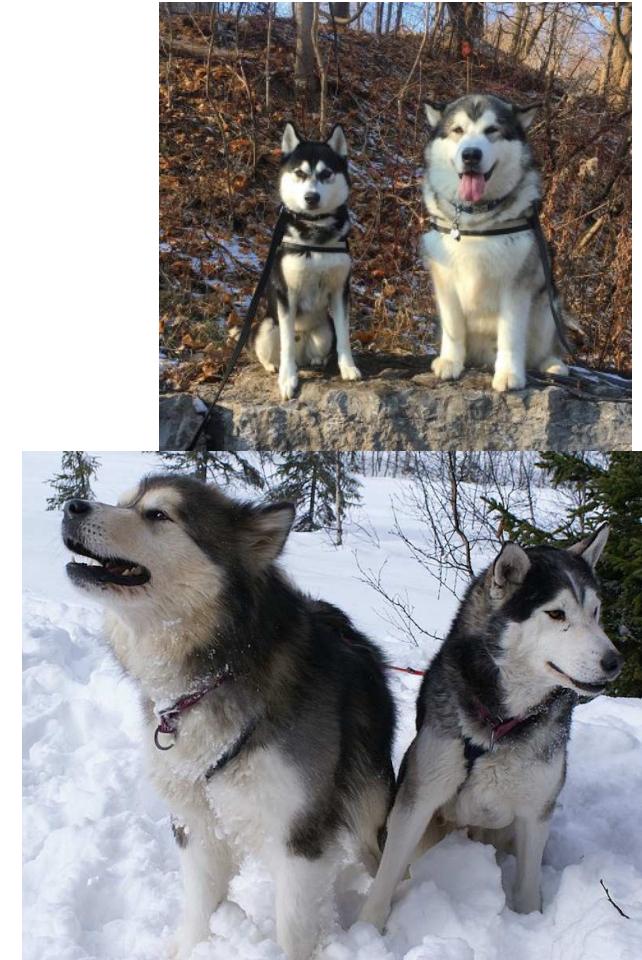
Face markings?

Husky vs Malamute



How to differentiate two breeds of dogs?

Size?



Husky vs Malamute



Alaskan Malamute vs Siberian Husky Comparison



	Alaskan Malamute	Siberian Husky
Size:	20-25 inches tall	20-22 inches tall
Weight:	85-100lbs	35-60lbs

How to differentiate two breeds of dogs?

Size?



Machine learning paradigm



Past data: features

How to choose/generate useful features?

Learning
model

How to determine this model?

Prediction: bus



Future data

Specific tasks: prediction/decision
(pre-defined)

Machine learning paradigm



Past data: **features**

How to choose/generate useful features?

Learning model

How to determine this model?

Prediction on future data



Task: prediction/decision

(pre-defined setting)

Pre-defined problem settings

- Supervised learning



bus



dog



bus



dog

complete labels: supervised learning

Pre-defined problem settings

- Non-supervised learning?



bus



dog

Labeled and unlabeled data: semi-supervised learning

Pre-defined problem settings

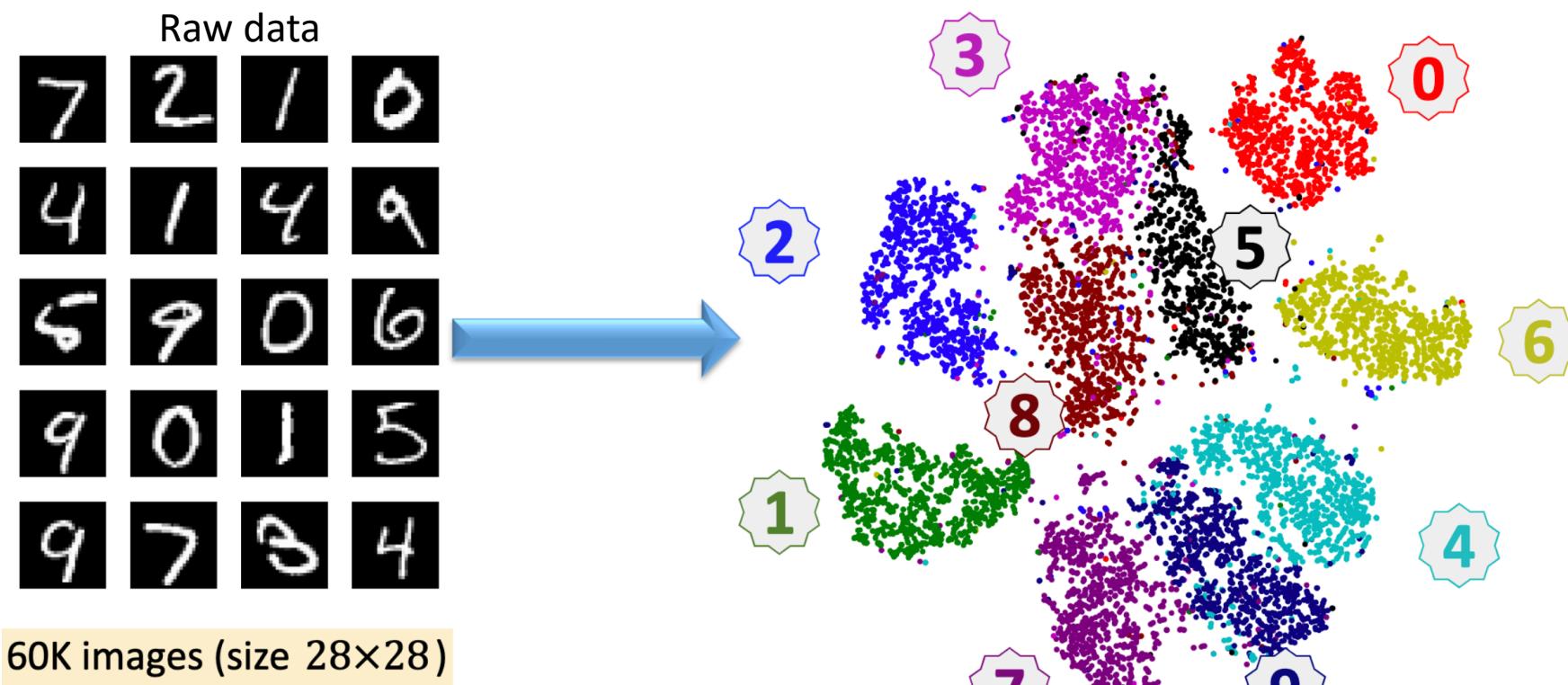
- Non-supervised learning?



Unlabeled data: unsupervised learning

Pre-defined problem settings

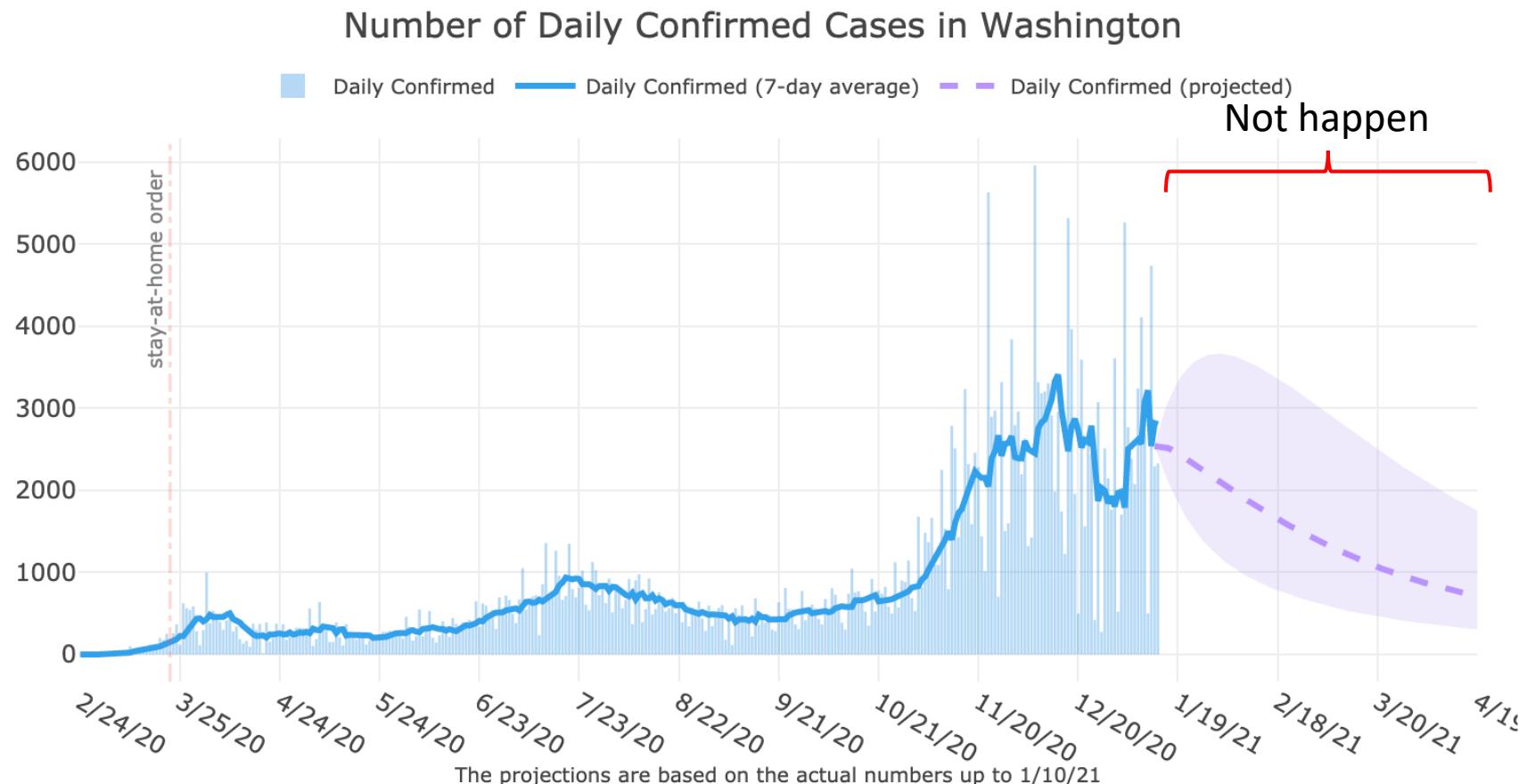
- Non-supervised learning?



Autoencoder for hand-written digit data, retrieved from Shusen Wang's slides at
https://github.com/wangshusen/DeepLearning/blob/master/Slides/1_ML_Basics.pdf

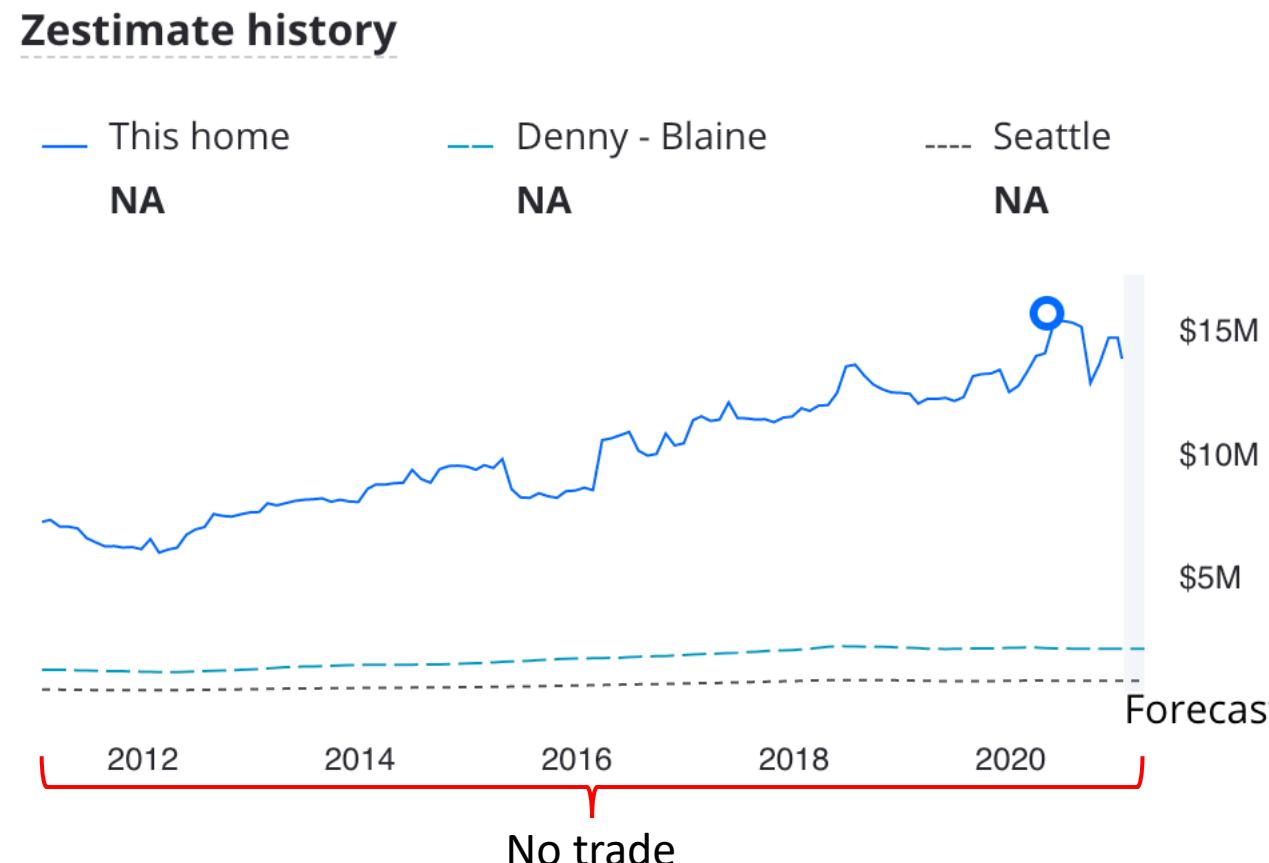
Pre-defined problem settings

- Prediction vs decision making



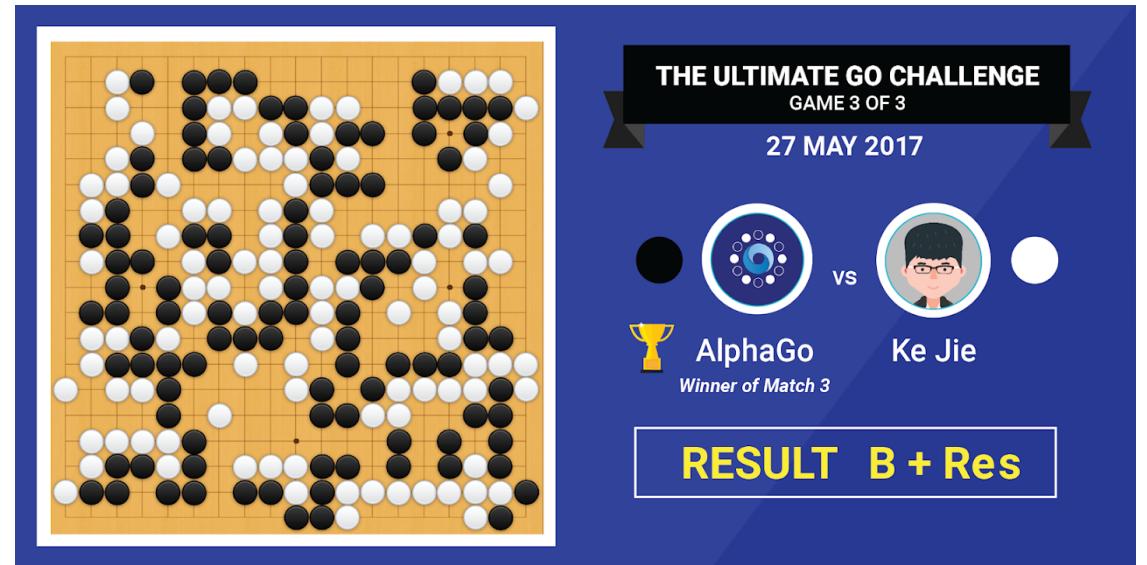
Pre-defined problem settings

- Prediction vs decision making



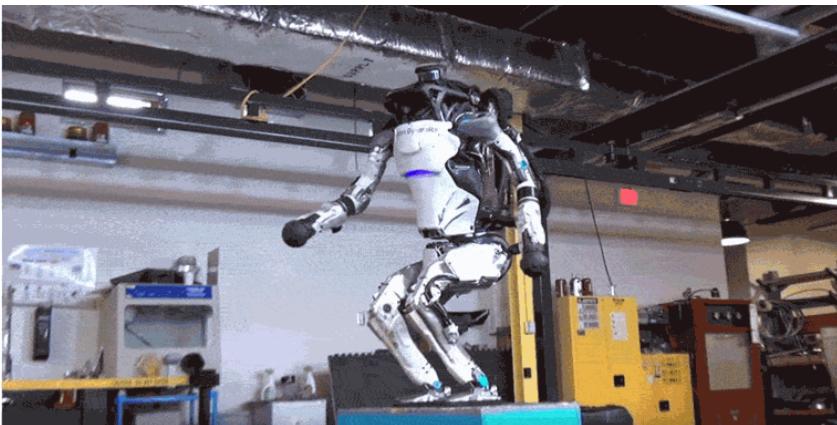
Pre-defined problem settings

- Prediction vs **decision making**



Pre-defined problem settings

- Prediction vs **decision making**



Machine learning paradigm



Past data: **features**

How to choose/generate useful features?

Learning
model

How to determine this model?

Prediction: bus



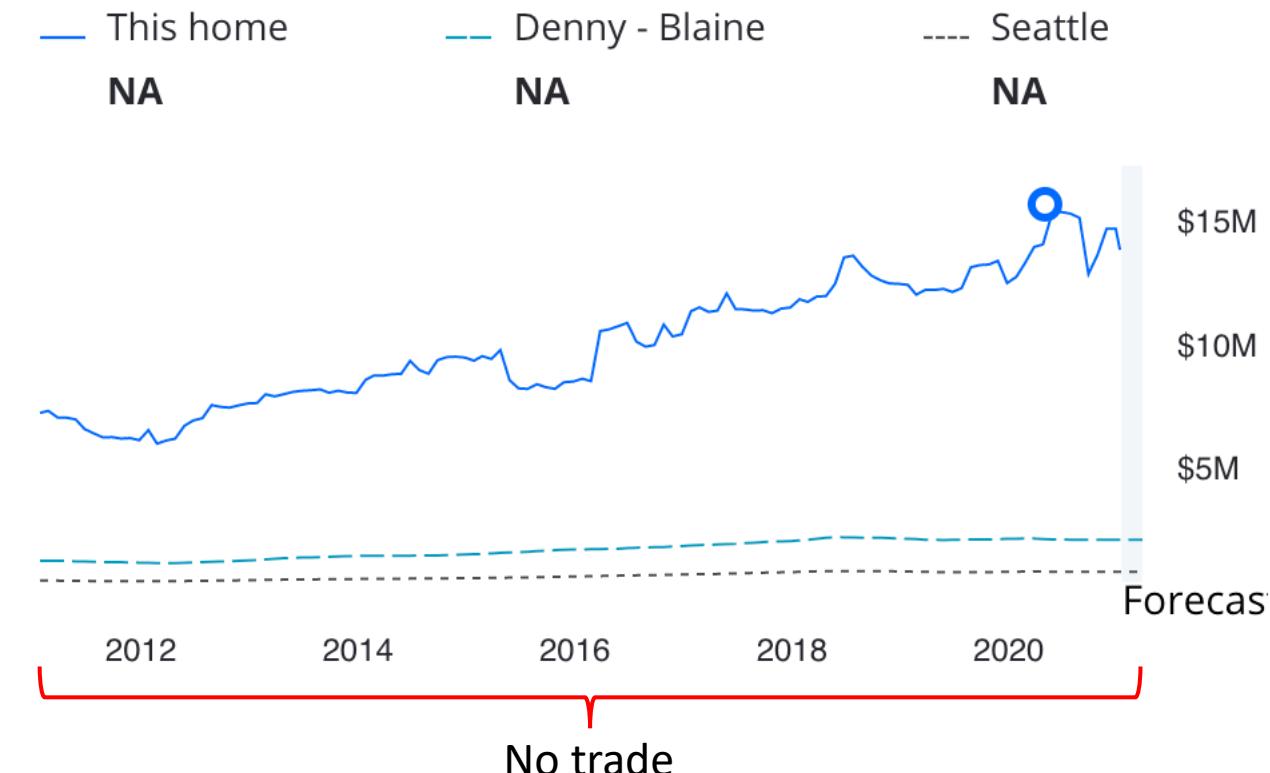
Future data

Specific tasks: prediction/decision
(pre-defined)

Features

What features can we use?

Zestimate history



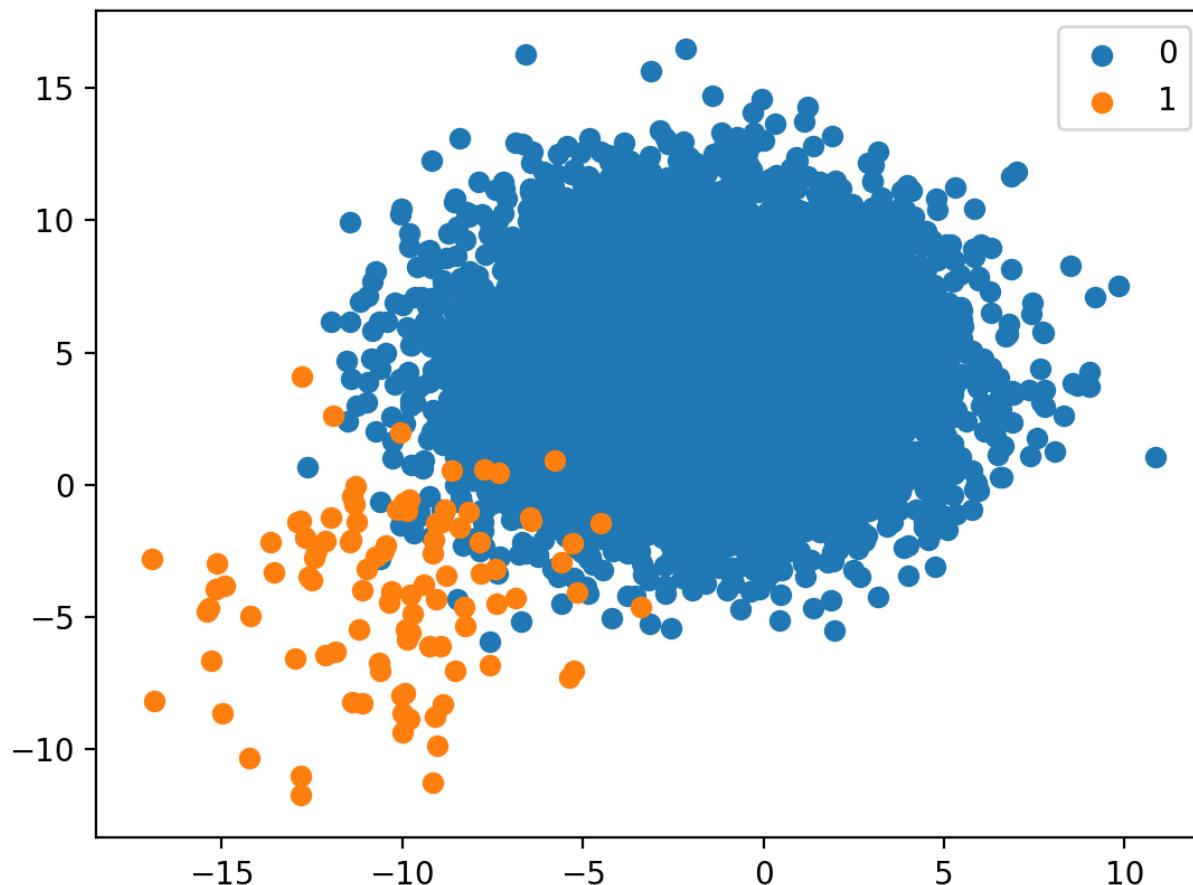
Features

- Home characteristics: lot size, location, #bedrooms
- Unique features: hardwood floors, granite countertops or a landscaped backyard
- On-market data: listing price, description, days on the market
- Off-market data: tax assessments, prior sales

Learning model

- What is a model

Try to separate two classes:

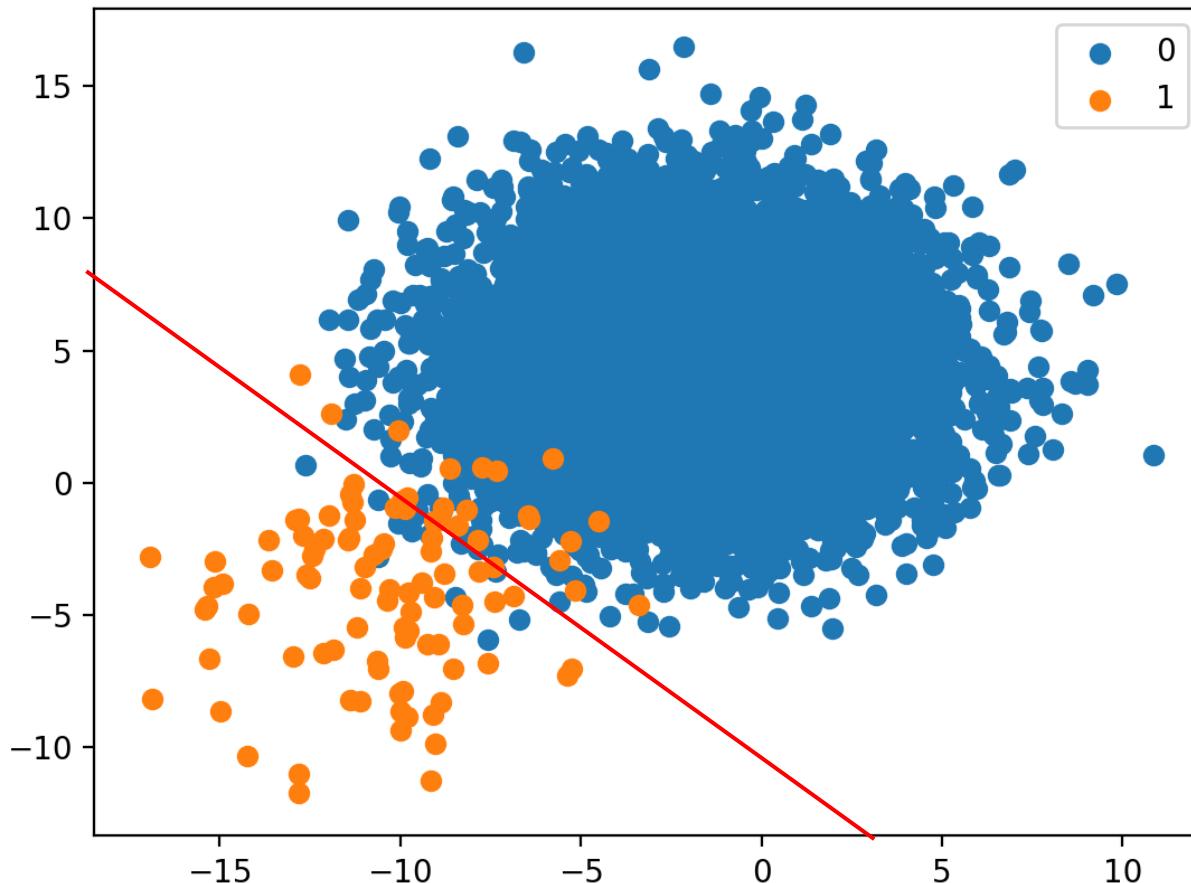


Learning model

- What is a model

Try to separate two classes:

A linear function



Learning model

- What is a model

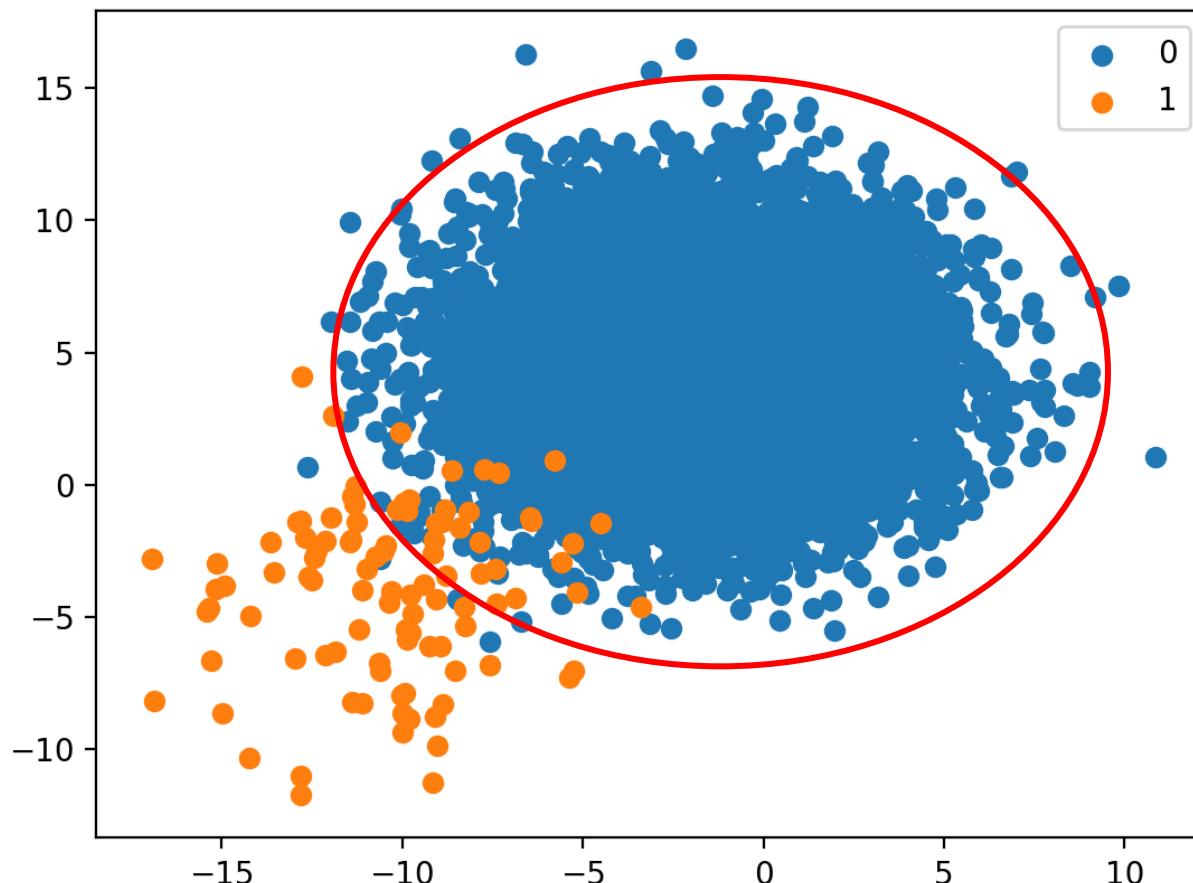
Try to separate two classes:

A linear function

or

A ball

Q: what is the feature used here?



Learning model

- What is a model

Try to separate two classes:

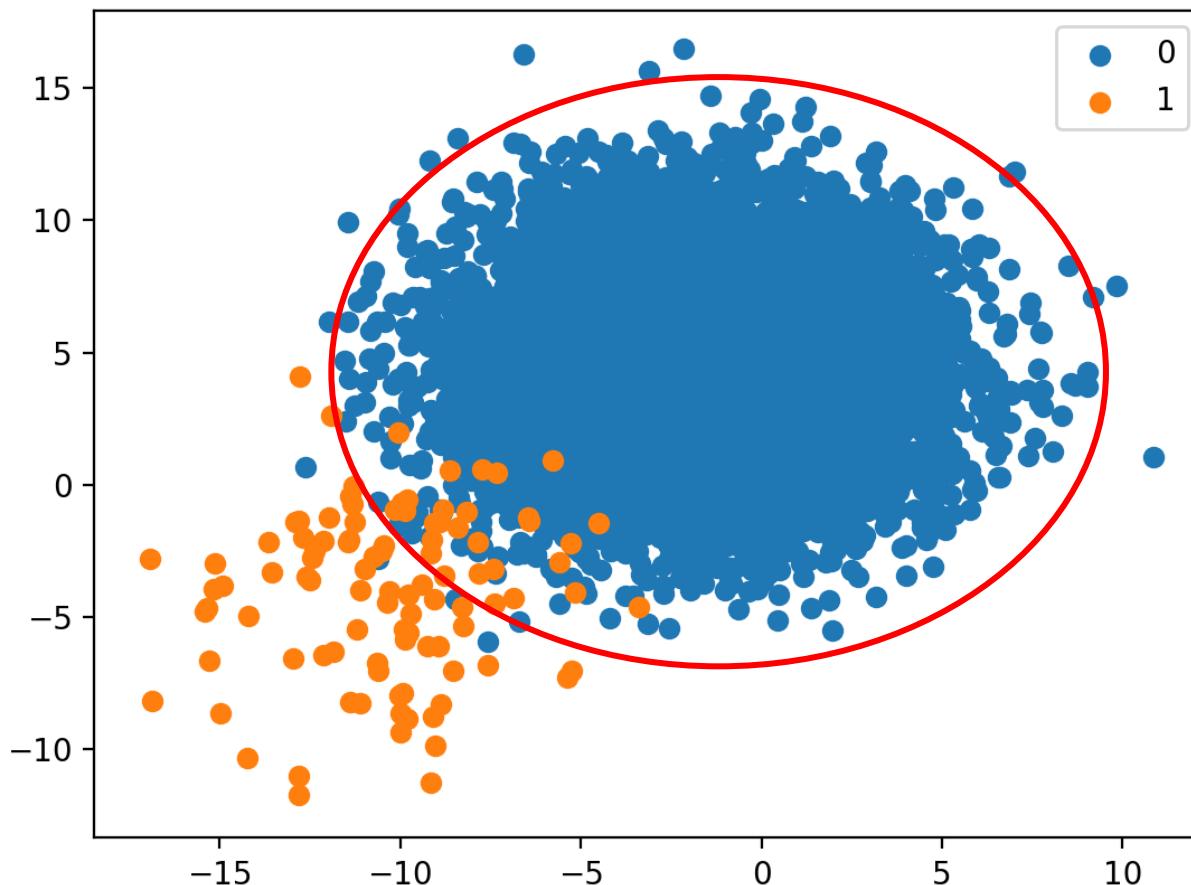
A linear function

or

A ball

Q: what is the feature used here?

x-y coordinates



Learning model

- What is a model

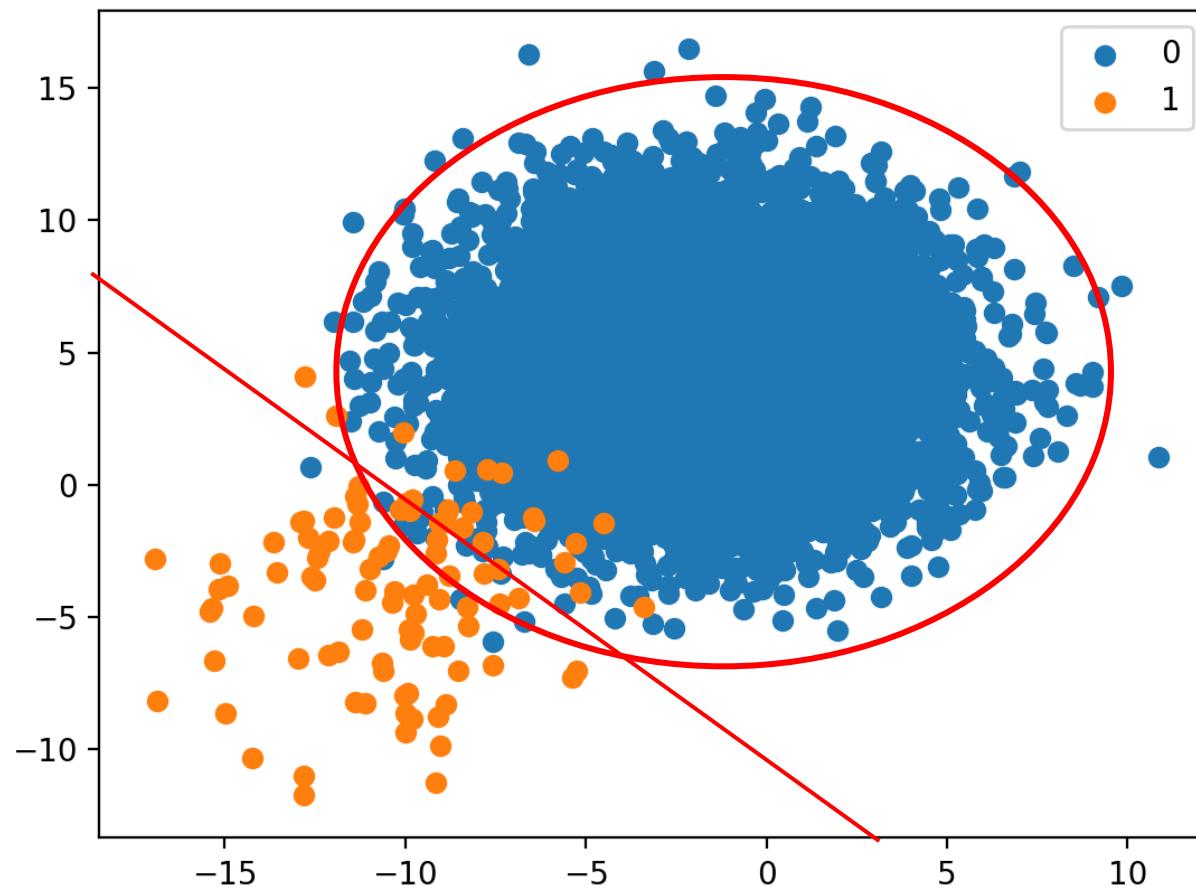
Try to separate two classes:

A linear function

or

A ball (nonlinear function)

Different structures



Learning model

- What is a model

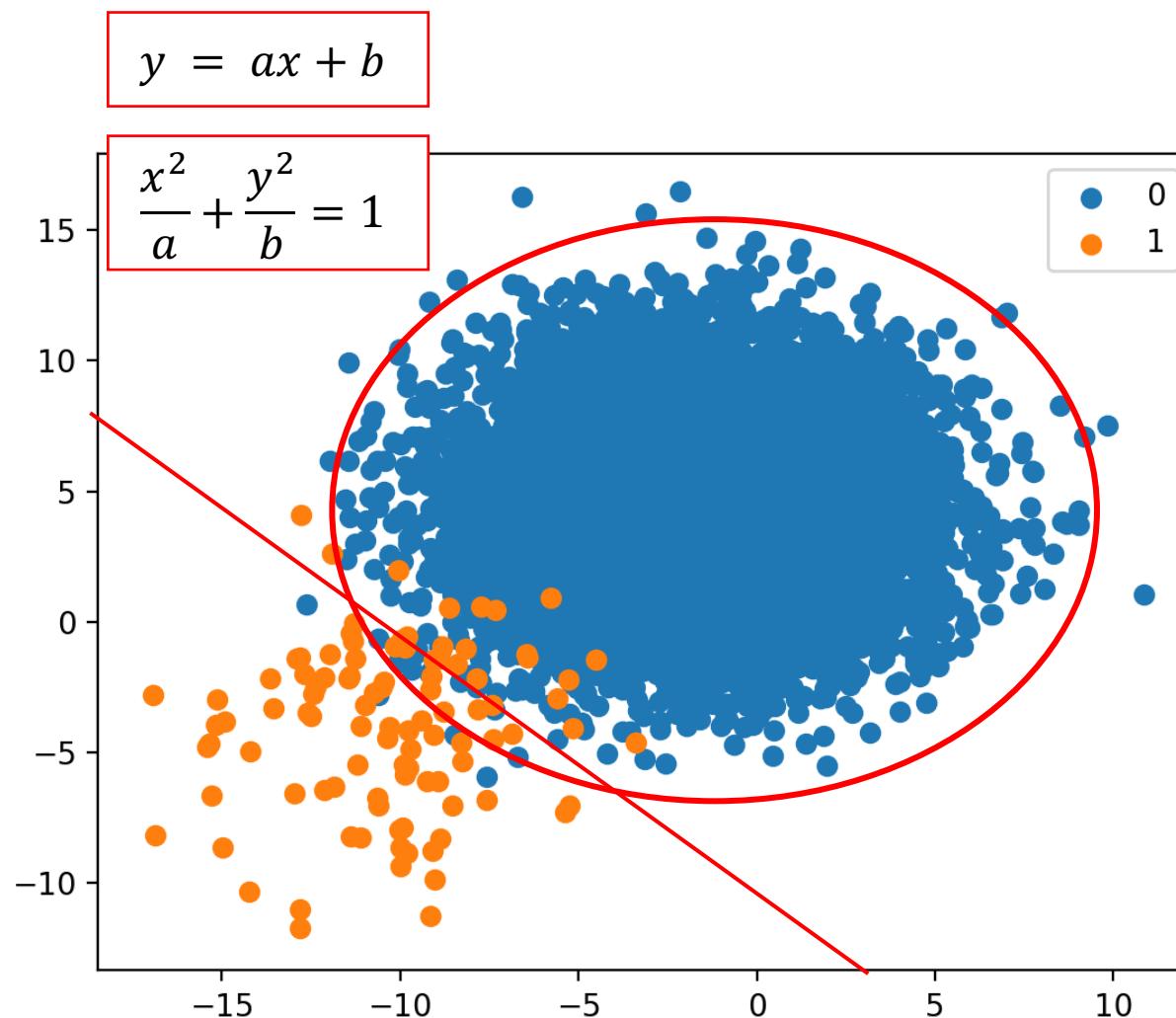
Try to separate two classes:

A linear function

or

A ball (nonlinear function)

Different structures



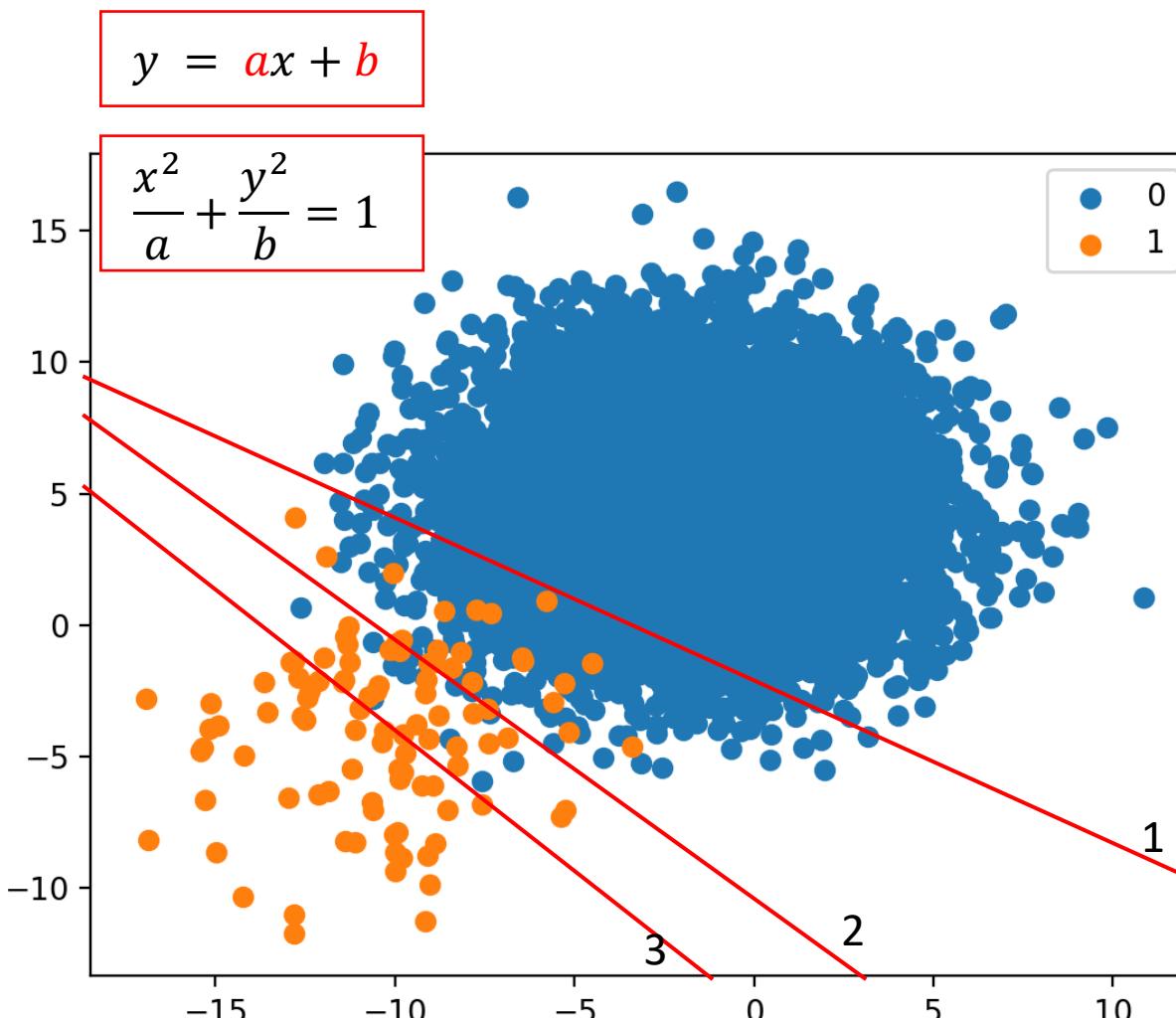
Learning model

- What is a model

After deciding model structure:

Adjust to satisfy our strategy

1. #error@1 = 0
2. #error@1 = #error@0
3. #error@0 = 0



*Melanoma recognition: case 1

Retrieved from <https://machinelearningmastery.com/how-to-develop-an-intuition-skewed-class-distributions/>

Summary

- What is learning
- What is machine learning paradigm
- What are pre-defined settings (tasks, supervision)
- What are features (discriminative properties)
- What are learning models (structures, adjustment)