COMP2261 ARTIFICIAL INTELLIGENCE / MACHINE LEARNING

Underfitting vs Overfitting

Dr SHI Lei

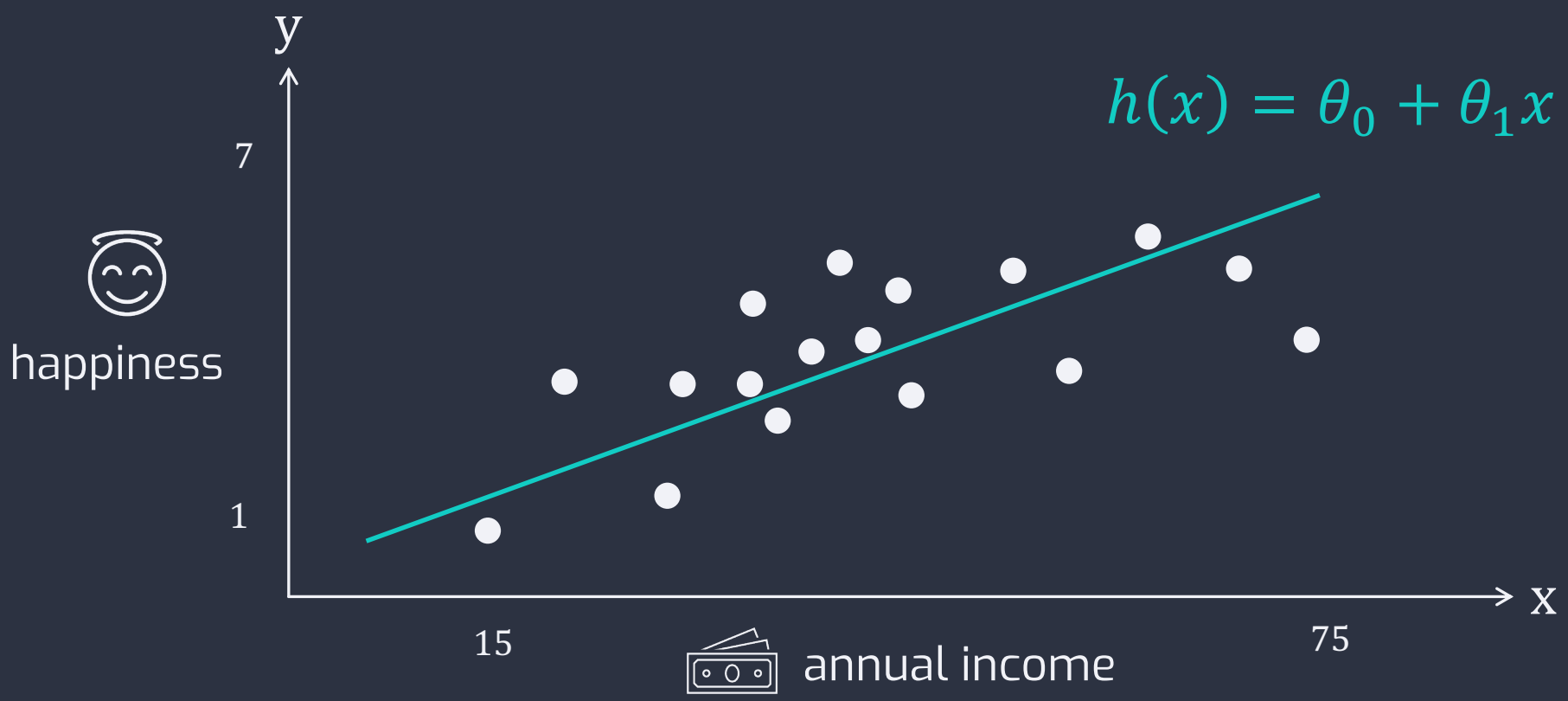
Durham University

learning lab

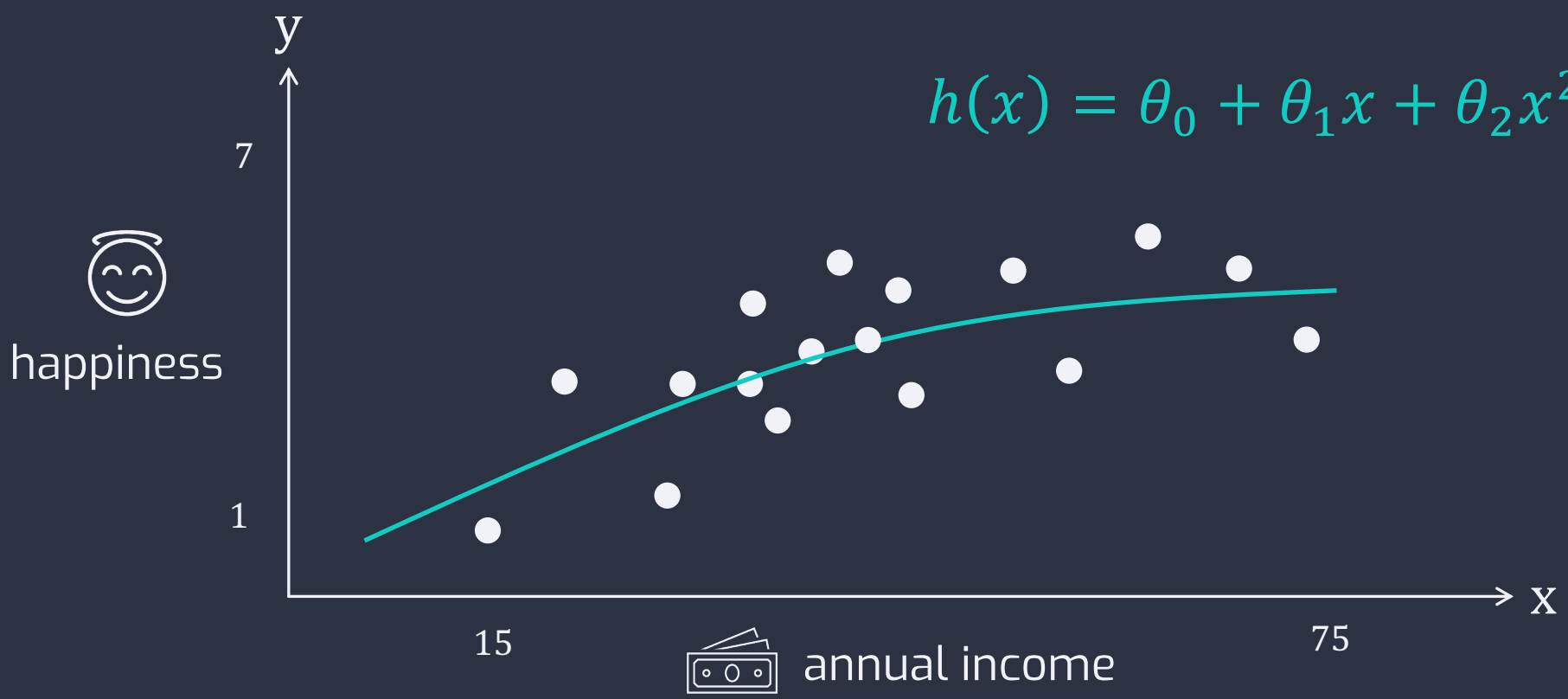
Learning Objectives

- Understand the definition of underfitting and overfitting
- Understand how underfitting and overfitting occur
- Understand how to prevent underfitting and overfitting

EXAMPLE. annual income to predict happiness

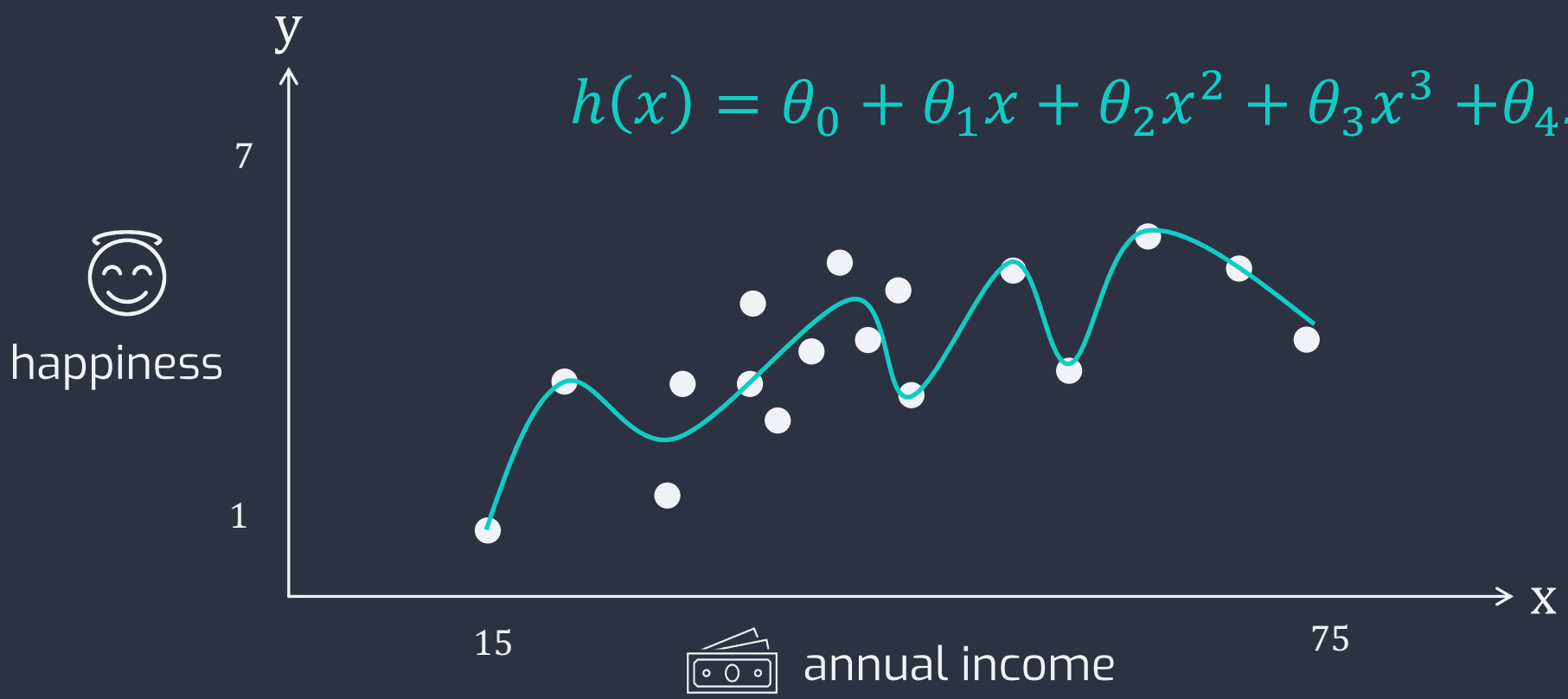


EXAMPLE. annual income to predict happiness



$$h(x) = \theta_0 + \theta_1 x + \theta_2 x^2$$

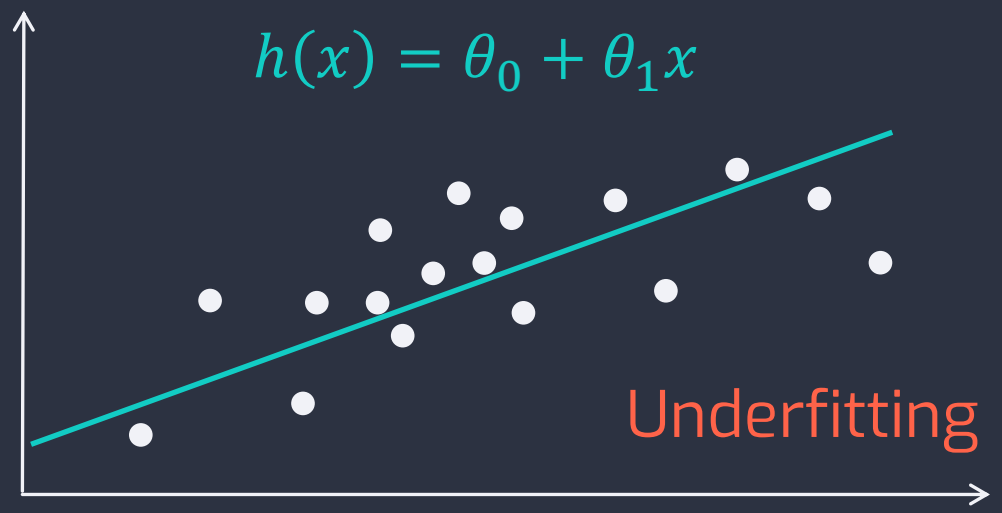
EXAMPLE. annual income to predict happiness



$$h(x) = \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \theta_4 x^4$$

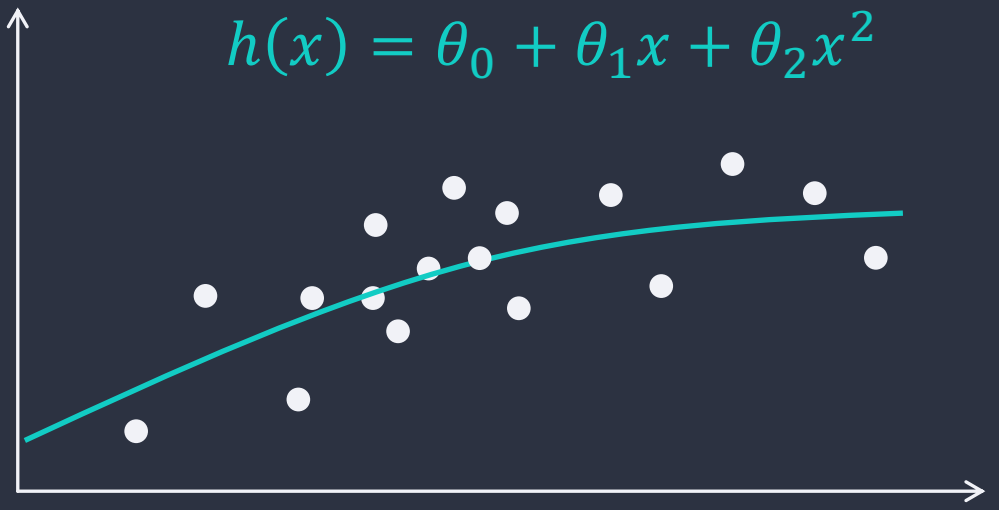
EXAMPLE. annual income to predict happiness

$h(x) = \theta_0 + \theta_1 x$

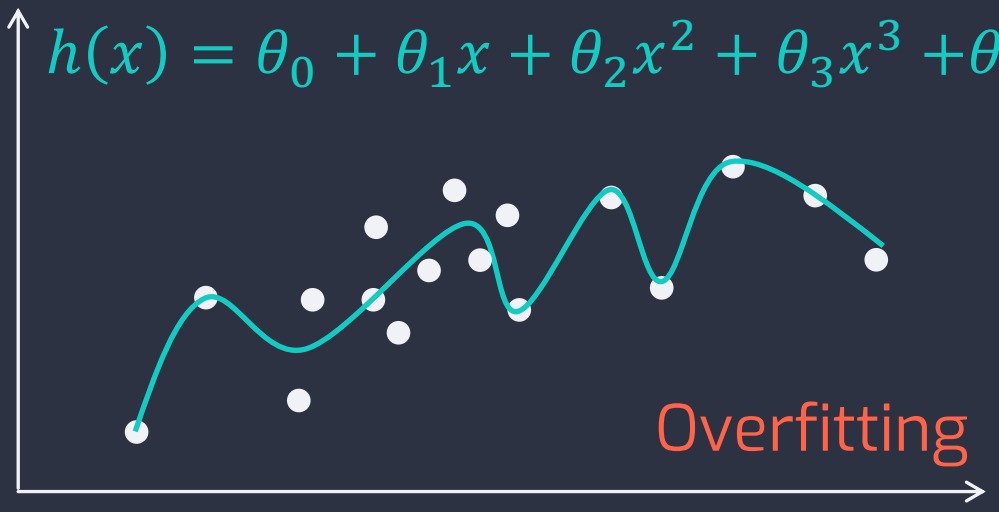


Underfitting

$h(x) = \theta_0 + \theta_1 x + \theta_2 x^2$



$h(x) = \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \theta_4 x^4$



Overfitting

Underfitting

Occurs when a learning algorithm can't capture underlying trend of the data. The model doesn't fit data well enough. This often happens if our model is excessively simple, e.g. polynomial order is too small.

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})^2 \gg 0$$

Overfitting

Occurs when a learning algorithm captures too much of noise from the data. The model fits the data too well. This often happens if the model is excessively complicated, e.g. polynomial order is too high.

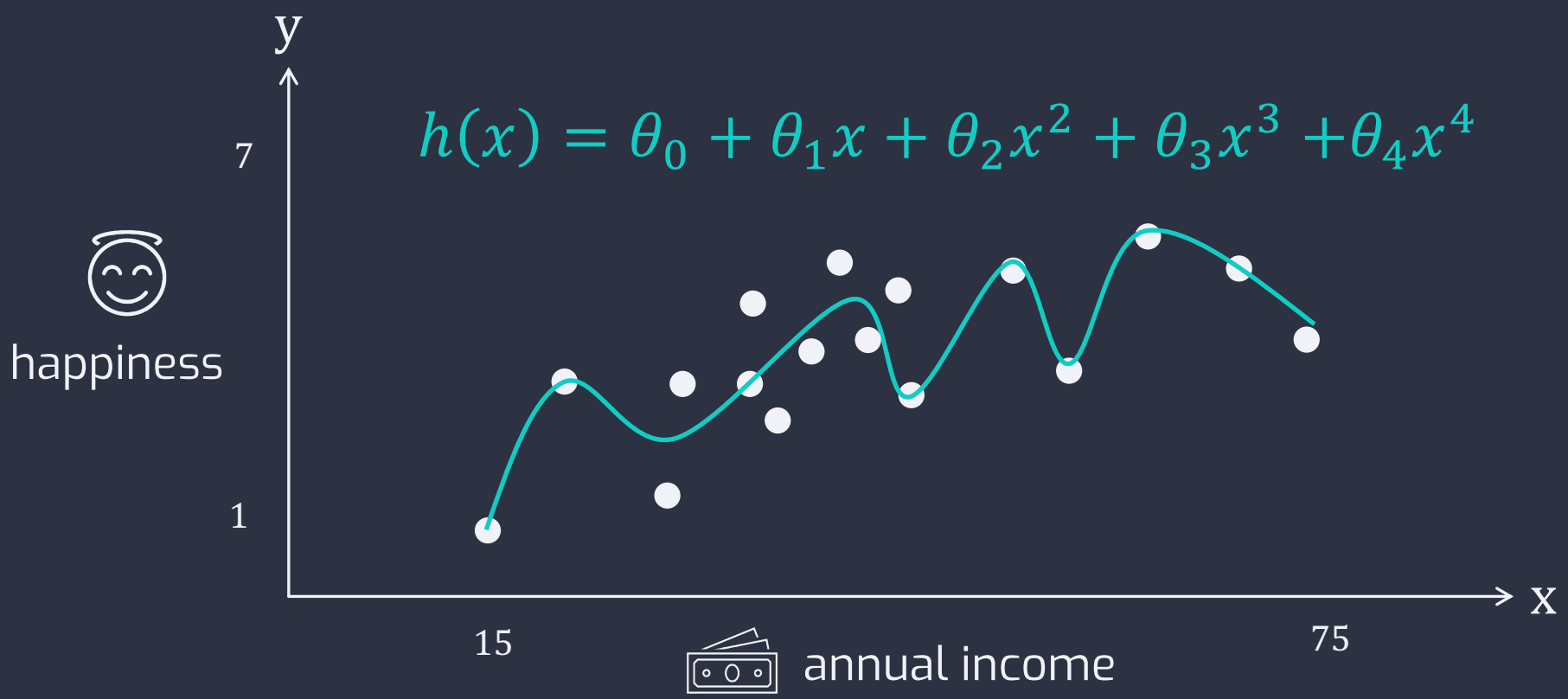
$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})^2 \approx 0$$

Both lead to poor predictions on new data.

Tackling Overfitting

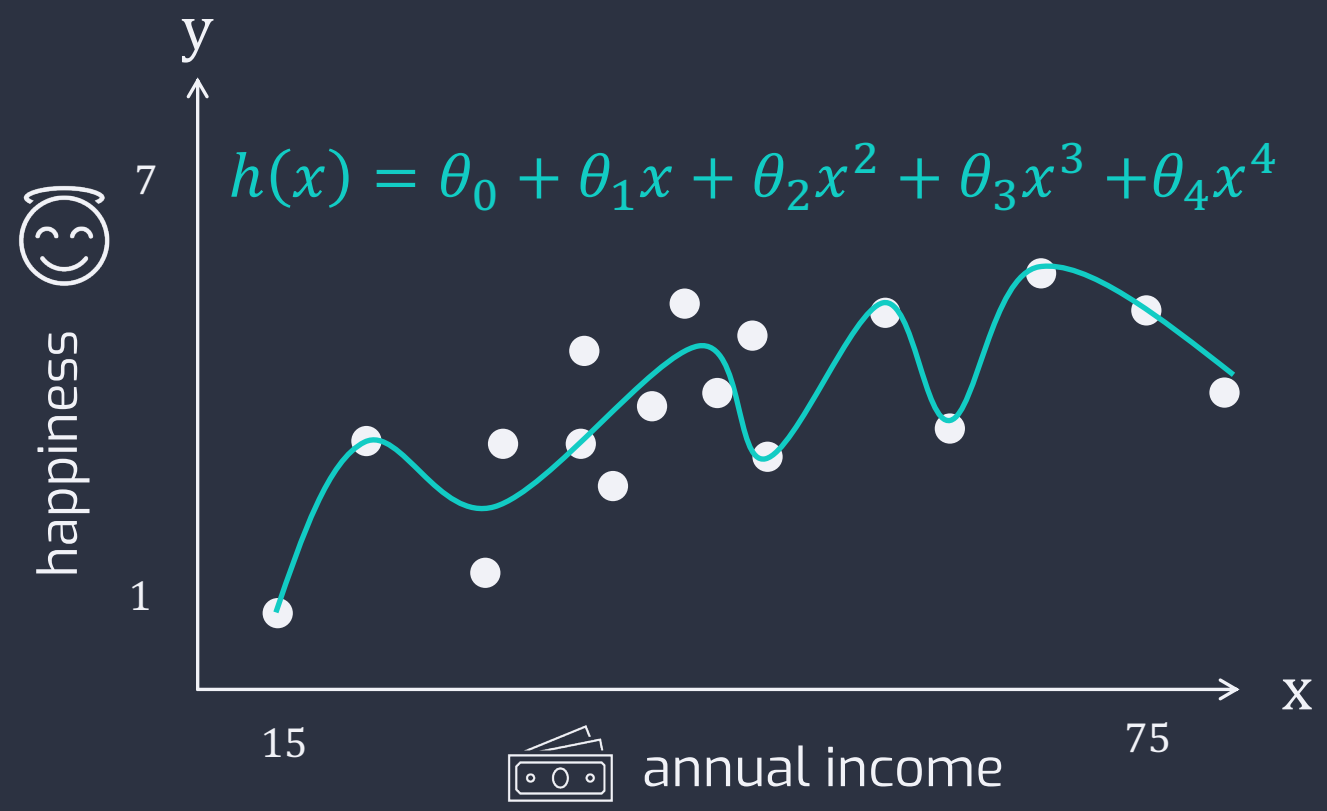
Tackling Overfitting

- plotting and observing



Tackling Overfitting

- plotting and observing



x_1: *annual income*

x_2: *age*

x_3: *number of children*

x_4: *cups of tea per week*

...

x_{50}

Too difficulty with too many features.

Tackling Overfitting

Include more data

- Collect more data
- Data augmentation

Cross-Validation

- K-fold cross validation
- Leave-one-out cross validation

Feature selection / reduction

- Manually
- Feature selection algorithms

Regularisation

- Keep all the features
- Regularise parameters

Tackling Underfitting

Include more features

- Increase complexity
- Relevant & decisive feature

Increase model complexity

- Higher order polynomial
- Linear to non-liner

Reduce regularisation

- Reduce penalty
- Reduce regularisation values

Increase training time

- Keep all the features
- Regularise parameters

Takeaway Points

- Definition of underfitting and overfitting
- Observe to identify underfitting and overfitting
- Methods to tackle underfitting and overfitting