Professional Certificate in Machine Learning and Artificial Intelligence

Office Hours with Pol Cuscó

Date: Thursday, October 13

Hyperparameters

Optimizing through cross-validation

The algorithm will find the optimal network parameters (weights and biases), but the **analyst** has to choose the model **hyperparameters**.

The optimal hyperparameters are found empirically by **cross-validation**.
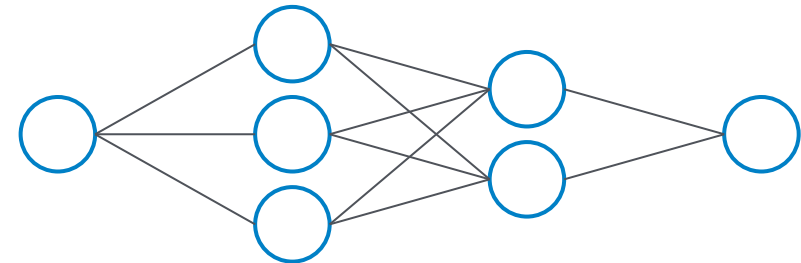
Hyperparameters

Parametrizing the network topology

The specific **number of layers** in the network, and the **number of neurons of each layer**, can be regarded as hyperparameters of the model.

Also, the connection **structure** (fully connected *vs* convolutional, feedforward *vs* recurrent, etc.) is a hyperparameter.
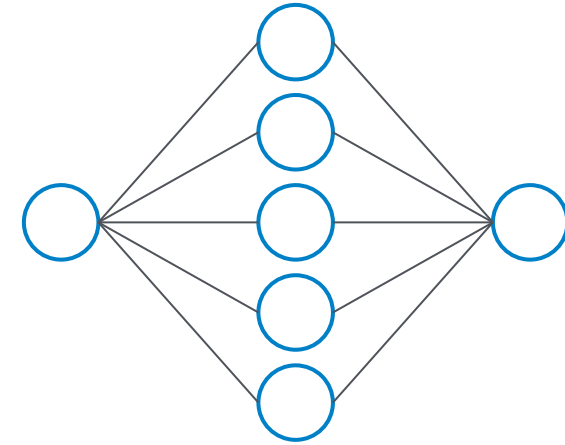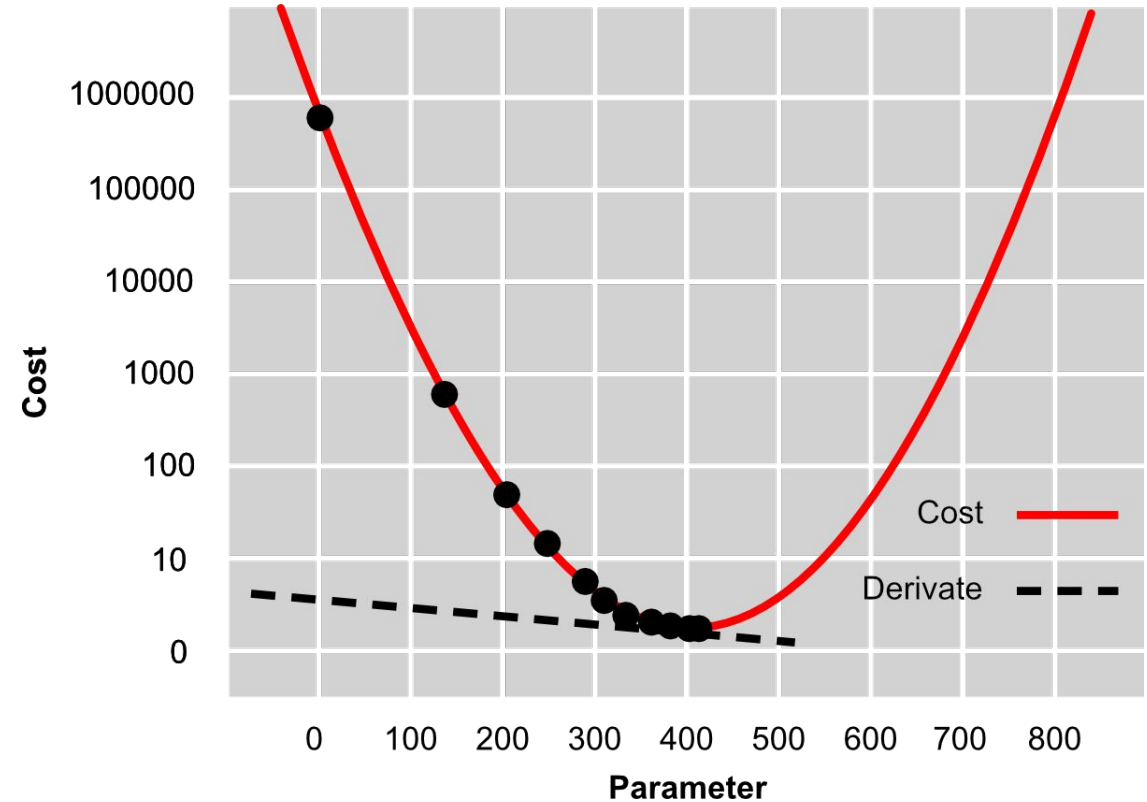

16 parameters

17 parameters

Hyperparameters

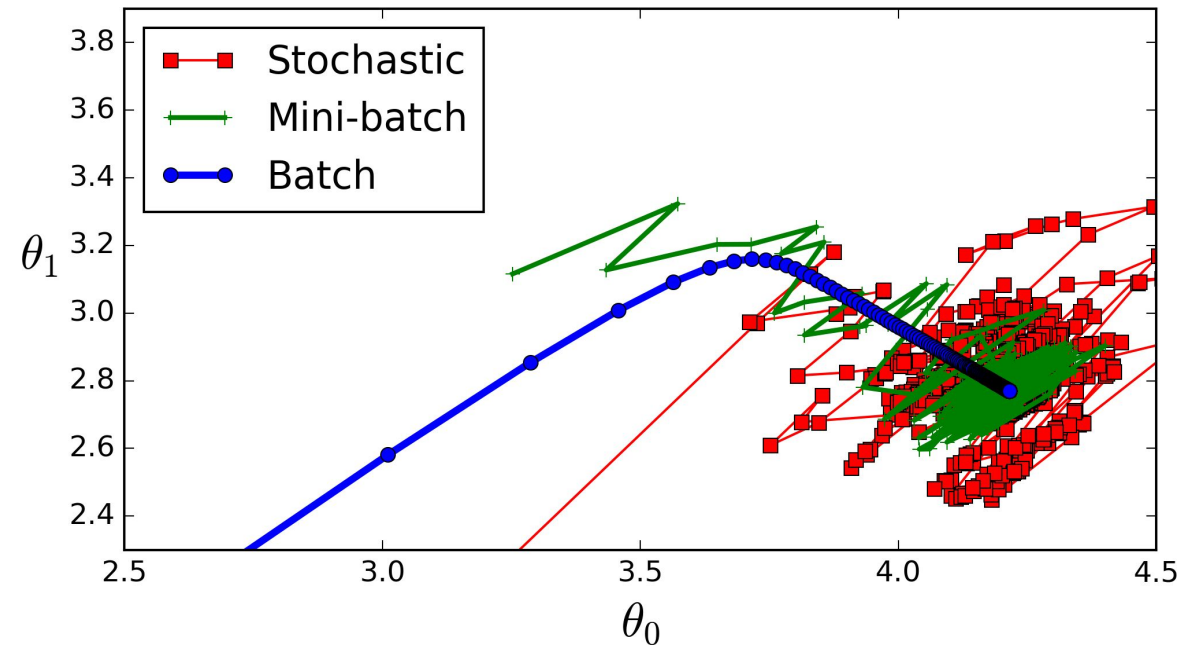## Other deep learning hyperparameters

- Step size
- Batch size
- Dropout rate
- Weight decay
  - L_1
  - L_2, lambdas

Hyperparameters

Other deep learning hyperparameters

- Step size
- Batch size
- Dropout rate
- Weight decay
  - L_1
  - L_2, lambdas

Hyperparameters

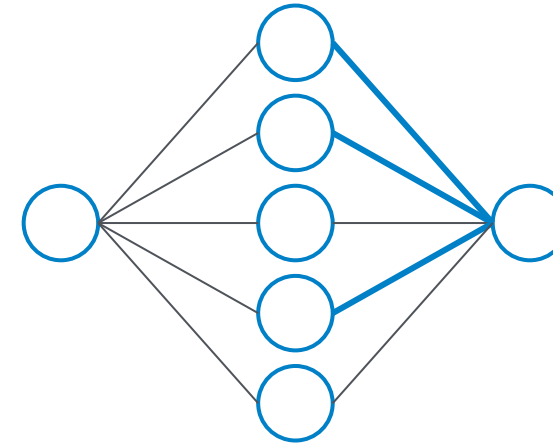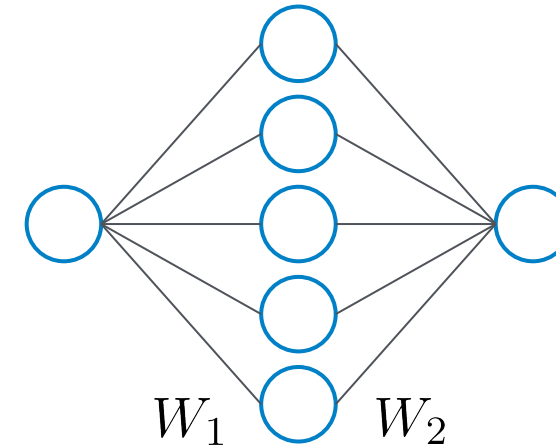**Other deep learning hyperparameters**

- Step size
- Batch size
- Dropout rate
- Weight decay
  - L_1
  - L_2, lambdas

Hyperparameters

Other deep learning hyperparameters

- Step size
- Batch size
- Dropout rate
- Weight decay
  - L_1 regularization
  - L_2 regularization

$$L_1 = \lambda_1 \|W_l\|_1 = \lambda_1 \sum_{i=1}^{n} |w_{il}|$$

$$L_2 = \lambda_2 \|W_l\|_2 = \lambda_2 \sum_{i=1}^{n} w_{il}^2$$

Q&A