Introduction to Machine Learning

Knowledge Sharing for CPE/SKE students

Sirakorn Lamyai

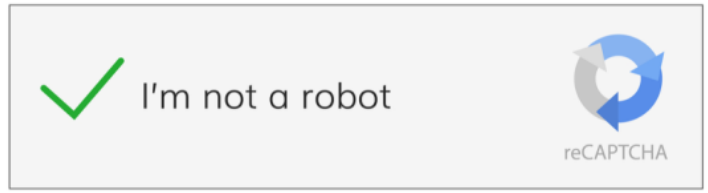
Student, Kasetsart U.

October 2, 2018

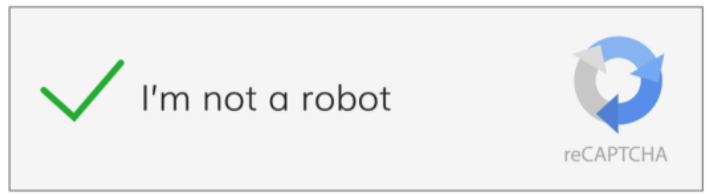
Outline

What is Machine Learning?

What is Machine Learning?

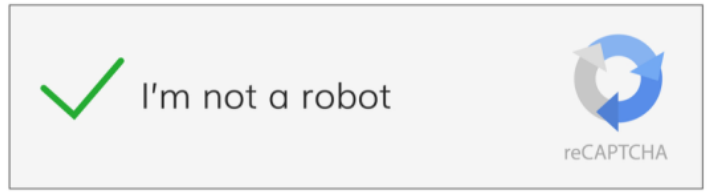


What is Machine Learning?



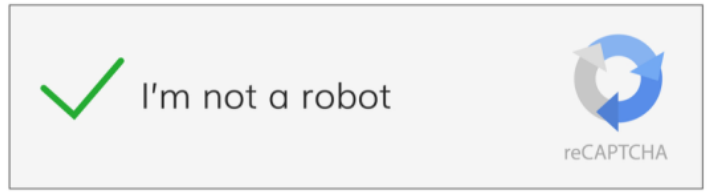
- This is Recaptcha.

What is Machine Learning?



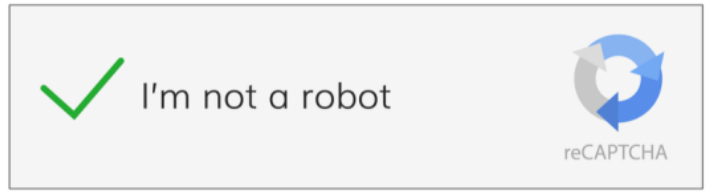
- This is Recaptcha.
  - Recaptcha helps stop millions of spam a day.

What is Machine Learning?



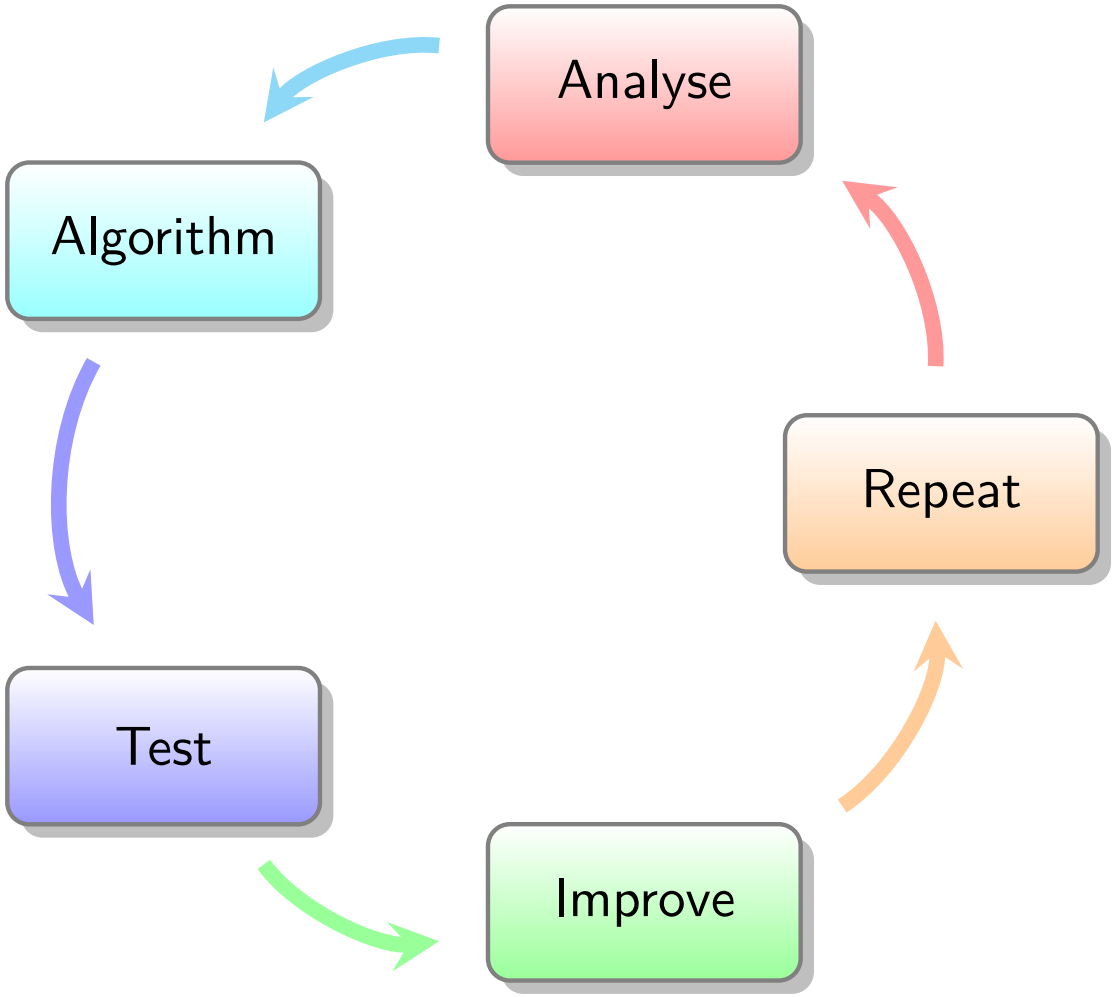
- This is Recaptcha.
  - Recaptcha helps stop millions of spam a day.
  - In some old days, we have to type Captcha texts to distinguish ourself from bots.

What is Machine Learning?

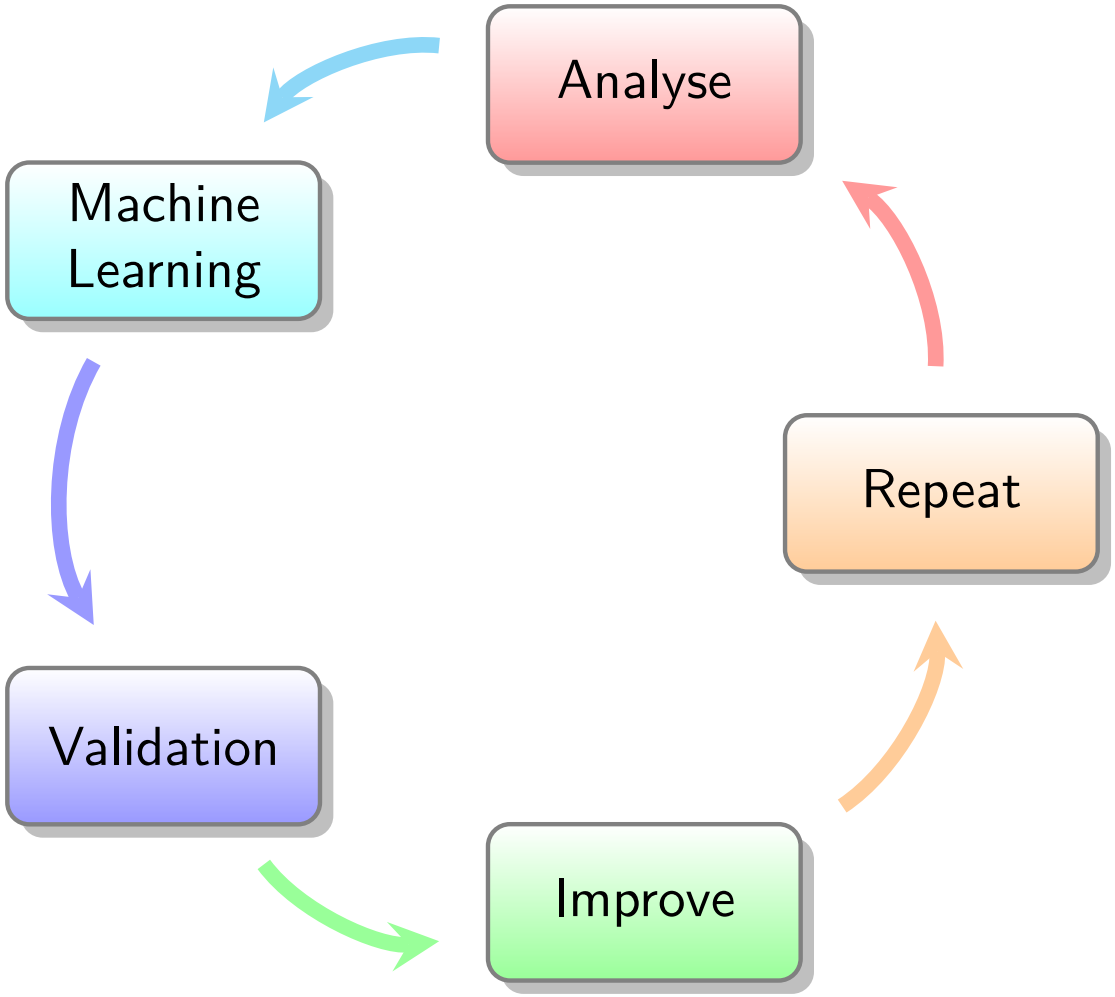


- This is Recaptcha.
  - Recaptcha helps stop millions of spam a day.
  - In some old days, we have to type Captcha texts to distinguish ourself from bots.
  - How is it possible that with a single click, an automated system can distinguish bots from humans?

Traditional programming approach



Machine learning approach



In other words...

Machine Learning

In other words...

Machine Learning

$$= \text{Data} + \text{Data analysis algorithm}$$

In other words...

Machine Learning

$$= \text{Data} + \text{Data analysis algorithm}$$

$$= \text{Adapt to change}$$

Types of Machine Learning problems

Types of Machine Learning problems

1. Supervised learning

Types of Machine Learning problems

1. Supervised learning
2. Unsupervised learning

Types of Machine Learning problems

1. Supervised learning
2. Unsupervised learning
3. Reinforcement learning

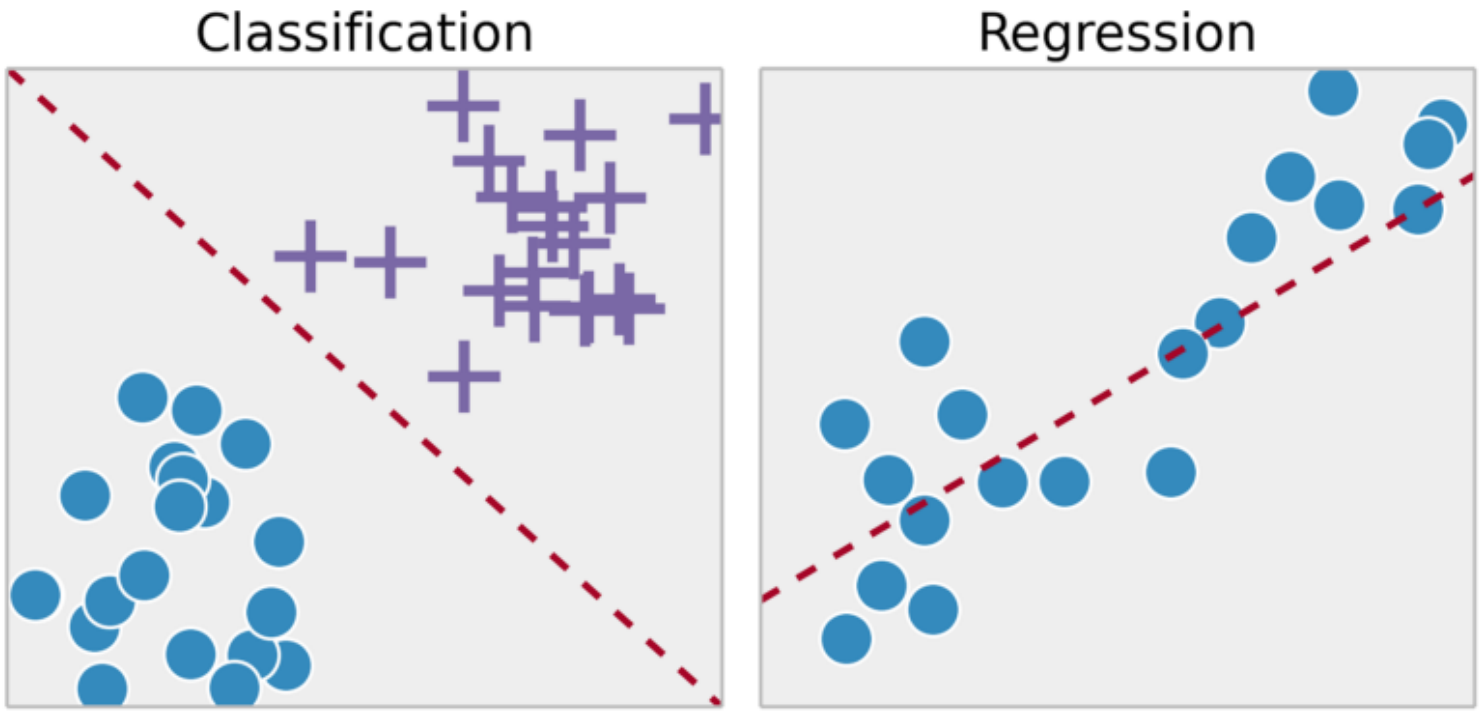
Types of Machine Learning problems

1. Supervised learning
2. Unsupervised learning
3. Reinforcement learning

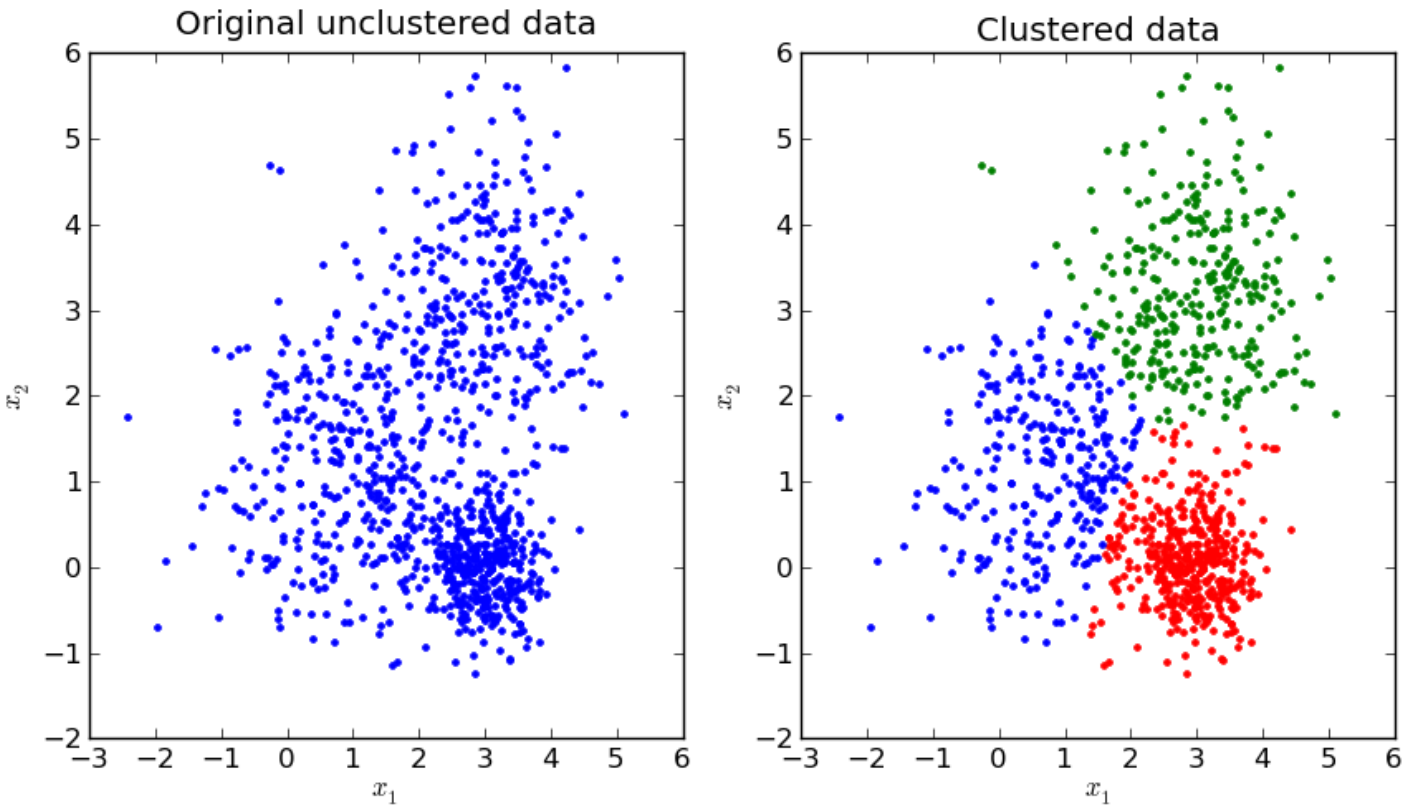
Determined by

Labels

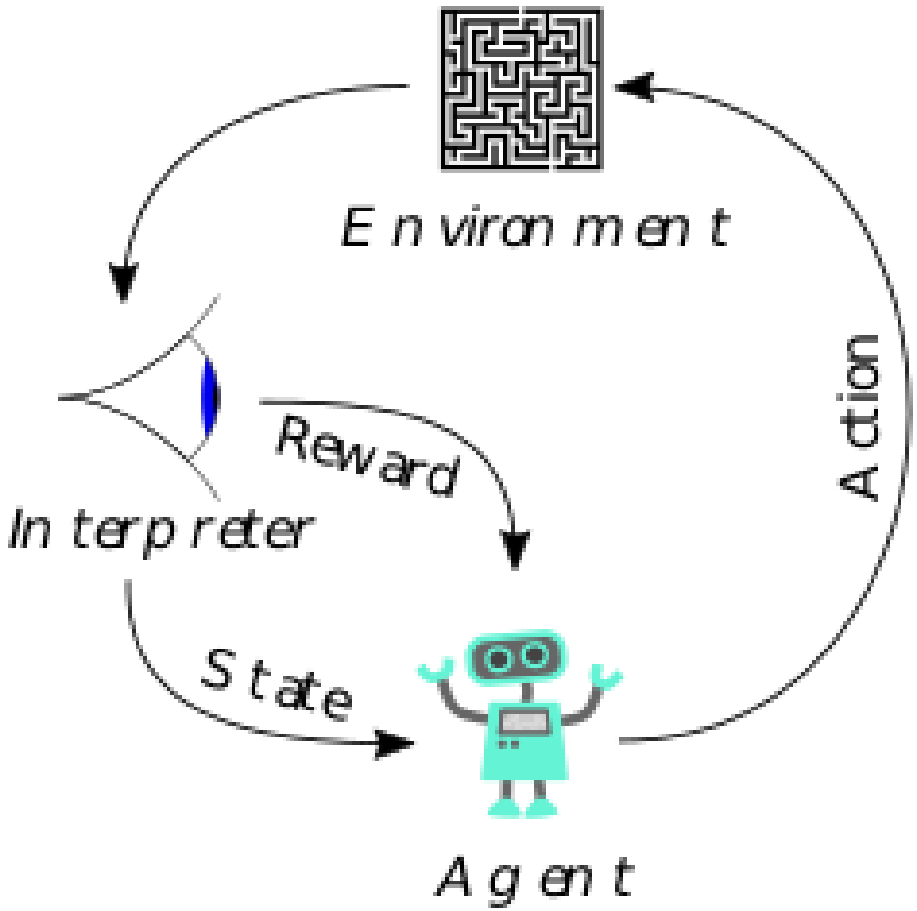
Supervised learning



Unsupervised learning



Reinforcement learning



Model

Model

Model

- A result of the combination between...

Model

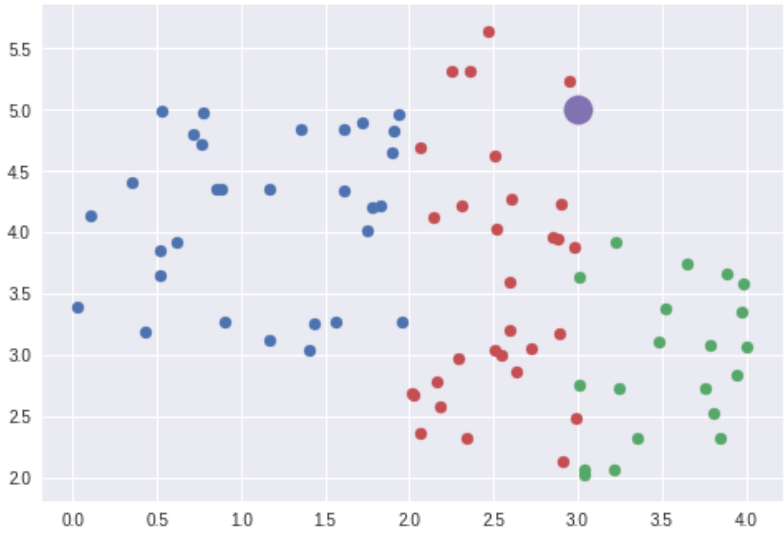
- A result of the combination between...
  - a **method** to recognise the data, and

Model

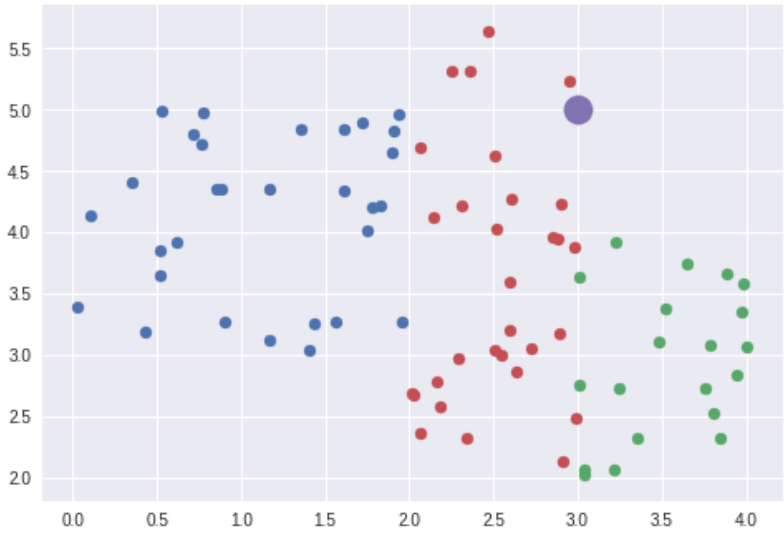
- A result of the combination between...
  - a **method** to recognise the data, and
  - **sample datas** for such the method

Model

Model

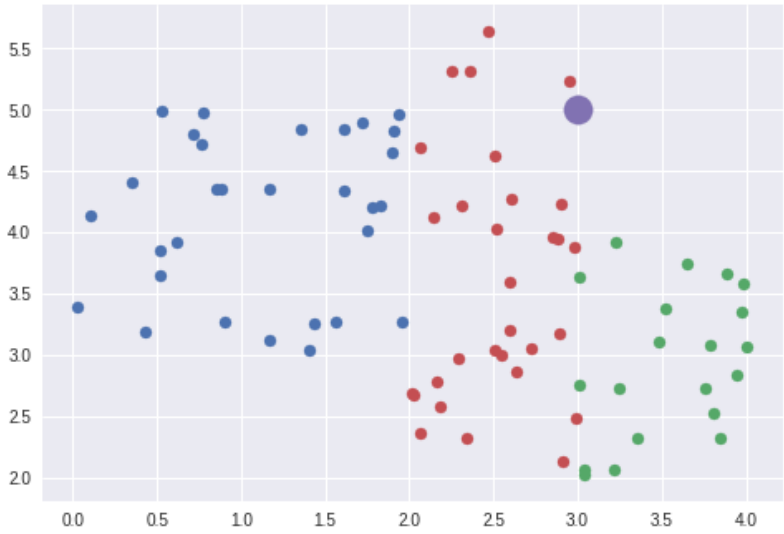


Model



Determine which group should the purple dot be in (red/green/blue) by **checking the colour of its nearest dot.**

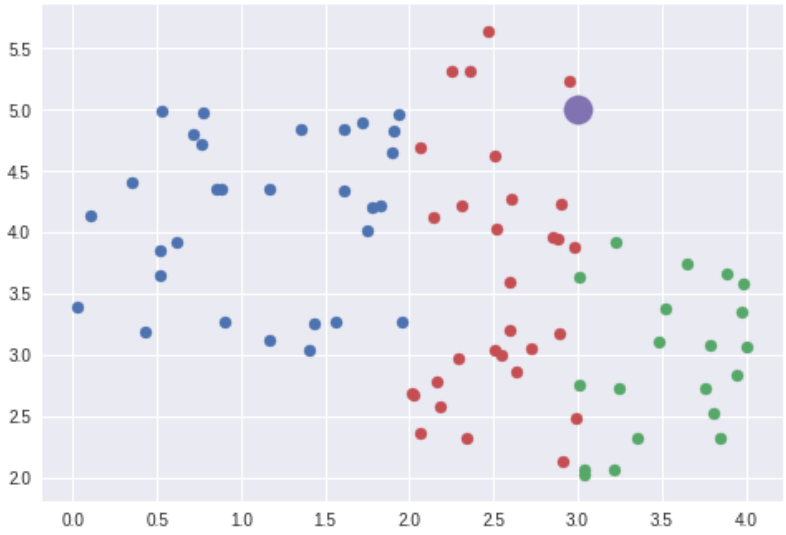
Model



Data

Determine which group should the purple dot be in (red/green/blue) by **checking the colour of its nearest dot.**

Model



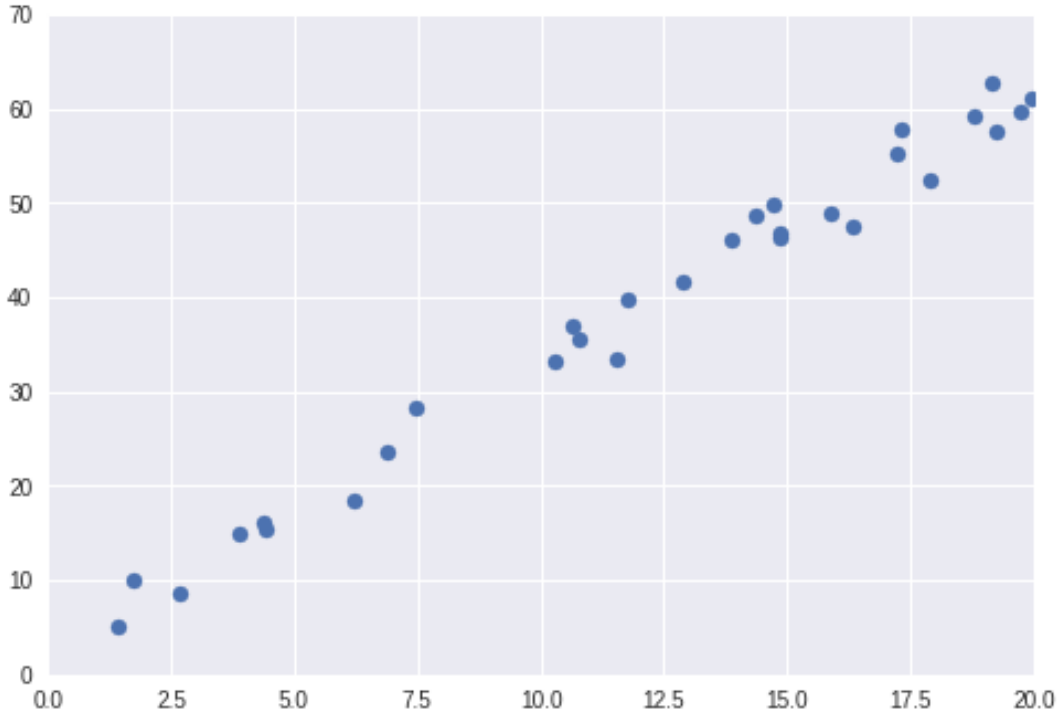
Data

Determine which group should the purple dot be in (red/green/blue) by **checking the colour of its nearest dot.**

Method

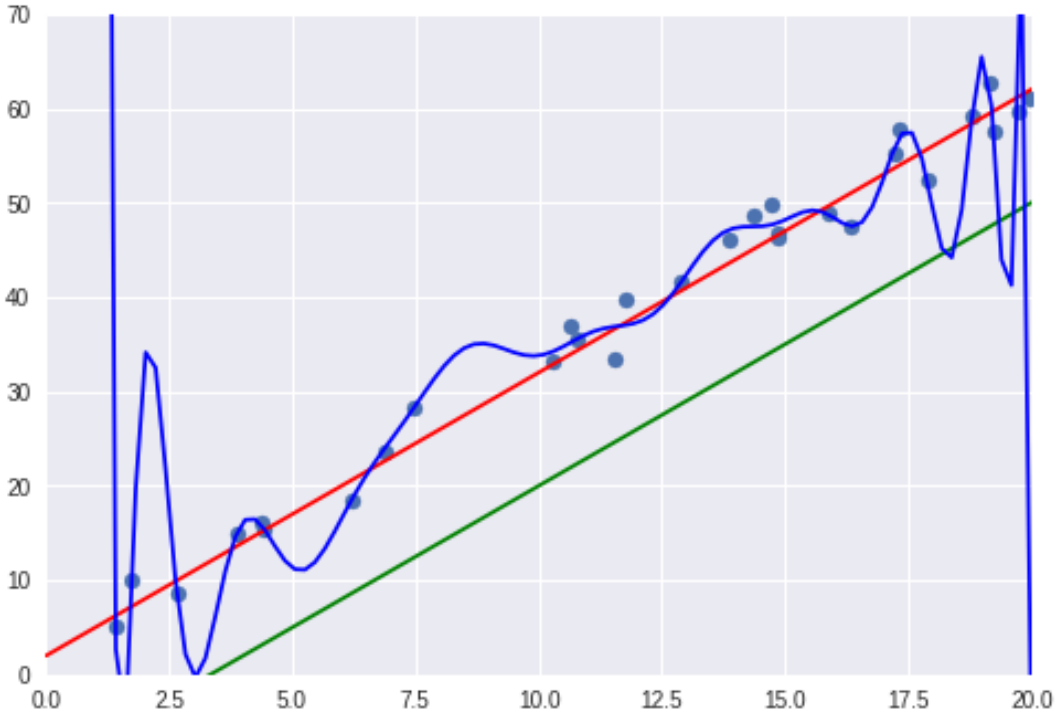
Good model?

Good model



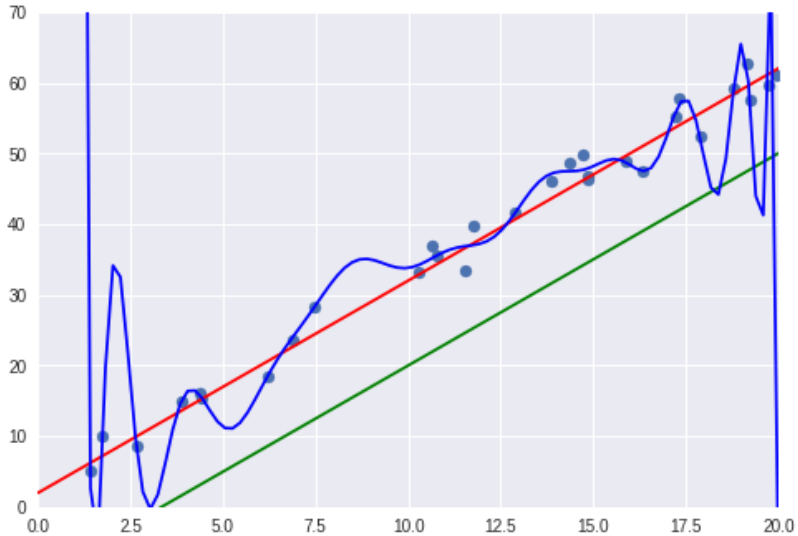
How should we *draw* the line to predict this data?

Good model

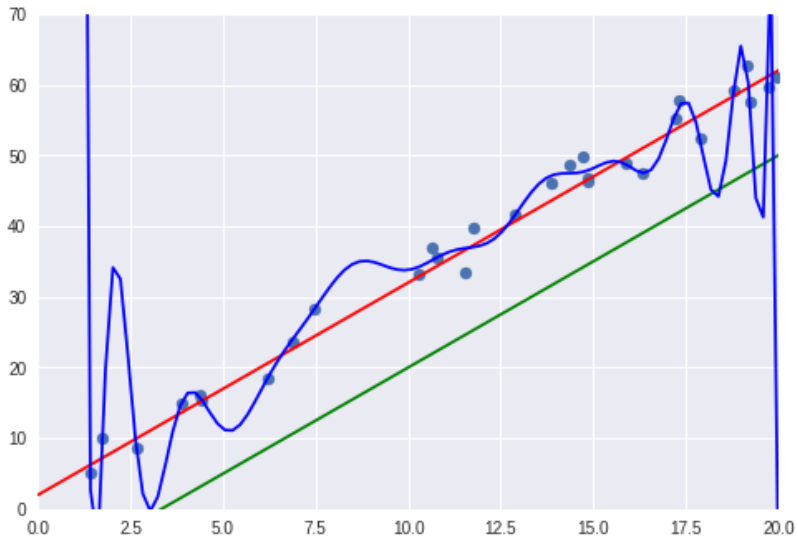


Blue, red, or green line?

Overfitting and underfitting

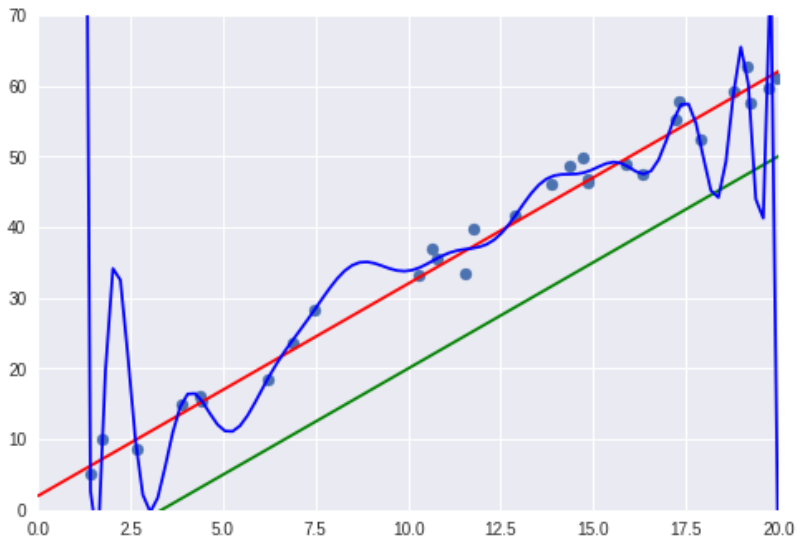


Overfitting and underfitting



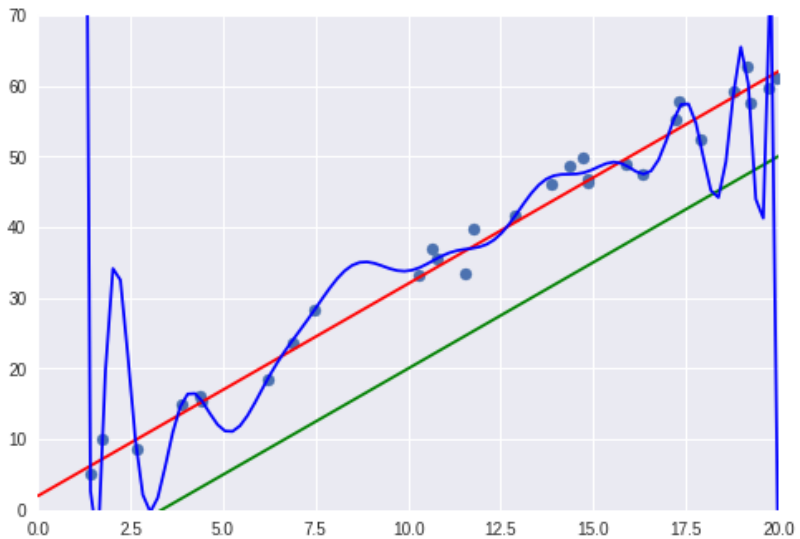
1. Underfitting

Overfitting and underfitting



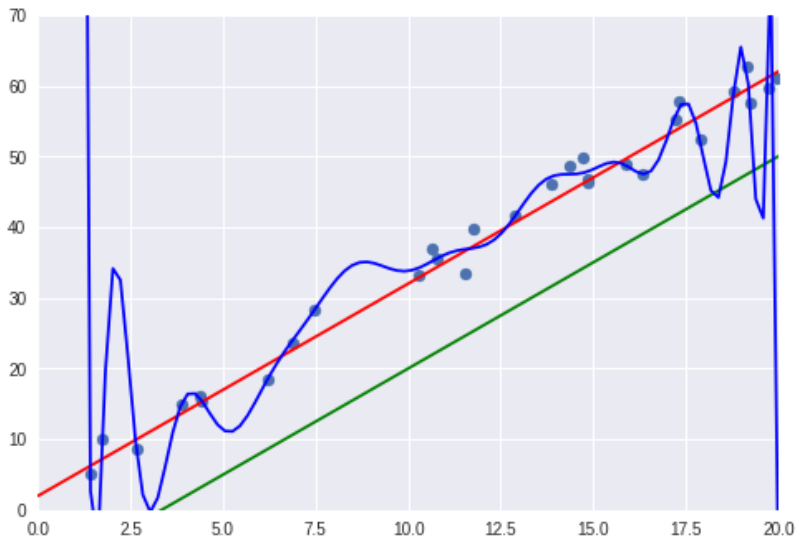
1. Underfitting
   - Our model **fails to know the data's trends**

Overfitting and underfitting



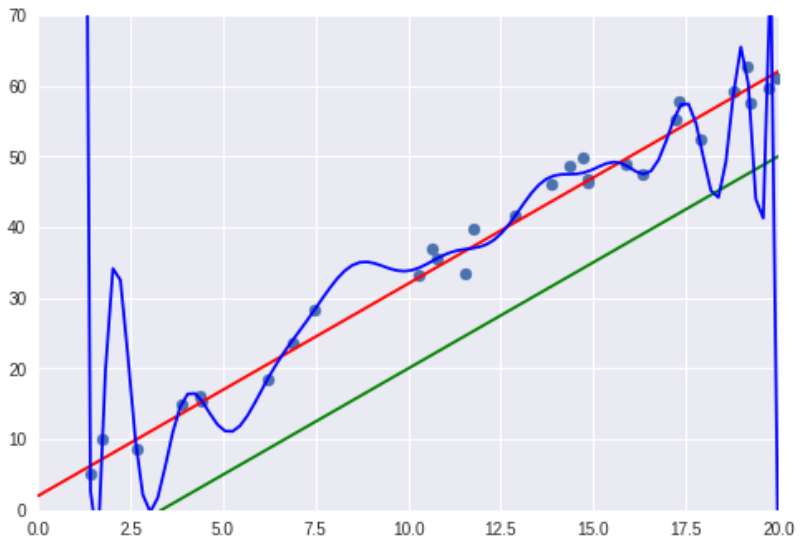
1. Underfitting
   - Our model **fails to know the data's trends**
   - Resulting in failure to predict further data

Overfitting and underfitting



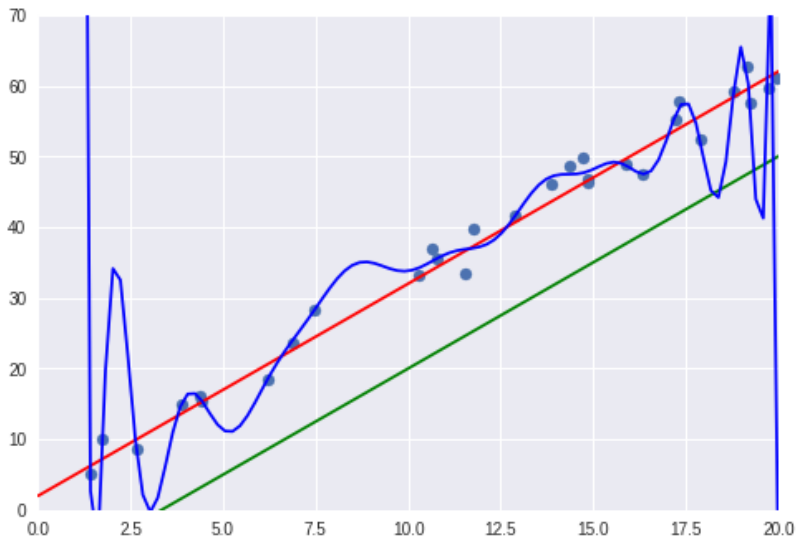
1. Underfitting
   - Our model **fails to know the data's trends**
   - Resulting in failure to predict further data
2. Overfitting

Overfitting and underfitting



1. Underfitting
   - Our model **fails to know the data's trends**
   - Resulting in failure to predict further data
2. Overfitting
   - Our model **memorise instead of generalise**

Overfitting and underfitting



1. Underfitting
   - Our model **fails to know the data's trends**
   - Resulting in failure to predict further data
2. Overfitting
   - Our model **memorise instead of generalise**
   - Resulting in failure to catch the trend

Good model

Good model
must generalise